MTH 562/662 **Spring 2019**
Abstract Algebra II **Drew Armstrong**

Contents

Week 13

Last semester I used the story of Galois Theory to motivate the study of **abstract groups**. This semester I will use the same story to motivate the study **abstract rings** and **fields**. As before, we will find that **linear algebra** is always hiding just beneath the surface.

To begin, let me refresh your memory. The classical (pre-1830) problem of algebra was to find explicit "formulas" for the roots of a polynomial equation. To be precise, suppose that some rational numbers (called "coefficients") are given:

$$e_1, e_2, \ldots, e_n \in \mathbb{Q}.$$

Then we want to find some numbers $r_1, r_2, \ldots, r_n$ (called "roots") such that

$$x^n - e_1 x^{n-1} + e_2 x^{n-2} - \cdots + (-1)^n e_n = (x - r_1)(x - r_2) \cdots (x - r_n).$$

A priori, it is not obvious what kind of "numbers" the roots should be, or whether they exist at all. Soon we will prove a result called the Fundamental Theorem of Algebra which says that the roots always exist in the field $\mathbb{C}$ of complex numbers. Unfortunately, this theorem will not tell us how to **find** the roots.

We would really like to have some formula or algorithm for computing the roots. To state the problem explicitly, we expand the right hand side and then equate coefficients to obtain a system of n non-linear equations in n unknowns:

$$\left\{ \begin{array}{rcl} e_1 & = & r_1 + r_2 + \cdots + r_n \\ e_2 & = & r_1 r_2 + r_1 r_3 + \cdots + r_{n-1} r_n \\ & \vdots & \\ e_k & = & \sum_{1 \le i_1 < i_2 < \cdots < i_k \le n} r_{i_1} r_{i_2} \cdots r_{i_k} \\ & \vdots & \\ e_n & = & r_1 r_2 \cdots r_n. \end{array} \right.$$

Our goal is to somehow "invert" this system. The best we could hope for is to find some explicit functions $f_1, f_2, \ldots, f_n$ from $\mathbb{Q}^n$ to $\mathbb{C}$ such that

$$\begin{cases} r_1 &= f_1(e_1, e_2, \ldots, e_n) \\ r_2 &= f_2(e_1, e_2, \ldots, e_n) \\ &\vdots \\ r_n &= f_n(e_1, e_2, \ldots, e_n). \end{cases}$$

But this hope is too naive. Indeed, no such functions can exist. To see why, observe that each coefficient e_k can be thought of as a function of the roots:

$$e_k = e_k(r_1, r_2, \ldots, r_n) = \sum_{1 \leq i_1 < i_2 < \cdots < i_k \leq n} r_{i_1} r_{i_2} \cdots r_{i_k}.$$

Furthermore, this function has the nice property of being "symmetric" under permutations of the roots. In other words, if $\sigma \in S_n$ is any permutation of the set $\{1, 2, \ldots, n\}$ then we have

$$e_k(r_1, r_2, \ldots, r_n) = e_k(r_{\sigma(1)}, r_{\sigma(2)}, \ldots, r_{\sigma(n)}).$$

The easiest way to see this is to observe that the product $(x - r_1)(x - r_2) \cdots (x - r_n)$ is symmetric under permutations of its factors. If we expand the product then each coefficient must also be a symmetric function.

[**Jargon.** The coefficients $e_1, \ldots, e_n$ are called the *elementary symmetric functions* of the roots. This explains my use of the letter "e."]

If $f : \mathbb{Q}^n \to \mathbb{C}$ is any function then we can think of the expression $f(e_1, \ldots, e_n)$ as a function of the roots, as follows:

$$f(e_1, \ldots, e_n)(r_1, \ldots, r_n) := f\left(e_1(r_1, \ldots, r_n), \ldots, e_n(r_1, \ldots, r_n)\right).$$

Furthermore, it is clear that $f(e_1, \ldots, e_n)$ is a symmetric function of the roots. Now we see why our first hope was too naive:

> There can be no function f_k such that $r_k = f_k(e_1, e_2, \ldots, e_n)$ because $f_k(e_1, e_2, \ldots, e_n)$ is always a **symmetric** function of the roots, whereas r_k is certainly **not** a symmetric function of the roots.

The solution is to weaken the requirement that f_k is a "function." Instead, we will allow "multi-valued functions"[1] such as square roots. We often talk about "the" square root as though it were a function

$$\sqrt{-} : \mathbb{C} \to \mathbb{C}.$$

But this is **not** a function. Indeed, for any $0 \neq \alpha \in \mathbb{C}$, the expression $\sqrt{\alpha}$ represents **two distinct complex numbers** and there is no natural way to choose between them. We can use this ambiguity to solve the quadratic equation.

[1] Recall that a "multi-valued function" is **not** a function. This is a terrible but common notation.

Example: The Quadratic Formula. For any rational coefficients $e_1, e_2 \in \mathbb{Q}$ we want to find some complex roots $r_1, r_2 \in \mathbb{C}$ such that

$$x^2 - e_1 x + e_2 = (x - r_1)(x - r_2) = x^2 - (r_1 + r_2)x + (r_1 r_2).$$

In other words, we want to find some "multi-valued functions" f_1, f_2 such that

$$\left\{ \begin{array}{rcl} e_1 &=& r_1 + r_2 \\ e_2 &=& r_1 r_2 \end{array} \right. \iff \left\{ \begin{array}{rcl} r_1 &=& f_1(e_1, e_2) \\ r_2 &=& f_2(e_1, e_2). \end{array} \right.$$

As you know, the solution is

$$\left\{ \begin{array}{rcl} f_1(e_1, e_2) &=& (e_1 + \sqrt{e_1^2 - 4e_2})/2 \\ f_2(e_1, e_2) &=& (e_1 - \sqrt{e_1^2 - 4e_2})/2. \end{array} \right.$$

The only subtlety here is that we must interpret the ambiguous expression "$\sqrt{e_1^2 - 4e_2}$" in the **same way** for both equations. In other words, we let "$\sqrt{e_1^2 - 4e_2}$" denote **one particular number** $\alpha \in \mathbb{C}$ such that $\alpha^2 = e_1^2 - 4e_2$. If $e_1^2 - 4e_2 \neq 0$ then there will be two choices and we just pick one at random.[2] ///

The process of choosing a random square root is called "breaking the symmetry." For higher degree equations we expect that we will need to break the symmetry by choosing random 3rd roots, 4th roots, etc. This leads us to the classical problem of algebra.

The Classical Problem of Algebra. Let $e_1, \ldots, e_n \in \mathbb{Q}$ be any rational numbers. By the Fundamental Theorem of Algebra there exist some unique complex numbers $r_1, \ldots, r_n \in \mathbb{C}$ such that

$$x^n - e_1 x^{n-1} + e_2 x^{n-2} - \cdots + (-1)^n e_n = (x - r_1)(x - r_2) \cdots (x - r_n).$$

Our goal is to find some way to compute these roots. Specifically, we want to find an "algebraic formula" expressing the roots in terms of the coefficients, using only the "algebraic operations"

$$+, -, \times, \div, \sqrt{}, \sqrt[3]{}, \sqrt[4]{}, \sqrt[5]{}, \ldots.$$

///

As you know, this problem turns out to be **impossible** when $n \geq 5$. Last semester we developed the group theory necessary for the proof of impossibility.[3] This semester we will fill in the other half of the proof.

[2] If the roots are imaginary then there is a deep sense in which they are indistinguishable. Indeed, we usually define the imaginary unit i as "the" square root of -1. But if -1 has any square root then it must have two. Which one do you want to call i?

[3] Specifically, we proved that the group S_n not "solvable" when $n \geq 5$.

The Classical Problem was definitively solved by Galois in the 1820s. However, he died too soon to really explain it to anyone. Galois' work was eventually published in 1846 by Joseph Liouville. The first textbook on Galois theory was Camille Jordan's *Traité des substitutions et des équations algébraiques* (1870). At this point "Galois theory" and "group theory" were the same subject, so Jordan's work can also be viewed as the first book about groups.

However, the subject was still difficult to understand. The next major advance was made by Richard Dedekind in the 1880s when he defined the concept of a *field*.[4]

Definition of Fields and Subfields/Extensions. A *field* is a set $\mathbb{F}$ together with two binary operations

$$+, \times : \mathbb{F} \times \mathbb{F} \to \mathbb{F}$$

(called *addition* and *multiplication*) and two special elements

$$0, 1 \in \mathbb{F}$$

(called *zero* and *one*), which satisfy the following three axioms:

(F1) $(\mathbb{F}, +, 0)$ is an abelian group.

(F2) $(\mathbb{F} - \{0\}, \times, 1)$ is an abelian group. We will use juxtaposition to denote multiplication:

$$ab := a \times b.$$

Furthermore, since multiplication is commutative we are free to use fractional notation to denote division:

$$ab^{-1} = b^{-1}a = \frac{a}{b}.$$

(F3) *Distribution.* For all $a, b, c \in \mathbb{F}$ we have

$$a(b + c) = ab + ac.$$

Now let $S \subseteq \mathbb{F}$ be any subset. We say that S is a *subfield* of $\mathbb{F}$ (equivalently, $\mathbb{F}$ is a *field extension* of S) if the following properties are satisfied:

- The special elements $0, 1$ are in S.
- For all $a, b \in S$ we have $a \pm b \in S$.

[4]Dedekind's name for this structure was *Körper*, short for *Zahlkörper* [body of numbers]. Dedekind's rival Leopold Kronecker used the term *Rationalitätsbereich* [domain of rationality]. The English term *field* was coined by E. H. Moore in 1893, possibly motivated by the word "domain." This creates a problem for English speakers: should we denote fields by the letter K or the letter F? To avoid confusion I will use the blackboard bold font (i.e., $\mathbb{K}$ or $\mathbb{F}$) to denote fields. Sadly, this is not a perfect solution because $\mathbb{Z}$ and $\mathbb{N}$ are **not** fields.

- For all $a, b \in S$ we have $ab \in S$.
- For all $a \in S - \{0\}$ we have $a^{-1} \in S$.

In other words: A subfield is a subset that is also a field with respect to the same operations and special elements. [Remark: We could shorten this definition by using the word "subgroup" in various places. Unfortunately, the subfield test cannot be reduced to one step, as the subgroup test can.] ///

The most basic examples of fields are

$$\mathbb{Q}, \mathbb{R}, \mathbb{C} \quad \text{and} \quad \mathbb{F}_p := \mathbb{Z}/p\mathbb{Z} \text{ for } p \text{ prime.}$$

Note that the inclusions $\mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}$ are field extensions. Prior to Dedekind no one felt the need to define the abstract concept of fields because it was synonymous with the concept of "numbers." However, Dedekind found that the abstract concept was helpful to simplify various ideas in number theory and Galois theory. Here is the first non-basic example.

The First Interesting Example. Let $\alpha = \sqrt{2}$ be any real number satisfying $\alpha^2 = 2$. If you want you want you can think of α as the **positive** square root of 2, but it doesn't really matter. Now consider the set

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

I claim that $\mathbb{Q}(\sqrt{2}) \subseteq \mathbb{R}$ is a subfield.

Proof. What needs to be checked?

- **Special Elements.** Note that $0 = 0 + 0\sqrt{2} \in \mathbb{Q}(\sqrt{2})$ and $1 = 1 + 0\sqrt{2} \in \mathbb{Q}(\sqrt{2})$.
- **Addition/Subtraction.** For all $a + b\sqrt{2}$ and $c + d\sqrt{2}$ in $\mathbb{Q}(\sqrt{2})$ note that

  $$(a + b\sqrt{2}) - (c + d\sqrt{2}) = (a - c) + (b - d)\sqrt{2} \in \mathbb{Q}(\sqrt{2})$$

  because $a - c \in \mathbb{Q}$ and $b - d \in \mathbb{Q}$.
- **Multiplication.** For all $a + b\sqrt{2}$ and $c + d\sqrt{2}$ in $\mathbb{Q}(\sqrt{2})$ we have

  $$(a + b\sqrt{2})(c + d\sqrt{2}) = (ac + 2bd) + (ad + bc)\sqrt{2} \in \mathbb{Q}(\sqrt{2})$$

  because $ac + 2bd \in \mathbb{Q}$ and $ad + bc \in \mathbb{Q}$.
- **Division.** This is the hardest step. For all nonzero elements $a + b\sqrt{2} \in \mathbb{Q}(\sqrt{2})$ we want to show that there exists some element $c + d\sqrt{2} \in \mathbb{Q}(\sqrt{2})$ such that

  $$(a + b\sqrt{2})(c + d\sqrt{2}) = 1 = 1 + 0\sqrt{2}.$$

  The solution is a trick called "rationalizing the denominator." First note that

  $$(a + b\sqrt{2})(a - b\sqrt{2}) = (a^2 - 2b^2) + (ab - ab)\sqrt{2} = (a^2 - 2b^2) + 0\sqrt{2} \in \mathbb{Q}.$$

Now assume that $a + b\sqrt{2} \neq 0$ (i.e., assume that a and b are not both zero). We are looking for rational numbers $c, d \in \mathbb{Q}$ such that

$$\begin{aligned} c + d\sqrt{2} &= \frac{1}{a + b\sqrt{2}} \\ &= \frac{1}{a + b\sqrt{2}} \cdot \frac{a - b\sqrt{2}}{a - b\sqrt{2}} \\ &= \frac{a - b\sqrt{2}}{a^2 - 2b^2} \\ &= \left(\frac{a}{a^2 - 2b^2}\right) + \left(\frac{-b}{a^2 - 2b^2}\right)\sqrt{2}. \end{aligned}$$

If $a^2 - 2b^2 \neq 0$ then we can take

$$c = \frac{a}{a^2 - 2b^2} \in \mathbb{Q} \qquad \text{and} \qquad d = \frac{-b}{a^2 - 2b^2} \in \mathbb{Q}.$$

So assume for contradiction that $a^2 - 2b^2 = 0$. If $b = 0$ then $a^2 = 2b^2 = 0$ implies $a = 0$, which contradicts the fact that a and b are not both zero. If $b \neq 0$ then we have

$$\begin{aligned} a^2 &= 2b^2 \\ a^2/b^2 &= 2 \\ (a/b)^2 &= 2. \end{aligned}$$

Since $a/b \in \mathbb{Q}$ this contradicts the well-known fact that $\pm\sqrt{2} \notin \mathbb{Q}$.[5]

□

[**Jargon.** The field $\mathbb{Q}(\sqrt{2})$ is called $\mathbb{Q}$ *adjoin* $\sqrt{2}$. We will see a generalization of this construction below.]

As with subgroups, It follows immediately from the definition that the intersection of subfields is a subfield.

Intersection of Subfields is a Subfield. Let $(\mathbb{F}, +, \times, 0, 1)$ be a field and let $\mathbb{K}_i \subseteq \mathbb{F}$ be any family of subfields (possibly infinite or even uncountable). Then the intersection

$$\bigcap_i \mathbb{K}_i \subseteq \mathbb{F}$$

[5] I'll put a proof of this on the first homework to remind you.

is also a subfield.

Proof. Since $0, 1 \in \mathbb{K}_i$ for all i we have $0, 1 \in \cap_i \mathbb{K}_i$. Now consider any two elements $a, b \in \cap_i \mathbb{K}_i$ with $a \neq 0$. By definition this means that $a, b \in \mathbb{K}_i$ for each i. But then since $\mathbb{K}_i \subseteq \mathbb{F}$ is a subfield we know that $a \pm b$, ab and a^{-1} are in $\mathbb{K}_i$. It follows that $a \pm b$, ab and a^{-1} are also in the intersection. $\square$

However, the union of subfields is not necessarily a subfield.

Union of Subfields is Not a Subfield. Consider the subfields $\mathbb{Q}(\sqrt{2}) \subseteq \mathbb{R}$ and $\mathbb{Q}(\sqrt{3}) \subseteq \mathbb{R}$.[6] I claim that the union $\mathbb{Q}(\sqrt{2}) \cup \mathbb{Q}(\sqrt{3})$ is **not** a subfield of $\mathbb{R}$.

Proof. Suppose for contradiction that $\mathbb{Q}(\sqrt{2}) \cup \mathbb{Q}(\sqrt{3}) \subseteq \mathbb{R}$ **is** a subfield. Since a subfield is closed under addition, we must have $\sqrt{2} + \sqrt{3} \in \mathbb{Q}(\sqrt{2}) \cup \mathbb{Q}(\sqrt{3})$, which implies that $\sqrt{2} + \sqrt{3}$ is in $\mathbb{Q}(\sqrt{2})$ or in $\mathbb{Q}(\sqrt{3})$. Let's assume that $\sqrt{2} + \sqrt{3} \in \mathbb{Q}(\sqrt{2})$. Then by definition we have

$$\sqrt{2} + \sqrt{3} = a + b\sqrt{2} \quad \text{for some } a, b \in \mathbb{Q}.$$

If $b = 1$ then we obtain the contradiction that $\sqrt{3} \in \mathbb{Q}$:

$$\sqrt{2} + \sqrt{3} = a + \sqrt{2} \quad \Longrightarrow \quad \sqrt{3} = a \in \mathbb{Q}.$$

Furthermore, if $a = 0$ then we obtain the contradiction that $\sqrt{6} \in \mathbb{Q}$:

$$\begin{aligned} \sqrt{2} + \sqrt{3} &= b\sqrt{2} \\ \sqrt{3} &= (b-1)\sqrt{2} \\ \sqrt{3} \cdot \sqrt{2} &= (b-1)\sqrt{2} \cdot \sqrt{2} \\ \sqrt{6} &= 2(b-1) \in \mathbb{Q}. \end{aligned}$$

But in all other cases we obtain the contradiction that $\sqrt{2} \in \mathbb{Q}$:

$$\begin{aligned} \sqrt{2} + \sqrt{3} &= a + b\sqrt{2} \\ \sqrt{3} &= a + (b-1)\sqrt{2} \\ 3 &= \left(a + (b-1)\sqrt{2}\right)^2 \\ 3 &= \left(a^2 + 2(b-1)^2\right) + 2a(b-1)\sqrt{2} \\ \sqrt{2} &= \frac{3 - \left(a^2 + 2(b-1)^2\right)}{2a(b-1)} \in \mathbb{Q}. \end{aligned}$$

In conclusion we have $\sqrt{2} + \sqrt{3} \notin \mathbb{Q}(\sqrt{2})$. A similar proof shows that $\sqrt{2} + \sqrt{3} \notin \mathbb{Q}(\sqrt{3})$. $\square$

Remarks:

[6] The proof that $\mathbb{Q}(\sqrt{3}) \subseteq \mathbb{R}$ is a subfield is exactly the same as the proof for $\mathbb{Q}(\sqrt{2}) \subseteq \mathbb{R}$.

- In the proof we needed the fact that $\sqrt{2}, \sqrt{3}, \sqrt{6} \in \mathbb{R}$ are *irrational* numbers. On the homework you will prove for all integers $D \in \mathbb{Z}$ that

$$\pm\sqrt{D} \notin \mathbb{Z} \quad \Longrightarrow \quad \pm\sqrt{D} \notin \mathbb{Q}.$$

- There is a more sophisticated way to phrase the theorem we just proved. One can view the real numbers $(\mathbb{R}, +, 0)$ as a *vector space* over $\mathbb{Q}$ in a very boring way. That is, for every "scalar" $a \in \mathbb{Q}$ and for every "vector" $b \in \mathbb{R}$ we define "scalar multiplication" via regular multiplication:

$$a(b) := ab \in \mathbb{R}.$$

  Then the vector space axioms are easily verified.[7] So what? In this language we can rephrase the above theorem by saying that

$$\text{the real numbers } 1, \sqrt{2}, \sqrt{3} \in \mathbb{R} \text{ are } \textit{linearly independent} \text{ over } \mathbb{Q}.$$

  In fact, it is true that any list of square roots of square-free integers is linearly independent over $\mathbb{Q}$, but this is quite tricky to prove. It seems that most algebra books pass over this fact without comment.

- More generally, if $\mathbb{E} \supseteq \mathbb{F}$ is any field extension then we can view $\mathbb{E}$ as a vector space over $\mathbb{F}$ in the same boring way. For this reason it turns out that linear algebra is very useful for the study of fields. In particular, we are interested in the **dimension**:

$$[\mathbb{E}/\mathbb{F}] := \dim_{\mathbb{F}}(\mathbb{E}) = \text{ the dimension of } \mathbb{E} \text{ as a vector space over } \mathbb{F}.$$

  On the homework you will verify that $[\mathbb{Q}(\sqrt{2})/\mathbb{Q}] = 2$. With more work (for example, by using the tricky theorem about square roots stated above) one can prove that $[\mathbb{R}/\mathbb{Q}] = \infty$, which is bad. In this course we prefer to study finite-dimensional field extensions.

///

As with subgroups, we should replace the union of subfields with the smallest subfield that contains the union. Here is the general construction.

The Subfield Generated by a Subset. Let $(\mathbb{F}, +, \times, 0, 1)$ be a field and let $S \subseteq \mathbb{F}$ be any subset. Let $\langle S \rangle \subseteq \mathbb{F}$ denote the intersection of all subfields $\mathbb{K} \subseteq \mathbb{F}$ that contain S:

$$\langle S \rangle := \bigcap_{S \subseteq \mathbb{K} \subseteq \mathbb{F}} \mathbb{K}.$$

We know from above that $\langle S \rangle \subseteq \mathbb{F}$ is a subfield. I claim that it is the **smallest** subfield of $\mathbb{F}$ that contains S. We call it the *subfield of* $\mathbb{F}$ *generated by* S.

Proof. The intersection is contained in any field that contains S. □

[7] This is a good time to remind yourself of the vector space axioms.

Actually, the notation $\langle S \rangle$ is not standard in field theory and we will only use it temporarily. The more common notation refers to the smallest subfield $\mathbb{F}' := \langle \emptyset \rangle \subseteq \mathbb{F}$, which is called the *prime subfield*. The reason for the notation is because the prime subfield of any field satisfies

$$\mathbb{F}' \cong \mathbb{Q} \qquad \text{or} \qquad \mathbb{F}' \cong \mathbb{Z}/p\mathbb{Z} \text{ for some prime } p.$$

You will prove this on a future homework, after we develop the necessary technology.

Here is the more standard notation for a subfield generated by a set.

The Definition of Adjunction. For any field extension $\mathbb{F} \subseteq \mathbb{E}$ and for any subset $S \subseteq \mathbb{E}$ we let $\mathbb{F} \subseteq \mathbb{F}(S) \subseteq \mathbb{E}$ be the intersection of all subfields that contain $\mathbb{F} \cup S$:

$$\mathbb{F}(S) := \langle \mathbb{F} \cup S \rangle = \bigcap_{(\mathbb{F} \cup S) \subseteq \mathbb{K} \subseteq \mathbb{E}} \mathbb{K}.$$

We call this field "$\mathbb{F}$ *adjoin* S." If we omit any mention of the base field $\mathbb{F}$ then we obtain

$$\langle S \rangle = \mathbb{E}'(S),$$

where $\mathbb{E}' \subseteq \mathbb{E}$ is the prime subfield. This explains the relationship between the standard and nonstandard terminology. In the case that S is a finite set we will write

$$\mathbb{F}(\{\alpha_1, \alpha_2, \ldots, \alpha_k\}) = \mathbb{F}(\alpha_1, \alpha_2, \ldots, \alpha_k).$$

This is the smallest field between $\mathbb{E}$ and $\mathbb{F}$ that contains the elements $\alpha_1, \ldots, \alpha_k$. ///

I'll ask you to verify some formal (i.e., "trivial") properties of adjunction on the homework. For example, you will verify that $\mathbb{F}(\alpha)(\beta) = \mathbb{F}(\beta)(\alpha) = \mathbb{F}(\alpha, \beta)$ for any $\alpha, \beta \in \mathbb{E}$.

Week 14

We saw last week that the union of two subfields is not necessarily a subfield. Instead, we will replace the union with a new operation.

The Lattice of Subfields. Let $\mathbb{E} \supseteq \mathbb{F}$ be a field extension and consider the set

$$\mathscr{L}(\mathbb{E}, \mathbb{F}) = \{ \text{ all subfields between } \mathbb{E} \text{ and } \mathbb{F} \ \} = \{\mathbb{K} : \mathbb{F} \subseteq \mathbb{K} \subseteq \mathbb{E}\}.$$

This set is partially ordered by containment, with bottom element $\mathbb{F}$ and top element $\mathbb{E}$. We have seen that any two intermediate fields $\mathbb{K}_1, \mathbb{K}_2 \in \mathscr{L}(\mathbb{E}, \mathbb{F})$ have a greatest lower bound ("meet") given by the intersection:

$$\mathbb{K}_1 \wedge \mathbb{K}_2 = \mathbb{K}_1 \cap \mathbb{K}_2.$$

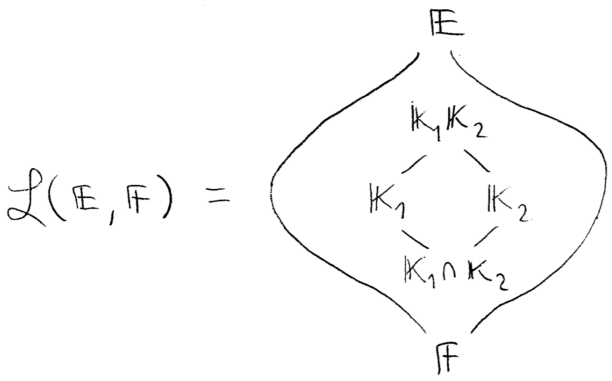
Furthermore, any two fields have a least upper bound ("join") which is defined as the intersection of all subfields that contain the union:

$$\mathbb{K}_1 \vee \mathbb{K}_2 := \langle \mathbb{K}_1 \cup \mathbb{K}_2 \rangle = \mathbb{K}_1(\mathbb{K}_2) = \mathbb{K}_2(\mathbb{K}_1).$$

More commonly this operation is called the *compositum* of subfields:

$$\mathbb{K}_1\mathbb{K}_2 := \mathbb{K}_1 \vee \mathbb{K}_2.$$

With these operations, the set $\mathscr{L}(\mathbb{E}, \mathbb{F})$ is called the *lattice of intermediate fields*. Here is the picture that I have in my mind:



FIX THIS PICTURE

If the base field $\mathbb{F}$ is not specified then we will write

$$\mathscr{L}(\mathbb{E}) = \mathscr{L}(\mathbb{E}, \mathbb{E}'),$$

where $\mathbb{E}' \subseteq \mathbb{E}$ is the prime subfield. This is called the *lattice of all subfields* of $\mathbb{E}$. ///

My goal for the rest of this week is to tell you Galois' Solvability Theorem in its modern form. I will also state the so-called Fundamental Theorem of Galois Theory, which is not due to Galois. It will take the rest of the semester to fill in the proofs of these theorems.

The main innovation of Galois was to associate a **group** to each polynomial equation $f(x) = 0$. If the coefficients of $f(x)$ lie in a field $\mathbb{F}$ then we will denote this group by $\mathrm{Gal}(f/\mathbb{F})$ and we will call it the *Galois group of f over $\mathbb{F}$*. Galois' original definition was a bit technical.

Galois' Definition of the Galois Group. The group $\mathrm{Gal}(f/\mathbb{F})$ is a certain subgroup of the group of permutations of the roots of $f(x)$. ///

It would take quite a few pages to tell you what "certain subgroup" means. Instead I will present the modern definition which is due to Dedekind. His main innovation was to translate the discussion of polynomials into the language of **fields**.

Dedekind's Definition of the Galois Group. Let $f(x)$ be a polynomial with coefficients in a field $\mathbb{F}$. There exists a certain "smallest" field extension $\mathbb{E} \supseteq \mathbb{F}$ (called the *splitting field*) in which $\mathbb{F}$ has all of its roots. For example, if $\mathbb{F} \subseteq \mathbb{C}$ then the FTA says that all the roots exist in $\mathbb{C}$, so $\mathbb{E}$ is just the intersection of all subfields that contain the roots. Then we define

$$\mathrm{Gal}(f/\mathbb{F}) := \{ \text{ field automorphisms } \sigma : \mathbb{E} \to \mathbb{E} \text{ such that } \sigma(a) = a \text{ for all } a \in \mathbb{F} \ \}.$$

By a *field automorphism* we mean any invertible function $\sigma : \mathbb{E} \to \mathbb{E}$ that satisfies

$$\sigma(a+b) = \sigma(a) + \sigma(b) \qquad \text{and} \qquad \sigma(ab) = \sigma(a)\sigma(b)$$

for all $a, b \in \mathbb{E}$. [Remark: You will prove on the homework that the invertibility hypothesis is redundant. That is, you will show that any function $\sigma : \mathbb{E} \to \mathbb{E}$ that fixes $\mathbb{F}$ and preserves addition and multiplication is **necessarily** invertible.] Cleary the collection of such functions is a group under composition. Since the definition doesn't refer to the polynomial $f(x)$ we will also use the notation

$$\mathrm{Gal}(\mathbb{E}/\mathbb{F}) := \mathrm{Gal}(f/\mathbb{F}).$$

///

And what do field automorphisms have to do with permutations of the roots? Suppose that $\alpha \in \mathbb{E}$ is a root of $f(x)$ and let $\sigma \in \mathrm{Gal}(\mathbb{E}/\mathbb{F})$ be any element of the Galois group. I claim that $\sigma(\alpha) \in \mathbb{E}$ is also a root of $f(x)$.

Proof. Since $f(x)$ has coefficients in $\mathbb{F}$ we can write

$$f(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n \qquad \text{for some } a_0, a_1, \ldots, a_n \in \mathbb{F}.$$

And since $\alpha \in \mathbb{E}$ is a root of $f(x)$ we have

$$f(\alpha) = a_0 + a_1\alpha + a_2\alpha^2 + \cdots + a_n\alpha^n = 0.$$

Now we apply the field automorphism σ to both sides of this equation and use the fact that $\sigma(a) = a$ for all $a \in \mathbb{F}$ to obtain

$$\begin{aligned}
\sigma(f(\alpha)) &= \sigma(0) \\
\sigma(a_0 + a_1\alpha + a_2\alpha^2 + \cdots + a_n\alpha^n) &= \sigma(0) \\
\sigma(a_0) + \sigma(a_1)\sigma(\alpha) + \sigma(a_2)\sigma(\alpha)^2 + \cdots + \sigma(a_n)\sigma(\alpha)^n &= \sigma(0) \\
a_0 + a_1\sigma(\alpha) + a_2\sigma(\alpha)^2 + \cdots + a_n\sigma(\alpha)^n &= 0 \\
f(\sigma(\alpha)) &= 0.
\end{aligned}$$

In other words, $\sigma(\alpha) \in \mathbb{E}$ is a root of $f(x)$. □

It follows that every element $\sigma \in \mathrm{Gal}(\mathbb{E}/\mathbb{F})$ restricts to a permutation of the roots of $f(x)$. Furthermore, it seems reasonable that since $\mathbb{E}$ is the "smallest" field containing the roots

then two different field automorphisms should restrict to two different permutations. Thus we obtain an **injective group homomorphism**:

$$\mathrm{Gal}(f/\mathbb{F}) \to \{ \text{ permutations of the roots of } f(x) \}.$$

I don't plan to get more specific than this because I am really more interested in Dedekind's version of the theory. If you want to see all the details of Galois' version I recommend Jean-Pierre Tignol's book *Galois' Theory of Algebraic Equations* (2001).

Before stating the Fundamental Theorem I want to show you a couple of basic examples.

Example: The Galois Group of $x^2 - 2$.

Consider the polynomial $x^2 - 2$ with coefficients in $\mathbb{Q}$. If $\sqrt{2} \in \mathbb{R}$ is the positive real square root of 2 then we know that this polynomial has exactly two roots: $+\sqrt{2}$ and $-\sqrt{2}$.[8] I claim that the splitting field $\mathbb{E} \supseteq \mathbb{Q}$ is the same field that we studied above:

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

Proof. Certainly we know that the field $\mathbb{Q}(\sqrt{2})$ contains the roots $+\sqrt{2}$ and $-\sqrt{2}$. Now let $\mathbb{K} \supseteq \mathbb{Q}$ be any field extension that contains these roots and consider any two rational numbers $a, b \in \mathbb{Q}$. Then since $\mathbb{K}$ is closed under addition and multiplication we have

$$a, b, \sqrt{2} \in \mathbb{K} \qquad \Longrightarrow \qquad a + b\sqrt{2} \in \mathbb{K},$$

and it follows that $\mathbb{Q}(\sqrt{2})$. □

Now let $\sigma \in \mathrm{Gal}(\mathbb{Q}(\sqrt{2})/\mathbb{Q})$ be an element of the Galois group. By definition this means that $\sigma : \mathbb{Q}(\sqrt{2}) \to \mathbb{Q}(\sqrt{2})$ is a field automorphism that fixes elements of $\mathbb{Q}$.[9] This means that for all $a + b\sqrt{2} \in \mathbb{Q}(\sqrt{2})$ we have

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

hence the automorphism σ is uniquely specified by the value of $\sigma(\sqrt{2})$. What are the options? Since $\sqrt{2}$ is a root of the polynomial $x^2 - 2$ we must have

$$(\sqrt{2})^2 - 2 = 0$$

[8] Wait, why do we know this? You will prove on the next homework that a polyomial of degree n can have at most n roots in any field extension. It may have more roots in other kinds of ring extensions. For example, the polynomial $x^2 - 2$ has **uncountably many** roots in the ring of quaternions $\mathbb{H} \supseteq \mathbb{Q}$.

[9] Actually, the requirement that σ fixes $\mathbb{Q}$ is redundant here because $\mathbb{Q}$ is the prime subfield of $\mathbb{Q}(\sqrt{2})$.

$$\sigma\left((\sqrt{2})^2 - 2\right) = \sigma(0)$$
$$\sigma(\sqrt{2})^2 - \sigma(2) = \sigma(0)$$
$$\sigma(\sqrt{2})^2 - 2 = 0,$$

and it follows that $\sigma(\sqrt{2}) \in \{\pm\sqrt{2}\}$. Therefore σ must be one of the following two functions:

$$\mathrm{id}(a + b\sqrt{2}) = a + b\sqrt{2},$$
$$\tau(a + b\sqrt{2}) = a - b\sqrt{2}.$$

The only remaining question is whether these two functions are indeed **field automorphisms**. Well, the identity clearly is, but it needs to be checked by hand that τ preserves addition and multiplication. You will do this on the homework.

In summary, we have found that the Galois group of the equation $x^2 - 2 = 0$ is the group of size 2 generated by the "conjugation automorphism" $\tau : \mathbb{Q}(\sqrt{2}) \to \mathbb{Q}(\sqrt{2})$:

$$\mathrm{Gal}((x^2 - 2)/\mathbb{Q}) = \mathrm{Gal}(\mathbb{Q}(\sqrt{2})/\mathbb{Q}) = \{\mathrm{id}, \tau\} \cong \mathbb{Z}/2\mathbb{Z}.$$

On the homework you will show that essentially the same results hold for any so-called "quadratic field extension" $\mathbb{F}(\sqrt{D}) \supseteq \mathbb{F}$. Another example is the complex field over the real field, which is the splitting field of the polynomial $x^2 + 1$. In this case we have

$$\mathrm{Gal}((x^2 + 1)/\mathbb{R}) = \mathrm{Gal}(\mathbb{C}/\mathbb{R}) = \{\mathrm{id}, \tau\} \cong \mathbb{Z}/2\mathbb{Z},$$

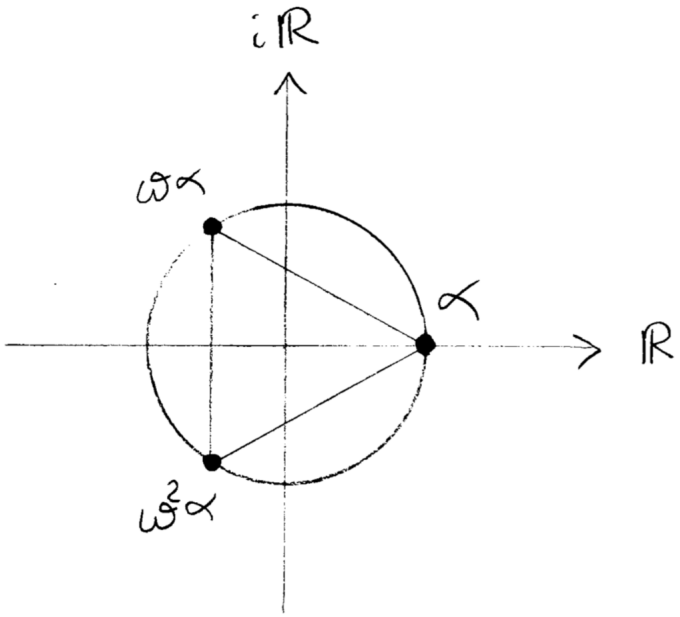
where the function $\tau : \mathbb{C} \to \mathbb{C}$ defined by

$$\tau(a + b\sqrt{-1}) := a - b\sqrt{-1}$$

is called "complex conjugation." ///

Example: The Galois Group of $x^3 - 2$.

Consider the polynomial $x^3 - 2$ with coefficients in $\mathbb{Q}$. We know that this polynomial has one real root and two complex roots. To be specific, let $\alpha := \sqrt[3]{2} \in \mathbb{R}$ be the real 3rd root of 2 and let $\omega := \exp(2\pi i/3)$ be a primitive 3rd root of 1. Then the roots $\alpha, \omega\alpha, \omega^2\alpha$ are the verticies of an equilateral triangle in the complex plane:



I claim that the splitting field $\mathbb{E} \supseteq \mathbb{Q}$ is obtained by adjoining the set $\{\alpha, \omega\}$:

$$\mathbb{E} = \mathbb{Q}(\alpha, \omega) = \text{ the smallest subfield of } \mathbb{C} \text{ that contains } \mathbb{Q} \cup \{\alpha, \omega\}.$$

Proof. Since the field $\mathbb{Q}(\alpha, \omega)$ contains the elements α, ω and is closed under multiplication, it must contain the roots $\alpha, \omega\alpha, \omega^2\alpha$. Now let $\mathbb{C} \supseteq \mathbb{K} \supseteq \mathbb{Q}$ be any field that contains the roots. Then since $\mathbb{K}$ is closed under inversion we must have

$$\alpha, \omega\alpha \in \mathbb{K} \qquad \Longrightarrow \qquad \omega = (\omega\alpha)(\alpha^{-1}) \in \mathbb{K}.$$

It follows that $\mathbb{Q} \cup \{\alpha, \omega\} \subseteq \mathbb{K}$ and hence $\mathbb{Q}(\alpha, \omega) \subseteq \mathbb{K}$. □

Unlike the previous example, we do not already know a basis for the vector space $\mathbb{Q}(\alpha, \omega)/\mathbb{Q}$ and this makes it harder to compute the Galois group. So let me just tell you without proof that the set $\{1, \alpha, \alpha^2, \omega, \omega\alpha, \omega\alpha^2\}$ is a basis. In other words, every element of the splitting field $\gamma \in \mathbb{Q}(\alpha, \omega)$ can be written in the form

$$\gamma = a + b\alpha + c\alpha^2 + d\omega + e\omega\alpha + f\omega\alpha^2 \qquad \text{for some } \textbf{unique } a, b, c, d, e, f \in \mathbb{Q}.$$

If $\sigma \in \mathrm{Gal}(\mathbb{Q}(\alpha, \omega)/\mathbb{Q})$ is any element of the Galois group then we find that

$$\sigma(\gamma) = a + b\sigma(\alpha)a + c\sigma(\alpha)^2 + d\sigma(\omega) + e\sigma(\omega)\sigma(\alpha) + f\sigma(\omega)\sigma(\alpha)^2,$$

and it follows that σ is uniquely specified by the values $\sigma(\alpha)$ and $\sigma(\omega)$. What are the options? Since $\alpha^3 - 2 = 0$ and $\omega^3 - 1 = 0$ we find that

$$\sigma(\alpha)^3 - 2 = 0 \qquad \text{and} \qquad \sigma(\omega)^3 - 1 = 0.$$

Furthermore, since σ is invertible with $\sigma(1) = 1$ and $\omega \neq 1$, we know that $\sigma(\omega) \neq 1$. It follows that there are at most $6 = 3 \cdot 2$ possibilities:

$$\sigma(\alpha) \in \{\alpha, \omega\alpha, \omega^2\alpha\} \qquad \text{and} \qquad \sigma(\omega) \in \{\omega, \omega^2\}.$$

Let me claim without proof that each of these six functions is indeed a **field automorphism**. It would extremely tedious to check this by hand. Later we will have an indirect method.

Thus we obtain a Galois group of size 6. Finally, I claim that this is the group of **all permutations** of the three roots, and hence

$$\mathrm{Gal}((x^3 - 2)/\mathbb{Q}) = \mathrm{Gal}(\mathbb{Q}(\alpha, \omega)/\mathbb{Q}) \cong S_3.$$

Proof. The function defined by $(\sigma(\alpha), \sigma(\omega)) := (\alpha, \omega^2)$ transposes the roots $\omega\alpha$ and $\omega^2\alpha$ and leaves α alone. (In fact this map is just complex conjugation.) Furthermore, the function defined by $(\sigma(\alpha), \sigma(\omega)) := (\omega\alpha, \omega^2)$ transposes the roots α and $\omega\alpha$ and leaves $\omega^2\alpha$ alone. Any other permutation can be obtained by composing these two transpositions. □

I apologize that there were some gaps in the second example. Sadly it will take some time to fill them in. But I wanted to have this example available next time when we discuss the Fundamental Theorem.

Now we have enough ingredients that I can state the main theorems of Galois theory. The modern definitions are really due to Dedekind, and the notation is heavily influenced by Emil Artin's 1942 lectures at the University of Notre Dame.[10]

First, here is Dedekind's translation of the notion of "solvability."

Definition of Solvable Field Extensions. Let $\mathbb{E} \supseteq \mathbb{F}$ be a field extension. We say that this extension is *solvable* if there exists a chain of field extensions

$$\mathbb{F} = \mathbb{F}_0 \subseteq \mathbb{F}_1 \subseteq \cdots \subseteq \mathbb{F}_k \supseteq \mathbb{E}$$

satisfying the following condition:

[10] Dedekind was the last student of Carl Friedrich Gauss at the University of Göttingen. The modern language of abstract algebra later emerged through the lectures of Émil Artin and Emmy Noether at Göttingen in the 1920s. Noether in particular viewed Dedekind as the spiritual father of the subject. When the German universities were decimated by the Nazis, many prominent mathematicians, including Artin and Noether, ended up in the United States.

For all i we have $\mathbb{F}_i = \mathbb{F}_{i-1}(\alpha_i)$ for some element $\alpha_i \in \mathbb{F}_i$ such that $\alpha_i \notin \mathbb{F}_{i-1}$ but $\alpha_i^{n_i} \in \mathbb{F}_{i-1}$ for some power $n_i \geq 2$. [**Jargon.** We say that $\mathbb{F}_i \supseteq \mathbb{F}_{i-1}$ is a *simple radical extension.*]

Essentially, this just means that every element of the field $\mathbb{E}$ can be expressed in terms of the elements of $\mathbb{F}$ using only the operations

$$+, -, \times, \div, \sqrt{}, \sqrt[3]{}, \sqrt[4]{}, \sqrt[5]{}, \ldots.$$

Beginning with $\mathbb{F} = \mathbb{F}_0$, if we apply the operations $+, -, \times, \div$ then we will stay inside the same field. But if we adjoin a specific n_i-th root α_i of some element $\alpha_i^{n_i} \in \mathbb{F}_{i-1}$ then we may jump up into a bigger field $\mathbb{F}_i = \mathbb{F}_{i-1}(\alpha_i)$. The goal is to obtain every element of $\mathbb{E}$ after a finite number of adjunctions. In most examples we will have $\mathbb{F}_k = \mathbb{E}$ but for technical reasons we cannot assume this. Sorry. ///

[Remark to myself: Look at Ian Stewart's *Galois Theory, Third Edition*, Natural Irrationalities.]

And here is the big theorem. This theorem is the ultimate motivation for many of the definitions in field theory and group theory. It took over 100 years to clean up all the details and still most mathematicians have never seen a full proof.

Galois' Solvability Theorem. Let $\mathbb{E} \supseteq \mathbb{F}$ be the splitting field for some polynomial $f(x)$ with coefficients in $\mathbb{F}$ and let $G = \mathrm{Gal}(\mathbb{E}/\mathbb{F}) = \mathrm{Gal}(f/\mathbb{F})$ be the Galois group. Then we have

$$\left\{ \begin{array}{c} f(x) = 0 \text{ is solvable} \\ \text{by radicals} \end{array} \right\} \iff \left\{ \begin{array}{c} \mathbb{E} \supseteq \mathbb{F} \text{ is a solvable} \\ \text{field extension} \end{array} \right\} \iff \left\{ \begin{array}{c} G \text{ is a solvable} \\ \text{group} \end{array} \right\}.$$

///

The key idea of the proof is a certain "abstract Galois connection" between the lattice of intermediate fields $\mathscr{L}(\mathbb{E}, \mathbb{F})$ and the lattice of subgroups $\mathscr{L}(G)$. Recall that $G = \mathrm{Gal}(\mathbb{E}/\mathbb{F})$ is the group of field automorphisms $\mathbb{E} \to \mathbb{E}$ that fix elements of the subfield $\mathbb{F}$. If $\mathbb{F} \subseteq \mathbb{K} \subseteq \mathbb{E}$ is any intermediate field then it follows by definition that $\mathrm{Gal}(\mathbb{E}/\mathbb{K})$ is a subgroup of G. Indeed, since $\mathbb{F} \subseteq \mathbb{K}$ any automorphism that fixes $\mathbb{K}$ must also fix $\mathbb{F}$. On the other hand, let $H \subseteq G$ be any subgroup and consider the set

$$\mathrm{Fix}_{\mathbb{E}}(H) := \{a \in \mathbb{E} : \sigma(a) = a \text{ for all } \sigma \in H\} \subseteq \mathbb{E}.$$

This set contains $\mathbb{F}$ because $\sigma(a) = a$ for all $\sigma \in G$ and because $H \subseteq G$. I claim that $\mathbb{F} \subseteq \mathrm{Fix}_{\mathbb{E}}(H) \subseteq \mathbb{E}$ is an intermediate field, called the *fixed subfield* of H.

Proof. Consider any $\sigma \in H$. Since $\sigma : \mathbb{E} \to \mathbb{E}$ is a field automorphism we must have $\sigma(0) = 0$ and $\sigma(1) = 1$, which implies that $0, 1 \in \mathrm{Fix}_{\mathbb{E}}(H)$. Then for all $a, b \in \mathrm{Fix}_{\mathbb{E}}(H)$ we have

$$\sigma(a + b) = \sigma(a) + \sigma(b) = a + b \qquad \text{and} \qquad \sigma(ab) = \sigma(a)\sigma(b) = ab,$$

which implies that $a+b, ab \in \mathrm{Fix}_{\mathbb{E}}(H)$. Finally, for all $a \in \mathbb{E} - \{0\}$ we have

$$\sigma(a^{-1}) = \sigma(a)^{-1} = a^{-1},$$

which implies that $a^{-1} \in \mathrm{Fix}_{\mathbb{E}}(H)$. □

In summary, we have a pair of functions between the lattices $\mathscr{L}(\mathbb{E}, \mathbb{F})$ and $\mathscr{L}(G)$. In the language of last semester, I claim that these two functions form an *abstract Galois connection*.[11] Actually it will be a Galois connection after we reverse the partial order on one of the posets. We will do this with the superscript "op" for "opposite:"

$$\mathrm{Gal}(\mathbb{E}/-) : \mathscr{L}(\mathbb{E}, \mathbb{F}) \rightleftarrows \mathscr{L}(G)^{\mathrm{op}} : \mathrm{Fix}_{\mathbb{E}}(-).$$

Proof. Consider any intermediate field $\mathbb{F} \subseteq \mathbb{K} \subseteq \mathbb{E}$ and any subgroup $H \subseteq G$. By the definition of Galois connection we need to show that

$$\mathbb{K} \subseteq \mathrm{Fix}_{\mathbb{E}}(H) \iff \mathrm{Gal}(\mathbb{E}/\mathbb{K}) \supseteq H.$$

And this is immediate from the definitions of $\mathrm{Gal}(\mathbb{E}/-)$ and $\mathrm{Fix}_{\mathbb{E}}(-)$:

$$\begin{aligned}
\mathbb{K} \subseteq \mathrm{Fix}_{\mathbb{E}}(H) &\iff \forall a \in \mathbb{K}, a \in \mathrm{Fix}_{\mathbb{E}}(H) \\
&\iff \forall a \in \mathbb{K}, \forall \sigma \in H, \sigma(a) = a \\
&\iff \forall \sigma \in H, \forall a \in \mathbb{K}, \sigma(a) = a \\
&\iff \forall \sigma \in H, \sigma \in \mathrm{Gal}(\mathbb{E}/\mathbb{K}) \\
&\iff H \subseteq \mathrm{Gal}(\mathbb{E}/\mathbb{K}).
\end{aligned}$$

□

Let me remind you what we get from this. It follows for purely formal (i.e., "trivial") reasons that these functions restrict to an isomorphism between certain subposets

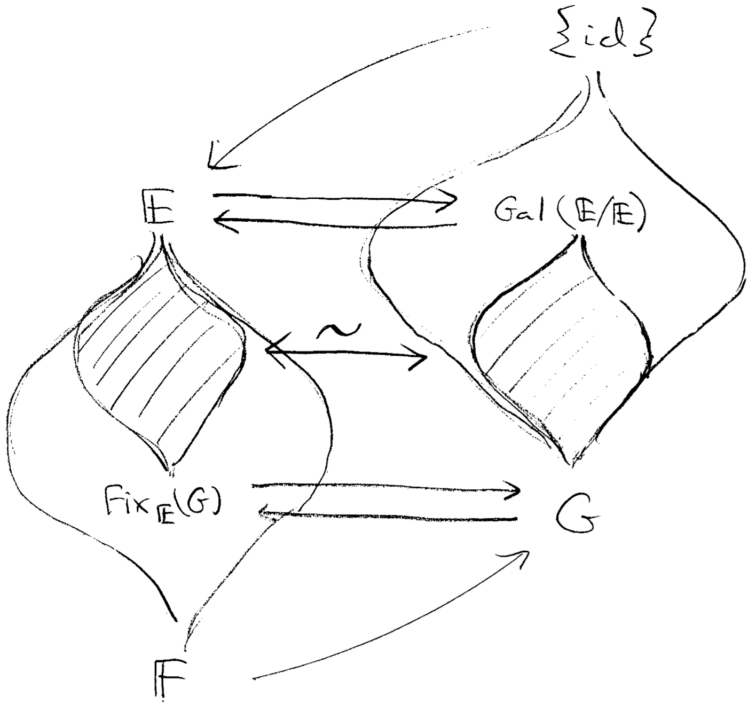
$$\mathrm{Gal}(\mathbb{E}/-) : \mathscr{L}(\mathbb{E}, \mathbb{F})' \xleftrightarrow{\sim} (\mathscr{L}(G)')^{\mathrm{op}} : \mathrm{Fix}_{\mathbb{E}}(-),$$

where the subposets are defined by

$$\begin{aligned}
\mathscr{L}(\mathbb{E}, \mathbb{F})' &:= \{\mathbb{F} \subseteq \mathbb{K} \subseteq \mathbb{E} : \mathrm{Fix}_{\mathbb{E}}(\mathrm{Gal}(\mathbb{E}/\mathbb{K})) = \mathbb{K}\}, \\
\mathscr{L}(G)' &:= \{H \subseteq G : \mathrm{Gal}(\mathbb{E}/\mathrm{Fix}_{\mathbb{E}}(H)) = H\}.
\end{aligned}$$

Here's a picture:

[11] This is a good time to remind yourself of the definition.



Actually this picture is a bit too loose because we always have $\mathrm{Gal}(\mathbb{E}/\mathbb{E}) = \{\mathrm{id}\}$. But never mind. The Fundamental Theorem says that under certain nice conditions (when $\mathbb{E} \supseteq \mathbb{F}$ is a splitting field for some polynomial) then the correspondence is as tight as possible.

The Fundamental Theorem of Galois Theory. Let $\mathbb{E} \supseteq \mathbb{F}$ be the splitting field for some polynomial $f(x)$ with coefficients in $\mathbb{F}$ and let $G = \mathrm{Gal}(\mathbb{E}/\mathbb{F}) = \mathrm{Gal}(f/\mathbb{F})$ be the Galois group. Then the following conditions hold:

(1) The Galois connection $\mathscr{L}(\mathbb{E}, \mathbb{F}) \leftrightarrows \mathscr{L}(G)$ is actually a **bijection**. That is, for all intermediate fields $\mathbb{F} \subseteq \mathbb{K} \subseteq \mathbb{E}$ and for all subgroups $H \subseteq G$ we have

$$\mathrm{Fix}_{\mathbb{E}}(\mathrm{Gal}(\mathbb{E}/\mathbb{K})) = \mathbb{K} \qquad \text{and} \qquad \mathrm{Gal}(\mathbb{E}/\mathrm{Fix}_{\mathbb{E}}(H)) = H.$$

(2) For any subgroup $H \subseteq G$ with fixed field $\mathbb{K} = \mathrm{Fix}_{\mathbb{E}}(H)$ we have

$$H \trianglelefteq G \text{ is normal} \quad \Longleftrightarrow \quad \mathbb{K} \supseteq \mathbb{F} \text{ is a splitting field for some polynomial,}$$

in which case the quotient group is isomorphic to the Galois group:

$$G/H \cong \mathrm{Gal}(\mathbb{K}/\mathbb{F}).$$

(3) For any subgroup $H \subseteq G$ with fixed field $\mathbb{K} = \mathrm{Fix}_{\mathbb{E}}(H)$ we have

$$\#\{ \text{ left cosets of } H \text{ in } G \} = \#(G/H) = [\mathbb{K}/\mathbb{F}] = \dim(\ \mathbb{K} \text{ as a vector space over } \mathbb{F}\ ).$$

///

[Remark: Recall that the notation "G/H" for the **set** of H-cosets is motivated by Lagrange's Theorem:

$$\#(G/H) = \#G \,/\, \#H.$$

The notation "$\mathbb{K}/\mathbb{F}$"[12] for $\mathbb{K}$ as an $\mathbb{F}$-**vector space** is motivated by a similar theorem, called Dedekind's Tower Law:

$$[\mathbb{E}/\mathbb{K}] = [\mathbb{E}/\mathbb{F}] \,/\, [\mathbb{K}/\mathbb{F}].$$

You will prove Dedekind's Law on the homework. In the situation where $\mathbb{E} \supseteq \mathbb{F}$ is a splitting field for some polynomial over $\mathbb{F}$, it follows from the Fundamental Theorem above that Lagrange's Theorem and Dedekind's Law are equivalent.]

It is easy to imagine how the proof of Galois' Solvability Theorem might following from the Fundamental Theorem, since the solvability of field extensions and groups are both defined by the existence of certain kinds of chains. Sadly, there is still one technicality. To go between solvable chains of subfields and solvable chains of subgroups we need to pass through a concept I will call "Kummer chains."[13] Then the proof of the Solvability Theorem goes as follows:

$$\left\{ \begin{array}{c} \exists \text{ a chain of simple} \\ \text{radical extensions} \end{array} \right\} \iff \left\{ \begin{array}{c} \exists \text{ a "Kummer} \\ \text{chain"} \end{array} \right\} \iff \left\{ \begin{array}{c} \exists \text{ a chain with} \\ \text{abelian quotients} \end{array} \right\}.$$

Never mind the details right now. Let me just show you the smallest interesting example.

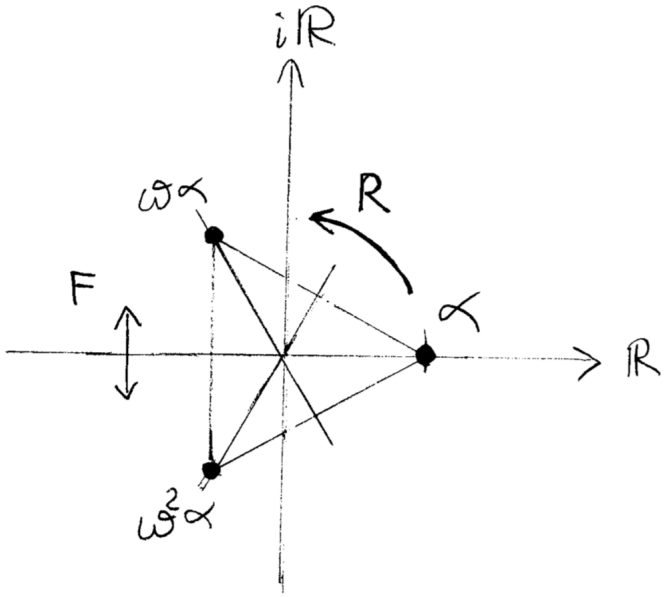
The Smallest Interesting Example. We saw last time that the splitting field $\mathbb{E} \supseteq \mathbb{Q}$ for the polynomial $x^3 - 2$ is given by

$$\mathbb{E} = \mathbb{Q}(\omega, \alpha) = \{a + b\alpha + c\alpha^2 + d\omega + e\omega\alpha + f\omega\alpha^2 : a, b, c, d, e, f \in \mathbb{Q}\} \supseteq \mathbb{Q}.$$

Furthermore, we saw that the Galois group is isomorphic to the group of all permutations of the roots $\{\alpha, \omega\alpha, \omega^2\alpha\}$. To be concrete, let's identify this group with the dihedral group of symmetries of the equilateral triangle, as in the following picture:

[12] The standard notation is $\mathbb{K} : \mathbb{F}$, but I think that the use of a colon to suggest a quotient doesn't read well to modern eyes. So I decided to update the notation.

[13] This is my own notation. I thought the concept needed a name.



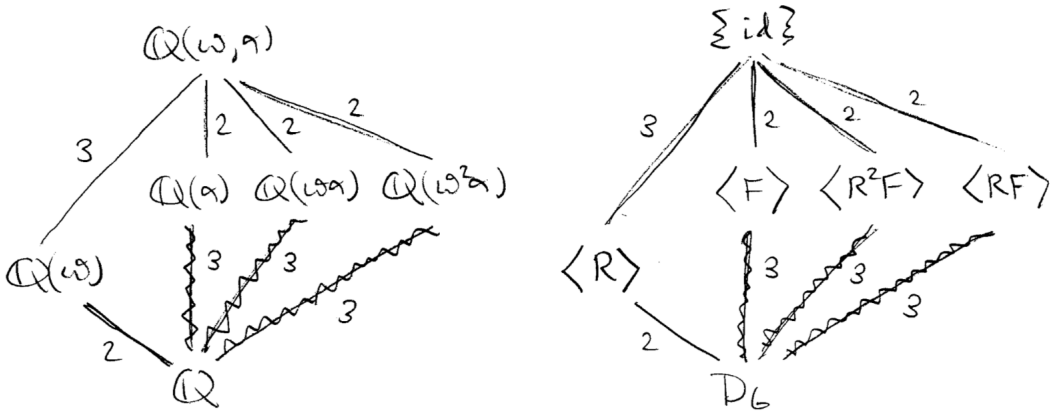
Recall that the six elements of this group can be expressed as follows:

$$D_6 = \{\mathrm{id}, R, R^2, F, RF, R^2F\}.$$

The "rotation" $R : \mathbb{Q}(\omega, \alpha) \to \mathbb{Q}(\omega, \alpha)$ is defined on the generators by $R(\alpha) = \omega\alpha$ and $R(\omega) = \omega$, while the "reflection" $F : \mathbb{Q}(\omega, \alpha) \to \mathbb{Q}(\omega, \alpha)$ is defined by $F(\alpha) = \alpha$ and $F(\omega) = \omega^2$. (In fact, F is the restriction of "complex conjugation" to the subfield $\mathbb{E} \subseteq \mathbb{C}$.) By working with the coordinates $a, b, c, d, e, f \in \mathbb{Q}$ one can compute the fixed fields of the cyclic subgroups generated by R and F:

$$\mathrm{Fix}_{\mathbb{E}}(\langle R \rangle) = \mathbb{Q}(\omega) \qquad \text{and} \qquad \mathrm{Fix}_{\mathbb{E}}(\langle F \rangle) = \mathbb{Q}(\alpha).$$

With a bit more work we obtain the following bijection between subfields and subgroups:



Since the finite group D_6 has finitely many subgroups, it follows from the Fundamental Theorem that the field extension $\mathbb{Q}(\omega, \alpha) \supseteq \mathbb{Q}$ has **finitely many intermediate fields**, which is certainly not obvious. I have labeled the edges with the degree of the field extension (left) or the number of cosets (right). The Fundamental Theorem says that these numbers are equal.

Finally, I have labeled the **non-normal subgroups** with squiggly lines. These correspond on the left to field extensions that are **not splitting fields** for any polynomial. We say that the group D_6 is *solvable* because of the existence of the chain of normal subgroups

$$D_6 \trianglerighteq \langle R \rangle \trianglerighteq \{\mathrm{id}\}$$

with abelian quotients $D_6/\langle R\rangle \cong \mathbb{Z}/2\mathbb{Z}$ and $\langle R\rangle/\{\mathrm{id}\} \cong \mathbb{Z}/3\mathbb{Z}$. It is less clear what is special about the corresponding chain of fields. We might ask:

Why is the chain $\mathbb{Q} \subseteq \mathbb{Q}(\omega) \subseteq \mathbb{Q}(\omega, \alpha)$ better than the chain $\mathbb{Q} \subseteq \mathbb{Q}(\alpha) \subseteq \mathbb{Q}(\omega, \alpha)$?

I'll just let you puzzle over this for now. The general rule for a "good" chain of field extensions (later we will call this a"Kummer chain") is to

adjoin the roots of unity first.

///

My goal is to prove all of this before the end of the course. We will do this by building everything up slowly from the basic theory of "commutative rings." The study of commutative rings (which is called "commutative algebra") is an absolutely huge subject,[14] so we will only cover the material that is relevant to Galois' theorem.

Problem Set 7

1. Degree of a Field Extension. Let $\mathbb{E} \supseteq \mathbb{F}$ be a field extension.

(a) There is an obvious "multiplication function" $\mathbb{F} \times \mathbb{E} \to \mathbb{E}$ defined by the rule $(a, b) \mapsto ab$. Verify that this multiplication makes $\mathbb{E}$ into a **vector space** over $\mathbb{F}$. We will denote this vector space by $\mathbb{E}/\mathbb{F}$. Its dimension is called the *degree* of the extension:

$$[\mathbb{E}/\mathbb{F}] := \dim(\mathbb{E}/\mathbb{F}).$$

[Remark: The fractional notation is a mnemonic device. Do not take it literally.]

(b) **Dedekind's Tower Law.** Now let $\mathbb{E} \supseteq \mathbb{K} \supseteq \mathbb{F}$ be any intermediate field. Prove that the degrees of the three extensions satisfy

$$[\mathbb{E}/\mathbb{F}] = [\mathbb{E}/\mathbb{K}] \cdot [\mathbb{K}/\mathbb{F}].$$

[14] Eisenbud's textbook on *Commutative Algebra (with a View Toward Algebraic Geometry)* is 800 pages long and it assumes everything that we will say in this course as a pre-requisite.

[Hint: Let $\{\alpha_i\}_i$ be a basis for $\mathbb{K}/\mathbb{F}$ and let $\{\beta_j\}_j$ be a basis for $\mathbb{E}/\mathbb{K}$. Prove that the set $\{\alpha_i\beta_j\}_{i,j}$ is a basis for $\mathbb{E}/\mathbb{F}$.] Does this remind you of Lagrange's Theorem?

(a) The vector space axioms follow immediately from the field axioms:

- For all $a \in \mathbb{E}$ we have $1a = a$.
- For all $a, b \in \mathbb{F}$ and $c \in \mathbb{E}$ we have $(ab)c = a(bc)$.
- For all $a, b \in \mathbb{F}$ and $c \in \mathbb{E}$ we have $(a + b)c = ac + bc$.
- For all $a \in \mathbb{F}$ and $b, c \in \mathbb{E}$ we have $a(b + c) = ab + ac$.

(b) Let $\{\alpha_i\}_i$ be a basis for $\mathbb{K}/\mathbb{F}$ and let $\{\beta_j\}_j$ be a basis for $\mathbb{E}/\mathbb{K}$. In this case I claim that the set of products $\{\alpha_i\beta_j\}_{i,j}$ is a basis for $\mathbb{E}/\mathbb{F}$.

- **Spanning.** Consider any element $c \in \mathbb{E}$. Since $\{\beta_j\}_j$ spans $\mathbb{E}/\mathbb{K}$ we can write

$$c = \sum_j b_j \beta_j \quad \text{for some } b_j \in \mathbb{K}.$$

  And since $\{\alpha_i\}_i$ spans $\mathbb{K}/\mathbb{F}$, every coefficient $b_j \in \mathbb{K}$ can be written as

$$b_j = \sum_i a_{i,j} \alpha_i \quad \text{for some } a_{i,j} \in \mathbb{F}.$$

  It follows that

$$c = \sum_j \left(\sum_i a_{i,j} \alpha_i \right) \beta_j = \sum_{i,j} a_{i,j} (\alpha_i \beta_j) \quad \text{for some } a_{i,j} \in \mathbb{F}.$$

- **Independence.** Suppose that we have

$$\sum_{i,j} a_{i,j} (\alpha_i \beta_j) = 0 \quad \text{for some } a_{i,j} \in \mathbb{F}.$$

  Now observe that

$$\sum_j \left(\sum_i a_{i,j} \alpha_i \right) \beta_j = \sum_{i,j} a_{i,j} (\alpha_i \beta_j) = 0 \quad \text{for some } \sum_i a_{i,j} \alpha_i \in \mathbb{K}.$$

  Since $\{\beta_j\}_j$ is independent over $\mathbb{K}$ we conclude that

$$\sum_i a_{i,j} \alpha_i = 0 \quad \text{for all } j.$$

  But then since $\{\alpha_i\}_i$ is independent over $\mathbb{F}$ we conclude that

$$a_{i,j} = 0 \quad \text{for all } i, j.$$

Finally, we want to count the elements of the basis $\{\alpha_i\beta_j\}_{i,j}$. So suppose that $\alpha_i\beta_j = \alpha_k\beta_\ell$ for some i, j, k, ℓ. Since $\alpha_i, \alpha_k \in \mathbb{K} - \{0\}$ (indeed, no element of a basis can be zero) and since the set $\{\beta_j\}_j$ is independent over $\mathbb{K}$ this will lead to a contradiction unless $i = k$ and $j = \ell$. It follows that

$$[\mathbb{E}/\mathbb{F}] = \#\{\alpha_i\beta_j\}_{i,j} = \#\{\beta_j\}_j \cdot \#\{\alpha_i\}_i = [\mathbb{E}/\mathbb{K}] \cdot [\mathbb{K}/\mathbb{F}].$$

□

2. Definition of the Galois Group. In this problem you will show that the hypothesis of invertibility is redundant in the definition of the Galois group. Let $\mathbb{E} \supseteq \mathbb{F}$ be a field extension and let $\sigma : \mathbb{E} \to \mathbb{E}$ be any function satisfying

$$\sigma(a+b) = \sigma(a) + \sigma(b) \quad \text{and} \quad \sigma(ab) = \sigma(a)\sigma(b) \quad \text{for all } a, b \in \mathbb{E}.$$

(a) Prove that σ is necessarily injective.

(b) If $\sigma(a) = a$ for all $a \in \mathbb{F}$, prove that $\sigma : \mathbb{E}/\mathbb{F} \to \mathbb{E}/\mathbb{F}$ is a linear function.

(c) If $\sigma(a) = a$ for all $a \in \mathbb{F}$ and if $[\mathbb{E}/\mathbb{F}] < \infty$,[15] combine parts (a) and (b) to prove that σ is necessarily bijective. [Hint: Use the Rank-Nullity Theorem.]

(a) First let me observe that $0a = 0$ for all elements $a \in \mathbb{E}$ in a field. Indeed, we have

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

and then subtracting $0a$ from both sides gives the result. Now observe that $\sigma : (\mathbb{E}, +, 0) \to (\mathbb{E}, +, 0)$ is a homomorphism of abelian groups. Thus in order to prove that σ is injective we only need to prove that the kernel is trivial, i.e., $\ker\sigma = \{0\}$. So assume for contradiction that we have $\sigma(a) = 0$ for some $a \neq 0$. Then we can use the fact that $\sigma : (\mathbb{E} - \{0\}, \times, 1) \to (\mathbb{E} - \{0\}, \times, 1)$ is a group homomorphism to obtain

$$1 = \sigma(1) = \sigma(aa^{-1}) = \sigma(a)\sigma(a)^{-1} = 0 \cdot \sigma(a)^{-1} = 0.$$

Did I mention that $0 \neq 1$ is one of the field axioms?

(b) Now suppose that $\sigma(a) = a$ for all $a \in \mathbb{F}$. Then for all $a, b \in \mathbb{F}$ and $\alpha, \beta \in \mathbb{E}$ we have

$$\sigma(a\alpha + b\beta) = \sigma(a)\sigma(\alpha) + \sigma(b)\sigma(\beta) = a\sigma(\alpha) + b\sigma(\beta).$$

(c) Suppose in addition that $[\mathbb{E}/\mathbb{F}] < \infty$. Since $\sigma : \mathbb{E}/\mathbb{F} \to \mathbb{E}/\mathbb{F}$ is linear, the Rank-Nullity theorem tells us that

$$\begin{aligned} \dim(\ker\sigma) + \dim(\operatorname{im}\sigma) &= \dim(\mathbb{E}/\mathbb{F}) \\ \dim(\{0\}) + \dim(\operatorname{im}\sigma) &= \dim(\mathbb{E}/\mathbb{F}) \end{aligned}$$

[15] This will be true if $\mathbb{E}$ is the splitting field of a polynomial over $\mathbb{F}$.

$$0 + \dim(\operatorname{im}\sigma) = \dim(\mathbb{E}/\mathbb{F}).$$

Finally, since $\operatorname{im}\sigma \subseteq \mathbb{E}$ is a vector subspace of the same dimension, we conclude that $\operatorname{im}\sigma = \mathbb{E}$.

Actually, let me prove this for general vectors spaces. Let V be a vector space over a field $\mathbb{F}$ and let $U \subseteq V$ be any vector subspace satisfying $\dim(U) = \dim(V)$. Then I claim that $U = V$.

Proof. Let $\mathbf{u}_1, \ldots, \mathbf{u}_n \in U$ be a basis and assume for contradiction that there exists a vector $\mathbf{v} \in V - U$. Since $\dim(V) = n$ we know that the set $\mathbf{u}_1, \ldots, \mathbf{u}_n, \mathbf{v} \in V$ of $n+1$ vectors is linearly **dependent**. [Recall from Steinitz Exchange that any linearly independent set can be enlarged to a basis.] Hence there exist coefficients $a_1, \ldots, a_n, b \in \mathbb{F}$, not all zero, such that

$$a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + \cdots + a_n\mathbf{u}_n + b\mathbf{v} = \mathbf{0}.$$

If $b = 0$ then this contradicts the fact that $\mathbf{u}_1, \ldots \mathbf{u}_n$ are linearly independent. But if $b \neq 0$ then we obtain the contradiction

$$\mathbf{v} = -\frac{a_1}{b}\mathbf{u}_1 - \frac{a_2}{b}\mathbf{u}_2 - \cdots - \frac{a_n}{b}\mathbf{u}_n \in U.$$

□

3. Adjoining a Subset to a Subfield. Let $\mathbb{F} \subseteq \mathbb{E}$ be a field extension and let $S \subseteq \mathbb{E}$ be any subset. We let $\mathbb{F}(S) \subseteq \mathbb{E}$ denote the smallest subfield of $\mathbb{E}$ that contains the set $\mathbb{F} \cup S$.

(a) Prove that $\mathbb{F}(S) = \mathbb{F}(S - \mathbb{F})$.

(b) For any two subsets $S, T \subseteq \mathbb{F}$ prove that $\mathbb{F}(S)(T) = \mathbb{F}(T)(S) = \mathbb{F}(S \cup T)$.

(c) If $\mathbb{K} \subseteq \mathbb{F}$ is a subfield, prove that $\mathbb{F}(\mathbb{K}) = \mathbb{K}(\mathbb{F}) = \mathbb{F} \vee \mathbb{K}$ is the join operation in the lattice of subfields. We also call this the *compositum* of subfields:

$$\mathbb{F}\mathbb{K} := \mathbb{F}(\mathbb{K}).$$

(a) Recall the definition of subtraction: $S - \mathbb{F} = S \cap \mathbb{F}^c$. Therefore we have

$$\mathbb{F} \cup (S - \mathbb{F}) = \mathbb{F} \cup (S \cap \mathbb{F}^c) = (\mathbb{F} \cup S) \cap (\mathbb{F} \cup \mathbb{F}^c) = (\mathbb{F} \cup S) \cap \mathbb{E} = \mathbb{F} \cup S$$

and hence $\langle \mathbb{F} \cup (S - \mathbb{F}) \rangle = \langle \mathbb{F} \cup S \rangle$. Maybe I said too much.

(b) Let $\mathbb{K} \subseteq \mathbb{E}$ be any subfield. To show that $\mathbb{F}(S)(T) = \mathbb{F}(S \cup T)$ we need to prove that

$$\mathbb{F}(S) \cup T \subseteq \mathbb{K} \quad \Longleftrightarrow \quad \mathbb{F} \cup (S \cup T) \subseteq \mathbb{K}.$$

Indeed, this equivalence follows from the fact that

$$\mathbb{F}(S) \subseteq \mathbb{K} \quad \Longleftrightarrow \quad (\mathbb{F} \cup S) \subseteq \mathbb{K}.$$

(c) Note that $\mathbb{F}(\mathbb{K})$ is the smallest subfield of $\mathbb{E}$ that contains $\mathbb{F} \cup \mathbb{K}$. For all subfields $\mathbb{L} \subseteq \mathbb{E}$ this implies that

$$\mathbb{F}(\mathbb{K}) \subseteq \mathbb{L} \iff (\mathbb{F} \cup \mathbb{K}) \subseteq \mathbb{L} \iff \mathbb{F} \subseteq \mathbb{L} \text{ and } \mathbb{K} \subseteq \mathbb{L}.$$

Thus $\mathbb{F}(\mathbb{K})$ satisfies the defining property of the join.

4. Quadratic Field Extensions. Let $\mathbb{E} \supseteq \mathbb{F}$ be a field extension and let $\alpha \in \mathbb{E}$ be any element such that $\alpha \notin \mathbb{F}$ and $\alpha^2 \in \mathbb{F}$. Consider the subfield $\mathbb{F} \subseteq \mathbb{F}(\alpha) \subseteq \mathbb{E}$ generated by α.

(a) Prove that the set $\{1, \alpha\} \subseteq \mathbb{F}(\alpha)/\mathbb{F}$ is linearly independent.

(b) Prove that $\{1, \alpha\} \subseteq \mathbb{F}(\alpha)/\mathbb{F}$ is a spanning set. [Hint: Prove that $\{a + b\alpha : a, b \in \mathbb{F}\} \subseteq \mathbb{E}$ is a subfield by "rationalizing the denominator."] It follows that $\{1, \alpha\}$ is a basis for $\mathbb{F}(\alpha)/\mathbb{F}$ and hence $[\mathbb{F}(\alpha)/\mathbb{F}] = 2$.

(c) Use Dedekind's Tower Law to prove that there **does not exist** any intermediate field:

$$\mathbb{F} \subsetneq \mathbb{K} \subsetneq \mathbb{F}(\alpha).$$

(d) Prove that the function $\tau : \mathbb{F}(\alpha) \to \mathbb{F}(\alpha)$ defined by $\tau(a + b\alpha) := a - b\alpha$ is a field automorphism. We call this operation *conjugation*.

(a) Suppose that $a + b\alpha = 0$ for some $a, b \in \mathbb{F}$. If $b \neq 0$ then we obtain $\alpha = -a/b \in \mathbb{F}$, contradicting the fact that $\alpha \notin \mathbb{F}$. Therefore we have $b = 0$ and hence $a = -b\alpha = -0\alpha = 0$.

(b) I claim that the set $\{a + b\alpha : a, b \in \mathbb{F}\} \subseteq \mathbb{E}$ is a subfield. Indeed, this set contains $0, 1$ and is closed under addition and multiplication. Now consider any $a + b\alpha \neq 0$ with $a, b \in \mathbb{F}$ not both zero. This implies that $a^2 - b^2\alpha^2$ is a nonzero element of $\mathbb{F}$. Indeed, suppose for contradiction that $a^2 - b^2\alpha^2 = 0$. If $b = 0$ then we obtain the contradiction that $a = b = 0$ and if $b \neq 0$ then we obtain $\alpha^2 = a^2/b^2 = (a/b)^2$ which gives the contradiction $\alpha = \pm a/b \in \mathbb{F}$.[16] We conclude that

$$\frac{1}{a + b\alpha} = \frac{1}{a + b\alpha} \cdot \frac{a - b\alpha}{a - b\alpha} = \left(\frac{a}{a^2 - b^2\alpha^2}\right) + \left(\frac{-b}{a^2 - b^2\alpha^2}\right)\alpha,$$

which is in the set $\{a + b\alpha : a, b \in \mathbb{F}\}$.

Clearly we have $\{a + b\alpha : a, b \in \mathbb{F}\} \subseteq \mathbb{F}(\alpha)$. Conversely, since $\{a + b\alpha : a, b \in \mathbb{F}\}$ is a subfield of $\mathbb{E}$ that contains the set $\mathbb{F} \cup \{\alpha\}$ we conclude that

$$\mathbb{F}(\alpha) \subseteq \{a + b\alpha : a, b \in \mathbb{F}\}.$$

In particular, $\{1, \alpha\}$ is a spanning set. Combining (a) and (b) gives $[\mathbb{F}(a)/\mathbb{F}] = 2$.

[16] Again, I am assuming the "obvious" fact that a quadratic equation has at most two roots. We will shortly give a rigorous proof of this.

(c) Consider any field $\mathbb{K}$ such that $\mathbb{F} \subseteq \mathbb{K} \subseteq \mathbb{F}(\alpha)$. From Dedekind's Tower Law and the fact that $\{1, \alpha\}$ is a basis for $\mathbb{F}(\alpha)/\mathbb{F}$ we obtain

$$[\mathbb{F}(\alpha)/\mathbb{K}] \cdot [\mathbb{K}/\mathbb{F}] = [\mathbb{F}(\alpha)/\mathbb{F}] = 2.$$

Then since 2 is a prime number we must have $[\mathbb{K}/\mathbb{F}] = 1$ and hence $\mathbb{K} = \mathbb{F}$, or $[\mathbb{F}(\alpha)/\mathbb{K}] = 1$ and hence $\mathbb{F}(\alpha) = \mathbb{K}$. [Recall that $U \subseteq V$ and $\dim(U) = \dim(V)$ imply $U = V$.]

(d) Consider the function $\tau : \mathbb{F}(\alpha) \to \mathbb{F}(\alpha)$ defined by $\tau(a + b\alpha) := a - b\alpha$. For all elements $a, b, c, d \in \mathbb{F}$ we have

$$\begin{aligned} \tau(a + b\alpha) + \tau(c + d\alpha) &= (a - b\alpha) + (c - d\alpha) \\ &= (a + c) - (b + d)\alpha \\ &= \tau\left((a + c) + (b + d)\alpha\right) \\ &= \tau\left((a + b\alpha) + (c + d\alpha)\right) \end{aligned}$$

and

$$\begin{aligned} \tau(a + b\alpha)\tau(c + d\alpha) &= (a - b\alpha)(c - d\alpha) \\ &= (ac + bd\alpha^2) - (bc + ad)\alpha \\ &= \tau\left((ac + bd\alpha^2) + (bc + ad)\alpha\right) \\ &= \tau\left((a + b\alpha)(c + d\alpha)\right). \end{aligned}$$

Furthermore, this function is invertible because $\tau^2 = \mathrm{id}$. Alternatively, we could quote the result of Problem 2(c).

5. Square Roots are Irrational. Let $D \in \mathbb{N}$ be a positive integer and let $\sqrt{D} \in \mathbb{R}$ be any real square root. In this problem you will show that

$$\sqrt{D} \notin \mathbb{Z} \quad \Longrightarrow \quad \sqrt{D} \notin \mathbb{Q}.$$

(a) Consider the set $S = \{n \in \mathbb{N} : n\sqrt{D} \in \mathbb{Z}\} \subseteq \mathbb{N}$. Observe that

$$S = \emptyset \quad \Longleftrightarrow \quad \sqrt{D} \notin \mathbb{Q}.$$

(b) Assuming that $\sqrt{D} \notin \mathbb{Z}$, use Well-Ordering to prove that there exists $a \in \mathbb{Z}$ such that

$$a < \sqrt{D} < a + 1.$$

(c) Suppose in addition that $\sqrt{D} \in \mathbb{Q}$. By part (a) and Well-Ordering, this means that the set S has a smallest element, say $m \in S$. Now use part (b) to obtain a contradiction. [Hint: Consider the number $m(\sqrt{D} - a)$.]

(a) Note that S is the set of (positive) **denominators** in rational expressions for $\sqrt{D}$. The number $\sqrt{D}$ is rational if and only if it has such a denominator, i.e., if and only if $S \neq \emptyset$.

(b) The set of integers below $\sqrt{D} \in \mathbb{R}$ has a greatest element, say $a < \sqrt{D}$. Since $a+1$ is greater than a we must have $a+1 \not< \sqrt{D}$, hence $\sqrt{D} \leq a+1$, and since $\sqrt{D} \notin \mathbb{Z}$ we must have $\sqrt{D} \neq a+1$, hence $\sqrt{D} < a+1$.

(c) Now let $\sqrt{D} \notin \mathbb{Z}$ and assume for contradiction that $\sqrt{D} \in \mathbb{Q}$. From part (a) there exists a smallest positive "denominator" $m \geq 1$ such that $m\sqrt{D} \in \mathbb{Z}$, and from from part (b) there exists an integer $a \in \mathbb{Z}$ such that $a < \sqrt{D} < a+1$. Then we have

$$\begin{array}{rcccl} a & < & \sqrt{D} & < & a+1 \\ 0 & < & \sqrt{D}-a & < & 1 \\ 0 & < & m(\sqrt{D}-a) & < & m. \end{array}$$

But note that $m(\sqrt{D}-a) = (m\sqrt{D}) - ma$ is a positive integer satisfying

$$m(\sqrt{D}-a)\sqrt{D} = mD - ma\sqrt{D} = mD - a(m\sqrt{D}) \in \mathbb{Z}.$$

In other words, $m(\sqrt{D}-a)$ an element of S. This contradicts the minimality of m. □

6. A Basic Example. Let $\sqrt{2}, \sqrt{3} \in \mathbb{R}$ be some specific square roots of 2 and 3, and consider the subfields $\mathbb{Q}(\sqrt{2}), \mathbb{Q}(\sqrt{3}) \subseteq \mathbb{R}$. We saw in class that the union $\mathbb{Q}(\sqrt{2}) \cup \mathbb{Q}(\sqrt{3}) \subseteq \mathbb{R}$ is **not** a subfield. So instead we will consider the join/compositum subfield:

$$\mathbb{Q}(\sqrt{2}) \vee \mathbb{Q}(\sqrt{3}) = \mathbb{Q}(\sqrt{2})\mathbb{Q}(\sqrt{3}) = \mathbb{Q}(\sqrt{2})(\sqrt{3}) = \mathbb{Q}(\sqrt{2},\sqrt{3}) \subseteq \mathbb{R}.$$

(a) **A Basis.** Prove that elements of this field have the following explicit form:

$$\mathbb{Q}(\sqrt{2},\sqrt{3}) = \{a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6} : a,b,c,d \in \mathbb{Q}\}.$$

[Hint: It's quite tricky to prove directly that the set on the right is a field. Use Dedekind's Tower Law for an indirect proof.]

(b) **The Galois Group.** Let $\sigma : \mathbb{Q}(\sqrt{2},\sqrt{3}) \to \mathbb{Q}(\sqrt{2},\sqrt{3})$ be any field automorphism. Prove that σ necessarily fixes the prime subfield $\mathbb{Q}$, and hence that σ is uniquely determined by the two values $\sigma(\sqrt{2})$ and $\sigma(\sqrt{3})$. Write down all of the possibilities and observe that you get a group isomorphic to $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$.

(c) **A Primitive Element.** Define the number $\gamma = \sqrt{2} + \sqrt{3}$ and prove that

$$\mathbb{Q}(\sqrt{2},\sqrt{3}) = \mathbb{Q}(\gamma).$$

[Hint: One inclusion is easy. For the other inclusion, expand γ^3 to show that $\sqrt{2}$ and $\sqrt{3}$ are in the field $\mathbb{Q}(\gamma)$.] You know from part (a) that $[\mathbb{Q}(\sqrt{2},\sqrt{3})/\mathbb{Q}] = 4$. It follows

that the five elements $1, \gamma, \gamma^2, \gamma^3, \gamma^4 \in \mathbb{Q}(\sqrt{2}, \sqrt{3})$ are **not** linearly independent over $\mathbb{Q}$, hence γ must satisfy a quartic equation of the form

$$a + b\gamma + c\gamma^2 + d\gamma^3 + e\gamma^4 = 0 \quad \text{for some nontrivial } a, b, c, d, e \in \mathbb{Q}.$$

Find this equation. [Hint: Expand γ^4 and work down.] If $\sigma : \mathbb{Q}(\sqrt{2}, \sqrt{3}) \to \mathbb{Q}(\sqrt{2}, \sqrt{3})$ is any field automorphism, prove that $\sigma(\gamma)$ is another solution of the same equation. Finally, use part (b) to obtain all four roots of the equation.

(a) Consider the chain of extensions $\mathbb{Q} \subseteq \mathbb{Q}(\sqrt{2}) \subseteq \mathbb{Q}(\sqrt{2})(\sqrt{3}) = \mathbb{Q}(\sqrt{2}, \sqrt{3})$. From Problem 4 we know that $[\mathbb{Q}(\sqrt{2})/\mathbb{Q}] = 2$ with basis $\{1, \sqrt{2}\}$. And we proved in class that $\sqrt{3} \notin \mathbb{Q}(\sqrt{2})$. Thus it also follows from Problem 4 that $[\mathbb{Q}(\sqrt{2}, \sqrt{3})/\mathbb{Q}(\sqrt{2})] = 2$ with basis $\{1, \sqrt{3}\}$. Finally, it follows from (the proof of) Dedekind's Tower Law that

$$[\mathbb{Q}(\sqrt{2}, \sqrt{3})/\mathbb{Q}] = 4 \quad \text{with basis} \quad \{1 \cdot 1, \sqrt{2} \cdot 1, 1 \cdot \sqrt{3}, \sqrt{2}\sqrt{3}\} = \{1, \sqrt{2}, \sqrt{3}, \sqrt{6}\}.$$

(b) Let $\sigma : \mathbb{Q}(\sqrt{2}, \sqrt{3}) \to \mathbb{Q}(\sqrt{2}, \sqrt{3})$ be any field automorphism. Since $\sigma(1) = 1$ and since σ preserves addition and subtraction we can show by induction that $\sigma(n) = n$ for all integers $n \in \mathbb{Z}$. Then since σ preserves multiplication and division it follows that

$$\sigma(m/n) = \sigma(m)/\sigma(n) = m/n \quad \text{for all } m, n \in \mathbb{Z} \text{ with } n \neq 0.$$

Now consider any element $a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6}$ of the field $\mathbb{Q}(\sqrt{2}, \sqrt{3})$. By the previous remarks we have

$$\sigma(a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6}) = a + b\sigma(\sqrt{2}) + c\sigma(\sqrt{3}) + d\sigma(\sqrt{2})\sigma(\sqrt{3})$$

and it follows that σ is uniquely determined by the values $\sigma(\sqrt{2})$ and $\sigma(\sqrt{3})$. Furthermore, we know that $\sigma(\sqrt{2}) \in \{\pm\sqrt{2}\}$ and $\sigma(\sqrt{3}) \in \{\pm\sqrt{3}\}$ because

$$(\sqrt{2})^2 = 2 \implies \sigma(\sqrt{2})^2 = 2 \quad \text{and} \quad (\sqrt{3})^2 = 3 \implies \sigma(\sqrt{3})^2 = 3.$$

It follows that the Galois group has size at most 4. Let us assume optimistically that all four choices lead to field automorphisms. (You can check the details if you want. We will have a general theorem for this later.) Define the functions

$$\sigma(a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6}) := a - b\sqrt{2} + c\sqrt{3} - d\sqrt{6},$$
$$\tau(a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6}) := a + b\sqrt{2} - c\sqrt{3} - d\sqrt{6}.$$

Then the Galois group is $\mathrm{Gal}(\mathbb{Q}(\sqrt{2}, \sqrt{3})/\mathbb{Q}) = \{\mathrm{id}, \sigma, \tau, \sigma\tau\}$. Since $\sigma\tau = \tau\sigma$ we know that this is an internal direct product:

$$\mathrm{Gal}(\mathbb{Q}(\sqrt{2}, \sqrt{3})/\mathbb{Q}) = \langle\sigma\rangle \times \langle\tau\rangle \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}.$$

[Remark: Incidentally, this gives us a (not very useful) way to "rationalize the denominator." For any element $\gamma \in \mathbb{Q}(\sqrt{2},\sqrt{3})$ it turns out that $\gamma \cdot \sigma(\gamma) \cdot \tau(\gamma) \cdot \sigma\tau(\gamma)$ is in $\mathbb{Q}$, and hence

$$\frac{1}{\gamma} = \frac{1}{\gamma} \cdot \frac{\sigma(\gamma) \cdot \tau(\gamma) \cdot \sigma\tau(\gamma)}{\sigma(\gamma) \cdot \tau(\gamma) \cdot \sigma\tau(\gamma)} = \frac{\text{something in } \mathbb{Q}(\sqrt{2},\sqrt{3})}{\text{something in } \mathbb{Q}}.$$

The actual formula is too terrible to write out.]

(c) Let $\gamma := \sqrt{2} + \sqrt{3}$. Then we clearly have $\mathbb{Q}(\gamma) \subseteq \mathbb{Q}(\sqrt{2},\sqrt{3})$. My computer tells me that

$$\begin{aligned} 1 &= 1 + 0\sqrt{2} + 0\sqrt{3} + 0\sqrt{6}, \\ \gamma &= 0 + 1\sqrt{2} + 1\sqrt{3} + 0\sqrt{6}, \\ \gamma^2 &= 5 + 0\sqrt{2} + 0\sqrt{3} + 2\sqrt{6}, \\ \gamma^3 &= 0 + 11\sqrt{2} + 9\sqrt{3} + 0\sqrt{6}, \\ \gamma^4 &= 49 + 0\sqrt{2} + 0\sqrt{3} + 20\sqrt{6}. \end{aligned}$$

It follows that $\sqrt{2} = (\gamma^3 - 9\gamma)/2 \in \mathbb{Q}(\gamma)$ and $\sqrt{3} = (\gamma^3 - 11\gamma)/(-2) \in \mathbb{Q}(\gamma)$ and hence

$$\mathbb{Q}(\sqrt{2},\sqrt{3}) \subseteq \mathbb{Q}(\gamma).$$

[Jargon: We say that γ is a *primitive element* for the field extension $\mathbb{Q}(\sqrt{2},\sqrt{3}) \supseteq \mathbb{Q}$.] From the above data we also find that

$$\gamma^4 - 10\gamma^2 + 1 = 0.$$

[Jargon: This is called the *minimal polynomial* of the element γ.] If $\sigma \in \mathrm{Gal}(\mathbb{Q}(\sqrt{2},\sqrt{3})/\mathbb{Q})$ is any element of the Galois group then we find that

$$\sigma(\gamma)^4 - 10\sigma(\gamma)^2 + 1 = 0,$$

and hence $\sigma(\gamma)$ is another root of the minimal polynomial. Finally, from part (b) we conclude that the polynomial $x^4 - 10x^2 + 1$ has the following four roots:

$$\gamma = \sqrt{2} + \sqrt{3}, \quad \sigma(\gamma) = -\sqrt{2} + \sqrt{3}, \quad \tau(\gamma) = \sqrt{2} - \sqrt{3}, \quad \sigma\tau(\gamma) = -\sqrt{2} - \sqrt{3}.$$

Again, we will prove shortly that these are the **only** roots.

[Remark: In this problem we observed that there exists a bijection between the elements of the Galois group and the roots of the "minimal polynomial" for a "primitive element." This is actually how Galois **defined** the Galois group. But then he had to prove that different primitive elements lead to isomorphic groups. See Tignol's book. In order to prove that Dedekind's version of the Galois group is well-defined we will show later that the splitting field of a polynomial is unique up to isomorphism.]

Week 15

In this course I have followed a mostly chronological development of abstract algebra. The study of groups began with Galois in the 1820s and the study of fields began with Dedekind in the 1870s. The first work to study algebra from a purely axiomatic point of view was Ernst Steinitz' *Algebraic Theory of Fields* (1910). Inspired by this, several authors considered a weaker structure called "rings."[17] The basic framework of ring theory emerged through the work of Emmy Noether in the 1920s. Here is the modern definition.

Definition of Rings and Subrings/Extensions. Let R be a set equipped with two binary operations $+, \times : R \times R \to R$ and two special elements $0, 1 \in R$. We call this structure a *(commutative) ring* if the following axioms hold:

(R1) $(R, +, 0)$ is an abelian group.

(R2) $(R, \times, 1)$ is a commutative monoid. This means that multiplication is an associative and commutative operation with identity element 1. There is also a notion of *non-commutative ring* in which multiplication is not assumed to be commutative.

(R3) For all $a, b, c \in R$ we have $a(b + c) = ab + ac$.

Now let $S \subseteq R$ be any subset. We say that S is a *subring* of R if the following properties hold:

- The special elements $0, 1$ are in S.
- For all $a, b \in S$ we have $a \pm b \in S$ and $ab \in S$.

Equivalently, we say that R is a *ring extension* of S. ///

Remarks:

- The distributive law (R3) tells us how the two binary operations $+$ and $\times$ interact. From this we obtain some basic rules mixing additive and multiplicative concepts:

$$\begin{gathered} 0a = 0, \\ a(-b) = (-a)b = -(ab), \\ (-a)(-b) = ab. \end{gathered}$$

  You will prove these on the homework.

- We are are allowed to have $1 = 0$ in a ring. But in this case we also have

$$a = 1a = 0a = 0 \quad \text{for all } a \in R.$$

  This structure is called the *zero ring* $R = 0$.

[17] The word *ring* (or *Zahlring*) comes from David Hilbert's *Zahlbericht* (1897). The word "Zahlbericht" means "number report" and "Zahlring" means "number ring." Nobody knows why he chose the word "ring."

- If $R \neq 0$ (i.e., if $1 \neq 0$) then we define the set

$$R^\times := \{a \in R : \exists b \in R, ab = 1\} \subseteq R - \{0\}.$$

  It follows that $(R^\times, \times, 1)$ is a group, called the *group of units* of the ring.

- If $R^\times = R - \{0\}$ then we say that R is a *field*. Note that this definition implies $1 \neq 0$.[18]

- As with subgroups and subfields, the intersection of any collection of subrings is again a subring, and we can use this to define the "subring genererated by a subset." For example, if $E \supseteq R$ is any ring extension and if $\alpha \in E$ is any element then we will use the following "square-bracket" notation:

$$R[\alpha] := \text{ the smallest subring of } E \text{ containing } R \cup \{\alpha\}.$$

  More on this later.

///

The main innovation of Emmy Noether was to recognize the importance of homomorphisms in the study of rings. These are much more interesting than homomorphisms between fields.

Definition of Ring Homomorphisms. Let R and S be rings and let $\varphi : R \to S$ be any function. We say that φ is a *ring homomorphism* if the following properties hold:

(H1) $\varphi(a + b) = \varphi(a) + \varphi(b)$,

(H2) $\varphi(ab) = \varphi(a)\varphi(b)$,

(H3) $\varphi(1) = 1$.

///

Remarks:

- The first property (H1) says that $\varphi : (R, +, 0) \to (S, +, 0)$ is a homomorphism of groups. Let me refresh your memory why this implies that

$$\varphi(0) = 0 \qquad \text{and} \qquad \varphi(-a) = -\varphi(a) \text{ for all } a \in R.$$

  Proof. First note that $\varphi(0) = \varphi(0 + 0) = \varphi(0) + \varphi(0)$. Now subtract $\varphi(0)$ from both sides to obtain $0 = \varphi(0)$. Then for any $a \in R$ we have

$$0 = \varphi(0) = \varphi(a - a) = \varphi(a) + \varphi(-a),$$

  which implies that $\varphi(-a) = -\varphi(a)$. □

[18] There is no "field with one element." I don't know why.

- Unfortunately, the second property (H2) does **not** imply that $\varphi(1) = 1$. Indeed, the proof from above does not work because we are not allowed to divide. For this reason we must include the property (H3). Alternatively, we could combine properties (H2) and (H3) by saying that $\varphi : (R, \times, 1) \to (S, \times, 1)$ is a *homomorphism of monoids*.

///

Recall from last semester that a homomorphism of groups $\varphi : G \to H$ leads to a Correspondence Theorem and three Isomorphism Theorems. Our next goal is to extend all of this structure to rings. However, we will find that the situation is a bit more complicated.

First of all, we note that a subring is the same thing as the image of a homomorphism.

Subring = Image of a Ring Homomorphism. Let S be a ring and let $S' \subseteq S$ be a subset. I claim that $S' \subseteq S$ is a subring if and only if there exists a ring homomorphism $\varphi : R \to S$ such that $\operatorname{im} \varphi = S'$.

Proof. First let $\varphi : R \to S$ be a ring homomorphism and consider the image

$$\operatorname{im} \varphi := \{\varphi(a) : a \in R\} \subseteq S.$$

Since $\varphi(0) = 0$ and $\varphi(1) = 1$ we find that $0, 1 \in \operatorname{im} \varphi$. Furthermore, if $\varphi(a), \varphi(b) \in \operatorname{im} \varphi$ are any two elements of the image then we have

$$\varphi(a) \pm \varphi(b) = \varphi(a \pm b) \in \operatorname{im} \varphi \qquad \text{and} \qquad \varphi(a)\varphi(b) = \varphi(ab) \in \operatorname{im} \varphi,$$

as desired. Conversely, let $S' \subseteq S$ be any subring and let $\mathrm{id}|_{S'} : S' \to S$ be the restriction of the identity function $\mathrm{id} : S \to S$. Clearly $\mathrm{id}_{S'}$ is a ring homomorphism with image S'. □

Digression: The urge to translate all concepts (such as "subring") into the language of homomorphisms ultimately leads to the subject of *category theory*.[19] I will not actually define categories in this class, but I will point out a few category-theoretic ideas. Here's one.

Initial and Final Rings. Let R be any ring. Then there exists a unique ring homomorphism from the ring of integers and a unique ring homomorphism to the zero ring:

$$\mathbb{Z} \xrightarrow{\exists!} R \qquad \text{and} \qquad R \xrightarrow{\exists!} 0.$$

[Jargon: We say that $\mathbb{Z}$ is the *initial object* and 0 is the *final object* in the category of rings.]

[19] The language of categories emerged in the 1940s and 1950s in order to clarify the subject of topology. Our course has focused on the years 1830 to 1930. The only aspect of category theory that I find suitable for a first course in algebra is the concept of an abstract Galois connection between lattices, which is a shadow of the concept of "adjoint functors."

Proof. There is a unique function $\varphi : R \to 0$ and this function is a ring homomorphism. On the other hand, recall that for any $a \in R$ and $n \in \mathbb{Z}$ we have defined the following notation:

$$n \cdot a := \begin{cases} \overbrace{a + a + \cdots + a}^{n \text{ times}} & \text{if } n \geq 1, \\ 0 & \text{if } n = 0, \\ \underbrace{-a - a - \cdots - a}_{-n \text{ times}} & \text{if } n \leq 1. \end{cases}$$

When discussing cyclic groups, we proved by induction that the map $n \mapsto n \cdot a$ is the unique group homomorphism $(\mathbb{Z}, +, 0) \to (R, +, 0)$ sending $1 \in \mathbb{Z}$ to $a \in R$. Now let $\varphi : \mathbb{Z} \to R$ be any ring homomorphism. In particular, since $\varphi : (\mathbb{Z}, +, 0) \to (R, +, 0)$ is a group homomorphism we must have $\varphi(n) = n \cdot 1$ for all $n \in \mathbb{Z}$. It only remains to prove that the function $\varphi(n) := n \cdot 1$ preserves multiplication, and this must be done by induction. Here is the key step:

$$\begin{aligned} \varphi(m)\varphi(n+1) &= (m \cdot 1)\left[(n+1) \cdot 1\right] \\ &= (m \cdot 1)(n \cdot 1 + 1) \\ &= (m \cdot 1)(n \cdot 1) + m \cdot 1 \\ &= (mn) \cdot 1 + m \cdot 1 && \text{induction on } n \\ &= (mn + m) \cdot 1 \\ &= [m(n+1)] \cdot 1 = \varphi(m(n+1)). \end{aligned}$$

□

End of digression.

Last time we saw that a subring is the same thing as the image of a ring homomorphism. As with group homomorphisms, there is also a notion of kernel for ring homomorphisms. This concept was implicit in the work of Kummer and Dedekind on unique factorization. (We will discuss this below.) Emmy Noether synthesized all of this into the modern definition.

Ideal = Additive Kernel of a Ring Homomorphism. Consider any ring homomorphism $\varphi : R \to S$. In particular this defines a homomorphism of additive groups $\varphi : (R, +, 0) \to (S, +, 0)$. Let $\ker \varphi \subseteq R$ denote the kernel of this group homomorphism:

$$\ker \varphi := \{a \in R : \varphi(a) = 0\}.$$

Then the First Isomorphism Theorem for groups tells us that the (well-defined) function $a + \ker \varphi \mapsto \varphi(a)$ is an isomorphism of additive groups:

$$\left(\frac{R}{\ker \varphi}, +, 0 + \ker \varphi\right) \cong (\operatorname{im} \varphi, +, 0).$$

However, since $\varphi : R \to S$ satisfies the extra properties $\varphi(ab) = \varphi(a)\varphi(b)$ and $\varphi(1) = 1$ we know from the above that $\operatorname{im}\varphi \subseteq S$ actually a ring. By pulling back this structure we conclude that the quotient **group** $R/\ker\varphi$ is also a **ring** with multiplication defined by

$$(a + \ker\varphi)(b + \ker\varphi) := (ab + \ker\varphi).$$

The fact that this operation is well-defined reflects a certain structural property of the kernel, which is analogous to the "normal subgroup" property of group kernels. In order to motivate the following definition I will state it as a theorem.

Theorem (Definition of Ideals and Quotient Rings). Let $(R, +, \times, 0, 1)$ be a ring and let $I \subseteq (R, +, 0)$ be any additive subgroup. Then the following are equivalent:

(I1) For all $a \in I$ and $b \in R$ we have $ab \in I$. In this case we say that I is an *ideal of* R.

(I2) There exists a ring homomorphism $\varphi : R \to S$ with $I = \ker\varphi$.

///

Proof. First we do the easy direction.

(I2)⇒(I1): Let $\varphi : R \to S$ be a ring homomorphism and consider the kernel $\ker\varphi := \{a \in R : \varphi(a) = 0\}$. Then for all $a \in \ker\varphi$ and $b \in R$ we have

$$\varphi(ab) = \varphi(a)\varphi(b) = 0\varphi(b) = 0,$$

and hence $ab \in \ker\varphi$. In other words, $\ker\varphi \subseteq R$ is an ideal.

(I1)⇒(I2): Let $I \subseteq R$ be an ideal and consider the additive quotient group $(R/I, +, 0 + I)$. I claim that the following "multiplication operation" on R/I is well-defined:

$$(a + I)(b + I) := (ab + I).$$

Indeed, suppose that we have $a + I = a' + I$ and $b + I = b' + I$. By definition this means that $a - a'$ and $b - b'$ are in I. But then we have

$$ab - a'b' = ab - ab' + ab' - a'b' = a(b - b') + (a - a')b' \in I$$

and hence $ab + I = a'b' + I$ as desired. It is easy to check that this defines a ring structure:

$$(R/I, +, \times, 0 + I, 1 + I).$$

Finally, consider the function $\pi : R \to R/I$ defined by $a \mapsto a + I$. It is easy to check that this function is a ring homomorphism with $\ker\pi = I$. □

Remarks:

- The construction of quotient rings appeared on a homework last semester. I decided to repeat it here to emphasize the connection with ring homomorphisms.

- In category theory we emphasize that the quotient ring is really a pair $(R/I, \pi)$ where $\pi : R \to R/I$ is the *canonical projection* with $\ker \pi = I$. It satisfies the following so-called universal property:

$$\begin{array}{ccccc} & & \overset{\forall \varphi}{\curvearrowright} & & \\ R & \xrightarrow[\pi]{} & R/I & \xrightarrow[\exists! \bar{\varphi}]{} & S \end{array}$$

  In words: For any ring homomorphism $\varphi : R \to S$ with $I \subseteq \ker \varphi$, there exists a unique ring homomorphism $\bar{\varphi} : R/I \to S$ satisfying $\varphi = \bar{\varphi} \circ \pi$. Feel free to ignore this remark.

- The word "ideal" comes from Ernst Kummer's concept of "ideal numbers." He introduced this concept in order to recover some version of unique prime factorization in rings such as $\mathbb{Z}[\sqrt{-5}]$ where the literal version fails. Dedekind shortened the name from "ideal number" to "ideal" and Emmy Noether later recognized the central importance of ideals in the study of abstract rings.

///

The whole point of the above theorem/definition was to generalize the First Isomorphism Theorem from additive groups to rings. Let me just state the result for posterity.

The First Isomorphism Theorem for Rings. Let $\varphi : R \to S$ be any ring homomorphism. Since $\ker \varphi \subseteq R$ is an ideal we may consider the quotient ring $R/\ker \varphi$. Then the natural map $a + \ker \varphi \mapsto \varphi(a)$ is a well-defined isomorphism of rings:

$$R/\ker \varphi \cong \operatorname{im} \varphi.$$

///

And what about the Correspondence Theorem and the Second/Third Isomorphism Theorems? This is a bit more complicated because there are **three** lattices naturally associated to a ring:

- The lattice of additive subgroups.
- The lattice of subrings.
- The lattice of ideals.

Since subrings and ideals are both examples of additive subgroups, it seems most reasonable to use the notation $\mathscr{L}$ for the lattice of additive subgroups. Thus for any ring R we define

$$\mathscr{L}(R) := \{\text{the lattice of subgroups of } (R, +, 0)\}.^{20}$$

For any subgroups $A, B \subseteq R$ recall from last semester that the meet and join are given by the intersection and the sum, respectively:

$$A \wedge B = A \cap B \qquad \text{and} \qquad A \vee B = A + B := \{a + b : a \in A, b \in B\}.$$

[20] I apologize that this choice conflicts with my use of $\mathscr{L}(\mathbb{F})$ for the lattice of subfields of a field $\mathbb{F}$. Hopefully this will cause no confusion.

Furthermore, if $I \subseteq R$ is any subgroup (probably an ideal) then we will use the notation

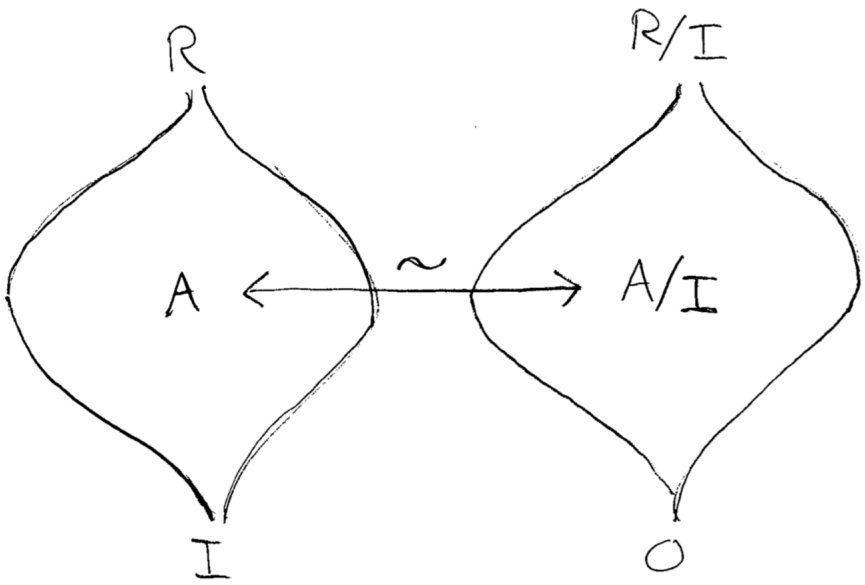
$$\mathscr{L}(R, I) := \{\text{subgroups } A \text{ such that } I \subseteq A \subseteq R\}.$$

Then we have the following theorem.

The Correspondence Theorem for Rings. Let R be a ring and let $I \subseteq R$ be an ideal. In particular, I is an additive subgroup, so the Correspondence Theorem for Groups gives us an isomorphism of lattices:

$$\begin{array}{ccc} \mathscr{L}(R, I) & \xrightarrow{\sim} & \mathscr{L}(R/I) \\ A & \mapsto & A/I. \end{array}$$

Here is a picture:



So far this is just group theory. To incorporate ring theory, we first observe that R/I is a ring because $I \subseteq R$ is an ideal. Then for any subgroup $I \subseteq A \subseteq R$ I claim that

$$A \subseteq R \text{ is a subring} \qquad \Longleftrightarrow \qquad A/I \subseteq R/I \text{ is a subring}$$

and

$$A \subseteq R \text{ is an ideal} \qquad \Longleftrightarrow \qquad A/I \subseteq R/I \text{ is an ideal.}$$

///

Proof. Last semester we gave a lengthy proof of the Correspondence Theorem for Groups. Luckily we don't have to prove it again. You will prove the final statements about subrings and ideals on the homework, where you will also prove ring versions of the Second and Third Isomorphism Theorems. □

Week 16

That was the general theory. This week we will start to focus on specific examples. But first, a bit of notational hygiene.

The Subring and the Ideal Generated by a Subset. Let R be a ring and let $S \subseteq R$ be any subset. Here's a question for you:

$$\text{What should the notation } \langle S \rangle \subseteq R \text{ represent?}$$

I can think of at least four possibilities:

- The smallest additive subgroup containing S.
- The smallest subring containing S.
- The smallest ideal containing S.
- The smallest subfield containing S. (If any exist.)

Because of this ambiguity I will try to avoid the notation $\langle S \rangle$ as much as possible. Instead I will use the following notations, which are fairly standard.

First, let $E \supseteq R$ be any ring extension and let $S \subseteq E$ be any subset. Then we define

$$R[S] := \bigcap \{\text{sub}\textbf{rings} \text{ of } E \text{ that contain the set } R \cup S\} \subseteq E.$$

In other words, $R[S]$ is the smallest subring of E that contains the set $R \cup S$. If E contains a subfield (for example, if E is a field) then we will also define

$$R(S) := \bigcap \{\text{sub}\textbf{fields} \text{ of } E \text{ that contain the set } R \cup S\} \subseteq E.$$

Observe that this is consistent with our previous notation for field extensions. Since every subfield is itself a subring, observe that we also have

$$R \subseteq R[S] \subseteq R(S) \subseteq E.$$

In general the inclusion $R[S] \subseteq R(S)$ is strict. However, we will meet a nice class of examples below (when R is a field and S is a finite set that is "algebraic" over R) for which $R[S] = R(S)$.

Next let R be a ring and let $S \subseteq R$ be any subset. Then we define

$$RS := \bigcap \{\text{ideals of } R \text{ that contain } S\} \subseteq R.$$

In other words, RS is the smallest ideal of R that contains the set S. You should be aware that many authors use the notations $\langle S \rangle$ or (S) for this ideal. But I prefer the "multiplicative" notation RS because of the following fact:

The ideal generated by S equals the set of *finite R-linear combinations*:[21]

$$RS = \{a_1\alpha_1 + a_2\alpha_2 + \cdots + a_k\alpha_k : a_1, \ldots, a_k \in R, \alpha_1, \ldots, \alpha_k \in S, k \geq 1\}.$$

Proof. Let I be the set of linear combinations above. Since RS is an ideal containing S we see that $I \subseteq RS$. On the other hand, for all $a_1\alpha_1 + \cdots a_k\alpha_k \in I$ and $b \in R$ we have

$$b(a_1\alpha_1 + \cdots a_k\alpha_k) = (ba_1)\alpha_1 + \cdots + (ba_k)\alpha_k \in I,$$

which implies that $I \subseteq R$ is an ideal. Furthermore, we have $S \subseteq I$ since $\alpha = 1\alpha \in I$ for all $\alpha \in S$. Finally, since RS is the **smallest** ideal containing S we conclude that $RS \subseteq I$. □

If R is a commutative ring (which for us it always will be), then we will also write $RS = SR$. However, when R is **non-commutative** then the notations RS and SR will denote the smallest **left ideal** and **right ideal** containing S, respectively. You can probably guess all the relevant definitions. ///

Ideals generated by a single element are very important so we give them a special name.

Definition of Principal Ideals. Let R be a commutative ring. The ideal generated by a single element is called a *principal ideal*. For $\alpha \in R$ we will use the notation

$$\alpha R = R\alpha := R\{\alpha\} = \{a\alpha : a \in R\}.$$

The smallest and largest ideals are both principal, generated by 0 and 1, respectively:

$$\begin{array}{rcll} 0R & = & \{0\} & \text{is called the } \textit{zero ideal}. \\ 1R & = & R & \text{is called the } \textit{unit ideal}. \end{array}$$

///

By the way, the name "unit ideal" is motivated by the following fact:

Let $I \subseteq R$ be an ideal. Then $I = R$ if and only if I contains a unit.

Proof. If $I = R$ then I contains the unit 1. Conversely, let $u \in I$ be a unit. By definition this means that $uu^{-1} = 1$ for some (unique) $u^{-1} \in R$. But then since I is an ideal we must have $1 = uu^{-1} \in I$, and it follows that

$$a = 1a \in I \text{ for all } a \in R.$$

□

Now for the examples.

Example: Fields. Let $\mathbb{F}$ be a field. Then I claim that $0\mathbb{F}$ and $1\mathbb{F}$ are the **only** ideals.

[21] This fact suggests an analogy between "ideals" and "vector subspaces." The analogy can be made precise with the definition of *R-modules*. But again, we don't have a need for that level of abstraction in this course.

Proof. Let $I \neq 0\mathbb{F}$ be a non-zero ideal. Then I contains a non-zero element, which must be a unit because $\mathbb{F}$ is a field. It follows from the previous remark that $I = 1\mathbb{F}$. □

Conversely, if R is any ring that contains **exactly two ideals**, then I claim that R is a field.

Proof. Assume that $0R \neq 1R$ are the only two ideals of R and let $0 \neq a \in R$ be any nonzero element. Now consider the principal ideal aR. Since $aR \neq 0R$ we must have $aR = 1R$. In particular, since $1 \in aR$ there exists an element $b \in R$ such that $1 = ab$. □

We have shown that a field is a ring with **exactly two ideals**. But recall that a group G with **exactly two normal subgroups** is called a *simple group*. Since ideals are analogous to normal subgroups (being kernels of the relevant homomorphisms) we might say that

a field is a "simple ring."

///

Example: The Integers. Fields are in some sense "too simple." The prototypical example of a ring is the ring of integers:

$$(\mathbb{Z}, +, \times, 0, 1).$$

But the ring $\mathbb{Z}$ is still "rather simple" because of the following fact:

Every ideal of $\mathbb{Z}$ is principal.

Proof. Let $I \subseteq \mathbb{Z}$ be any ideal. If $I = 0\mathbb{Z}$ then it is principal. Otherwise, let $0 \neq n \in I$ be a nonzero element with minimal absolute value (which exists by the well-ordering). Since I is an ideal we have $n\mathbb{Z} \subseteq I$. I claim that in fact $I = n\mathbb{Z}$. To see this, consider any element $a \in I$ and divide by n to obtain

$$a = qn + r \text{ for some (unique) } q, r \in \mathbb{Z} \text{ with } 0 \leq r < |n|.^{22}$$

If $r \neq 0$ then the facts $r = a - qn \in I$ and $0 < |r| = r < |n|$ contradict the minimality of n. Thus we must have $r = 0$ and hence $a = qn \in n\mathbb{Z}$. Finally, since this is true for all $a \in I$ we conclude that $I \subseteq n\mathbb{Z}$. □

You may recall that we already proved this last semester under the guise of "cyclic groups." In fact, we now see that every cyclic group $\mathbb{Z}/n\mathbb{Z}$ has a natural ring structure defined by

$$(a + n\mathbb{Z})(b + n\mathbb{Z}) = (ab + n\mathbb{Z}).$$

I had to hold my tongue many times.[23] Here is another important property of integers:

[22] This is not usually regarded as an axiom, but it is close to being the defining property of the integers.

[23] And the tongue holding will continue, because abelian groups are the same as "$\mathbb{Z}$-modules" and rings are the same as "$\mathbb{Z}$-algebras." Modules and algebras are important types of algebraic structures but I consider them more suitable for a graduate course.

For all $a, b \in \mathbb{Z}$, if $ab = 0$ then we must have $a = 0$ or $b = 0$.

This could be taken as an axiom, but it is usually proved using induction and the fact that $0 \neq 1$. Equivalently we could say that

the ring $\mathbb{Z}$ has no *zero-divisors*.

The technical term for a ring without zero-divisors is an *integral domain*.[24] You might at first assume that **every** ring is an integral domain, but consider the following fact:

If $n \in \mathbb{Z}$ is not zero or prime then the ring $\mathbb{Z}/n\mathbb{Z}$ is not an integral domain.

Proof. If n is not prime then by definition there exist $a, b \in \mathbb{Z}$ where $n = ab$ and neither of a, b is in the ideal $n\mathbb{Z}$. Then we have

$$(a + n\mathbb{Z})(b + n\mathbb{Z}) = (n + n\mathbb{Z}) = (0 + n\mathbb{Z}),$$

where neither of $a + n\mathbb{Z}$ or $b + n\mathbb{Z}$ is equal to the zero element $0 + n\mathbb{Z}$. □

On the other hand, you proved on a previous homework that the ring $\mathbb{Z}/p\mathbb{Z}$ for prime p is actually a **field** (hence also an integral domain). This fact is closely related to the concept of

division with remainder.

///

The examples of fields and integers will motivate much of the theory going forward. Next time we will meet the third fundamental example of a ring: polynomials.

Today's lecture is a bit philosophical, but don't worry — we'll discuss examples soon.

What is a Polynomial? For any subset of a ring $S \subseteq R$ we saw that the the smallest ideal $S \subseteq RS \subseteq R$ containing S consists of all the finite R-linear combinations:

$$RS = \{a_1\alpha_1 + a_2\alpha_2 + \cdots + a_k\alpha_k : a_1, \ldots, a_k \in R, \alpha_1, \ldots, \alpha_k \in S, k \geq 0\}.$$

[24] This notation is a near-literal translation of Kronecker's term *Integralitätsbereich* [domain of integrality]. Compare this to the term *Rationalitätsbereich* [domain of rationality], which was his name for fields. Why are these terms so closely related? You will prove on a future homework that the concepts of "integral domain" and "subring of a field" are equivalent.

We can give a similar explicit characterization of the sub**ring** generated by a subset. To be specific, let $E \supseteq R$ be a ring extension and let $S \subseteq E$ be any subset. Then the smallest subring $(R \cup S) \subseteq R[S] \subseteq E$ containing $R \cup S$ equals the set of all "finite polynomial expressions:"

$$R[S] = \left\{ \sum_{n_1,\ldots,n_k} a_{n_1,\ldots,n_k} \alpha_1^{n_1} \cdots \alpha_k^{n_k} : a_{n_1,\ldots,n_k} \in R, \alpha_i \in S, k, n_i \geq 0 \right\}.$$

The word "finite" means that only finitely many of the coefficients $a_{n_1,\ldots,n_k}$ are nonzero. I won't bother to prove this fact because the notation is atrocious. Instead we'll prove the special case when $S = \{\alpha\} \subseteq E$ has just one element. In this case I claim that we have

$$R[\alpha] = \{a_0 + a_1\alpha + a_2\alpha^2 + \cdots + a_n\alpha^n : a_0, \ldots, a_n \in R, n \geq 0\}.$$

Proof. Let $P = \{a_0 + a_2\alpha^2 + \cdots + a_n\alpha^n\}$ be the set of "finite R-polynomial expressions in α." Since $R[\alpha] \subseteq E$ is a subring containing $R \cup \{\alpha\}$ we note that $P \subseteq R[\alpha]$. On the other hand, one can see that P contains $0, 1 \in E$ and is closed under addition and multiplication, hence $P \subseteq E$ is a subring. Then since P contains $R \cup \{\alpha\}$, and since $R[\alpha]$ is the **smallest** subring containing $R \cup \{\alpha\}$, we conclude that $R[\alpha] \subseteq P$. □

This fact motivates the following definition.

Definition of Polynomials in One Variable. Let R be a ring and let x be a formal symbol, called a "variable." We use this to define a sequence of formal symbols "x^n" for $n \geq 0$:

$$x^0 := 1, \quad x^1 := x, \quad x^2, \quad x^3, \quad x^4, \quad \ldots.$$

Then we define the set of *formal polynomials in* x:

$$R[x] := \{a_0 + a_1x + a_2x^2 + \cdots a_nx^n : a_0, \ldots, a_n \in R, n \geq 0\}.$$

This set has a natural ring structure which we define by pretending that

$$a_nx^n = a_n \cdot \underbrace{x \cdot x \cdots x}_{n \text{ times}},$$

even though x is not a number so this multiplication is imaginary. In other words, for all polynomials $f(x) = \sum_i a_ix^i$ and $g(x) = \sum_i b_ix^i$ we define

$$\left(\sum_i a_ix^i\right) + \left(\sum_i b_ix^i\right) := \sum_i (a_i + b_i)x^i$$

and

$$\left(\sum_i a_ix^i\right)\left(\sum_i b_ix^i\right) := \sum_k \left(\sum_{i+j=k} a_ib_j\right) x^k.$$

To interpret these expressions one should assume that the indices run over all non-negative integers and that only finitely many coefficients are non-zero. ///

Remarks:

- Someone should prove that addition and multiplication of polynomials satisfy the ring axioms, but I won't do it here because the notation is too ugly. The proof that multiplication is associative comes down to the following identity:

$$\sum_{k+\ell=m} \left(\sum_{i+j=\ell} a_i b_j \right) c_k = \sum_{i+j+k=m} a_i b_j c_k = \sum_{i+\ell=m} a_i \left(\sum_{j+k=\ell} b_j c_k \right).$$

- What is the relationship between "formal polynomials" and "polynomial functions?" For any formal polynomial $f(x) = \sum_i a_i x^i \in R[x]$ we define a function $f : R \to R$ by evaluating $f(x)$ at $\alpha \in R$:

$$f(\alpha) := \sum_i a_i \alpha^i \in R.$$

  More generally, for any ring homomorphism $\varphi : R \to S$ we define a polynomial $f^\varphi(x) \in S[x]$ by applying φ to the coefficients and then we define a function $f^\varphi : S \to S$ by evaluating at $\alpha \in S$:

$$f^\varphi(\alpha) := \sum_i \varphi(a_i) \alpha^i \in S.$$

  For any fixed $\alpha \in S$ it turns out that the "evaluate at α" function $f(x) \mapsto f^\varphi(\alpha)$ is a **ring homomorphism** $R[x] \to S$, which plays an important role in Galois theory. (Don't worry, I'll remind you of the definitions later.)

  ///

The key to the structure of polynomial rings is the following "division algorithm," which is analogous to the division algorithm for integers. In order to state and prove this result we need a couple of definitions.

Definition: Degree of a Polynomial. Let R be a ring and consider any polynomial $f(x) \in R[x]$. The *degree* of $f(x)$ is defined as the highest power of x that occurs with a nonzero coefficient. In other words, we say that $\deg(f) = n$ when

$$f(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n \quad \text{with } a_n \neq 0.$$

Note that a polynomial of degree 0 is the same thing as a nonzero constant:

$$\deg(f) = 0 \iff f(x) = a_0 \text{ for some } 0 \neq a_0 \in R.$$

It is more difficult to define the degree of the *zero polynomial* $0 \in R[x]$. We could just say that deg(0) is undefined, but I prefer the following convention:

$$\deg(0) := -\infty.$$

[After all, this **is** the highest power of x that occurs with a nonzero coefficient.] Note that the degree of a sum always satisfies the following property:

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

What about the degree of a product? We would like to say that $\deg(fg) = \deg(f)+\deg(g)$, but this is not always true. For example, consider the polynomials $f(x) = 1+2x$ and $g(x) = 1+2x^2$ with coefficients in $\mathbb{Z}/4\mathbb{Z}$. Then we have

$$(1+2x)(1+2x^2) = 1+2x+2x^2+4x^3 = 1+2x+2x^2+0x^3 = 1+2x+2x^2,$$

so that $2 = \deg(fg) \neq \deg(f)+\deg(g) = 1+2$. The problem here is that the leading coefficients of f and g are zero-divisors. If we assume that the leading coefficients of $f(x)$ and $g(x)$ are **not zero-divisors** then we will have

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

///

Now we are ready for the theorem.

The Division Theorem for Polynomials. Let R be a ring and consider polynomials $f(x), g(x) \in R[x]$. If the leading coefficient of $g(x)$ is a unit then there exist unique polynomials $q(x), r(x) \in R[x]$ such that

$$\begin{cases} f(x) = q(x)g(x) + r(x), \\ \deg(r) < \deg(g). \end{cases}$$

These $q(x)$ and $r(x)$ are called the *quotient* and the *remainder* of $f(x)$ mod $g(x)$. ///

Proof. First we will prove existence of $q(x)$ and $r(x)$, then we will prove uniqueness.

Existence. Consider the set of potential remainders:

$$S = \{f(x) - q(x)g(x) : q(x) \in R[x]\}.$$

Let $r(x) \in S$ be a potential remainder of **minimal degree**, which exists by well-ordering of the set of degrees $\{-\infty < 0 < 1 < 2 < \cdots\}$. By definition we have $f(x) = q(x)g(x) + r(x)$ for some $q(x) \in R[x]$ and it remains to prove that $\deg(r) < \deg(g)$. To do this, let us assume for contradiction that $\deg(r) \geq \deg(g)$. Then we must have

$$g(x) = a_0 + \cdots + a_m x^m \quad \text{and} \quad r(x) = b_0 + \cdots b_n x^n$$

where a_m is a unit, b_n is nonzero, and $0 \leq m \leq n$. We can use these facts to cook up a remainder with strictly lower degree. Specifically, we define the polynomial

$$h(x) := r(x) - \frac{b_n}{a_m} x^{n-m} \cdot g(x) = \left(b_n - \frac{b_n}{a_m} a_m \right) x^n + \text{ lower terms.}$$

Note that $\deg(h) < \deg(r)$ by construction. But we also have

$$h(x) = f(x) - q(x)g(x) - \frac{b_n}{a_m} x^{n-m} \cdot g(x) = f(x) - \left(q(x) + \frac{b_n}{a_m} x^{n-m} \right) g(x) \in S,$$

which contradicts the minimality of $r(x)$. It follows that $\deg(r) < \deg(g)$ as desired.

Uniqueness. Suppose that we have polynomials $q_1, q_2, r_1, r_2 \in R[x]$ satisfying

$$\begin{cases} f(x) = q_1(x)g(x) + r_1(x), \\ \deg(r_1) < \deg(g), \end{cases} \qquad \begin{cases} f(x) = q_2(x)g(x) + r_2(x), \\ \deg(r_2) < \deg(g). \end{cases}$$

To prove that $r_1 = r_2$ and $q_1 = q_2$ we first equate expressions for f to obtain

$$\begin{aligned} q_1 g + r_1 &= q_2 g + r_2 \\ (q_1 - q_2)g &= (r_2 - r_1). \end{aligned}$$

If $r_2 - r_1 \neq 0$ then since the leading coefficient of g is a unit (in particular, not a zero-divisor) we also have $q_1 - q_2 \neq 0$, which implies that

$$\deg(r_2 - r_1) = \deg((q_1 - q_2)g) = \deg(q_1 - q_2) + \deg(g) \geq \deg(g).$$

But this contradicts the fact that $\deg(r_2 - r_1) \leq \max\{\deg(r_1), \deg(r_2)\} < \deg(g)$, so we must have $r_2 - r_1 = 0$. Finally, since $r_2 - r_1 = 0$ and since g has a unit leading coefficient we conclude that

$$(q_1 - q_2)g = 0 \qquad \Longrightarrow \qquad q_1 - q_2 = 0.$$

□

This theorem is most interesting when $R = \mathbb{F}$ is a field, because of the following fact:

Every nonzero polynomial in $\mathbb{F}[x]$ has a unit leading coefficient.

The Division Theorem for Polynomials is really an algorithm. Here are two examples.

Example: Long Division Over $\mathbb{Z}$. Consider the following polynomials over $\mathbb{Z}$:

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

Since the leading coefficient of $g(x)$ is the unit $1 \in \mathbb{Z}$ we we know that there exist (unique) polynomials $q(x), r(x)$ satisfying

$$\begin{cases} (2x^3 + 3x + 1) = q(x)(x+1) + r(x), \\ \deg(r) < \deg(x+1) = 1. \end{cases}$$

The proof of existence above leads to the following algorithm for computing $q(x)$ and $r(x)$:

$$\begin{array}{rr|llll} & & 2x^2 & -2x & +5 & \\ \hline x+1 & & 2x^3 & & +3x & +1 \\ & & -2x^3 & -2x & & \\ \hline & & & -2x^2 & +3x & +1 \\ & & & -2x^2 & -2x & \\ \hline & & & & 5x & +1 \\ & & & & -5x & -5 \\ \hline & & & & & -4 \end{array}$$

We conclude that

$$(2x^3 + 3x + 1) = q(x)(x+1) + r(x) = (2x^2 - 2x + 5)(x+1) - 4.$$

Note that the remainder $r(x) = -4$ satisfies $0 = \deg(-4) < \deg(x+1) = 1$ as expected. ///

Example: Long Division Over $\mathbb{Q}$. Let's change the example slightly:

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

This time the polynomial $g(x) = 2x + 1 \in \mathbb{Z}[x]$ has a non-unit leading coefficient. Thus there is no guarantee that the quotient and remainder exist in $\mathbb{Z}[x]$. In fact, we see that the division algorithm fails at the second step:

$$\begin{array}{rr|llll} & & 2x^2 & ? & & \\ \hline 2x+1 & & 2x^3 & & +3x & +1 \\ & & -2x^3 & -x^2 & & \\ \hline & & & -x^2 & +3x & +1 \\ & & & ? & & \end{array}$$

In order to cancel the "leading term" $-x^2$ we would need to multiply $2x+1$ by the "monomial" $-\frac{1}{2}x$, which does not exist in $\mathbb{Z}[x]$. However, since $\mathbb{Z} \subseteq \mathbb{Q}$ is a subring we could also think of $f(x)$ and $g(x)$ as elements of $\mathbb{Q}[x]$. Then since 2 **is a unit** in $\mathbb{Q}$ (with inverse 1/2) the algorithm is guaranteed to succeed:

$$\begin{array}{rr|lll} & & 2x^2 & -\frac{1}{2}x & +\frac{7}{4} \\ \cline{2-5} 2x+1 & & 2x^3 & & +3x \quad +1 \\ & & -2x^3 & -x^2 & \\ \cline{3-5} & & & -x^2 & +3x \quad +1 \\ & & & x^2 & +\frac{1}{2}x \\ \cline{4-5} & & & & \frac{7}{2}x \quad +1 \\ & & & & -\frac{7}{2}x \quad -\frac{7}{4} \\ \cline{5-5} & & & & -\frac{3}{4} \end{array}$$

We conclude that the unique quotient and remainder in $\mathbb{Q}[x]$ are

$$q(x) = 2x^2 - \frac{1}{2}x + \frac{7}{4} \qquad \text{and} \qquad r(x) = -\frac{3}{4}.$$

Actually we don't have to extend all the way to $\mathbb{Q}$. The quotient and remainder already exist in the smaller ring $\mathbb{Z}[1/2][x] \subseteq \mathbb{Q}[x]$, where $\mathbb{Z}[1/2] \subseteq \mathbb{Q}$ is the subring of fractions whose denominators are powers of 2. This is the smallest subring of $\mathbb{Q}$ in which 2 is a unit. ///

Next I will show you two important corollaries of the division algorithm. The first result goes back to René Descartes in his work *La Géométrie* (1637). This is the same work in which he introduced the concepts of "analytic geometry" and "Cartesian coordinates." Here is a direct quote from an English translation (1954, Dover, pg. 159–160):

> *It is evident from the above that [a polynomial equation] having several roots is always divisible by a binomial consisting of the unknown quantity diminished by the value of one of the true roots, or plus the value of one of the true roots. In this way, the degree of an equation can be lowered. On the other hand, if [a polynomial] is not divisible by a binomial consisting of the unknown quantity plus or minus some other quantity, then this latter quantity is not a root of the equation.*

And here is the modern version.

Corollary (Descartes' Factor Theorem). Let $E \supseteq R$ be an extension of (commutative) rings. Then for any polynomial $f(x) \in R[x]$ and for any element $\alpha \in E$ there exists a (unique) polynomial $g(x) \in E[x]$ of degree $\deg(f) - 1$ such that

$$f(x) = (x - \alpha)g(x) + f(\alpha).$$

It follows from this that

$$\alpha \in E \text{ is a root of } f(x) \quad \Longleftrightarrow \quad (x-\alpha) \text{ divides } f(x) \text{ in the ring } E[x].$$

///

Proof. Since $x - \alpha = 1x - \alpha \in E[x]$ has a unit leading coefficient we know that there exist (unique) polynomials $q(x), r(x) \in E[x]$ satisfying

$$\begin{cases} f(x) = (x-\alpha)q(x) + r(x), \\ \deg(r) < \deg(x-\alpha). \end{cases}$$

Since $\deg(x-\alpha) = 1$ this implies that $\deg(r) = 0$ or $\deg(r) = -\infty$. In other words, $r(x) = c \in E$ is a constant. To compute this constant we simply plug in $x = \alpha$ to obtain

$$f(\alpha) = (\alpha - \alpha)q(\alpha) + c = 0 \cdot q(\alpha) + c = 0 + c = c.$$

[Remark: For this step we needed the fact that α commutes with all of the coefficients of $q(x)$. This is why we assumed that E is a **commutative** ring.] Then to compute the degree of $q(x)$ we use the fact that 1 is not a zero-divisor to obtain

$$\deg(f) = \deg((x-\alpha)q + c) = \deg((x-\alpha)q) = \deg(x-\alpha) + \deg(q) = 1 + \deg(q).$$

Finally, if $f(x) = (x-\alpha)g(x)$ for some polynomial $g(x) \in E[x]$ then we have $f(\alpha) = (\alpha - \alpha)g(\alpha) = 0 \cdot g(\alpha) = 0$. [Again, we assume that E is commutative.] Conversely, if $f(\alpha) = 0$ then the above result implies that

$$f(x) = (x-\alpha)q(x) + f(\alpha) = (x-\alpha)q(x) + 0 = (x-\alpha)q(x)$$

for some $q(x) \in E[x]$. □

For example, recall from above that -4 is the remainder of $2x^3 + 3x + 1 \bmod x + 1$. On the other hand, by plugging $x = -1$ into $2x^3 + 3x + 1$ we obtain

$$2(-1)^3 + 3(-1) + 1 = -4.$$

Remarks:

- You will give a more constructive proof of this result on the homework.
- You will also use induction to prove the following corollary: A polynomial of degree n has **at most** n **distinct roots** in any integral domain. This finally justifies some of our remarks from weeks 13 and 14.
- A polynomial of degree n may have **more than** n **roots** in a noncommutative ring. For example, you may be familiar with the ring of *quaternions*:

  $$\mathbb{H} = \{a + bi + cj + dk : a, b, c, d \in \mathbb{R}, i^2 = j^2 = k^2 = ijk = -1\}.$$

  This ring is noncommutative because, for example, $ij = -ji$. Note that the polynomial $x^2 + 1 \in \mathbb{H}[x]$ has at least three roots: i, j, k. [In fact, one can show that $x^2 + 1$ has **uncountably many** roots in $\mathbb{H}$.]

- On the homework you will also show that a polynomial of degree n may have more than n roots in a non-integral domain.

///

To end this section I will present another important corollary of the division algorithm.

Corollary (Every Ideal of $\mathbb{F}[x]$ is Principal). Let $\mathbb{F}$ be a field and consider the ring of polynomials $\mathbb{F}[x]$. Then every ideal $I \subseteq \mathbb{F}[x]$ is principal.

Proof. Let $I \subseteq \mathbb{F}[x]$ be an ideal. If $I = 0\mathbb{F}[x]$ then it is principal. Otherwise, let $0 \neq m(x) \in I$ be a nonzero element of minimal degree (which exists by well-ordering of degrees). Since I is an ideal we have $m(x)\mathbb{F}[x] \subseteq I$. I claim that in fact $I = m(x)\mathbb{F}[x]$. To see this, consider any element $f(x) \in I$. Since $\mathbb{F}$ is a field and $m(x) \neq 0$ we know that the leading coefficient of $m(x)$ is a unit, hence there exist $q(x), r(x) \in \mathbb{F}[x]$ satisfying

$$\begin{cases} f(x) = q(x)m(x) + r(x), \\ \deg(r) < \deg(m). \end{cases}$$

If $r(x) \neq 0$ then the fact that $r(x) = f(x) - q(x)m(x) \in I$ contradicts the minimality of $m(x)$. Hence we must have $r(x)$ and it follows that $f(x) = m(x)q(x) \in m(x)\mathbb{F}[x]$. Finally, since this is true for all $f(x) \in I$ we conclude that $I \subseteq m(x)\mathbb{F}[x]$ as desired. □

Remarks:

- Note that this proof is almost identical to our proof that every ideal of $\mathbb{Z}$ is principal. They both depended on the existence of a division algorithm.
- In addition to having only principal ideals, both of the rings $\mathbb{Z}$ and $\mathbb{F}[x]$ are integral domains. For this reason we will call them PIDs (Principal Ideal Domains).
- The definition of PIDs is not completely obvious, but it turns out that this is a very natural class of rings with many nice properties. In particular, every PID satisfies "unique prime factorization." We will discuss this next week.

Problem Set 8

0. One Step Ideal Test.

1. Addition vs. Multiplication. Prove that following properties hold in any ring.

(a) $0a = 0$,

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

(c) $(-a)(-b) = ab$.

(a) Let $(R, +, \times, 0, 1)$ be a ring. For any element $a \in R$ we have

$$0a + 0a = (0 + 0)a = 0a.$$

Then adding the inverse $-0a$ to both sides gives $0a = 0$.

(b) Consider any elements $a, b \in R$. From part (a) we have

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

Adding the inverse $-(ab)$ to both sides gives $(-a)b = -(ab)$. And since multiplication is commutative we also have

$$a(-b) = (-b)a = -(ba) = -(ab).$$

(c) Finally, observe from part (b) that

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

2. Characteristic of a Ring. Let R be a ring and let $R' \subseteq R$ be the smallest subring. Recall that there exists a unique ring homomorphism $\iota_R : \mathbb{Z} \to R$ from the integers.

(a) Prove that $R' \cong \mathbb{Z}/n\mathbb{Z}$ for some integer $n \geq 0$, which we call the *characteristic* of R:

$$\operatorname{char}(R) = n.$$

[Hint: Apply the First Isomorphism Theorem to ι_R.]

(b) If $\varphi : R \to S$ is any ring homomorphism prove that $\operatorname{char}(S)$ divides $\operatorname{char}(R)$. [Hint: By uniqueness we know that $\iota_S = \varphi \circ \iota_R$. Consider the kernel.]

(c) Next let R be an *integral domain*, which means that R has no *zero-divisors*:

$$\forall a, b \in R, (ab = 0) \Rightarrow (a = 0 \text{ or } b = 0).$$

In this case prove that $\operatorname{char}(R) = 0$ or $\operatorname{char}(R) = p$ for some prime p.

(d) Finally, let $\mathbb{F}$ ba a field and let $\mathbb{F}' \subseteq \mathbb{F}$ be the smallest subfield. Prove that

$$\mathbb{F}' \cong \mathbb{Q} \qquad \text{or} \qquad \mathbb{F}' \cong \mathbb{Z}/p\mathbb{Z} \text{ for some prime } p.$$

(a) Consider the ring homomorphism $\iota'_R : \mathbb{Z} \to R' \subseteq R$. Since this also defines a ring homomorphism $\iota_{R'} : \mathbb{Z} \to R$ we must have $\iota_{R'} = \iota_R$ by uniqueness. Furthermore, since $\operatorname{im} \iota_R \subseteq R' \subseteq R$ are subrings and since R' is the **smallest** subring of R we must have $\operatorname{im} \iota_R = R'$. Finally, since $\mathbb{Z}$ is a PID we know that $\ker \iota_R = n\mathbb{Z}$ for some unique integer $n \geq 0$ and it follows that

$$\mathbb{Z}/n\mathbb{Z} = \mathbb{Z}/\ker \iota_R \cong \operatorname{im} \iota_R = R'.$$

(b) For any rings R and S part (a) says that

$$\ker \iota_R = \text{char}(R)\mathbb{Z} \quad \text{and} \quad \ker \iota_S = \text{char}(S)\mathbb{Z}.$$

Now let $\varphi : R \to S$ be any ring homomorphism. Since the composition $\varphi \circ \iota_R$ is a homomorphism from $\mathbb{Z}$ to S we conclude that $\varphi \circ \iota_R = \iota_S$ by uniqueness. But then we have

$$\text{char}(R)\mathbb{Z} = \ker \iota_R \subseteq \ker \iota_S = \text{char}(S)\mathbb{Z},$$

which says that $\text{char}(S)$ divides $\text{char}(R)$.

[Remark: This result places restrictions on the existence of ring homomorphisms. For example, if $n \neq 0$ then there does not exist any ring homomorphism $\mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}$ because $0 = \text{char}(\mathbb{Z})$ does not divide $n = \text{char}(\mathbb{Z}/n\mathbb{Z})$.]

(c) Let R be an integral domain of characteristic $n \neq 0$ and assume for contradiction that n is **not** prime. By definition this means that there exist integers $a, b \in \mathbb{Z} - n\mathbb{Z}$ such that $n = ab$. Now consider the unique ring homomorphism $\iota_R : \mathbb{Z} \to R$. By assumption we have $\ker \iota_R = n\mathbb{Z}$ and $a, b \notin n\mathbb{Z}$, which implies that

$$\iota_R(a) \neq 0 \quad \text{and} \quad \iota_R(b) \neq 0.$$

But since ι_R is a homomorphism we also have

$$\iota_R(a)\iota_R(b) = \iota_R(ab) = \iota_R(n) = 0,$$

which contradicts the fact that R is an integral domain.

(d) Finally, let $\mathbb{F}$ be a field. If $\mathbb{F}' \subseteq \mathbb{F}$ be the smallest sub**field** and if $R \subseteq \mathbb{F}$ is the smallest sub**ring**, then since every subfield is a subring we must have $R \subseteq \mathbb{F}'$. Furthermore, since every subring of a field is an integral domain we know from parts (a) and (c) that $R \cong \mathbb{Z}/0\mathbb{Z} = \mathbb{Z}$ or $R \cong \mathbb{Z}/p\mathbb{Z}$ for some prime p. In the first case, since $\mathbb{Z} \cong R \subseteq \mathbb{F}'$ and since $\mathbb{F}'$ is closed under division we know that $\mathbb{Q} \cong \text{Frac}(R) \subseteq \mathbb{F}'$. Then since $\text{Frac}(R) \subseteq \mathbb{F}$ is a subfield and $\mathbb{F}' \subseteq \mathbb{F}$ is the **smallest** subfield we must have $\text{Frac}(R) = \mathbb{F}'$. In the second case, since $\mathbb{Z}/p\mathbb{Z} \cong R \subseteq \mathbb{F}$ is already a subfield, the same argument shows that $R = \mathbb{F}'$. □

[Remark: I was a bit sloppy at the end. The notation $\text{Frac}(R)$ represents the "field of fractions" of an integral domain R. You will study this on the next homework.]

3. Chinese Remainder Theorem, Part II. Let R be a ring. For any ideals $I, J \subseteq R$ we define the *product ideal*:

$$IJ := \text{intersection of all ideals that contain } \{ab : a \in I, b \in J\}.$$

(a) Prove that $IJ \subseteq I \cap J$.

(b) We say that $I, J \subseteq R$ are *coprime* if $I + J = R$. In this case show that $I \cap J \subseteq IJ$, and hence $IJ = I \cap J$. [Hint: Since $1 \in I + J$ we have $1 = x + y$ for some $x \in I$ and $y \in J$.]

(c) If $I, J \subseteq R$ are coprime, prove that the obvious map $(a + IJ) \mapsto (a + I, a + J)$ defines an isomorphism of rings:

$$\frac{R}{IJ} \cong \frac{R}{I} \times \frac{R}{J}.$$

[Hint: The hardest part is surjectivity. Use the same trick that you used when $R = \mathbb{Z}$.]

(a) Consider any $a \in I$ and $b \in J$. Since I is an ideal we have $ab \in I$ and since J is an ideal we have $ab \in J$, hence $ab \in I \cap J$. Then since the ideal $I \cap J$ contains the set $\{ab : a \in I, b \in J\}$ it must contain the **smallest** ideal that contains this set:

$$IJ \subseteq I \cap J.$$

(b) Next assume that the ideals I, J are coprime, i.e., that $I + J = R$. This implies that we can write $1 = x + y$ for some $x \in I$ and $y \in J$. But then for any $a \in I \cap J$ we have $xa \in IJ$ and $ay \in IJ$. Since IJ is closed under addition this implies that

$$a = 1a = (x + y)a = xa + ya = xa + ay \in IJ,$$

and hence $I \cap J \subseteq IJ$. [Remark: This proof uses the fact that R is commutative.]

(c) Note that the *direct product* of rings is defined componentwise. For any two ideals $I, J \subseteq R$ I claim that the map $\varphi(a + IJ) := (a + I, a + J)$ defines a ring homomorphism from R/IJ to the direct product of R/I and R/J:

$$\begin{array}{rcl} \varphi : R/IJ & \to & R/I \times R/J \\ a + IJ & \mapsto & (a + I, a + J). \end{array}$$

If this map is well-defined then it is obviously a ring homomorphism. So assume that we have $a + IJ = b + IJ$. By definition this means that $a - b \in IJ$. Then from part (a) we have $a - b \in I$ and $a - b \in J$, which implies that $a + I = b + I$ and $a + J = b + J$ as desired.

In the case that I and J are **coprime** I claim moreover that φ is a ring **isomorphism**.

Injective. Suppose that we have $(a + I, a + J) = (b + I, b + J)$. By definition this means that $a + I = b + I$ and $a + J = b + J$, hence $a - b \in I$ and $a - b \in J$. Then since I, J are coprime, part (b) implies that $a - b \in IJ$ and hence $a + IJ = b + IJ$.

Surjective. Given elements $a, b \in R$ we want to find some element $c \in R$ such that

$$\varphi(c + IJ) = (c + I, c + J) = (a + I, b + J).$$

Since I and J are coprime we can write $1 = x + y$ for some (non-unique) $x \in I$ and $y \in I$. Now define the element $c := ay + bx \in R$ and observe that we have

$$c - a = (ay + bx) - a = a(y - 1) + bx = -ax + bx = (-a + b)x \in I$$

and

$$c - b = (ay + bx) - b = ay + b(x - 1) = ay - by = (a - b)y \in J,$$

which implies that $c + I = a + I$ and $c + J = b + J$ as desired. □

4. Ring Isomorphism Theorems. Let R be a ring and let $I \subseteq R$ be an ideal.

(a) For any additive subgroup $I \subseteq S \subseteq R$ prove that

$$S \subseteq R \text{ is a subring} \quad \iff \quad S/I \subseteq R/I \text{ is a subring.}$$

(b) For any subring $S \subseteq R$ prove that we have an isomorphism of rings:

$$\frac{S}{S \cap I} \cong \frac{S + I}{I}.$$

[Hint: Consider the ring homomorphism $\varphi : S \to R/I$ defined by $\varphi(a) = a + I$.]

(c) For any additive subgroup $I \subseteq J \subseteq R$ prove that

$$J \subseteq R \text{ is an ideal} \quad \iff \quad J/I \subseteq R/I \text{ is an ideal,}$$

in which case we have an isomorphism of rings:

$$\frac{R/I}{J/I} \cong \frac{R}{J}.$$

[Hint: Consider the ring homomorphism $\varphi : R/I \to R/J$ defined by $\varphi(a + I) = a + J$.]

(a) Let $I \subseteq R$ be an ideal and let S be an additive subgroup satisfying $I \subseteq S \subseteq R$. If $S \subseteq R$ is a subring then I claim that the additive subgroup $S/I \subset R/I$ is also a subring. Indeed, since $1 \in S$ we have $1 + I \in S/I$. And for any $a, b \in S$ we have $ab \in S$ which implies that

$$(a + I)(b + I) = ab + I \in S/I.$$

Conversely, suppose that $S/I \subseteq R/I$ is a subring. Then I claim that the additive subgroup $S \subseteq R$ is also a subring. Indeed, since $1 + I \in S/I$ we have $1 + I = a + I$ for some $a \in S$. Then since $1 - a \in I \subseteq S$ and since S is an additive subgroup we have

$$1 = (1 - a) + a \in S.$$

Finally, let $a, b \in S$. Since S/I is a subring we must have $ab + I = (a + I)(b + I) \in S/I$ which implies that $ab + I = c + I$ for some $c \in S$. Then since $ab - c \in I \subseteq S$ and since S is closed under addition we conclude that

$$ab = (ab - c) + c \in S.$$

(b) Next let $I \subseteq R$ be an ideal and let $S \subseteq R$ be a subring. This time we do **not** assume that $I \subseteq S$. The function $\varphi : S \to R/I$ defined by $\varphi(a) = a + I$ is clearly a ring homomorphism with $\ker \varphi = S \cap I$. I claim that $\operatorname{im} \varphi = (S + I)/I$ as additive groups. Indeed, we have $\operatorname{im} \varphi \subseteq S/I \subseteq (S+I)/I$ by definition. Conversely, consider any coset $(a+b)+I \in (S+I)/I$ with $a \in S$ and $b \in I$. Then since $(a+b)-a = b \in I$ it follows that $(a+b)+I = a+I = \varphi(a) \in \operatorname{im} \varphi$. Finally, we conclude from the First Isomorphism Theorem that

$$\frac{S}{S \cap I} = \frac{S}{\ker \varphi} \cong \operatorname{im} \varphi = \frac{S + I}{I}.$$

[Remark: It follows automatically from this that $(S \cap I) \subseteq S$ is an ideal, and that $I \subseteq (S+I) \subseteq R$ is an ideal in a subring. Feel free to check these by hand if you want.]

(c) Let $I \subseteq R$ be an ideal and let $J \subseteq R$ be an additive subgroup satisfying $I \subseteq J$. If $J \subseteq R$ is an ideal then I claim that the additive subgroup $J/I \subseteq R/I$ is also an ideal. Indeed, consider any $a + I \in J/I$ and $b + I \in R/I$. Then since $a \in J$ we have $ab \in J$ and it follows that

$$(a + I)(b + I) = ab + I \in J/I.$$

Conversely, suppose that $J/I \subseteq R/I$ is an ideal. Then for all $a \in J$ and $b \in R$ we must have $(a + I)(b + I) = ab + I \in J/I$. This means that $ab + I = c + I$ (and hence $ab - c \in I$) for some $c \in J$. Finally, since $I \subseteq J$ and since $J \subseteq R$ is an additive subgroup we have

$$ab = (ab - c) + c \in J.$$

Next suppose that $I \subseteq J \subseteq R$ are both ideals. I claim that the rule $\varphi(a + I) := a + J$ defines a function from R/I to R/J. Indeed, for all $a, b \in R$ we have

$$a + I = b + I \quad \Longrightarrow \quad a - b \in I \quad \Longrightarrow \quad a - b \in J \quad \Longrightarrow \quad a + J = b + J.$$

Then the function $\varphi : R/I \to R/J$ is automatically a surjective ring homomorphism. Finally, I claim that $\ker \varphi = J/I$. Indeed, we have $J/I \subseteq \ker \varphi$. And if $\varphi(a + I) = a + J = 0 + J$ for some $a \in R$ then we must have $a \in J$ and hence $a + I \in J/I$. We conclude from the First Isomorphism Theorem that

$$\frac{R/I}{J/I} = \frac{R/I}{\ker \varphi} \cong \operatorname{im} \varphi = \frac{R}{J}.$$

□

5. Descartes' Factor Theorem Again. Let $E \supseteq R$ be any ring extension and let $f(x) \in R[x]$ be any polynomial with coefficients in R.

(a) For any element $\alpha \in E$ prove that $f(\alpha) = 0$ if and only if there exists a polynomial $h(x) \in E[x]$ with coefficients in E such that $f(x) = (x-\alpha)h(x)$ and $\deg(h) = \deg(f)-1$. [Hint: For all integers $n \geq 2$ observe that

$$x^n - \alpha^n = (x-\alpha)(x^{n-1} + x^{n-2}\alpha + \cdots + x\alpha^{n-2} + \alpha^{n-1}) \in E[x].$$

Now consider the polynomial $f(x) - f(\alpha) \in E[x]$.]

(b) **Counting Roots.** If E is an *integral domain*, use the result of part (a) to prove that any polynomial $f(x) \in R[x]$ has at most $\deg(f)$ distinct roots in E.

(c) **A Non-Example.** Let $E = R = \mathbb{Z}/8\mathbb{Z}$ and consider the polynomial $x^2 - 1$. How many roots does this polynomial have? Why does this not contradict part (b)?

(a) We already proved this using the division algorithm. Here's a different proof.

For all elements $\alpha \in E$ and integers $n \geq 0$ we define the polynomial $[n]_{x,\alpha} \in E[x]$ by

$$[n]_{x,a} := \begin{cases} 1 & \text{if } n = 0, \\ x^n + x^{n-1}\alpha + \cdots + x\alpha^{n-1} + \alpha^n & \text{if } n \geq 1. \end{cases}$$

Then for all $n \geq 1$ we have $x^n - \alpha^n = (x-\alpha) \cdot [n-1]_{x,\alpha}$.[25] Now consider any polynomial $f(x) = a_0 + a_1 x + \cdots + a_n x^n \in R[x]$ of degree n and observe that

$$\begin{aligned} f(x) - f(\alpha) &= (a_0 + a_1 x + \cdots + a_n x^n) - (a_0 + a_1\alpha + \cdots + a_n\alpha^n) \\ &= a_0(1-1) + a_1(x-\alpha) + a_2(x^2 - \alpha^2) + \cdots + a_n(x^n - \alpha^n) \\ &= (x-\alpha) \cdot (a_1[0]_{x,\alpha} + a_2[1]_{x,\alpha} + \cdots a_n[n-1]_{x,\alpha}) \\ &= (x-\alpha) \cdot g(x). \end{aligned}$$

One can see directly that $g(x) = (a_n x^{n-1} + \text{ lower terms}) \in E[x]$ so that $g(x)$ has degree $n-1$.

(b) Now let $E \supseteq R$ be an integral domain, i.e., a commutative ring with no zero-divisors. We will prove by induction on n that

" any polynomial $f(x) \in E[x]$ of degree n has at most n roots in E, "

which will imply that

" any polynomial $f(x) \in R[x]$ of degree n has at most n roots in E. "

For the base case, let $f(x) \in E[x]$ have degree 0. This means that $f(x) = c$ is a nonzero constant, hence we have $f(\alpha) = c \neq 0$ for all $\alpha \in E$. It follows that $f(x)$ has **no roots** in E.

It's not logically necessary, but let's also consider the case when $f(x) \in E[x]$ has degree 1. This means that $f(x) = ax + b$ for some $a, b \in R$ with $a \neq 0$. Now suppose that $\alpha, \beta \in E$ are

[25] This is a generalization of the "difference of squares" formula.

roots of $f(x)$, so that $a\alpha + b = 0 = a\beta + b$ and hence $a(\alpha - \beta) = 0$. Then since E is a domain and since $a \neq 0$ we conclude that $\alpha - \beta = 0$ and hence $\alpha = \beta$. This implies that $f(x)$ has **at most one root** in E.

For the induction step, fix $n \geq 1$ and assume that any polynomial in $E[x]$ of degree n has at most n roots in E. Now let $f(x) \in E[x]$ be any polynomial of degree $n+1$. If $f(x)$ has no roots in E then we are done because $0 \leq n+1$. So assume that $f(x)$ has a root $\alpha \in E$. Then from part (a) we have

$$f(x) = (x - \alpha)g(x) \text{ for some } g(x) \in E[x] \text{ of degree } n.$$

If $\alpha \neq \beta \in E$ is any other root of $f(x)$ then since E is commutative we have

$$0 = f(\beta) = (\alpha - \beta)g(\beta),$$

and since E is a domain this implies that $g(\beta) = 0$. But we know by assumption that $g(x)$ has at most n roots in E. Therefore $f(x)$ has at most $n+1$ roots in E. □

(c) Consider the polynomial $x^2 - 1 \in \mathbb{Z}/8\mathbb{Z}[x]$. The following table shows that this polynomial has exactly 4 roots in $\mathbb{Z}/8\mathbb{Z}$:

x	0	1	2	3	4	5	6	7
x^2-1	7	0	3	0	7	0	3	0

This doesn't contradict part (b) because $\mathbb{Z}/8\mathbb{Z}$ is not an integral domain. (It contains the zero-divisors $2, 4, 6 \in \mathbb{Z}/8\mathbb{Z}$.)

6. Prime and Maximal Ideals. Let R be a ring and let $I \subseteq R$ be an ideal.

(a) We say that I is a *maximal ideal* if

$$\text{for any ideal } J \subseteq R \text{ we have } (I \subsetneq J) \Rightarrow (J = R).$$

Prove that R/I is a field if and only if I is maximal.

(b) We say that I is a *prime ideal*

$$\text{for any } a, b \in R \text{ we have } (ab \in I) \Rightarrow (a \in I \text{ or } b \in I).$$

Prove that R/I is an integral domain if and only if I is prime.

(c) Prove that every maximal ideal is prime.

(d) Let $\mathbb{Z}[x]$ be the ring of polynomials over $\mathbb{Z}$ and consider the principal ideal

$$x\mathbb{Z}[x] = \{xf(x) : f(x) \in \mathbb{Z}[x]\}.$$

Prove that $x\mathbb{Z}[x]$ is prime but not maximal. [Hint: $\mathbb{Z}[x]/x\mathbb{Z}[x] \cong \mathbb{Z}$.]

(a) We proved in class that a (commutative) ring is a field if and only if it has exactly two ideals. Now let $I \subseteq R$ be an ideal in an arbitrary (commutative) ring, let $\mathscr{L}_{\text{idl}}(R, I)$ be the lattice of ideals $I \subseteq J \subseteq R$ and let $\mathscr{L}_{\text{idl}}(R/I)$ be the lattice of all ideals of R/I. You proved on Problem 4(c) that there is a bijection $\mathscr{L}_{\text{idl}}(R, I) \leftrightarrow \mathscr{L}_{\text{idl}}(R/I)$. Therefore we have

$$\begin{aligned} I \subseteq R \text{ is maximal} &\iff \#\mathscr{L}_{\text{idl}}(R, I) = 2 \\ &\iff \#\mathscr{L}_{\text{idl}}(R/I) = 2 \\ &\iff R/I \text{ is a field.} \end{aligned}$$

(b) Let $I \subseteq R$ be an ideal in an arbitrary (commutative) ring. First let I be prime. To show that R/I is a domain, consider any two nonzero elements $a + I \neq 0 + I$ and $b + I \neq 0 + I$. By definition this means that $a \notin I$ and $b \notin I$. Then since I is prime we have $ab \notin I$ and it follows that

$$(a + I)(b + I) = ab + I \neq 0 + I.$$

Conversely, let R/I be a domain. To show that I is prime consider any $a, b \notin I$. By definition this means that $a + I \neq 0 + I$ and $b + I \neq 0 + I$. Then since R/I is a domain we have $ab + I = (a + I)(b + I) \neq 0 + I$ and it follows that $ab \notin I$ as desired.

(c) Let $I \subseteq R$ be an ideal in an arbitrary (commutative) ring. Then from (a) and (b) we have

$$\begin{aligned} I \subseteq R \text{ is maximal} &\implies R/I \text{ is a field} \\ &\implies R/I \text{ is a domain} \\ &\implies I \subseteq R \text{ is prime.} \end{aligned}$$

(d) To show that the converse is **not** true in general, consider the ideal $x\mathbb{Z}[x] \subseteq \mathbb{Z}[x]$. I claim that this ideal is the kernel the surjective ring homomorphism $\varphi : \mathbb{Z}[x] \to \mathbb{Z}$ defined by substituting $x = 0$ into each polynomial:

$$\begin{array}{rccc} \varphi : & \mathbb{Z}[x] & \xrightarrow{x=0} & \mathbb{Z} \\ & \sum_i a_i x^i & \mapsto & a_0. \end{array}$$

Indeed, a polynomial $\sum_i a_i x^i$ is in the kernel of φ if and only if $a_0 = 0$, in which case we have

$$\sum_{i \geq 0} a_i x^i = x \left(\sum_{i \geq 1} a_i x^{i-1} \right) \in x\mathbb{Z}[x].$$

It follows from the First Isomorphism Theorem that

$$\frac{\mathbb{Z}[x]}{x\mathbb{Z}[x]} = \frac{\mathbb{Z}[x]}{\ker \varphi} \cong \operatorname{im} \varphi = \mathbb{Z}.$$

Then since $\mathbb{Z}$ is an integral domain that is not a field, it follows from (a) and (b) that $x\mathbb{Z}[x] \subseteq \mathbb{Z}[x]$ is a prime ideal that is not maximal. $\square$

[Remark: We just proved that the ideal $x\mathbb{Z}[x] \subseteq \mathbb{Z}[x]$ is not maximal, but we did not exhibit any ideals strictly between $x\mathbb{Z}[x]$ and $\mathbb{Z}[x]$. In fact, there are infinitely many. The Correspondence Theorem gives a bijection between ideals $x\mathbb{Z}[x] \subseteq I \subseteq \mathbb{Z}[x]$ and ideals $\varphi[I] \subseteq \mathbb{Z}$. If $\varphi[I] = n\mathbb{Z}$ then one can show that $I \subseteq \mathbb{Z}[x]$ consists of all polynomials in which the constant term (the coefficient of x^0) is a multiple of n.]

Week 17

Abstract ring theory is a big subject with too many definitions. But I believe that most of the theory is motivated by an analogy between the following two kinds of rings:

$$\textit{integers } \mathbb{Z} \quad \approx \quad \textit{polynomials in one variable over a field } \mathbb{F}[x]$$

This week we will explore the basics of this analogy up to the theory of unique prime factorization. I will try not to get carried away with pathological counterexamples.

The first similarity between $\mathbb{Z}$ and $\mathbb{F}[x]$ is the property of being "integral domains." We have already seen some of the following concepts but I want to collect them in one place for posterity.

Theorem (Definition of Integral Domains). Let R be a ring. Then the following three conditions are equivalent:

(D1) The zero ideal $0R \subseteq R$ is prime:

$$(a \notin 0R \text{ and } b \notin 0R) \Rightarrow (ab \notin 0R).$$

(D2) The ring R has no zero-divisors:

$$(a \neq 0 \text{ and } b \neq 0) \Rightarrow (ab \neq 0).$$

(D3) The ring R satisfies multiplicative cancellation:

$$(a \neq 0 \text{ and } ab = ac) \Rightarrow (b = c).$$

Any ring satisfying one (and hence all) of these conditions is called an *integral domain.* ///

Proof. Note that (D1)$\Leftrightarrow$(D2) because $a \in 0R \Leftrightarrow a = 0$.

(D2)⇒(D3): Suppose that R has no zero-divisors and consider $a, b, c \in R$ with $a \neq 0$ and $ab = ac$. Then we have

$$\begin{aligned} ab &= ac \\ ab - ac &= 0 \\ a(b-c) &= 0 \\ (b-c) &= 0 \qquad \text{because } a \neq 0 \\ b &= c. \end{aligned}$$

(D3)⇒ (D2): Let R satisfy multiplicative cancellation and assume for contradiction that there exists a pair of zero-divisors: $a \neq 0$, $b \neq 0$ and $ab = 0$. Then we have $ab = a0$ which since $a \neq 0$ implies that $b = 0$. Contradiction. □

Remarks:

- Recall from the homework that a general ideal $I \subseteq R$ is called *prime* when its complement is closed under multiplication:

  $$(a \notin I \text{ and } b \notin I) \Rightarrow (ab \notin I).$$

  If $I \subseteq J \subseteq R$ are any ideals, then one can show that "primeness" is preserved by the Correspondence Theorem:

  $$(J/I \subseteq R/I \text{ is prime}) \Leftrightarrow (J \subseteq R \text{ is prime}).$$

  It follows that

  $$(R/I \text{ is a domain}) \Leftrightarrow (I/I \subseteq R/I \text{ is prime}) \Leftrightarrow (I \subseteq R \text{ is prime}).$$

- As for the name "prime," suppose that $p, a, b \in \mathbb{Z}$ are integers with p prime. Then Euclid's Lemma says that

  $$(p|ab) \Rightarrow (p|a \text{ or } p|b).$$

  Note that this is the same as

  $$(ab \in p\mathbb{Z}) \Rightarrow (a \in p\mathbb{Z} \text{ or } b \in p\mathbb{Z}).$$

  In other words, $p\mathbb{Z} \subseteq \mathbb{Z}$ is a prime ideal. We will generalize this idea below.

- There are only two basic ways[26] that a ring can **fail** to be an integral domain, both of which are illustrated by the rings $\mathbb{Z}/n\mathbb{Z}$:

[26] I'm lying a bit here. The real theorem says that a ring with zero-divisors has a nilpotent or more than one minimal prime ideal. Having more than one minimal prime ideal is closely related to the existence of idempotents, and both of these are related to the idea of being "disconnected."

(1) We say that $a \in R$ is *nilpotent* if $a \neq 0$ and $a^m = 0$ for some minimal $m \geq 2$. Then $a \cdot a^{m-1} = 0$ shows that R is not an integral domain. For example, the element $2 \in \mathbb{Z}/2^k\mathbb{Z}$ is nilpotent.

(2) We say that $e \in R$ is *idempotent* if $e \notin \{0, 1\}$ and $e^2 = e$. Then $e(1-e) = 0$ shows that R is not an integral domain. For example, note that $e = 3 \in \mathbb{Z}/6\mathbb{Z}$ is idempotent with $1 - e = 4$ and $e(1-e) = 3 \cdot 4 = 0$. Ultimately this comes from the Chinese Remainder isomorphism

$$\begin{array}{ccc} \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z} & \to & \mathbb{Z}/6\mathbb{Z} \\ (a, b) & \mapsto & 3a + 4b \end{array}$$

since the elements $(1, 0), (0, 1)$ of the direct product $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z}$ are idempotents.

And that's probably enough about that.

///

As the name suggests, the ring $\mathbb{Z}$ is an integral domain. Some authors adopt (D2) or (D3) as an axiom for the integers, but usually these are proved from the following axioms of order:

- For all $a, b \in \mathbb{Z}$ exactly one of following holds: $a < b$, $a = b$ or $a > b$.
- For all $a, b, c \in \mathbb{Z}$ with $a < b$ we have $(c > 0) \Rightarrow (ac < bc)$ and $(c < 0) \Rightarrow (ac > bc)$.

These axioms, in turn, can be derived from Peano's Axioms. Basically, every property of the ring $\mathbb{Z}$ is a property of induction.

The fact that $\mathbb{F}[x]$ is an integral domain is implied by the following more general fact:

$$(R \text{ is an integral domain}) \Rightarrow (R[x] \text{ is an integral domain}).$$

Proof. Let R be an integral domain let $f(x), g(x) \in R[x]$ be nonzero polynomials. By definition this means that

$$f(x) = a_0 + a_1 x + \cdots + a_m x^m \quad \text{and} \quad g(x) = b_0 + b_1 x + \cdots + b_n x^n,$$

where the leading coefficients $a_m, b_n \in R$ are nonzero. But then we have

$$f(x)g(x) = a_m b_n x^{n+m} + \text{ lower terms.}$$

Since R is a domain it follows that $a_m b_n \neq 0$ and hence $f(x)g(x) \neq 0$. □

The second similarity between $\mathbb{Z}$ and $\mathbb{F}[x]$ is the fact that each has a "division algorithm." The following definition is a bit awkward and we only really care about it as a stepping-stone to the next theorem.

Definition of Euclidean Ring. We say that a ring R is *Euclidean* if there exists a well-ordered[27] set $(\Omega, \leq)$ and a function $\nu : R \to \Omega$ satisfying the following property:

[27] This means that every non-empty subset of Ω has a smallest element.

For all $a, b \in R$ with $b \neq 0$, there exist $q, r \in R$ such that

$$\begin{cases} a = qb + r, \\ \nu(r) < \nu(b). \end{cases}$$

We do not assume that the elements q, r are unique.

///

Note that $\mathbb{Z}$ is Euclidean with $\nu : \mathbb{Z} \to \{0 < 1 < 2 < \cdots\}$ given by the absolute value and that $\mathbb{F}[x]$ is Euclidean with $\nu : \mathbb{F}[x] \to \{-\infty < 0 < 1 < \cdots\}$ given by the degree. Here is the only reason we care about Euclidean rings.

Theorem (Euclidean Implies Principal Ideals). Let R be a ring. Then

$$(R \text{ is Euclidean}) \Rightarrow (\text{Every ideal } I \subseteq R \text{ is principal}).$$

///

Proof. We already proved this twice but let's do it one last time. If $I = 0R$ then we're done. Otherwise, choose $0 \neq m \in I$ with minimal $\nu(m)$. Then $mR \subseteq I$ and I claim that $mR = I$. Indeed, for any $a \in I$ we have

$$\begin{cases} a = qm + r, \\ \nu(r) < \nu(m). \end{cases}$$

If $r \neq 0$ then since $r \in I$ we get a contradiction to minimality. It follows that $r = 0$ and hence $a = qm \in mR$. Since this is true for all $a \in I$ we get $I \subseteq mR$. □

We summarize all of these properties with the following definition.

Definition of PIDs. Let R be a (commutative) ring. We say that R is a *principal ideal domain* (PID) if the following two properties hold:

- R is an integral domain,
- every ideal of R is principal.

///

In a subject with too many bad definitions, I believe that the definition of PIDs is good.[28]

---

[28] What do I mean by this? There are many theorems of the form PID ⇒ X for which we surprisingly also have X ⇒ PID or (X + something small) ⇒ PID. (The most basic example says that if $R[x]$ is a PID then R is a field.) This is rare in commutative algebra. When it happens you know you have a good definition.

Commutative algebra began with the study of number theory. In particular, many concepts of the subject were motivated by attempts to prove the following theorem:

> for all integers $x, y, z, n \in \mathbb{Z}$ with $n \geq 3$ and $xyz \neq 0$ we have
>
> $$x^n + y^n \neq z^n.$$

In 1637 (the same year as Descartes' *Géométrie*), Pierre de Fermat scribbled this result in the margin of his copy of Diophantus' *Arithmetica*, together with the following remark:

> *I have a truly marvelous demonstration of this proposition, which this margin is too narrow to contain.*

Through his correspondence Fermat tried to interest his contemporaries in number theory and he often challenged them by stating results without proof. However, this was the during the heart of the scientific revolution and it is likely that his contemporaries were more interested in applied areas of mathematics. For example, Christian Huygens made the following remark about Fermat's challenges in a 1658 letter to John Wallis:

> *There is no lack of better things for us to do.*[29]

It was approximately 100 years later when Leonhard Euler became interested in Fermat's number-theoretic ideas. Euler provided proofs for some of Fermat's unproved theorems (e.g., Fermat's Little Theorem) and he disproved others (e.g., Fermat's assertion that every number of the form $2^{2^n}+1$ is prime).[30] But the result stated above resisted Euler's attempts and hence became known as "Fermat's Last Theorem" (FLT). In 1847 Gabriel Lamé gave a false proof of FLT in which he assumed that the ring $\mathbb{Z}[e^{2\pi i/n}]$ always has unique prime factorization. However, in 1844 Ernst Kummer had discovered the surprising fact that

$$\textit{the ring } \mathbb{Z}[e^{2\pi i/23}] \textit{ does } \textbf{not} \textit{ have unique prime factorization.}$$

This motivated Kummer to develop a theory of "ideal prime factorization," which sadly did not repair Lamé's proof of FLT[31] but it did lead to the abstract theory of ideals.

In modern language, the motivation for Kummer's theory is to replace each **element** a in a ring R by the **principal ideal** $aR \subseteq R$ that it generates. The first observation is that

$$aR \supseteq bR \quad \Longleftrightarrow \quad a|b.$$

Proof. Suppose that $aR \supseteq bR$. Then since $b \in aR$ we have $ac = b$ for some $c \in R$. Conversely, suppose that $ac = b$ for some $c \in R$. Then for all $d \in R$ we have $bd = (ac)d = a(cd) \in aR$, and hence $bR \subseteq aR$. □

Many books use the following mnemonic:

[29] See Weil, *Number Theory: An approach through history from Hammurapi to Legendre*, page 119.

[30] On the homework you will investigate another result of Fermat on integers that can be expressed as a sum of two squares.

[31] It was eventually proved by Andrew Wiles in 1994.

to contain is to divide.

For the next observation we assume that R is an **integral domain**. Then we have

$$aR = bR \quad \iff \quad au = b \text{ for some unit } u \in R^\times.$$

Proof. One direction is done. For the other direction, suppose that $aR = bR$. If $a = 0$ or $b = 0$ then we have $a \cdot 1 = 0 = b$ with $1 \in R^\times$ as desired. So assume that $a \neq 0$. From the previous result we have $a|b$ and $b|a$, which implies that $ak = b$ and $b\ell = a$ for some $k, \ell \in R$. Finally, since R is an integral domain we have

$$\begin{aligned} b\ell &= a \\ ak\ell &= a \\ a(k\ell - 1) &= 0 \\ k\ell - 1 &= 0 \qquad\qquad (a \neq 0) \\ k\ell &= 1, \end{aligned}$$

and it follows that $k, \ell \in R^\times$. □

In general, we define a relation on the elements of a ring called "association."

Definition of Association. Let R be a ring. For all elements $a, b \in R$ we define

$$a \sim b \quad \iff \quad au = b \text{ for some unit } u \in R^\times.$$

[Exercise: Check that this is an equivalence relation.] When $a \sim b$ holds we say that a and b are *associates*. We will write $aR^\times := \{au : u \in R^\times\}$ for the equivalence class of $a \in R$, so that

$$a \sim b \quad \iff \quad aR^\times = bR^\times,$$

and we will use the notation $R/R^\times := \{aR^\times : a \in R\}$ for the set of equivalence classes. Warning: This is **not** a set of cosets because $R^\times \subseteq R$ is **not** a multiplicative subgroup. ///

Remarks:

- We always have $0R^\times = \{0\}$ and $1R^\times = R^\times$. If $a \in R$ is **non-zero-divisor** then the function $R^\times \to aR^\times$ defined by $u \mapsto au$ is a bijection $R^\times \leftrightarrow aR^\times$.
- If R is an **integral domain** then the previous result says that

  $$aR = bR \quad \iff \quad aR^\times = bR^\times.$$

  Hence we obtain a bijection between principal ideals and classes of associates.

- The most basic example is the ring of integers $\mathbb{Z}$ with group of units $\mathbb{Z}^\times$. In this case the classes of associates are

$$\mathbb{Z}/\mathbb{Z}^\times = \{\{0\}, \{1, -1\}, \{2, -2\}, \{3, -3\}, \ldots\}.$$

  By choosing the non-negative integer from each class we obtain the well-known bijection

$$\begin{array}{ccccccc} \mathbb{N} & \leftrightarrow & \mathbb{Z}/\mathbb{Z}^\times & \leftrightarrow & \{\text{principal ideals of } \mathbb{Z}\} & = & \{\text{ideals of } \mathbb{Z}\} \\ n & & & & & & n\mathbb{Z}. \end{array}$$

- If R is not an integral domain then strange things can happen. For example, consider the ring $\mathbb{Z}/12\mathbb{Z}$ with group of units $(\mathbb{Z}/12\mathbb{Z})^\times = \{1, 5, 7, 11\}$. One can check that

$$(\mathbb{Z}/12\mathbb{Z})/(\mathbb{Z}/12\mathbb{Z})^\times = \{\{0\}, \{1, 5, 7, 11\}, \{2, 10\}, \{3, 9\}, \{4, 8\}, \{6\}\}.$$

  Note that each class contains a unique divisor of 12. [Challenge Problem: Prove that the same holds in general for $\mathbb{Z}/n\mathbb{Z}$.]

///

Next we want to classify the principal ideals (hence all ideals) of the ring $\mathbb{F}[x]$. More generally, let R be any integral domain. Then I claim that

$$R[x]^\times = R^\times.$$

Proof. We think of $R \subseteq R[x]$ as the subring of constant polynomials. Note that this implies $R^\times \subseteq R[x]^\times$. Conversely, consider any elements $f(x), g(x) \in R[x]$ with $f(x)g(x) = 1$. Since R is a domain this implies that $\deg(f) + \deg(g) = \deg(1) = 0$. Then since degrees are non-negative we conclude that $\deg(f) = \deg(g) = 0$ and hence $f(x), g(x) \in R^\times$. □

[Remark: If the ring R has a nilpotent element then the inclusion $R^\times \subsetneq R[x]^\times$ is strict. For example, suppose we have $a \neq 0$ and $a^n = 0$ for some $n \geq 2$. Then the polynomial $1 - ax \in R[x]$ is a unit because

$$1 = 1 - 0x = 1 - a^n x^n = (1 - ax)(1 + ax + a^2x^2 + \cdots + a^{n-1}x^{n-1}).$$

This is yet another reason to prefer integral domains.]

This leads to the following theorem/definition.

Theorem (Definition of Monic Polynomials). We say that a polynomial is *monic* when its leading coefficient equals 1. If $\mathbb{F}$ is a field then we have a bijection

$$\{\text{ideals of } \mathbb{F}[x]\} \longleftrightarrow \{0\} \cup \{\text{monic polynomials in } \mathbb{F}[x]\}.$$

Proof. We know that $f(x)\mathbb{F}[x]$ is a PID, hence every ideal has the form $f(x)\mathbb{F}[x]$ for some polynomial $f(x) \in \mathbb{F}[x]$. If $f(x) \neq 0$ then we have $f(x) = a_0 + a_1x + \cdots a_nx^n$ for some $a_n \neq 0$. Since $\mathbb{F}$ is a field we can divide by a_n to obtain a **monic** polynomial $g(x) := f(x)/a_n$. Then since $f(x)|g(x)$ and $g(x)|f(x)$ we have $f(x)\mathbb{F}[x] = g(x)\mathbb{F}[x]$. Conversely, suppose that $g(x)\mathbb{F}[x] = h(x)\mathbb{F}[x]$ where $g(x)$ and $h(x)$ are both monic. Since $\mathbb{F}[x]$ is a domain this implies that $g(x)u(x) = h(x)$ for some unit $u(x) \in \mathbb{F}[x]^\times$, which from the previous result equals a **nonzero constant** $u(x) = u \in \mathbb{F}^\times$. But then

$$u = (\text{leading coefficient of } gu) = (\text{leading coefficient of } h) = 1,$$

and we conclude that $g(x) = h(x)$. □

In summary, for any field $\mathbb{F}$ we have an isomorphism of lattices:

$$\left\{ \begin{array}{c} \text{ideals of } \mathbb{F}[x] \text{ under} \\ \text{reverse containment} \end{array} \right\} \cong \left\{ \begin{array}{c} \text{monic polynomials} \cup 0 \\ \text{under divisibility} \end{array} \right\}.$$

To end this week I will prove that each of the rings $\mathbb{Z}$ and $\mathbb{F}[x]$ has "unique prime factorization." The hardest part of the proof is to find the correct definition of "prime."

Let's begin with some examples. In the ring $\mathbb{Z}$ we can factor 12 in many ways:

$$\begin{aligned} 12 &= 2 \cdot 2 \cdot 3 \\ &= 2 \cdot 3 \cdot 2 \\ &= (-2) \cdot (-3) \cdot 2 \\ &= 3 \cdot 2 \cdot (-1)(-1) \cdot 2 \cdot 1 \cdot 1 \cdot 1 \\ &\text{etc.} \end{aligned}$$

We say that 12 has prime factors ± 2 and ± 3 with multiplicities 2 and 1, respectively. We don't want to say that ± 1 are prime because this will ruin uniqueness.

In the ring $\mathbb{Q}[x]$ the polynomial $x^2 + 1$ has many trivial factorizations:

$$\begin{aligned} (x^2 + 1) &= \frac{1}{2}(2x^2 + 2) \\ &= 3\left(\frac{1}{3}x^2 + \frac{1}{3}\right) \\ &\text{etc.} \end{aligned}$$

But I claim that $x^2 + 1$ cannot be factored in a non-trivial way.

Proof. Recall that a polynomial $f(x) \in \mathbb{Q}[x]$ is non-constant if and only if $\deg(f) \geq 1$. Now suppose for contradiction that we have $x^2 + 1 = f(x)g(x)$ for some non-constant polynomials $f(x)g(x) \in \mathbb{Q}[x]$. Since $\deg(f) + \deg(g) = \deg(fg) = 2$ this implies that $\deg(f) = \deg(g) = 1$. In particular we must have $f(x) = \alpha x + \beta$ for some $\alpha, \beta \in \mathbb{Q}$ with $\alpha \neq 0$. But then $f(-\beta/\alpha) = 0$, which implies that

$$(-\beta/\alpha)^2 + 1 = f(-\beta/\alpha)g(-\beta/\alpha) = 0g(-\beta/\alpha) = 0.$$

But I claim that this is impossible. Indeed, since $-\beta/\alpha \in \mathbb{Q}$ we must have $-\beta/\alpha = c/d$ for some $c, d \in \mathbb{Z}$ with $d \neq 0$. But then we have

$$\begin{aligned} (-c/d)^2 + 1 &= 0 \\ c^2/d^2 &= -1 \\ c^2 &= -d^2 < 0, \end{aligned}$$

which contradicts the fact that $c^2 \geq 0$. □

We say that the polynomial $x^2 + 1$ is *irreducible over* $\mathbb{Q}$. However, the same polynomial is *reducible* over the field extension $\mathbb{C} \supseteq \mathbb{Q}$ because

$$x^2 + 1 = (x - i)(x + i).$$

Hopefully these examples will motivate the following definition. The definition is a bit awkward because it only applies to integral domains, and because it rather arbitrarily excludes the zero element and the units. We exclude units because they lead to silly non-unique factorizations. As for the zero element: I say that 0 is irreducible in a domain, but that's just my opinion.

Definition of Irreducible Elements. Let R be an integral domain. We say that an element $a \in R$ is *irreducible* if the principal ideal $aR \subseteq R$ is "nontrivial and maximal among principal ideals." In other words, when the following two conditions hold:

- $0R \subsetneq aR \subsetneq 1R$,
- for all $b \in R$ we have $(aR \subseteq bR \subseteq 1R) \Rightarrow (aR = bR \text{ or } bR = 1R)$.

Equivalently, these two conditions say that

- a is not zero and not a unit,
- if $a = bc$ for some $b, c \in R$ then we have $c \in R^\times$ $(aR = bR)$ or $b \in R^\times$ $(bR = 1R)$.

///

Our first goal is to show that every element in a domain can be factored as a (finite) product of irreducible elements, times a unit. Sadly, there exist pathological examples where this is false. For example, consider the ring of polynomials over $\mathbb{Q}$ with constant term in $\mathbb{Z}$:

$$\mathbb{Z} + x\mathbb{Q}[x] = \{f(x) \in \mathbb{Q}[x] : f(0) \in \mathbb{Z}\} \subseteq \mathbb{Q}[x].$$

The units of this ring are just $\mathbb{Z}^\times = \{\pm 1\}$. But the polynomial x can never be factored into irreducibles because

$$x = 2 \cdot \frac{x}{2} = 2 \cdot 2 \cdot \frac{x}{4} = 2 \cdot 2 \cdot 2 \cdot \frac{x}{8} = \cdots .$$

Ultimately, the problem is that we have an infinite increasing chain of principal ideals:

$$\langle x \rangle \subsetneq \langle x/2 \rangle \subsetneq \langle x/4 \rangle \subsetneq \langle x/8 \rangle \subsetneq \cdots .$$

We will show that this problem does not occur in a PID.

Theorem (PID ⇒ Factorization Terminates). Every element in a PID can be expressed as a (finite)[32] product of irreducible elements, times a unit.

Proof. Let $a_0 \in R$ be a non-zero non-unit[33] and assume for contradiction that the factoring process does not terminate. Then we obtain an infinite increasing chain of principal ideals:

$$a_0R \subsetneq a_1R \subsetneq a_2R \subsetneq a_3R \subsetneq \cdots .$$

I claim that the infinite union $I = \cup_i a_i R \subseteq R$ is an ideal. To see this, consider any $b, c \in I$ and $r \in R$. By definition there exist indices i, j such that $b \in a_iR$ and $c \in a_jR$. If $k = \max\{i, j\}$ then we have $b, c \in a_kR$ and it follows that

$$b - rc \in a_kR \subseteq I.$$

Since R is a PID we must have $I = aR$ for some $a \in I$. But then by definition we have $a \in a_kR$ for some k and it follows that

$$I = aR \subseteq a_kR \subsetneq a_{k+1}R \subseteq I.$$

Contradiction. □

[**Jargon.** We say that a ring R is *Noetherian* if it does not contain an infinite strictly increasing chain of ideals. We can rephrase the above theorem by saying that every PID is Noetherian. Emmy Noether showed that this condition is a very convenient abstract substitute for the well-ordering principle.]

And what about uniqueness? Consider the ring $\mathbb{Z}[\sqrt{-3}] = \{a + b\sqrt{-3} : a, b \in \mathbb{Z}\}$ and note that the element 4 has two seemingly different factorizations:

$$4 = 2 \cdot 2 = (1 + \sqrt{-3})(1 - \sqrt{-3}).$$

[32] Products in a general ring are necessarily finite. To definite an infinite product one would need some notion of "convergence," which we do not have.

[33] I say that 0 is irreducible (times 1) and that every unit is a product of itself times no irreducibles.

Indeed, you will show on the homework that the elements 2, $(1+\sqrt{-3})$ and $(1-\sqrt{-3})$ are irreducible but that 2 is not associate to either of $(1+\sqrt{-3})$ or $(1-\sqrt{-3})$. The problem is that none of these irreducible elements is "prime" in the following sense.

Definition of Prime Elements. Let $p \in R$ be a non-unit element of a ring. We say that

$$p \text{ is prime} \quad \iff \quad pR \subseteq R \text{ is a prime ideal.}$$

In other words, we say that $p \notin R^\times$ is prime if for all $a, b \in R$ we have

$$(p|ab) \quad \implies \quad (p|a \text{ or } p|b).$$

///

If R is a **domain** then for all elements $p \in R$ we have

$$p \text{ is prime} \quad \implies \quad p \text{ is irreducible.}$$

Proof. Let $p \in R$ be prime. If $p = 0$ then we're done because (in my opinion) 0 is irreducible. So let $p \neq 0$ and assume for contradiction that we have $p = ab$, where a and b are non-zero non-units. Since p is prime we have $p|a$ or $p|b$. Without loss of generality, suppose that $p|a$. Then the facts that $p|a$ and $a|p$ imply that $p = au$ for some unit $u \in R^\times$. Finally, we have

$$\begin{aligned} \not{p}b &= \not{p}u \\ b &= u, \end{aligned}$$

which contradicts the fact that $b \notin R^\times$. □

But irreducible elements are not prime in general. For example, consider again the domain $\mathbb{Z} + x\mathbb{Q}[x]$. I claim that the element x is irreducible but not prime. Indeed, a polynomial of degree 1 over a domain is always irreducible. To see that x is not prime, first note that x divides the product $2 \cdot (x/2)$. But $x \nmid 2$ (for reasons of degree) and $x \nmid (x/2)$ because $1/2$ is not in the ring. This example was necessarily rather artificial, because of the following theorem.

I call this Euclid's Lemma because he proved it for integers.

Theorem (Euclid's Lemma). Let R be a PID. Then for all $p \in R$ we have

$$p \text{ is prime} \quad \iff \quad p \text{ is irreducible.}$$

Fancy Proof. We already proved that every prime element in a domain is irreducible. For the other direction, let $p \in R$ be irreducible. By definition this means that the ideal $pR \subsetneq R$ is maximal among principal ideals. Since R is a PID this means that pR is maximal among all

ideals. Finally, since every maximal ideal is prime we conclude that pR is a prime ideal, hence $p \in R$ is a prime element. $\square$

Euclid's Proof. Let $p \in R$ be irreducible and assume that $p|ab$ for some $a, b, \in R$, say $pk = ab$. We will show that $p \nmid a$ implies $p|b$. So suppose that $a \notin pR$, which means that $pR \subsetneq pR + aR$. Since R is a PID we know that pR is a maximal ideal, hence $pR + aR = R$. In other words, there exist elements $x, y \in R$ such that $px + ay = 1$. Now multiply both sides by b to obtain

$$\begin{aligned} px + ay &= 1 \\ pbx + (ab)y &= b \\ pbx + (pk)y &= b \\ p(bx + ky) &= b. \end{aligned}$$

We conclude that $p|b$ as desired. $\square$

[**Jargon.** We say that elements $a, b \in R$ are *coprime* if $aR + bR = R$, or, in other words, if there exist elements $x, y \in R$ such that $ax + by = 1$. If the ring R is Euclidean then one can use the so-called Euclidean Algorithm to find some specific elements x, y.]

Finally, we can prove that every element in a PID has a unique factorization into irreducibles. The proof of this result should be taken as the motivation for all of the previous definitions.

Theorem (PID $\Rightarrow$ UFD). Let R be a PID. We showed previously that every element can be factored as a product of irreducibles, times a unit. Now suppose that we have

$$p_1 p_2 \cdots p_k = u \cdot q_1 q_2 \cdots q_\ell$$

where u is a unit and where the elements $p_1, \ldots, p_k, q_1, \ldots, q_\ell$ are irreducible. Then I claim that $k = \ell$ and we can relabel the factors so that $p_i \sim q_i$ are associate for all i.

Proof. We use induction on $\min\{k, \ell\}$. For the base case, let $\ell = 0$, so we have $p_1 \cdots p_k = u$. If $k \neq 0$ then $p_1|u$ and $u|p_1$ imply that p_1 is a unit, contradicting the fact that p_1 is irreducible. For the general case, assume that

$$p_1 p_2 \cdots p_k = u \cdot q_1 q_2 \cdots q_\ell.$$

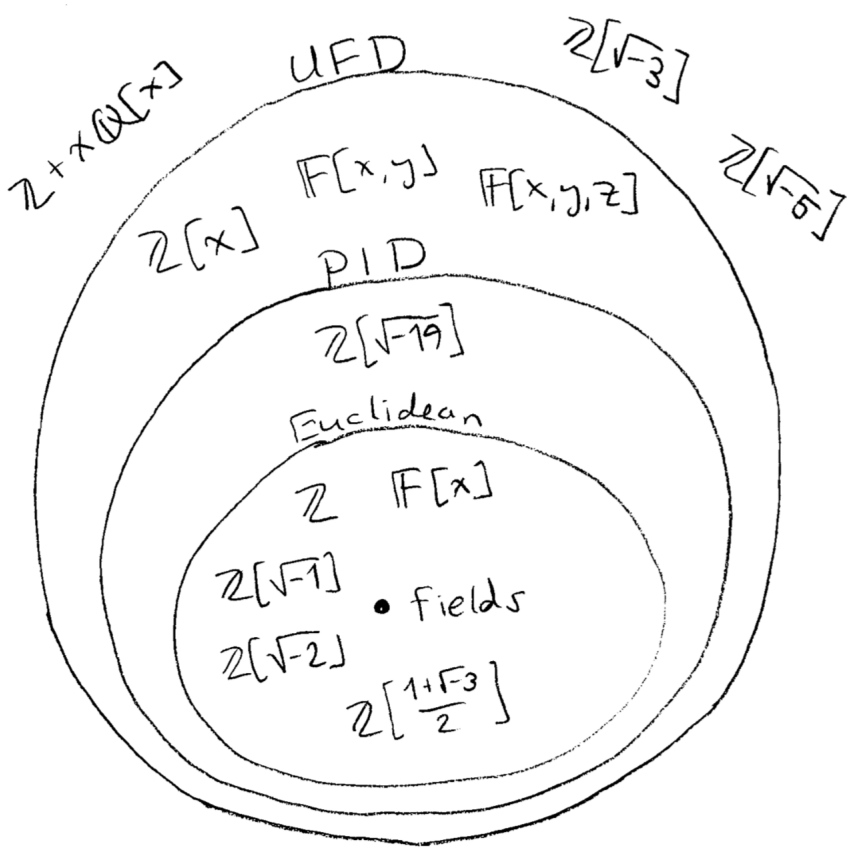
Since $p_1|q_1 \cdots q_\ell$ it follows from Euclid's Lemma that $p_1|q_i$ for some i. Without loss, suppose that $p_1|q_1$. Since q_1 is irreducible and p_1 is not a unit we must have $q_1 = p_1 u'$ for some unit $u' \in R^\times$. Then since $p_1 \neq 0$ we can cancel p_1 from both sides to obtain

$$\begin{aligned} \cancel{p_1} p_2 \cdots p_k &= uu' \cdot \cancel{p_1} q_2 \cdots q_\ell \\ p_2 \cdots p_k &= uu' \cdot q_2 \cdots q_\ell \end{aligned}$$

Since $\min\{k-1, \ell-1\} < \min\{k, \ell\}$, we have by induction that $k-1 = \ell - 1$ and we can reorder the factors so that $p_i \sim q_i$ are associate for all $i \geq 2$. $\square$

By the way, any domain that satisfies the conclusion of this theorem is called a *unique factorization domain* (UFD).

In summary, here is a sketch of the different kinds of integral domains:



Remarks:

- As you see, each of the inclusions is strict:

$$\{\text{Fields}\} \subsetneq \{\text{Euclidean Domains}\} \subsetneq \{\text{PIDs}\} \subsetneq \{\text{UFDs}\} \subsetneq \{\text{Domains}\}.$$

- Carl Friedrich Gauss proved that if R is a UFD then $R[x]$ is also a UFD. (Actually he proved that $\mathbb{Z}[x]$ is a UFD, but his proof works in general. The proof is not very difficult but it is too long to include in this course.) Since $\mathbb{F}[x]$ is a UFD we conclude that $\mathbb{F}[x,y] = \mathbb{F}[x][y]$ is also a UFD, and by induction the ring of polynomials in any (finite) number of variables over a field is a UFD.
- The rings $\mathbb{Z}[\sqrt{d}]$ for negative integers $d < 0$ are well understood. (Technically: If $d = 1 \mod 4$ then we should replace $\mathbb{Z}[\sqrt{d}]$ by the ring $\mathbb{Z}[(1+\sqrt{d})/2]$, which has nicer properties. One such nice property is that PID $\Leftrightarrow$ UFD.) Gauss proved that these rings have unique factorization when

  $$d \in \{-1, -2, -3, -7, -11, -19, -43, -67, -163\}.$$

  It is a modern theorem of Baker, Heenger and Stark that for all other $d < 0$ the ring of integers does **not** have unique factorization.
- For $d > 0$ it is an unsolved problem to determine when $\mathbb{Z}[\sqrt{d}]$ has unique factorization. Number theory is hard.
- Understanding polynomials in one variable over a field is easier, so we return to that topic next week.

Week 18

In the 1700s, Enlightment mathematicians such as Leonhard Euler took for granted the nature and existence of the basic number systems:

$$\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}.$$

These mathematical objects were regarded as real in the same sense that the physical world is real. In the 1800s, mathematicians began to doubt their intuition[34] and they started to ask for more rigorous definitions of basic concepts.

The first major post-Enlightment mathematician was Augustin-Louis Cauchy (1789–1857). His textbook *Cours d'Analyse* (1821) gave the first rigorous treatment of calculus. Later, in 1847, he gave the first rigorous definition of the complex numbers. Assuming that the real numbers exist, Cauchy defined the complex numbers as a quotient ring:

$$\mathbb{C} := \frac{\mathbb{R}[x]}{(x^2+1)\mathbb{R}[x]}.$$

Since the polynomial $x^2 + 1$ is an irreducible element of the PID $\mathbb{R}[x]$ we know that the ideal $(x^2+1)\mathbb{R}[x] \subseteq \mathbb{R}[x]$ is maximal, hence the quotient ring is a field. The role of the imaginary unit "$\sqrt{-1}$" is played by the coset of x:

$$\sqrt{-1} := x + (x^2+1)\mathbb{R}[x] = \{x + (x^2+1)f(x) : f(x) \in \mathbb{R}[x]\}.$$

[34]There were many reasons, but perhaps the most important was the discovery of self-consistent "non-Euclidean geometries" by Gauss, Bolyai and Lobachevsky.

Indeed, this coset is nonzero and we can check that it is a square root of the coset of -1:

$$\left(x + (x^2+1)\mathbb{R}[x]\right)^2 = \left(x^2 + (x^2+1)\mathbb{R}[x]\right) = \left(-1 + (x^2+1)\mathbb{R}[x]\right).$$

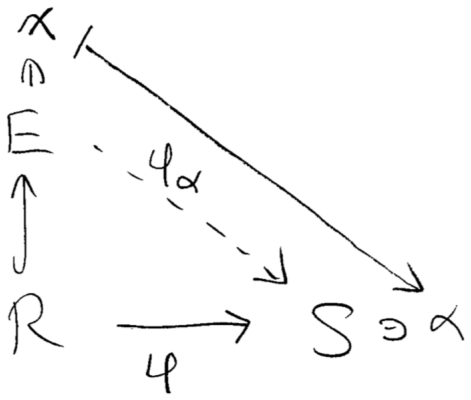
This level of abstraction was too much for Cauchy's contemporaries but it was later taken up in the 1880s by Leopold Kronecker. This week I will present Kronecker's proof that every polynomial over a field has a root in some extension field.

In order to do this we need a more modern definition of polynomials.

Theorem/Definition (Universal Property of Polynomials). Let R be a ring and let x be an indeterminate. We say that a ring E is a/the *free R-algebra generated by x* if the following two properties hold:

(1) We have $x \in E$ and we have a subring $R \subseteq E$ isomorphic to R.

(2) For any ring homomorphism $\varphi : R \to S$ and for any element $\alpha \in S$ there exists a unique ring homomorphism $\varphi_\alpha : E \to S$ satisfying $\varphi_\alpha(x) = \alpha$ and $\varphi_\alpha(a) = a$ for all $a \in R$. In other words, there exists a unique φ_α making the following diagram commute:



Note that the polynomial ring $R[x]$ satisfies (1) and (2). Furthermore, if E and E' are any two rings satisfying (1) and (2) then there exists a unique ring isomorphism $E \cong E'$ fixing the subset $R \cup \{x\}$. In this sense, we can say that

$$R[x] \text{ is the unique free } R\text{-algebra generated by } x$$

///

Proof. The fact that $R[x]$ satisfies (1) and (2) is easy. For (1) we can think of $R \subseteq R[x]$ as the subring of constant polynomials. For (2) suppose that $\varphi_\alpha : R[x] \to S$ is any ring homomorphism satisfying $\varphi_\alpha(x) = \alpha$ and $\varphi_\alpha(a) = a$ for all $a \in R$. Then we must have

$$\varphi_\alpha\left(\sum_i a_i x^i\right) = \sum_i \varphi_\alpha(a_i)\varphi_\alpha(x)^i = \sum_i \varphi(a_i)\alpha^i.$$

And it is easy to check that this function is indeed a ring homomorphism, as long as $\alpha \in S$ commutes with the elements of the subring $\operatorname{im}\varphi \subseteq S$.[35]

The surprising thing is that properties (1) and (2) determine the ring $R[x]$ up to isomorphism. I will only sketch the proof of this and you are free to skip it. So let E and E' be two rings satisfying (1) and (2). From (1) we note that $R \cup \{x\}$ is a subset of E and E' and from (2) we note that the identity maps $\mathrm{id}_E : E \to E$ and $\mathrm{id}_{E'} : E' \to E'$ are the **unique** ring homomorphisms $E \to E$ and $E' \to E'$ fixing the subset $R \cup \{x\}$. Also from (2) we know that there **exist** ring homomorphisms $\phi : E \to E'$ and $\psi : E' \to E$ fixing $R \cup \{x\}$. Since the compositions fix $R \cup \{x\}$, we conclude from the previous remark that $\phi \circ \psi = \mathrm{id}_E$ and $\psi \circ \phi = \mathrm{id}_{E'}$, hence $E \cong E'$. □

Remarks:

- This definition is an example of a "universal property." You will see another example on the homework when you study the field of fractions of a domain. The concept of universal properties was promoted by Saunders Mac Lane in the 1940s and 1950s, so it is strictly speaking a bit too modern for this course.[36] Nevertheless, the universal property of polynomials is important for Galois theory.
- Instead of telling us what a polynomial "is," the universal property tells us what a polynomial "does." Sometimes this is more important.
- To be explicit, the purpose of a polynomial is to be "evaluated." Given a ring homomorphism $\varphi : R \to S$, there exists a unique ring homomorphism $\varphi : R[x] \to S[x]$ acting by φ on the coefficients and sending $x \mapsto \alpha$. We denote this map by $f(x) \mapsto f^{\varphi}(x)$. Then for any element $\alpha \in S$ there exists a unique ring homomorphism $\varphi_\alpha : R[x] \to S[x] \to S$ defined by "evaluating the polynomial $f^{\varphi}(x)$ at the argument $x = \alpha$."

///

Last time we proved that for any ring homomorphism $\varphi : R \to S$ and for any element $\alpha \in S$ there exists a ring homomorphism $\varphi_\alpha : R[x] \to S$ acting on the coefficients by φ and sending $x \mapsto \alpha$. Today we will focus on the special case when φ is just the identity homomorphism on a subring $R \subseteq S$.

Definition of Evaluation. Let $R \subseteq S$ be a subring and let $\mathrm{id} : R \hookrightarrow S$ be the restriction of the identity homomorphism $S \to S$. Then for any element $\alpha \in S$ we have a homomorphism

[35] For us this condition is automatic because we assume that S is a commutative ring.

[36] Saunders Mac Lane was an American mathematician who studied at Göttingen in the 1930s. After the war he co-founded with Samuel Eilenberg the subject of *category theory*. The extreme abstraction of categories was shortly taken up by French mathematicians and became part of the mathematical mainstream in the 1960s.

$\mathrm{id}_\alpha : R[x] \to S$ defined by fixing the coefficients and sending $x \mapsto \alpha$. For any polynomial $f(x) = \sum_i a_i x^i \in R[x]$ we will use the notation

$$f(\alpha) := \mathrm{id}_\alpha(f(x)) = \mathrm{id}_\alpha\left(\sum_i a_i x^i\right) = \sum_i a_i \alpha^i.$$

We call id_α the *evaluation homomorphism at* $x = \alpha$. ///

I claim that the image of the evaluation $\mathrm{id}_\alpha : R[x] \to S$ is equal to the smallest subring of S that contains the set $R \cup \{\alpha\}$:

$$\mathrm{im}(\mathrm{id}_\alpha) = R[\alpha] \subseteq S.$$

Proof. Since $R[\alpha]$ is a subring containing $R \cup \{\alpha\}$ and since every element of $\mathrm{im}(\mathrm{id}_\alpha)$ is formed from $R \cup \{\alpha\}$ using a finite number of ring operations, we have $\mathrm{im}(\mathrm{id}_\alpha) \subseteq R[\alpha]$. Conversely, since $\mathrm{im}(\mathrm{id}_\alpha) \subseteq S$ is a subring containing $R \cup \{\alpha\}$ and since $R[\alpha] \subseteq S$ is the **smallest** subring containing $R \cup \{\alpha\}$ we must have $R[\alpha] \subseteq \mathrm{im}(\mathrm{id}_\alpha)$. □

The kernel of an evaluation is more complicated, so we will restrict our attention to polynomials over a **field**. Then there are two basic cases, called *transcendental* and *algebraic*.

Definition of the Minimal Polynomial. Let $\alpha \in \mathbb{E} \supseteq \mathbb{F}$ be an element of a field extension and consider the evaluation homomorphism $\mathrm{id}_\alpha : \mathbb{F}[x] \to \mathbb{E}$. By the First Isomorphism Theorem we have

$$\frac{\mathbb{F}[x]}{\ker(\mathrm{id}_\alpha)} \cong \mathrm{im}(\mathrm{id}_\alpha) = \mathbb{F}[\alpha] \subseteq \mathbb{E}.$$

If $\ker(\mathrm{id}_\alpha) = 0$ then we say that α *is transcendental over* $\mathbb{F}$ and we obtain an isomorphism

$$\mathbb{F}[x] \cong \mathbb{F}[\alpha].$$

In other words, we can think of a transcendental element as a "variable."

If $\ker(\mathrm{id}_\alpha) \neq 0$ then we say that α *is algebraic over* $\mathbb{F}$. In this case, since $\mathbb{F}[x]$ is a PID there exists a unique monic polynomial $m_{\alpha/\mathbb{F}}(x) \in \mathbb{F}[x]$ satisfying

$$\ker(\mathrm{id}_\alpha) = \langle m_{\alpha/\mathbb{F}}(x)\rangle = m_{\alpha/\mathbb{F}}(x)\mathbb{F}[x] = \{m_{\alpha/\mathbb{F}}(x)g(x) : g(x) \in \mathbb{F}[x]\}.$$

We call this $m_{\alpha/\mathbb{F}}(x) \in \mathbb{F}[x]$ the *minimal polynomial of* α *over* $\mathbb{F}$.

In less algebraic terms we can say that $m_{\alpha/\mathbb{F}}(x) \in \mathbb{F}[x]$ is the unique monic polynomial of minimal degree that has α as a root. Indeed, we have $m_{\alpha/\mathbb{F}}(\alpha) = 0$ by definition. And for any polynomial $f(x) \in \mathbb{F}[x]$ we have

$$f(\alpha) = 0 \quad \Leftrightarrow \quad f(x) \in \ker(\mathrm{id}_\alpha) \quad \Leftrightarrow \quad f(x) \in \langle m_{\alpha/\mathbb{F}}(x)\rangle \quad \Leftrightarrow \quad m_{\alpha/\mathbb{F}}(x) | f(x).$$

If $f(\alpha) = 0$ with $f(x) \in \mathbb{F}[x]$ monic, then since $m_{\alpha/\mathbb{F}}(x)|f(x)$ and $f(x) \neq 0$ we conclude that $\deg(m_{\alpha/\mathbb{F}}) \leq \deg(f)$. ///

The most basic case occurs when $\alpha \in \mathbb{F}$. In this case I claim that the minimal polynomial is just $m_{\alpha/\mathbb{F}}(x) = x - \alpha \in \mathbb{F}[x]$. Indeed, we already know from Descartes' Factor Theorem that for all $f(x) \in \mathbb{F}[x]$ we have

$$f(\alpha) = 0 \quad \Longleftrightarrow \quad (x - \alpha)|f(x).$$

Thus we can view the concept of the minimal polynomial as some kind of generalization of Descartes' Theorem.

The following theorem is one of our main tools for studying field extensions.

The Minimal Polynomial Theorem. Let $\mathbb{E} \supseteq \mathbb{F}$ be a field extension and let $\alpha \in \mathbb{E}$ be algebraic over $\mathbb{F}$ with minimal polynomial $m_{\alpha/\mathbb{F}}(x) \in \mathbb{F}[x]$.

(1) If $f(x) \in \mathbb{F}[x]$ is any irreducible monic polynomial with $f(\alpha) = 0$ then $f(x) = m_{\alpha/\mathbb{F}}(x)$.

(2) The minimal polynomial $m_{\alpha/\mathbb{F}}(x) \in \mathbb{F}[x]$ is irreducible and it follows that $\mathbb{F}[\alpha]$ is a field. In other words, we have
$$\mathbb{F}[\alpha] = \mathbb{F}(\alpha).$$

(3) If $\deg(m_{\alpha/\mathbb{F}}) = n$ then I claim that $\{1, \alpha, \ldots, \alpha^{n-1}\}$ is a basis for the vector space $\mathbb{F}(\alpha)$ over $\mathbb{F}$. It follows that
$$[\mathbb{F}(\alpha)/\mathbb{F}] = \deg(m_{\alpha/\mathbb{F}}).$$

///

Proof. (1) Suppose that we have $f(\alpha) = 0$ for some $f(x) \in \mathbb{F}[x]$. By definition this means that $m_{\alpha/\mathbb{F}}(x)|f(x)$. If $f(x)$ is irreducible then this implies that $f(x) = \lambda \cdot m_{\alpha/\mathbb{F}}(x)$ for some nonzero constant $\lambda \in \mathbb{F}$ and if $f(x)$ is monic then we must have $\lambda = 1$, hence $f(x) = m_{\alpha/\mathbb{F}}(x)$.

(2) To prove that $m_{\alpha/\mathbb{F}}(x)$ is irreducible, suppose that we have $m_{\alpha/\mathbb{F}}(x) = f(x)g(x)$ for some non-constant polynomials $f(x), g(x) \in \mathbb{F}[x]$. In particular, we have $\deg(f) < \deg(m_{\alpha/\mathbb{F}})$ and $\deg(g) < \deg(m_{\alpha/\mathbb{F}})$. Then evaluating at $x = \alpha$ gives

$$f(\alpha)g(\alpha) = m_{\alpha/\mathbb{F}}(\alpha) = 0.$$

Since $\mathbb{F}$ is a field this implies that $f(\alpha) = 0$ or $g(\alpha) = 0$. Without loss of generality suppose that $f(\alpha) = 0$, so that $m_{\alpha/\mathbb{F}}(x)|f(x)$. But then since $f(x) \neq 0$ we must have $\deg(m_{\alpha/\mathbb{F}}) \leq \deg(f)$, which is a contradiction.

Now recall that $\mathbb{F}[\alpha]$ and $\mathbb{F}(\alpha)$ are by definition the smallest sub**ring** and sub**field** of $\mathbb{E}$ that contain the set $\mathbb{F} \cup \{\alpha\}$. Since every subfield is a subring we have $\mathbb{F}[\alpha] \subseteq \mathbb{F}(\alpha)$. Conversely,

since $m_{\alpha/\mathbb{F}}(x) \in \mathbb{F}[x]$ is irreducible in a PID we know that $\langle m_{\alpha/\mathbb{F}}(x)\rangle \subseteq \mathbb{F}[x]$ is a maximal ideal and hence

$$\frac{\mathbb{F}[x]}{\langle m_{\alpha/\mathbb{F}}(x)\rangle} = \frac{\mathbb{F}[x]}{\ker(\mathrm{id}_\alpha)} \cong \mathrm{im}(\mathrm{id}_\alpha) = \mathbb{F}[\alpha] \text{ is a field.}$$

Then since $\mathbb{F}[\alpha] \subseteq \mathbb{E}$ is a subfield that contains $\mathbb{F} \cup \{\alpha\}$ we conclude that $\mathbb{F}(\alpha) \subseteq \mathbb{F}[\alpha]$.

(3) Let $\deg(m_{\alpha/\mathbb{F}}) = n$ and consider the set $\{1, \alpha, \alpha^2, \ldots, \alpha^{n-1}\} \subseteq \mathbb{F}(\alpha)$. To show that this set **spans** $\mathbb{F}(\alpha)$ over $\mathbb{F}$ we observe that every element of $\mathbb{F}(\alpha) = \mathbb{F}[\alpha] = \mathrm{im}(\mathrm{id}_\alpha)$ has the form $f(\alpha)$ for some polynomial $f(x) \in \mathbb{F}[x]$. From the Division Theorem there exist polynomials $q(x), r(x) \in \mathbb{F}[x]$ with

$$f(x) = m_{\alpha/\mathbb{F}}(x)q(x) + r(x) \quad \text{and} \quad \deg(r) < \deg(m_{\alpha/\mathbb{F}}).$$

Since $\deg(r) < \deg(m_{\alpha/\mathbb{F}}) = n$ we can write $r(x) = a_0 + a_1x + \cdots + a_{n-1}x^{n-1}$ for some $a_0, \ldots, a_{n-1} \in \mathbb{F}$ and then evaluating at $x = \alpha$ gives

$$\begin{aligned} f(\alpha) &= m_{\alpha/\mathbb{F}}(\alpha)q(\alpha) + r(\alpha) \\ &= 0 \cdot q(\alpha) + r(\alpha) \\ &= r(\alpha) \\ &= a_0 + a_1\alpha + \cdots + a_{n-1}\alpha^{n-1}. \end{aligned}$$

Finally, to show that $\{1, \alpha, \alpha^2, \ldots, \alpha^{n-1}\}$ is **independent** over $\mathbb{F}$, suppose that we have

$$a_0 + a_1\alpha + \cdots + a_{n-1}\alpha^{n-1} = 0 \quad \text{for some } a_0, \ldots, a_{n-1} \in \mathbb{F}.$$

In other words, suppose we have $f(\alpha) = 0$ for some polynomial $f(x) = a_0 + a_1x + \cdots + a_{n-1}x^{n-1} \in \mathbb{F}[x]$. In this case I claim that $f(x) = 0$ is the zero polynomial and hence $a_0 = a_1 = \cdots = a_{n-1} = 0$. Indeed, since $f(\alpha) = 0$ we have $m_{\alpha/\mathbb{F}}(x) | f(x)$. If $f(x) \neq 0$ then this implies that $\deg(m_{\alpha/\mathbb{F}}) \leq \deg(f)$, which contradicts the fact that $\deg(f) < n$. □

You investigated a special case of this theorem on a previous homework. Now we will relate this example to the theory of minimal polynomials.

Example: Quadratic Field Extensions. Let $\mathbb{E} \supseteq \mathbb{F}$ be a field extension and consider an element $\alpha \in \mathbb{E}$ with $\alpha^2 \in \mathbb{F}$ and $\alpha \notin \mathbb{F}$. Then I claim that $x^2 - \alpha^2 \in \mathbb{F}[x]$ is the minimal polynomial of α over $\mathbb{F}$. Since $x^2 - \alpha^2$ is monic and has α as a root, it suffices to show that this polynomial is irreducible. So assume for contradiction that we have $x^2 - \alpha^2 = f(x)g(x)$ for some non-constant polynomials $f(x), g(x) \in \mathbb{F}[x]$. This implies that $\deg(f) = \deg(g) = 1$. If $f(x) = ax + b$ with $a, b \in \mathbb{F}$ and $a \neq 0$ then we see that $-b/a \in \mathbb{F}$ is a root of $f(x)$, hence also a root of $x^2 - \alpha^2$. But from Descartes' Factor Theorem we know that $\pm\alpha$ are the **only** roots of $x^2 - \alpha^2$ and by assumption these are not in $\mathbb{F}$.

Then since $m_{\alpha/\mathbb{F}}(x) = x^2 - \alpha^2$ has degree 2 we conclude from the Minimal Polynomial Theorem that $[\mathbb{F}(\alpha)/\mathbb{F}] = 2$ with basis $\{1, \alpha\}$, and it follows that

$$\mathbb{F}(\alpha) = \{a + b\alpha : a, b \in \mathbb{F}\}.$$

Recall that we originally proved this result by "rationalizing the denominator." The new method is better because it extends to more general situations. ///

And here is one of those more general situations.

Example: The Minimal Polynomial of $\sqrt[3]{2}$ Over $\mathbb{Q}$. Let $\alpha = \sqrt[3]{2} \in \mathbb{R}$ be the unique real cube root of 2. I claim that the minimal polynomial of α over $\mathbb{Q}$ is

$$m_{\alpha/\mathbb{Q}}(x) = x^3 - 2.$$

Proof. Since α is a root of $x^3 - 2$, we only need to check that $x^3 - 2 \in \mathbb{Q}[x]$ is irreducible. So suppose for contradiction that $x^3 - 2 = f(x)g(x)$ for some non-constant polynomials $f(x), g(x) \in \mathbb{Q}[x]$. By considering degrees we must have $\deg(f) = 1$ or $\deg(g) = 1$, and then it follows as in the previous example that $x^3 - 2$ has a root in $\mathbb{Q}$. To be specific, suppose that $(a/b)^3 - 2 = 0$ for some $a, b \in \mathbb{Z}$ with $\gcd(a, b) = 1$. Then we have

$$\begin{aligned} a^3/b^3 - 2 &= 0 \\ a^3/b^3 &= 2 \\ a^3 &= 2b^3. \end{aligned}$$

Since $b|a^3$ with $\gcd(a, b) = 1$ we must have $b \in \{\pm 1\}$. And since $a|2b^3$ with $\gcd(a, b) = 1$ we must have $a \in \{\pm 1, \pm 2\}$. It follows that

$$a/b \in \{\pm 1, \pm 2\},$$

and one can check that that none of these is a root of $x^3 - 2$. [Remark: This method is called the Rational Root Test. You will examine the general version on the homework.] □

Finally, since $m_{\alpha/\mathbb{Q}}(x) = x^3 - 2$ has degree 3 we conclude from the Minimal Polynomial Theorem that $\{1, \alpha, \alpha^2\}$ is a basis for the field $\mathbb{Q}(\alpha)$ over $\mathbb{Q}$, and it follows that

$$\mathbb{Q}(\alpha) = \{a + b\alpha + c\alpha^2 : a, b, c \in \mathbb{Q}\}.$$

In particular, we conclude that the set on the right is a **field**. However, you would find it very difficult to "rationalize the denominator" by hand:

$$\frac{1}{a + b\alpha + c\alpha^2} = (?) + (?)\alpha + (?)\alpha^2.$$

I set up a 3×3 linear system and used my computer to find that

$$\frac{1}{a + b\alpha + c\alpha^2} = \left(\frac{a^2 - 2bc}{\Delta}\right) + \left(\frac{2c^2 - ab}{\Delta}\right)\alpha + \left(\frac{b^2 - ac}{\Delta}\right)\alpha^2,$$

with $\Delta = a^3 + 2b^3 + 4c^3 - 6abc$. This formula is probably not useful for anything. ///

Let $f(x) \in \mathbb{F}[x]$ be any irreducible polynomial and suppose that there exists a field extension $\mathbb{E} \supseteq \mathbb{F}$ and an element $\alpha \in \mathbb{E}$ such that $f(\alpha) = 0$. Last time we proved that $\langle f(x) \rangle$ is the kernel of the evaluation homomorphism $\mathrm{id}_\alpha : \mathbb{F}[x] \to \mathbb{E}$. Then since $\langle f(x) \rangle \subseteq \mathbb{F}[x]$ is a maximal ideal we obtain an isomorphism of fields:

$$\mathbb{F}[x]/\langle f(x) \rangle = \mathbb{F}[x]/\ker(\mathrm{id}_\alpha) \cong \mathrm{im}(\mathrm{id}_\alpha) = \mathbb{F}(\alpha) \subseteq \mathbb{E}.$$

Furthermore, we observe that the coset $x + \langle f(x) \rangle \in \mathbb{F}[x]/\langle f(x) \rangle$ gets identified with the root $\alpha \in \mathbb{E}$. But what if we don't know any roots of $f(x)$? Today we will reverse this construction and use it to **create a root** for any given polynomial over a field.

Leopold Kronecker is known as a "constructivist" mathematician, meaning that he would not accept the existence of a mathematical entity unless he could give a finite algorithm for constructing it. His contemporary Dedekind, on the other hand, was happy to accept infinite sets defined by implicit conditions. Dedekind's point of view eventually became standard but the following construction of Kronecker is still imporant. Kronecker's original goal was to give a concrete way to work with algebraic irrational numbers such as $\sqrt{2}$.[37]

Kronecker's Theorem (Every Polynomial Has a Root Somewhere). Let $\mathbb{F}$ be a field[38] and let $f(x) \in \mathbb{F}[x]$ be a polynomial of degree ≥ 1. Then there exists a field extension $\mathbb{E} \supseteq \mathbb{F}$ in which $f(x)$ has a root.

Proof. Let $f(x) \in \mathbb{F}[x]$ be a polynomial of degree ≥ 1. Since $\mathbb{F}[x]$ is a PID we know that we can write $f(x) = p(x)g(x)$ with $p(x), g(x) \in \mathbb{F}[x]$ and with $p(x)$ **irreducible**. Suppose that we can find an extension $\mathbb{E} \supseteq \mathbb{F}$ and an element $\alpha \in \mathbb{E}$ such that $p(\alpha) = 0$. Then this α is also a root of $f(x)$ because "evaluation at $x = \alpha$" is a ring homomorphism:

$$f(\alpha) = p(\alpha)g(\alpha) = 0g(\alpha) = 0.$$

In order to construct such a field $\mathbb{E}$ and element $\alpha \in \mathbb{E}$ we consider the principal ideal

$$\langle p(x) \rangle = p(x)\mathbb{F}[x] = \{p(x)g(x) : g(x) \in \mathbb{F}[x]\}.$$

Again, since $p(x) \in \mathbb{F}[x]$ is irreducible in a PID we know that $\langle p(x) \rangle \subseteq \mathbb{F}[x]$ is a **maximal ideal**, hence the quotient ring is a **field**. We will call it $\mathbb{E}$:

$$\mathbb{E} := \mathbb{F}[x]/\langle p(x) \rangle.$$

[37] The method doesn't help with transcendental numbers such as π. It is said that Kronecker did not believe in such numbers.

[38] The result also applies to polynomials over an integral domain R by taking $\mathbb{F} = \mathrm{Frac}(R) \supseteq R$. As always, non-domains are a different story. I'm starting to think that the concept of "rings" is too broad.

I claim that the coset $\alpha = x + \langle p(x) \rangle \in \mathbb{E}$ is the desired root of $f(x)$. Wait a minute, that sounds ridiculous. How can a coset be a root?

First we need to view $\mathbb{F}$ as a subfield of $\mathbb{E}$. So consider the following ring homomorphism:

$$\begin{array}{rcl} \iota : \mathbb{F} & \to & \mathbb{E} \\ a & \mapsto & a + \langle p(x) \rangle. \end{array}$$

Note that this function is injective. Indeed, if $\iota(a) = \iota(b)$ for some $a, b \in \mathbb{F}$ then we have

$$\begin{aligned} \iota(a) &= \iota(b) \\ a + \langle p(x) \rangle &= b + \langle p(x) \rangle \\ a - b &\in \langle p(x) \rangle \\ a - b &= p(x)g(x) \text{ for some } g(x) \in \mathbb{F}[x]. \end{aligned}$$

But note that $\deg(p) \geq 1$ (since $p(x)$ is an irreducible polynomial) and $\deg(a - b) \leq 0$ (since $a - b$ is a constant). Then since $\deg(pg) = \deg(p) + \deg(g)$, the only possible solution is $g(x) = 0$, which implies that $a - b = 0$ as desired.

From the First Isomorphism Theorem we conclude that $\operatorname{im} \iota = \{a + \langle p(x) \rangle : a \in \mathbb{F}\} \subseteq \mathbb{E}$ is a subfield isomorphic to $\mathbb{F}$. Now pay close attention to the following remark:

we choose to **identify** $\mathbb{F}$ *with the subfield* $\{a + \langle p(x) \rangle : a \in \mathbb{F}\} \subseteq \mathbb{E}$.

Now it only remains to show that the element $\alpha = x + \langle p(x) \rangle \in \mathbb{E}$ is a root of the polynomial $p(x) \in \mathbb{F}[x]$. More generally, consider any polynomial $h(x) = \sum_i a_i x^i = \sum_i (a_i + \langle p(x) \rangle) x^i$. Then by definition we have

$$\begin{aligned} h(\alpha) &= h(x + \langle p(x) \rangle) \\ &= \sum_i (a_i + \langle p(x) \rangle)(x + \langle p(x) \rangle)^i \\ &= (\sum_i a_i x^i) + \langle p(x) \rangle \\ &= h(x) + \langle p(x) \rangle. \end{aligned}$$

In particular, this implies that $p(\alpha) = p(x) + \langle p(x) \rangle = 0 + \langle p(x) \rangle$ as desired.[39] □

I apologize for level of abstraction in that proof. Kronecker's Theorem is similar in spirit to the construction of fractions from integers, or the construction of real numbers from fractions. At first we think of a fraction as an infinite equivalence class of ordered pairs of integers.

[39] If you don't like that, here's a different argument. From the universal property of polynomials we know that there exists a unique ring homomorphism $\mathbb{F}[x] \to \mathbb{E}$ sending $x \mapsto x + \langle p(x) \rangle$, called "evaluation at $x + \langle p(x) \rangle$." But note that the quotient map $\mathbb{F}[x] \to \mathbb{F}[x]/\langle p(x) \rangle$ also satisfies this condition! Hence by uniqueness we conclude that the evaluation map equals the quotient map: $h(x) \mapsto h(x) + \langle p(x) \rangle$.

Similarly, we first think of a real number as either a "Dedekind cut" (ordered pair of infinite sets of fractions) or an infinite equivalence class of "Cauchy sequences." However, after we are satisfied with the existence of these objects we always revert to a more concrete point of view.

For example, see the following corollary.

Corollary/Definition (Every Polynomial Has a Splitting Field). Let $\mathbb{F}$ be a field and let $f(x) \in \mathbb{F}[x]$ be a polynomial of degree $n \geq 1$. Then there exists a field $\mathbb{E} \supseteq \mathbb{F}$ and elements $\alpha_1, \ldots, \alpha_n \in \mathbb{E}$ such that

$$f(x) = (x - \alpha_1)(x - \alpha_2) \cdots (x - \alpha_n) \text{ in } \mathbb{E}[x].$$

Recall that we define $\mathbb{F} \subseteq \mathbb{F}(\alpha_1, \ldots, \alpha_n) \subseteq \mathbb{E}$ as the smallest subfield of $\mathbb{E}$ containing the set $\mathbb{F} \cup \{\alpha_1, \ldots, \alpha_n\}$. In the case that

$$\mathbb{F}(\alpha_1, \ldots, \alpha_n) = \mathbb{E}$$

we will say that $\mathbb{E}$ is *a splitting field for* $f(x)$. ///

Proof by Induction. The base case is Kronecker's Theorem. So let $n \geq 2$ and consider a polynomial $f(x) \in \mathbb{F}[x]$ of degree $n \geq 2$. By Kronecker's Theorem there exists a field $\mathbb{K}_1 \supseteq \mathbb{F}$ and an element $\alpha_1 \in \mathbb{K}_1$ such that $f(\alpha_1) = 0$, and then by Descartes' Factor Theorem we have

$$f(x) = (x - \alpha_1)g(x) \text{ for some polynomial } g(x) \in \mathbb{K}_1[x] \text{ of degree } n - 1.$$

Now by induction there exists a field $\mathbb{K} \supseteq \mathbb{K}_1 \supseteq \mathbb{F}$ and elements $\alpha_2, \ldots, \alpha_n \in \mathbb{K}$ such that

$$g(x) = (x - \alpha_2) \cdots (x - \alpha_n)$$
$$f(x) = (x - \alpha_1)g(x) = (x - \alpha_1)(x - \alpha_2) \cdots (x - \alpha_n) \text{ in } \mathbb{K}[x].$$

Finally, note that $\mathbb{E} := \mathbb{F}(\alpha_1, \ldots, \alpha_n) \subseteq \mathbb{K}$ is a splitting field for $f(x)$. □

You may recall from the introduction of this course that splitting fields are central to the Fundamental Theorem of Galois Theory. Later we will show that if $\mathbb{E} \supseteq \mathbb{F}$ and $\mathbb{E}' \supseteq \mathbb{F}$ are two splitting fields for the same polynomial $f(x) \in \mathbb{F}[x]$ then there exists an isomorphism $\mathbb{E} \cong \mathbb{E}'$ fixing $\mathbb{F}$. However, this isomorphism is **not** unique.[40] In fact, you already know this. The collection of such isomorphisms $\mathbb{E} \cong \mathbb{E}'$ is called the Galois group of $f(x)$ over $\mathbb{F}$.

[40] In other words, the concept of splitting fields does not satisfy a "universal property."

Problem Set 9

1. The Definition of PIDs is Good. For any ring R prove that

$$(R \text{ is a field}) \iff (R[x] \text{ is a PID}).$$

Proof. If R is a field then we have already seen that $R[x]$ is a Euclidean domain, hence a PID. Conversely, suppose that $R[x]$ is a PID. Since $R \subseteq R[x]$ is a subring this also implies that R is a domain. Now consider the evaluation homomorphism at $x = 0$:

$$\begin{array}{rcl} \mathrm{id}_0 : R[x] & \to & R \\ f(x) & \mapsto & f(0). \end{array}$$

Note that this is a surjective ring homomorphism with kernel $\langle x \rangle = xR[x]$, hence we have

$$R[x]/\langle x \rangle \cong R.$$

Since R is a domain this implies that $\langle x \rangle \subseteq R[x]$ is a prime ideal, hence $x \in R[x]$ is a prime element, hence $x \in R[x]$ is an irreducible element, hence $\langle x \rangle \subseteq R[x]$ is maximal among principal ideals. Then since $R[x]$ is a PID we conclude that $\langle x \rangle \subseteq R[x]$ is maximal among **all** ideals, hence $R \cong R[x]/\langle x \rangle$ is a field. □

2. Quadratic Field Extensions, Part II. Let $\mathbb{E} = \mathbb{F}(\iota) \supseteq \mathbb{F}$ for some element $\iota \in \mathbb{E}$ satisfying $\iota \notin \mathbb{F}$ and $\iota^2 \in \mathbb{F}$. Recall that the vector space $\mathbb{E}/\mathbb{F}$ has basis $\{1, \iota\}$ and the Galois group $\mathrm{Gal}(\mathbb{E}/\mathbb{F})$ is generated by the "conjugation" automorphism $(a + b\iota)^* := a - b\iota$.

(a) For any $\alpha \in \mathbb{E}$ show that $\alpha \in \mathbb{F}$ if and only if $\alpha^* = \alpha$. Use this to show that $\alpha\alpha^*$ and $\alpha + \alpha^*$ are in $\mathbb{F}$ for all $\alpha \in \mathbb{E}$.

(b) For any polynomial $f(x) = \sum_i \alpha_i x^i \in \mathbb{E}[x]$ we define $f^*(x) := \sum_i \alpha_i^* x^i$. Show that this is a ring automorphism $* : \mathbb{E}[x] \to \mathbb{E}[x]$. Use this to prove that $f(x)f^*(x)$ and $f(x) + f^*(x)$ are in $\mathbb{F}[x]$ for all $f(x) \in \mathbb{E}[x]$.

(c) For all $f(x) \in \mathbb{F}[x]$ show that the roots of $f(x)$ in $\mathbb{E} - \mathbb{F}$ come in conjugate pairs.

(d) **Application.** Let $f(x) \in \mathbb{F}[x]$ have degree 3. If f has a root in $\mathbb{E}$, prove that f also has a root in $\mathbb{F}$. [Hint: Use Descartes' Factor Theorem.]

(a) If $\alpha = a \in \mathbb{F}$ then we can write $\alpha = a + 0\iota$ and hence $\alpha^* = a - 0\iota = \alpha$. Conversely, consider any $\alpha = a + b\iota$ and suppose that $\alpha = \alpha^* = a - b\iota$. Then since $\{1, \iota\}$ are linearly independent over $\mathbb{F}$ we have $b = -b$ which implies that $b = 0$ and hence $\alpha = a \in \mathbb{F}$.

Finally, to see that $\alpha + \alpha^* \in \mathbb{F}$ and $\alpha\alpha^* \in \mathbb{F}$ we note that

$$(\alpha + \alpha^*) = \alpha^* + \alpha^{**} = \alpha^* + \alpha = (\alpha + \alpha^*)$$

and

$$(\alpha\alpha^*)^* = \alpha^*\alpha^{**} = \alpha^*\alpha = (\alpha\alpha^*).$$

(b) To show that $* : \mathbb{E}[x] \to \mathbb{E}[x]$ is a ring automorphism, first note that $1^* = 1$. Now consider any polynomials $f(x) = \sum_i \alpha_i x^i$ and $g(x) = \sum_i \beta_i x^i$ in $\mathbb{E}[x]$. Applying $*$ to $f(x) + g(x)$ gives

$$\begin{aligned}
(f(x) + g(x))^* &= \left(\sum_i \alpha_i x^i + \sum_i \beta_i x^i \right)^* \\
&= \left(\sum_i (\alpha_i + \beta_i) x^i \right)^* \\
&= \sum_i (\alpha_i + \beta_i)^* x^i \\
&= \sum_i (\alpha_i^* + \beta_i^*) x^i \\
&= \sum_i \alpha_i^* x^i + \sum_i \beta_i^* x^* \\
&= f^*(x) + g^*(x).
\end{aligned}$$

And applying $*$ to $f(x)g(x)$ gives

$$\begin{aligned}
(f(x)g(x))^* &= \left(\sum_k \left(\sum_{i+j=k} \alpha_i \beta_j \right) x^i \right)^* \\
&= \sum_k \left(\sum_{i+j=k} \alpha_i \beta_j \right)^* x^i \\
&= \sum_k \left(\sum_{i+j=k} \alpha_i^* \beta_j^* \right) x^i \\
&= \left(\sum_i \alpha_i^* x^i \right) \left(\sum_i \beta_i^* x^i \right) \\
&= f^*(x) g^*(x).
\end{aligned}$$

Thus we have shown that $*$ is a ring homomorphism. To see that $*$ is invertible note that $f^{**}(x) = f(x)$ for all $f(x) \in \mathbb{E}[x]$. Note from part (a) that we also have

$$f^*(x) = f(x) \quad \Leftrightarrow \quad \alpha_i^* = \alpha_i \text{ for all } i \quad \Leftrightarrow \quad \alpha_i \in \mathbb{F} \text{ for all } i \quad \Leftrightarrow \quad f(x) \in \mathbb{F}[x].$$

Finally, to see that $f(x) + f^*(x) \in \mathbb{F}[x]$ and $f(x)f^*(x) \in \mathbb{F}[x]$ we note that

$$(f(x) + f^*(x))^* = f^*(x) + f^{**}(x) = f^*(x) + f(x) = (f(x) + f^*(x))$$

and

$$(f(x)f^*(x))^* = f^*(x)f^{**}(x) = f^*(x)f(x) = (f(x)f^*(x)).$$

(c) Let $f(x) \in \mathbb{F}[x]$ and consider any root $\alpha \in \mathbb{E} - \mathbb{F}$. Then we have

$$0 = 0^* = f(\alpha)^* = f^*(\alpha^*) = f(\alpha^*),$$

which implies that $\alpha^* \neq \alpha$ is another root.

(d) **Application.** Let $f(x) \in \mathbb{F}[x]$ have degree 3 and suppose that $f(x)$ has a root $\alpha \in \mathbb{E}$. We will show that $f(x)$ also has a root in $\mathbb{F}$. If $\alpha \in \mathbb{F}$ then we are done. Otherwise, from part (c) we know that $\alpha^* \neq \alpha$ is another root and from Descartes' Theorem we have

$$f(x) = (x - \alpha)(x - \alpha^*)g(x) \quad \text{for some } g(x) \in \mathbb{E}[x] \text{ of degree 1.}$$

Then since $(x - \alpha)(x - \alpha^*) \in \mathbb{F}[x]$ it follows from the uniqueness of the quotient that $g(x)$ is also in $\mathbb{F}[x]$, say $g(x) = ax + b$ for some $a, b \in \mathbb{F}$ with $a \neq 0$. Finally, we conclude that $f(x)$ has the root $-b/a \in \mathbb{F}$. □

3. Wilson's Theorem. We saw in the previous problem that any ring homomorphism $\varphi : R \to S$ extends to a ring homomorphism $\varphi : R[x] \to S[x]$ by acting on coefficients. Now let $p \in \mathbb{Z}$ be prime and consider the following polynomial with integer coefficients:

$$f(x) := (x^{p-1} - 1) - \prod_{k=1}^{p-1}(x - k) \in \mathbb{Z}[x].$$

(a) Let $\pi : \mathbb{Z} \to \mathbb{Z}/p\mathbb{Z}$ be the quotient homomorphism. Prove that the polynomial $f^\pi(x) \in (\mathbb{Z}/p\mathbb{Z})[x]$ has $p - 1$ distinct roots and degree $< p - 1$. [Hint: Fermat's Little Theorem.]

(b) Use Descartes' Factor Theorem to show that every coefficient of $f(x) \in \mathbb{Z}[x]$ is a multiple of p. Show that this implies $(p - 1)! = -1 \bmod p$.

(a) The polynomial $f^\pi(x)$ can be written as

$$f^\pi(x) = (x^{p-1} - 1) - \prod_{k=1}^{p-1}(x - k) \in (\mathbb{Z}/p\mathbb{Z})[x],$$

where we interpret all numbers as elements of $\mathbb{Z}/p\mathbb{Z}$. For any $a \in \mathbb{Z}/p\mathbb{Z}$ with $a \neq 0$, we have

$$\prod_{k=1}^{p-1}(a - k) = 0 \text{ because } a \in \{1, 2, \ldots, p - 1\}$$

and

$$a^{p-1} - 1 = 0 \text{ from Fermat's Little Theorem,}$$

and it follows that $f^\pi(a) = 0$. We conclude that $f^\pi(x) \in (\mathbb{Z}/p\mathbb{Z})[x]$ is a polynomial of degree $< p-1$ with $p-1$ distinct roots in $\mathbb{Z}/p\mathbb{Z}$.

(b) But then since $\mathbb{Z}/p\mathbb{Z}$ is a domain this implies that $f^\pi(x) \in (\mathbb{Z}/p\mathbb{Z})[x]$ is the zero polynomial, and hence every coefficient of $f(x) \in \mathbb{Z}[x]$ is a multiple of p. In particular, if $p \neq 2$ then since the constant term of $f(x)$ is $-1-(p-1)!$ we conclude that

$$p|\,(-1-(p-1)!) \text{ and hence } (p-1)! = -1 \bmod p.$$

And if $p = 2$ then the result is still true because $1 = -1 \bmod 2$. □

[Remark: Actually this proof gives us more. By examining the other coefficients we find that

$$\sum_{1 \le n_1 < n_2 < \cdots < n_k \le p-1} n_1 n_2 \cdots n_k = 0 \bmod p$$

for any $1 \le k \le p-2$. For example, when $p = 5$ we have

$$\begin{aligned} 1+2+3+4 &= 10, \\ 1\cdot 2 + 1\cdot 3 + 1\cdot 4 + 2\cdot 3 + 2\cdot 4 + 3\cdot 4 &= 35, \\ 1\cdot 2\cdot 3 + 1\cdot 2\cdot 4 + 1\cdot 3\cdot 4 + 2\cdot 3\cdot 4 &= 50, \end{aligned}$$

all of which are equal to 0 mod 5.]

4. Gaussian Integers. The following theorem is mostly due to Fermat:

> An integer $n \in \mathbb{N}$ is a sum of two squares if and only if any prime factor $p|n$ satisfying $p = 3 \bmod 4$ occurs to an even power.

In this problem we will give a mostly algebraic proof due to Gauss. Let $i \in \mathbb{C}$ be any square root of -1 and consider the following ring extension of $\mathbb{Z}$, called the ring of *Gaussian integers*:

$$\mathbb{Z} \subseteq \mathbb{Z}[i] = \{a + bi : a, b \in \mathbb{Z}\} \subseteq \mathbb{C}.$$

(a) Let $N : \mathbb{Z}[i] \to \mathbb{N}$ be the "norm" function defined by $N(a+ib) := a^2 + b^2$. Prove that $(\mathbb{Z}[i], N)$ is a Euclidean domain, hence $\mathbb{Z}[i]$ is a UFD. [Hint: For any $\alpha, \beta \in \mathbb{Z}[i]$ with $\beta \neq 0$, the ideal $\beta\mathbb{Z}[i]$ is a "square lattice" in $\mathbb{C}$ with (squared) side length $N(\beta)$. Let $\beta\zeta$ be the closest element of $\beta\mathbb{Z}[i]$ to α and observe that $N(\alpha - \beta\zeta) < N(\beta)$.]

(b) For all $\alpha, \beta \in \mathbb{Z}[i]$ prove that $N(\alpha\beta) = N(\alpha)N(\beta)$. Use this to show that

$$\mathbb{Z}[i]^\times = \{\alpha \in \mathbb{Z}[i] : N(\alpha) = 1\} = \{\pm 1, \pm i\}.$$

(c) For all $n \in \mathbb{N}$ show that $n = 3 \bmod 4$ implies $n \notin \operatorname{im} N$. [Hint: What are the square elements of the ring $\mathbb{Z}/4\mathbb{Z}$?]

(d) Use induction on n to prove the following statement:

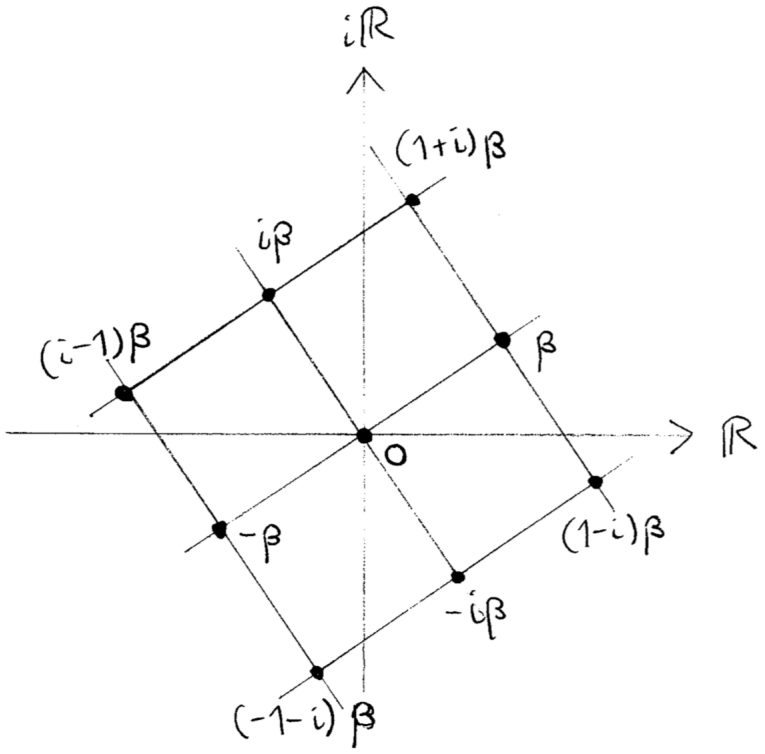
$$n \in \operatorname{im} N \Rightarrow (\text{every prime } p|n \text{ with } p = 3 \bmod 4 \text{ occurs to an even power}).$$

[Hint: Let $n = a^2 + b^2 \in \operatorname{im} N$ and let $p \in \mathbb{Z}$ be prime. If $p = 3 \bmod 4$ use (b) and (c) to show that p is irreducible in $\mathbb{Z}[i]$. Then if $p|n$ use (a) to show that $p|(a+bi)$ or $p|(a-bi)$ in $\mathbb{Z}[i]$. In either case show that $p|a$ and $p|b$, hence $n/p^2 \in \operatorname{im} N$.]

(e) Conversely, for prime $p \in \mathbb{N}$ show that $p = 1 \bmod 4$ implies $p \in \operatorname{im} N$. [Hint: Let $p = 4k + 1$ and assume for contradiction that $p \notin \operatorname{im} N$. Use (a) and (b) to show that p is irreducible and hence prime in $\mathbb{Z}[i]$. On the other hand, set $m := (2k)!$ and use Wilson's Theorem to show that $p|(m - i)(m + i)$.]

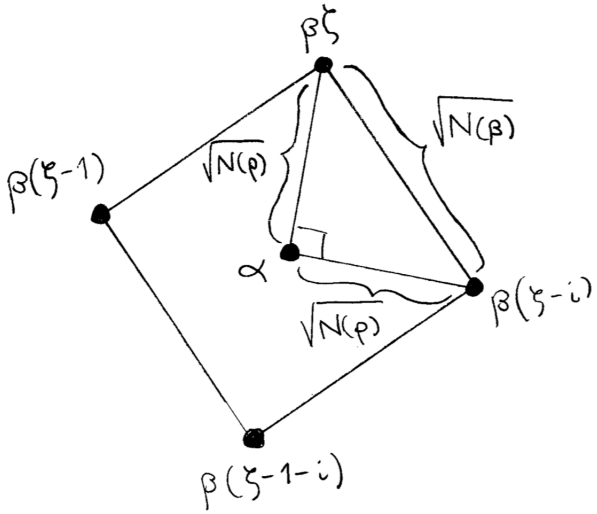
(f) Finish the proof.

(a) Consider any $\alpha, \beta \in \mathbb{Z}[i]$ with $\beta \neq 0$, and note that the principal ideal $\beta\mathbb{Z}[i]$ is a "square lattice" in the complex plane:



Let $\beta\zeta \in \beta\mathbb{Z}[i]$ be any point of this lattice such that the squared distance $N(\alpha - \beta\zeta)$ is minimal, and set $\rho := \alpha - \beta\zeta$.[41] In the worst-case-scenario, α will lie at the exact center of one of the squares, as in the following picture:

[41] In a Euclidean domain we do not require uniqueness of remainders.



Then from the Pythagorean Theorem we obtain

$$N(\beta) = N(\rho) + N(\rho) = 2N(\rho),$$

which implies that $N(\rho) < N(\beta)$ as desired. It follows from the existence of such an element $\rho \in \mathbb{Z}[i]$ that $\mathbb{Z}[i]$ is a Euclidean doman, hence a PID, hence a UFD.

(b) Let $* : \mathbb{Z}[i] \to \mathbb{Z}[i]$ be the restriction of complex conjugation to the subring $\mathbb{Z}[i] \subseteq \mathbb{C}$, and note that $N(\alpha) = \alpha\alpha^*$. It follows for any $\alpha, \beta \in \mathbb{Z}[i]$ that

$$N(\alpha)N(\beta) = (\alpha\alpha^*)(\beta\beta^*) = (\alpha\beta)(\alpha\beta)^* = N(\alpha\beta).$$

Now suppose that $\alpha \in \mathbb{Z}[i]^\times$ is a unit, so that $\alpha\beta = 1$ for some $\beta \in \mathbb{Z}[i]$. Applying the norm gives $N(\alpha)N(\beta) = N(\alpha\beta) = N(1) = 1$, which implies that $N(\alpha) = 1$ because $N(\alpha)$ and $N(\beta)$ are positive integers. It follows that $\alpha \in \{\pm 1, \pm i\}$, and we observe that each of these four elements is indeed a unit:

$$1 \cdot 1 = 1, \quad (-1)(-1) = 1, \quad i(-i) = 1.$$

(c) If $n = a^2 + b^2$ for some integers $n, a, b \in \mathbb{Z}$ then we also have $n = a^2 + b^2 \bmod 4$. But observe that 0 and 1 are the only square elements of $\mathbb{Z}/4\mathbb{Z}$:

x	0	1	2	3
x^2	0	1	0	1

For any $a, b \in \mathbb{Z}$ it follows that $a^2 + b^2 = 0$, 1 or 2 mod 4. In other words, if $n \in \mathbb{N}$ is a sum of two square integers then $n \neq 3 \bmod 4$.

(d) Now let $n \in \mathbb{N}$ and consider the following statement:

"if $n \in \operatorname{im} N$ then every prime $p|n$ with $p = 3 \bmod 4$ occurs to an even power."

First, observe that this statement is true for $n = 1$. Now fix some $n \geq 1$ and assume for induction that the statement holds for all integers $1 \leq k < n$. To prove that the statement holds for n, assume that $n = a^2 + b^2$ for some $a, b \in \mathbb{Z}$. If n has no prime factor of the form $p = 3 \bmod 4$ then we are done because zero is an even number. Otherwise, let $p|n$ be such a prime factor. I claim that $p \in \mathbb{Z}[i]$ is an irreducible element. Indeed, if $p = \alpha\beta$ for some non-units $\alpha, \beta \in \mathbb{Z}[i]$ then applying the norm gives $p^2 = N(p) = N(\alpha)N(\beta)$. Since α, β are non-units we know from (b) that $N(\alpha), N(\beta) \geq 2$ and hence we must have $N(\alpha) = N(\beta) = p$. But this contradicts (c) because $p = 3 \bmod 4$.

We have shown that $p \in \mathbb{Z}[i]$ is an **irreducible** element. But from (a) we know that $\mathbb{Z}[i]$ is a UFD, hence $p \in \mathbb{Z}[i]$ is also a **prime** element. Since we have assumed that $p|n$ and $n = a^2 + b^2 = (a + ib)(a - ib)$, it follows that

$$p|(a + ib) \quad \text{or} \quad p|(a - ib).$$

Without loss of generality, suppose that $p|(a + ib)$ so that $(a + ib) = p(c + id) = (pc) + i(pd)$ for some $c, d \in \mathbb{Z}$. Comparing real and imaginary parts gives $a = pc$ and $b = pd$, hence

$$n = a^2 + b^2 = (pc)^2 + (pd)^2 = p^2(c^2 + d^2).$$

But now $k := c^2 + d^2 < n$ is a smaller sum of squares, so by induction we know that p occurs to an even power e in k, hence p occurs to the even power $e + 2$ in n. We have shown that any prime factor $p|n$ of the form $p = 3 \bmod 4$ occurs to an even power in n. □

(e) Next we will show that every prime of the form $p = 1 \bmod 4$ **is** a sum of squares. (This smaller statement by itself is sometimes called Fermat's Theorem.) We will give Lagrange's 1773 proof, which follows from Wilson's Theorem. So let $p = 4k + 1$ be prime and assume for contradiction that $p \notin \operatorname{im} N$. It follows that $p \in \mathbb{Z}[i]$ is irreducible since if $p = \alpha\beta$ for non-units $\alpha, \beta \in \mathbb{Z}[i]$ then $p^2 = N(\alpha)N(\beta)$ implies that $N(\alpha) = N(\beta) = p$, which contradicts the fact that $p \notin \operatorname{im} N$. Then since $\mathbb{Z}[i]$ is a UFD we know that $p \in \mathbb{Z}[i]$ is also a **prime** element.

Now let $m := (2k)! \in \mathbb{Z}$. Then since $p + i = i \bmod p$ for all i, Wilson's Theorem gives

$$\begin{aligned}
-1 &= (p-1)! \\
&= (4k)! \\
&= 1 \cdot 2 \cdots 2k \cdot (2k+1)(2k+2) \cdots (4k) \\
&= 1 \cdot 2 \cdots 2k \cdot (p-2k)(p-2k+1) \cdots (p-1) \\
&= 1 \cdot 2 \cdots 2k \cdot (-2k)(-2k+1) \cdots (-1) && p + i = i \\
&= (-1)^{2k} 1 \cdot 2 \cdots 2k \cdot (2k)(2k-1) \cdots (1) \\
&= (1 \cdot 2 \cdots (2k))^2 \\
&= m^2 \bmod p.
\end{aligned}$$

In other words, we have $p|(m^2+1)$. Then since $m^2+1=(m+i)(m-i)$ and since p is prime in $\mathbb{Z}[i]$ we conclude that

$$p|(m+i) \quad \text{or} \quad p|(m-i).$$

Without loss of generality, suppose that $p|(m+i)$ so that $(m+i)=p(a+ib)=(pa)+i(pb)$ for some $a,b\in\mathbb{Z}$. But then comparing imaginary parts gives $1=pb$, which contradicts the fact that p is prime. □

(f) **Summary.** I claim that $n\in\mathbb{N}$ is a sum of two square integers if and only if each prime factor $p|n$ of the form $p=3 \bmod 4$ occurs with even multiplicity. Part (d) proves one direction. For the other direction, suppose that

$$n=2^k p_1^{2e_1}\cdots p_r^{2e_r} q_1^{f_1}\cdots q_s^{f_s},$$

where p_i, q_j are distinct primes satisfying $p_i=3 \bmod 4$ and $q_j=1 \bmod 4$. Now observe that

- $2=1^2+1^2\in \operatorname{im} N$,
- $p_i^{2e_i}=(p_i^{e_i})^2+0^2\in \operatorname{im} N$ for all i,
- $q_j\in \operatorname{im} N$ for all j from part (e).

Finally, since $\operatorname{im} N$ is closed under multiplication we conclude that $n\in\operatorname{im} N$. □

[Remark: The expression as a sum of squares is not necessarily unique. Lagrange actually gave a formula for the number of distinct representations of $n\in\mathbb{N}$ as a sum of squares:

$$2\left(1+\left(\frac{-1}{n}\right)\right)\sum_{d|n}\left(\frac{-1}{d}\right).$$

Here the notation $\left(\frac{a}{b}\right)$ is called the *Jacobi symbol* and I am not going to define it.]

5. $\mathbb{Z}[\sqrt{-3}]$ is not a UFD. Let $\sqrt{-3}\in\mathbb{C}$ be a fixed square root of -3 and consider the ring

$$\mathbb{Z}\subseteq\mathbb{Z}[\sqrt{-3}]=\{a+b\sqrt{-3}: a,b\in\mathbb{Z}\}\subseteq\mathbb{C}.$$

(a) Let $N:\mathbb{Z}[\sqrt{-3}]\to\mathbb{N}$ be defined by $N(a+b\sqrt{-3}):=a^2+3b^2$. For all $\alpha,\beta\in\mathbb{Z}[\sqrt{-3}]$ prove that $N(\alpha\beta)=N(\alpha)N(\beta)$ and use this to show that

$$\mathbb{Z}[\sqrt{-3}]^\times=\{\alpha\in\mathbb{Z}[\sqrt{-3}]: N(\alpha)=1\}=\{\pm 1\}.$$

(b) Prove that there is no element $\alpha\in\mathbb{Z}[\sqrt{-3}]$ with $N(\alpha)=2$. Use this to show that any element with $N(\alpha)=4$ is irreducible. In particular, $2\in\mathbb{Z}[\sqrt{-3}]$ is irreducible.

(c) But show that $2\in\mathbb{Z}[\sqrt{-3}]$ is **not prime** because

$$2|(1+\sqrt{-3})(1-\sqrt{-3}) \text{ and } 2\nmid(1+\sqrt{-3}) \text{ and } 2\nmid(1-\sqrt{-3}).$$

(d) Use this to prove that the following ideal is **not principal**:

$$\{2\alpha + (1+\sqrt{-3})\beta : \alpha, \beta \in \mathbb{Z}[\sqrt{-3}]\} \subseteq \mathbb{Z}[\sqrt{-3}].$$

(a) As in the previous problem, we observe that

$$N(a+b\sqrt{-3}) = a^2 + 3b^2 = (a+b\sqrt{-3})(a-b\sqrt{-3}) = (a+b\sqrt{-3})(a+b\sqrt{-3})^*,$$

where $* : \mathbb{Z}[\sqrt{-3}] \to \mathbb{Z}[\sqrt{-3}]$ is the restriction of complex conjugation to the subring $\mathbb{Z}[\sqrt{-3}] \subseteq \mathbb{C}$. It follows that $N(\alpha\beta) = N(\alpha)N(\beta)$ for all $\alpha, \beta \in \mathbb{Z}[\sqrt{-3}]$.

We will use this property to compute the units. First suppose that $\alpha \in \mathbb{Z}[\sqrt{-3}]^\times$ is a unit so that $\alpha\beta = 1$ for some $\beta \in \mathbb{Z}[\sqrt{-3}]$. Applying the norm gives $N(\alpha)N(\beta) = N(\alpha\beta) = N(1) = 1$, which implies that $N(\alpha) = 1$ because $N(\alpha)$ and $N(\beta)$ are positive integers. Now let $\alpha = a + b\sqrt{-3}$ with $N(\alpha) = a^2 + 3b^2 = 1$. If $b \neq 0$ then we would obtain the contradiction $N(\alpha) \geq 3$ so we must have $b = 0$ and hence $a^2 = 1$. It follows that $\alpha = a \in \{\pm 1\}$. Finally, we observe that $\{\pm 1\} \subseteq \mathbb{Z}[\sqrt{-3}]^\times$ because $1 \cdot 1 = 1$ and $(-1)(-1) = 1$. In summary, we have

$$\mathbb{Z}[\sqrt{-3}]^\times = \{\alpha \in \mathbb{Z}[\sqrt{-3}] : N(\alpha) = 1\} = \{\pm 1\}.$$

(b) Suppose for contradiction that there exists $\alpha = a+b\sqrt{-3} \in \mathbb{Z}[\sqrt{-3}]$ with $N(\alpha) = a^2+3b^2 = 2$. If $b \neq 0$ then we obtain the contradiction that $N(\alpha) \geq 3$, so we must have $b = 0$. But then the equation $a^2 = 2$ is a contradiction because $a \in \mathbb{Z}$.

We will use this fact to show that any element $\gamma \in \mathbb{Z}[\sqrt{-3}]$ of norm 4 is irreducible. So let $N(\gamma) = 4$ and assume for contradiction that $\gamma = \alpha\beta$ for some non-zero non-units $\alpha, \beta \in \mathbb{Z}[\sqrt{-3}]$. Since α and β are non-zero non-units we know from part (a) that $N(\alpha) \geq 2$ and $N(\beta) \geq 2$. But then since $N(\alpha)N(\beta) = N(\alpha\beta) = N(\gamma) = 4$ we must have $N(\alpha) = N(\beta) = 2$, which contradicts the fact that there are no elements of norm 2.

(c) Part (b) shows that the elements 2, $1+\sqrt{-3}$ and $1-\sqrt{-3}$ are irreducible in $\mathbb{Z}[\sqrt{-3}]$ because they each have norm 4. Since the units are $\mathbb{Z}[\sqrt{-3}]^\times = \{\pm 1\}$ we also know that 2 is not associate to either of $1 \pm \sqrt{-3}$, hence $2 \nmid (1+\sqrt{-3})$ and $2 \nmid (1-\sqrt{-3})$. Finally, since 2 divides the product $(1+\sqrt{-3})(1-\sqrt{-3}) = 4$ we see that **the irreducible element 2 is not prime**. It follows from this that $\mathbb{Z}[\sqrt{-3}]$ is not a UFD, hence not a PID, hence not a Euclidean domain.

[Remark: It is difficult to imagine any direct way to prove that a domain is not Euclidean.]

(d) We showed in part (c) that non-principal ideals exist, but we didn't exhibit any. I claim that the following ideal is not principal:

$$2\mathbb{Z}[\sqrt{-3}] + (1+\sqrt{-3})\mathbb{Z}[\sqrt{-3}] = \left\{2\alpha + (1+\sqrt{-3})\beta : \alpha, \beta \in \mathbb{Z}[\sqrt{-3}]\right\}.$$

To prove this, assume for contradiction that we have $2\mathbb{Z}[\sqrt{-3}]+(1+\sqrt{-3})\mathbb{Z}[\sqrt{-3}] = \gamma\mathbb{Z}[\sqrt{-3}]$ for some $\gamma \in \mathbb{Z}[\sqrt{-3}]$. In part (c) we showed that $1+\sqrt{-3} \notin 2\mathbb{Z}[\sqrt{-3}]$, which implies that

$$2\mathbb{Z}[\sqrt{-3}] \subsetneq 2\mathbb{Z}[\sqrt{-3}] + (1+\sqrt{-3})\mathbb{Z}[\sqrt{-3}] = \gamma\mathbb{Z}[\sqrt{-3}].$$

Since 2 is irreducible we must have $\gamma\mathbb{Z}[\sqrt{-3}] = 1\mathbb{Z}[\sqrt{-3}]$. But then we have $1 \in 2\mathbb{Z}[\sqrt{-3}] + (1+\sqrt{-3})\mathbb{Z}[\sqrt{-3}]$ which implies that

$$2\alpha + (1+\sqrt{-3})\beta = 1 \quad \text{for some } \alpha, \beta \in \mathbb{Z}[\sqrt{-3}].$$

Finally, multiplying both sides by $1-\sqrt{-3}$ gives

$$\begin{aligned} 2\alpha + (1+\sqrt{-3})\beta &= 1 \\ 2(1-\sqrt{-3})\alpha + 4\beta &= (1-\sqrt{-3}) \\ 2\left[(1-\sqrt{-3})\alpha + 2\beta\right] &= (1-\sqrt{-3}), \end{aligned}$$

which contradicts the fact that $2 \nmid (1-\sqrt{-3})$. □

[Remark: The previous two problems are part of a tricky subject called *algebraic number theory*. We will now leave this subject behind, since any further discussion would lead us away from the main goals of this course.]

6. Field of Fractions. In this problem you will show that "integral domain" and "subring of a field" are the same concept. Let R be an integral domain and consider the following set of abstract symbols, called *fractions*:

$$\mathrm{Frac}(R) := \left\{ \frac{a}{b} : a, b \in R, b \neq 0 \right\}.$$

(a) Prove that the following relation is an equivalence on the set of fractions:

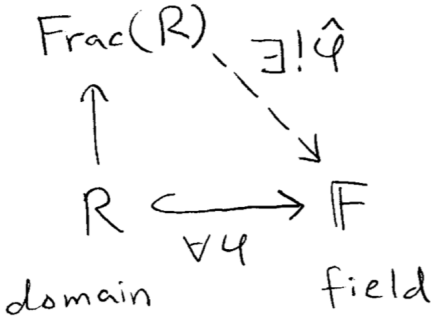
$$\frac{a}{b} \sim \frac{a'}{b'} \quad \Longleftrightarrow \quad ab' = a'b.$$

(b) Prove that the following operations are well-defined on equivalence classes:

$$\frac{a}{b} \cdot \frac{c}{d} := \frac{ac}{bd} \qquad \text{and} \qquad \frac{a}{b} + \frac{c}{d} := \frac{ad+bc}{bd}.$$

It follows that the set of equivalence classes $\mathrm{Frac}(R)/\sim$ is a field. Following tradition, we will just call it $\mathrm{Frac}(R)$ and we will write $=$ instead of $\sim$. Furthermore, we will write $R \subseteq \mathrm{Frac}(R)$ for the image of the injective ring homomorphism $a \mapsto a/1$.

(c) **Universal Property.** Let $\mathbb{F}$ be a field and let $\varphi : R \to \mathbb{F}$ be an **injective** ring homomorphism. Prove that this extends to a unique ring homomorphism $\varphi : \mathrm{Frac}(R) \to \mathbb{F}$, which is also injective. [Hint: Show that $\hat{\varphi}(a/b) := \varphi(a)/\varphi(b)$ is well-defined.] Here is a picture:



(d) **Application.** If a field $\mathbb{F}$ contains a subring isomorphic to $\mathbb{Z}$, prove that $\mathbb{F}$ also contains a subfield isomorphic to $\mathbb{Q}$.

(a) It is easy to show that the relation $\sim$ is reflexive and symmetric. To prove that $\sim$ is transitive, suppose that $a/b \sim c/d$ and $c/d \sim e/f$ for some fractions $a/b, c/d, e/f \in \mathrm{Frac}(R)$. By definition this means that $ad = bc$ and $cf = de$ for some $a, b, c, d, e, f \in R$ with $b, d, f \neq 0$. But then since R is a domain and $d \neq 0$ we have

$$\begin{aligned} ad &= bc \\ adf &= b(cf) \\ adf &= b(de) \\ \cancel{d}(af) &= \cancel{d}(be) \\ af &= be, \end{aligned}$$

which implies that $a/b \sim e/f$.

(b) To prove that multiplication and addition of fractions are well-defined, suppose that $a/b \sim a'/b'$ and $c/d \sim c'/d'$. By definition this means that $a'b = ab'$ and $cd' = c'd$ for some $a, b, c, d, a', b', c', d' \in R$ with $b, d, b', d' \neq 0$. Since R is a domain we have $bd, b'd' \neq 0$, hence the products ac/bd and $a'c'/b'c'$ exist. To prove that they are equivalent we observe that

$$\begin{aligned} (ac)(b'd') &= (ab')(cd') \\ &= (a'b)(c'd) \\ &(a'c')(bd). \end{aligned}$$

Since $bd, b'd' \neq 0$ we also see that the sums $(ad+bc)/bd$ and $(a'd'+b'c')/(b'd')$ exist. To prove that they are equivalent we observe that

$$\begin{aligned} (ad + bc)(b'd') &= (ad)(b'd') + (bc)(b'd') \\ &= (ab')(dd') + (cd')(bb') \\ &= (a'b)(dd') + (c'd)(bb') \end{aligned}$$

$$= (a'd')(bd) + (b'c')(bd)$$
$$= (a'd' + b'c')(bd).$$

(c) **Universal Property.** Let $\mathbb{F}$ be a field and let $\varphi : R \to \mathbb{F}$ be an **injective** ring homomorphism. Our goal is to define a ring homomorphism $\hat{\varphi} : \mathrm{Frac}(R) \to \mathbb{F}$ with the property that $\hat{\varphi}(a/1) = \varphi(a)$ for all $a \in R$. If any such homomorphism exists then it must satisfy $\hat{\varphi}(a/b) = \varphi(a)/\varphi(b)$ for all $a, b \in R$ with $b \neq 0$. Since φ is injective we have $\varphi(b) \neq 0$ and hence the element $\varphi(a)/\varphi(b) \in \mathbb{F}$ exists. It only remains to show that the function $\hat{\varphi}$ is well-defined. So consider any fractions a/b and c/d. Then we have

$$\begin{aligned} a/b = c/d &\iff ad = bd \\ &\iff \varphi(ad) = \varphi(bd) \\ &\iff \varphi(a)\varphi(d) = \varphi(b)\varphi(d) \\ &\iff \varphi(a)/\varphi(b) = \varphi(c)/\varphi(d) \\ &\iff \hat{\varphi}(a/b) = \hat{\varphi}(c/d). \end{aligned}$$

The right arrows show that $\hat{\varphi}$ is well-defined and the left arrows show that $\hat{\varphi}$ is injective.

[Remark: Note that the assumption of injectivity is necessary. Indeed, note that the quotient map $\varphi : \mathbb{Z} \to \mathbb{Z}/p\mathbb{Z}$ is a non-injective ring homomorphism from a domain to a field. Suppose that this lifts to a ring homomorphism $\hat{\varphi} : \mathbb{Q} \to \mathbb{Z}/p\mathbb{Z}$. Then we must have

$$1 = \hat{\varphi}(1) = \hat{\varphi}\left(\frac{p}{1} \cdot \frac{1}{p}\right) = \varphi(p)\hat{\varphi}(1/p),$$

which is a contradiction because $\varphi(p) = 0$.]

(d) **Application.** Suppose that $R \subseteq \mathbb{F}$ is a subring with $R \cong \mathbb{Z}$ and recall that $\mathbb{Q} := \mathrm{Frac}(\mathbb{Z})$. First we will prove that $\mathrm{Frac}(R) \cong \mathrm{Frac}(\mathbb{Z})$. One could give a direct argument, but I prefer to sketch a proof that explicitly uses the universal property. First note that the identity $\mathrm{id} : \mathrm{Frac}(R) \to \mathrm{Frac}(R)$ is the **unique** ring homomorphism that fixes the subring $R \subseteq \mathrm{Frac}(R)$, and similarly the identity $\mathrm{id} : \mathbb{Q} \to \mathbb{Q}$ is the **unique** ring homomorphism that fixes the subring $\mathbb{Z} \subseteq \mathbb{Q}$. Now consider any pair of inverse isomorphisms $\varphi : R \leftrightarrows \mathbb{Z} : \psi$. Then $\varphi : R \to \mathbb{Q}$ lifts to $\hat{\varphi} : \mathrm{Frac}(R) \to \mathbb{Q}$ and $\psi : \mathbb{Z} \to \mathrm{Frac}(R)$ lifts to $\hat{\psi} : \mathbb{Q} \to \mathrm{Frac}(R)$. Observe that $\hat{\varphi} \circ \hat{\psi} : \mathbb{Q} \to \mathbb{Q}$ fixes $\mathbb{Z}$, hence **by uniqueness** we must have $\hat{\varphi} \circ \hat{\psi} = \mathrm{id}$. For the same reason we also have $\hat{\psi} \circ \hat{\varphi} = \mathrm{id}$, and it follows that $\mathrm{Frac}(R) \cong \mathbb{Q}$.

Finally, let $\iota : R \hookrightarrow \mathbb{F}$ be the inclusion of the subring $R \subseteq \mathbb{F}$. From the universal property of fractions we have an injective ring homomorphism $\hat{\iota} : \mathrm{Frac}(R) \hookrightarrow \mathbb{F}$ and we conclude that the image of $\hat{\iota}$ is a subring of $\mathbb{F}$ that is isomorphic to $\mathbb{Q}$. $\square$

[Remark: Part (d) fills a gap in our earlier proof characterizing prime subfields. This problem illustrates that the rigorous theory of fractions is subtle.[42] We will usually just follow our intuition.]

7. Newton's Theorem on Symmetric Polynomials. Given a ring R and a set of "independent variables" $\mathbf{x} = \{x_1, \ldots, x_n\}$ we define *multivariate polynomials* by induction:

$$R[\mathbf{x}] = R[x_1, \ldots, x_n] := R[x_1, \ldots, x_{n-1}][x_n] = \left\{ f(\mathbf{x}) = \sum_{\mathbf{k} \in \mathbb{N}^n} a_{\mathbf{k}} \mathbf{x}^{\mathbf{k}} : a_{\mathbf{k}} \in R \right\}.$$

To save space we use the notations $\mathbf{k} = (k_1, \ldots, k_n) \in \mathbb{N}^k$ and $\mathbf{x}^{\mathbf{k}} = x_1^{k_1} \cdots x_n^{k_n}$. We assume that all but finitely many of the coefficients $a_{\mathbf{k}} \in R$ are zero.

(a) We say that a polynomial $f(\mathbf{x}) = R[\mathbf{x}]$ is *symmetric* if for all $\sigma \in S_n$ we have

$$f(x_{\sigma(1)}, \ldots, x_{\sigma(n)}) = f(x_1, \ldots, x_n).$$

Observe that the symmetric polynomials are a subring of $R[\mathbf{x}]$.

(b) **Newton's Theorem.** Recall the definition of the *elementary symmetric polynomials*:

$$e_k(x_1, \ldots, x_n) := \sum_{1 \leq i_1 < \cdots < i_k \leq n} x_{i_1} \cdots x_{i_k}.$$

For convenience, let's define $\mathbf{e}^{\mathbf{k}} := e_1^{k_1} \cdots e_n^{k_n}$. For any symmetric polynomial $f(\mathbf{x}) = \sum_{\mathbf{k}} a_{\mathbf{k}} \mathbf{x}^{\mathbf{k}} \in R[\mathbf{x}]$, prove that there exist some $b_{\mathbf{k}} \in R$ such that $f(x) = \sum_{\mathbf{k}} b_{\mathbf{k}} \mathbf{e}^{\mathbf{k}}$. [Hint: Order the degree vectors $\mathbf{k} \in \mathbb{N}^n$ by "dictionary order" and let $a_{\mathbf{k}} \mathbf{x}^{\mathbf{k}}$ be the "leading term." By symmetry of f we must have $k_1 \geq k_2 \geq \cdots \geq k_n$. Show that there exists $\mathbf{k}' \in \mathbb{N}^k$ so that $a_{\mathbf{k}} \mathbf{e}^{\mathbf{k}'}$ has the same leading term, hence $f(\mathbf{x}) - a_{\mathbf{k}} \mathbf{e}^{\mathbf{k}'}$ is a symmetric polynomial of "smaller degree."]

(c) **Important Corollary.** Suppose that a polynomial $f(x) \in R[x]$ of degree n splits in some ring extension $E \supseteq R$. That is, suppose that we have

$$f(x) = x^n - e_1 x^{n-1} + e_2 x^{n-2} - \cdots + (-1)^n e_n = (x - \alpha_1) \cdots (x - \alpha_n) \in E[x].$$

Prove that any "symmetric expression of the roots" is in R.

(d) **Application: Discriminant of a Cubic.** Let $f(x) = x^3 + ax^2 + bx + c \in R[x]$ and let $E \supseteq R$ be a ring extension such that

$$x^3 + ax^2 + bx + c = (x - \alpha)(x - \beta)(x - \gamma) \in E[x].$$

From part (c) we know that the following element of E (called the *discriminant* of f) is actually in R:

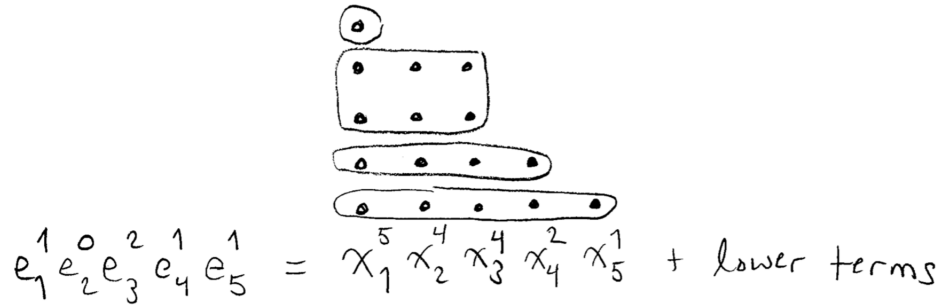
$$\mathrm{Disc}(f) := (\alpha - \beta)^2 (\alpha - \gamma)^2 (\beta - \gamma)^2.$$

[42] And it is deeper than it looks. The general construction of fractions is called "localization." It somehow corresponds to "zooming in" on a point of an algebraic variety.

Use the algorithm from part (b) to express $\mathrm{Disc}(f)$ as a specific polynomial in the coefficients. [I'll get you started: Note that $\mathrm{Disc}(f) = (\alpha^4\beta^2 + \text{lower terms})$ and $a^2b^2 = (\alpha^4\beta^2 + \text{lower terms})$. Now find the leading term of $\mathrm{Disc}(f) - a^2b^2$.]

(a) I have observed it.

(b) **Newton's Theorem.** Consider any symmetric polynomial $f(\mathbf{x}) = \sum_{\mathbf{k}} a_{\mathbf{k}}\mathbf{x}^{\mathbf{k}} \in R[\mathbf{x}]$ and let $a_{k_1,\ldots,k_n}x_1^{k_1}\cdots x_n^{k_n}$ be the leading term in dictionary order. If there exist any $i < j$ such that $k_i < k_j$ then since $f(\mathbf{x})$ is symmetric we may apply the transposition (ij) to obtain a term that is earlier in the dictionary, contradiction. Hence we must have $k_1 \geq k_2 \geq \cdots \geq k_n$. We will view the numbers k_i as the column heights of an array of dots. If k_i' is the number of rows of length i in this array then I claim that $a_{k_1,\ldots,k_n}\mathbf{e}^{\mathbf{k}'}$ has the same leading term as $f(\mathbf{x})$. Instead of proving the claim, I'll just show you a picture:[43]



Then since $f(\mathbf{x}) - a_{k_1,\ldots,k_n}\mathbf{e}^{\mathbf{k}'}$ is a symmetric polynomial with strictly smaller "degree," the result follows by induction.

[Remark: I call this Newton's Theorem but the method of proof is due to Gauss, who used these concepts in his second proof of the Fundamental Theorem of Algebra.]

(c) **Important Corollary.** Let $f(x) \in R[x]$ be a polynomial of degree n and let $E \supseteq R$ be a ring extension in which $f(x)$ splits:

$$x^n - e_1x^{n-1} + e_2x^{n-2} - \cdots + (-1)^n e_n = (x-\alpha_1)\cdots(x-\alpha_n)$$

Note that the coefficients e_i are in the base ring R and the roots α_i are in the extension ring E. However, from Newton's theorem we conclude that **any symmetric R-polynomial of the roots** can be expressed as an R-polynomial[44] of the coefficients, hence must be in R.

[43] You may have noticed that the algebraic notation in this problem is atrocious. Combinatorial problems of this type are easier to understand with pictures. Sometimes we don't even bother with an algebraic proof.

[44] probably non-symmetric

(d) **Application: Discriminant of a Cubic.** For example, consider the general cubic

$$f(x) = x^3 + ax^2 + bx + c = (x - \alpha)(x - \beta)(x - \gamma)$$

with coefficients $a, b, c \in R$ and roots $\alpha, \beta, \gamma \in E \supseteq R$. Note that the following expression is a symmetric R-polynomial in the roots:

$$\mathrm{Disc}(f) := (\alpha - \beta)^2(\alpha - \gamma)^2(\beta - \gamma)^2.$$

We will apply Newton's algorithm to express this as a polynomial in the **elementary** symmetric polynomials:

$$\begin{aligned} -a &= \alpha + \beta + \gamma, \\ b &= \alpha\beta + \alpha\gamma + \beta\gamma, \\ -c &= \alpha\beta\gamma. \end{aligned}$$

To begin, note that

$$\begin{aligned} \mathrm{Disc}(f) &= \alpha^4\beta^2 + \text{lower terms}, \\ a^2b^2 &= \alpha^4\beta^2 + \text{lower terms}. \end{aligned}$$

Then we have

$$\begin{aligned} \mathrm{Disc}(f) - a^2b^2 &= -4\alpha^3\beta^3 + \text{lower terms}, \\ -4b^3 &= -4\alpha^3\beta^3 + \text{lower terms}. \end{aligned}$$

And then

$$\begin{aligned} \mathrm{Disc}(f) - a^2b^2 + 4b^3 &= -4\alpha^4\beta\gamma + \text{lower terms}, \\ -4a^3c &= -4\alpha^4\beta\gamma + \text{lower terms}. \end{aligned}$$

Then

$$\begin{aligned} \mathrm{Disc}(f) - a^2b^2 + 4b^3 + 4a^3c &= 18\alpha^3\beta^2\gamma + \text{lower terms}, \\ 18abc &= 18\alpha^3\beta^2\gamma + \text{lower terms}. \end{aligned}$$

And finally

$$\begin{aligned} \mathrm{Disc}(f) - a^2b^2 + 4b^3 + 4a^3c - 18abc &= -27\alpha^2\beta^2\gamma^2 + \text{zero}, \\ -27c^2 &= -27\alpha^2\beta^2\gamma^2 + \text{zero}. \end{aligned}$$

We conclude that

$$\boxed{\mathrm{Disc}(f) = a^2b^2 - 4b^3 - 4a^3c + 18abc - 27c^2.}$$

[Remark: Note that $\mathrm{Disc}(f) = 0$ if and only if the polynomial $f(x)$ has a multiple root. However, this fact might not be very useful since the formula for the discriminant is hard to remember. To clean it up a bit we can make the *Tschirnhaus substitution* $g(x) := f(x - a/3)$ to obtain a cubic of the form $g(x) = x^3 + px + q$ with same discriminant:

$$\mathrm{Disc}(f) = \mathrm{Disc}(g) = -4p^3 - 27q^2.$$

Does this remind you of Cardano's Formula?]

Week 19

Last week we developed the following tool for studying field extensions:

> Let $\alpha \in \mathbb{E} \supseteq \mathbb{F}$ be an element of a field extension and consider the intermediate field $\mathbb{E} \supseteq \mathbb{F}(\alpha) \supseteq \mathbb{F}$. If $f(x) \in \mathbb{F}[x]$ is an **irreducible** polynomial with $f(\alpha) = 0$ then $1, \alpha, \alpha^2, \ldots, \alpha^{\deg(f)-1}$ is a basis for $\mathbb{F}(\alpha)$ as a vector space over $\mathbb{F}$, hence
>
> $$[\mathbb{F}(\alpha)/\mathbb{F}] = \deg(f).$$

But this tool is only useful if we have some way to prove that a given polynomial is irreducible. Here are a couple of basic tricks.

Low-Degree Trick. Let $f(x) \in \mathbb{F}[x]$ have degree 2 or 3. Then

$$f(x) \in \mathbb{F}[x] \text{ is reducible} \quad \Longleftrightarrow \quad f(x) \text{ has a root in } \mathbb{F}.$$

Proof. Since $\mathbb{F}$ is a field we know from Descartes' Theorem that

$$f(x) \in \mathbb{F}[x] \text{ has a factor of degree 1} \quad \Longleftrightarrow \quad f(x) \text{ has a root in } \mathbb{F}.$$

Indeed, if $\alpha \in \mathbb{F}$ is a root then we have $f(x) = (x - \alpha)g(x)$ for some $g(x) \in \mathbb{F}[x]$. Conversely, if $f(x) = (ax + b)g(x)$ for some $a, b \in \mathbb{F}$ with $a \neq 0$ then $-b/a \in \mathbb{F}$ is a root. Finally, if $f(x) \in \mathbb{F}[x]$ has degree 2 or 3 then we observe that $f(x)$ is reducible if and only if $f(x)$ has a factor of degree 1. □

Here is an example to show that the trick does not work for polynomials of degree four.

Example: Leibniz' Mistake. One of the first problems of Calculus was to compute the antiderivative for any given rational function $f(x)/g(x)$ with $f(x), g(x) \in \mathbb{R}[x]$. By 1675, Leibniz knew that

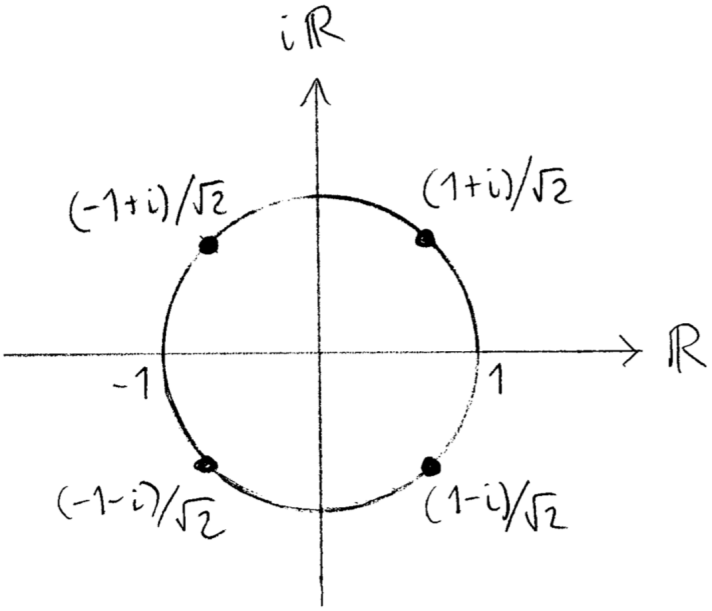
$$\int x^n dx = \frac{x^{n+1}}{n+1} \text{ (if } n \neq 1), \quad \int \frac{dx}{x} = \log(x), \quad \int \frac{dx}{x^2+1} = \arctan(x).$$

He also knew that if the denominator $g(x)$ can be factored into polynomials of degree 1 and 2, then the function $f(x)/g(x)$ can be expanded by partial fractions and hence the antiderivative can be computed from the above three formulas.

Today we know that every irreducible polynomial in $\mathbb{R}[x]$ has degree 1 or 2. (This is one way to state the Fundamental Theorem of Algebra.) But this fact is certainly not obvious. Indeed, Leibniz made a famous mistake in 1702 when he claimed that polynomials of the form $x^4 + a^4$ with $a \in \mathbb{R}$ and $a \neq 0$ are irreducible over $\mathbb{R}$. Here is his exact quote:

Therefore, $\int \frac{dx}{x^4+a^4}$ cannot be reduced to the squaring of the circle or the hyperbola by our analysis above, but founds a new kind of its own.[45]

Leibniz' problem was that he didn't have a good understanding of the complex 4th roots of -1. Today we know that these roots are the vertices of a square in the complex plane:



Then by grouping the complex roots of $x^4 + a^4$ into conjugate pairs we obtain

$$\begin{aligned} x^4 + a^4 &= \left(x - \frac{a(1+i)}{\sqrt{2}}\right)\left(x - \frac{a(1-i)}{\sqrt{2}}\right)\left(x - \frac{a(-1+i)}{\sqrt{2}}\right)\left(x - \frac{a(-1-i)}{\sqrt{2}}\right) \\ &= \left(x^2 - a\sqrt{2}x + a^2\right)\left(x^2 + a\sqrt{2}x + a^2\right), \end{aligned}$$

and it follows from this that the antiderivative of $1/(x^4 + a^4)$ **can** be expressed in terms of log and arctan (but I won't write the formula because it's too terrible). However, for our purposes, the main point of this example is that the polynomial $x^4 + a^4 \in \mathbb{R}[x]$ (for $a \neq 0$) is **reducible over** $\mathbb{R}$ but has **no roots in** $\mathbb{R}$. ///

To apply the Low-Degree Trick we still need some method to prove that a polynomial has no roots in a certain field. The following trick works when we are looking for roots in the field of fractions of a UFD.

The Rational Root Test. Let R be a UFD (for example, $\mathbb{Z}$) and consider a polynomial

$$f(x) = a_0 + a_1 x + \cdots + a_n x^n \in R[x].$$

[45] See Tignol, page 75. By "squaring of the circle" Leibniz means $\int \frac{dx}{x^2+1} = \arctan(x)$ and "squaring of the hyperbola" he means $\int \frac{dx}{x} = \log(x)$.

Since R is a UFD we can write any fraction $p/q \in \mathrm{Frac}(R)$ in "lowest terms," i.e., with $\gcd(p,q) = 1$. If $f(p/q) = 0$ then we must have

$$p|a_0 \quad \text{and} \quad q|a_n.$$

And these restrictions give us a finite list of possible roots p/q that we can check by hand.

Proof. Multiplying both sides of the equation $f(p/q) = 0$ by q^n gives

$$\begin{aligned} a_0 + a_1(p/q) + \cdots + a_{n-1}(p/q)^{n-1} + a_n(p/q)^n &= 0 \\ a_0 q^n + a_1 p q^{n-1} + \cdots + a_{n-1} p^{n-1} q + a_n p^n &= 0. \end{aligned}$$

Pulling $a_0 q^n$ to one side gives

$$a_0 q^n = -p(a_1 q^{n-1} + \cdots + a_{n-1} p^{n-2} q + a_n p^{n-1}) \quad \Longrightarrow \quad p|a_0 q^n,$$

which implies that $p|a_0$ because $\gcd(p,q) = 1$. Similarly, pulling $a_n p^n$ to one side gives

$$a_n p^n = -q(a_0 q^{n-1} + a_1 p q^{n-2} + \cdots + a_{n-1} p^{n-1}) \quad \Longrightarrow \quad q|a_n p^n,$$

which implies that $q|a_n$ because $\gcd(p,q) = 1$. □

[Remark: We just used the fact that $a|bc$ and $\gcd(a,b) = 1$ imply $a|c$. In a PID we can prove this by writing $ax + by = 1$ and then multiplying both sides by c. In a general UFD these x, y might not exist, but we can still prove the result by comparing prime factorizations. The details are not important.]

The tricks from the previous lecture are surprisingly useful. Here is an example that fills in a gap from our discussion in the introduction.

Example: The Splitting Field of $x^3 - 2$. The polynomial $x^3 - 2 \in \mathbb{Q}[x]$ has three distinct complex roots. To be specific, if $\alpha := \sqrt[3]{2} \in \mathbb{R}$ and $\omega := e^{2\pi i/3} \in \mathbb{C}$ then we can write

$$x^3 - 2 = (x - \alpha)(x - \omega\alpha)(x - \omega^2\alpha) \in \mathbb{C}[x].$$

Let $\mathbb{E} := \mathbb{Q}(\alpha, \omega\alpha, \omega^2\alpha) \subseteq \mathbb{C}$ be the splitting field. In the introduction I claimed that $[\mathbb{E}/\mathbb{Q}] = 6$ with basis $\{1, \alpha, \alpha^2, \omega, \omega\alpha, \omega\alpha^2\}$ but we were not able to prove this at the time. Now we can.

Proof. First observe that $\mathbb{Q}(\alpha, \omega\alpha, \omega^2\alpha) = \mathbb{Q}(\alpha, \omega)$ because $\{\alpha, \omega\alpha, \omega^2\alpha\}$ can be obtained from $\{\alpha, \omega\}$ through field operations and, conversely, $\{\alpha, \omega\}$ can be obtained from $\{\alpha, \omega\alpha, \omega^2\alpha\}$ through field operations. Consider the following chain of field extensions:

$$\mathbb{Q} \subseteq \mathbb{Q}(\alpha) \subseteq \mathbb{Q}(\alpha)(\omega) = \mathbb{E}.$$

Our goal is to compute a vector space basis for each extension and then combine them using Dedekind's Tower Law. We have already seen that $\mathbb{Q}(\alpha)/\mathbb{Q}$ has basis $\{1, \alpha, \alpha^2\}$ but let me prove this again quickly. Note that α is a root of $f(x) := x^3 - 2 \in \mathbb{Q}$. If $f(p/q) = 0$ is a rational root in lowest terms then the Rational Root Trick says that $p|2$ and $q|1$. But we can check by hand that ± 2 are **not** roots of $f(x)$. Since $f(x)$ has degree 3 and no rational roots we conclude that $f(x) = m_{\alpha/\mathbb{Q}}(x)$ is the minimal polynomial for α over $\mathbb{Q}$, and since $\deg(m_{\alpha/\mathbb{Q}}) = 3$ we conclude that $\{1, \alpha, \alpha^2\}$ is a basis for $\mathbb{Q}(\alpha)/\mathbb{Q}$.

To find a basis for $\mathbb{Q}(\alpha)(\omega)/\mathbb{Q}(\alpha)$ we need to compute the minimal polynomial:

$$m_{\omega/\mathbb{Q}(\alpha)}(x) \in \mathbb{Q}(\alpha)[x].$$

First of all, note that ω is a root of $x^3 - 1 \in \mathbb{Q}[x] \subseteq \mathbb{Q}(\alpha)[x]$. But this polynomial is **not irreducible** because

$$x^3 - 1 = (x-1)(x^2 + x + 1) \in \mathbb{Q}[x] \subseteq \mathbb{Q}(\alpha)[x].$$

Furthermore, since

$$(\omega - 1)(\omega^2 + \omega + 1) = \omega^3 - 1 = 0$$

we must have $\omega^2 + \omega + 1 = 0$. I claim that $g(x) := x^2 + x + 1 \in \mathbb{Q}(\alpha)[x]$ is irreducible and hence is the minimal polynomial for ω over $\mathbb{Q}(\alpha)$. Indeed, since $g(x)$ has degree 2 we only need to check that it has **no roots in the field** $\mathbb{Q}(\alpha)$. But we know that $g(x)$ has two **non-real roots** $\omega, \omega^2 \in \mathbb{C} - \mathbb{R}$ and since $\alpha \in \mathbb{R}$ we know that $\mathbb{Q}(\alpha)$ is contained in $\mathbb{R}$, hence

$$m_{\omega/\mathbb{Q}(\alpha)}(x) = x^2 + x + 1.$$

Then since $\deg(m_{\omega/\mathbb{Q}(\alpha)}) = 2$ we conclude that $\{1, \omega\}$ is a basis for $\mathbb{E} = \mathbb{Q}(\alpha)(\omega)/\mathbb{Q}(\alpha)$. Finally, by applying Dedekind's Tower Law we obtain the basis

$$\{1, \alpha, \alpha^2\} \cdot \{1, \omega\} = \{1 \cdot 1, \alpha \cdot 1, \alpha^2 \cdot 1, 1 \cdot \omega, \alpha \cdot \omega, \alpha^2 \cdot \omega\} = \{1, \alpha, \alpha^2, \omega, \omega\alpha, \omega\alpha^2\}$$

for the splitting field $\mathbb{E}/\mathbb{Q}$, and it follows that $[\mathbb{E}/\mathbb{Q}] = 6$. □

Remarks:

- After seeing that $\mathbb{Q}(\alpha, \omega\alpha, \omega^2\alpha) = \mathbb{Q}(\alpha, \omega)$ you might wonder if the splitting field can be generated by a single element:

  $$\mathbb{Q}(\alpha, \omega) = \mathbb{Q}(\gamma) \text{ for some } \gamma \in \mathbb{Q}(\alpha, \omega)?$$

  If this is possible then we will call γ a *primitive element* for the field extension.[46] Later we will prove that any finite dimensional extension over $\mathbb{Q}$ has a primitive element (in fact, infinitely many). However, it is not easy to find one by hand. For this example

[46] Another name for a primitive element is a *Galois resolvent*, hence the letter γ.

I used my computer to verify that $\gamma := \alpha + \omega$ is a primitive element with minimal polynomial[47]

$$m_{\gamma/\mathbb{Q}}(x) = x^6 + 3x^5 + 6x^4 + 3x^3 + 0x^2 + 9x + 9 \in \mathbb{Q}[x].$$

Note this polynomial has degree 6 as expected.

- Suppose that we have field extensions $\mathbb{E} \supseteq \mathbb{K} \supseteq \mathbb{F}$ and an element $\alpha \in \mathbb{E}$. If the minimal polynomial $m_{\alpha/\mathbb{K}}(x) \in \mathbb{K}[x]$ has coefficients in $\mathbb{F}$ then it necessarily follows that

$$m_{\alpha/\mathbb{K}}(x) = m_{\alpha/\mathbb{F}}(x) \in \mathbb{F}[x].$$

Proof. Any polynomial that is irreducible over $\mathbb{K}$ is still irreducible over $\mathbb{F}$. Then the result follows since $m_{\alpha/\mathbb{K}}(x) \in \mathbb{F}[x]$ is monic, irreducible and has α as a root. □

Thus from the above example we have

$$m_{\omega/\mathbb{Q}(\alpha)}(x) = m_{\omega/\mathbb{Q}}(x) = x^2 + x + 1.$$

///

We have seen that $x^2 + x + 1$ is the minimal polynomial over $\mathbb{Q}$ for the primitive third roots of unity. More generally, I claim that the following definition gives the minimal polynomial over $\mathbb{Q}$ for any primitive n-th root of unity.

Definition of Cyclotomic Polynomials. For any integer $n \geq 1$ we define

$$\Phi_n(x) := \prod_{\substack{0 \leq k < n \\ \gcd(k,n)=1}} (x - e^{2\pi i k/n}) \in \mathbb{C}[x].$$

///

At first it seems that cyclotomic polynomials have complex coefficients. However, on the next homework you will prove by induction that $\Phi_n(x) \in \mathbb{Z}[x] \subseteq \mathbb{Q}[x]$ for all $n \geq 1$. We have already seen the first few examples:

$$\begin{aligned}
\Phi_1(x) &= x - 1, \\
\Phi_2(x) &= x + 1, \\
\Phi_3(x) &= x^2 + x + 1.
\end{aligned}$$

For the next case, observe that the primitive 4th roots of unity are $\{\pm i\}$, hence

$$\Phi_4(x) = (x - i)(x + i) = x^2 + 1.$$

[47] In particular, this polynomial is irreducible over $\mathbb{Q}$. But I would never know that if you showed it to me out of context.

So far it is clear that each of these polynomials is irreducible over $\mathbb{Q}$. But you will show that

$$\Phi_5(x) = x^4 + x^3 + x^2 + x + 1,$$

and more generally that

$$\Phi_p(x) = x^{p-1} + x^{p-2} + \cdots + x^2 + x + 1 \quad \text{for any prime } p.$$

It is not clear why these polynomials should be irreducible over $\mathbb{Q}$. Gauss proved in the *Disquisitiones* that $\Phi_p(x) \in \mathbb{Z}[x]$ is irreducible for any prime p. The typical textbook proof of this uses a clever trick called "Eisenstein's Criterion," which was communicated in an 1850 letter from Gotthold Eisenstein to Gauss. It is also true, but quite tricky to prove, that $\Phi_n(x) \in \mathbb{Q}[x]$ is irreducible for any n. Then it follows that $\Phi_n(x) \in \mathbb{Q}[x]$ is the minimal polynomial for any primitive root of n over $\mathbb{Q}$, and hence the dimension of the *cyclotomic field* $\mathbb{Q}(e^{2\pi i/n})/\mathbb{Q}$ is equal to Euler's totient function:

$$[\mathbb{Q}(e^{2\pi i/n})/\mathbb{Q}] = \phi(n) = \#\{0 \leq k < n : \gcd(k, n) = 1\}.$$

Any why did Gauss care about this? His original goal was to investigate whether the n-th roots of unity can be expressed in terms of square roots.

Definition of Constructible Numbers. We say that a complex number $\alpha \in \mathbb{C}$ is *constructible* if it can be obtained from $\mathbb{Q}$ by solving a sequence of quadratic equations, i.e., if there exists a chain of fields

$$\alpha \in \mathbb{F}_k \supseteq \mathbb{F}_{k-1} \supseteq \cdots \supseteq \mathbb{F}_1 \supseteq \mathbb{F}_0 = \mathbb{Q}$$

satisfying $[\mathbb{F}_{i+1}/\mathbb{F}_i] = 2$ for all i. ///

The motivation for the word "constructible" comes from Euclidean geometry. Suppose that we start with the points $(0, 0)$ and $(1, 0)$ in the Cartesian plane $\mathbb{R}^2$. From these two points we are allowed to construct new points via Euclid's Postulates:

- We are allowed to draw the straight line through any two points.
- Given points $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$ we are allowed to draw the circle through $\mathbf{y}$ with center at $\mathbf{x}$.
- We are allowed to draw the points of intersection for any constructed lines and circles.

One can check that the points of intersection of any two lines and circles can always be computed by a quadratic equation,[48] hence any point $(\alpha, \beta) \in \mathbb{R}^2$ that is constructible in the geometric sense will have coordinates $\alpha, \beta \in \mathbb{R}$ that are constructible in the algebraic sense.[49]

[48] The hardest case is the intersection of two circles.

[49] The converse is also true but I feel no need to discuss this.

The young Gauss applied this reasoning to the construction of regular polygons. He completed the *Disquisitiones Arithmeticae* in 1798, at the age of 21. The final chapter of this work contains a study of "cyclotomy." We can summarize the main points as follows:

$$\begin{aligned}
&\text{the regular } n\text{-gon is constructible} \\
&\iff \text{the point } (\cos(2\pi/n), \sin(2\pi/n)) \in \mathbb{R}^2 \text{ is constructible} \\
&\iff \text{the numbers } \cos(2\pi/n), \sin(2\pi/n) \in \mathbb{R} \text{ are constructible} \\
&\iff \text{the number } e^{2\pi i/n} \in \mathbb{C} \text{ is constructible} \\
&\iff \text{the dimension } [\mathbb{Q}(e^{2\pi i/n})/\mathbb{Q}] \text{ is a power of 2} \\
&\iff e^{2\pi i/n} \in \mathbb{C} \text{ has minimal polynomial with degree a power of 2} \\
&\iff \text{the cyclotomic polynomial } \Phi_n(x) \text{ has degree a power of 2} \\
&\iff \text{Euler's totient } \phi(n) \text{ is a power of 2.}
\end{aligned}$$

Some of these implications were filled in by Pierre Wantzel in 1837, when he was 23 years old.[50] Hence this result is sometimes called the **Gauss-Wantzel Theorem**.

For example, by observing that $\phi(7) = 6$ is not a power of 2 we can explain why the ancient Greeks were never able to construct a regular heptagon with straightedge and compass. (You will give a more elementary proof of this fact on the next homework.) However, the more surprising result is the existence of constructible polygons that the ancient Greeks missed. By observing that $\phi(17) = 16$ is a power of 2, Gauss was able to prove (indirectly) that

the regular 17*-gon is constructible with straightedge and compass!*

Week 20

This week we will apply our knowledge of irreducible polynomials to the construction of finite fields. We already know that finite fields exist since $\mathbb{F}_p := \mathbb{Z}/p\mathbb{Z}$ is a field for any prime $p \in \mathbb{Z}$. But are there any other finite fields?

Suppose that $\mathbb{E}$ is a finite field. This implies that $\mathbb{E}$ has characteristic $p > 0$ since otherwise the prime subfield would be $\mathbb{Q}$, which is infinite. So let $\mathbb{F}_p \subseteq \mathbb{E}$ be the prime subfield and consider the vector space $\mathbb{E}/\mathbb{F}_p$. Since $\mathbb{E}$ is **finite** we know that this vector space is **finite-dimensional**, say $[\mathbb{E}/\mathbb{F}_p] = k$. In this case I claim that $\#\mathbb{E} = p^k$.

Proof. Let $\alpha_1, \alpha_2, \ldots, \alpha_k \in \mathbb{E}$ be a basis for $\mathbb{E}$ as a vector space over $\mathbb{F}_p$. By definition, every element of $\mathbb{E}$ can be expressed uniquely in the form

$$a_1\alpha_1 + a_2\alpha_2 + \cdots + a_k\alpha_k \text{ for some } a_1, a_2, \ldots, a_k \in \mathbb{F}_p.$$

Then since there are p ways to choose each coefficient we conclude that

$$\#E = (\#\text{ choices for } a_1)(\#\text{ choices for } a_2)\cdots(\#\text{ choices for } a_k) = p^k.$$

[50] Abel died in 1829 at age 26 and Galois died in 1832 at age 20. For some reason there were a lot of precocious mathematicians in the early 1800s.

□

But we still have not seen any fields of size p^k with $k \geq 2$. Here is our first example.

Example: A Field of Size Four. Consider the polynomial $x^2+x+1 \in \mathbb{F}_2[x]$ with coefficients in the field of two elements $\mathbb{F}_2 = \{0, 1\}$. It is easy to see that this polynomial is irreducible over $\mathbb{F}_2$ because it has no roots in $\mathbb{F}_2$:

x	0	1
x^2+x+1	1	1

Since $\text{char}(\mathbb{F}_2) = 2 \neq 0$ the Fundamental Theorem of Algebra doesn't tell us anything about the existence of roots, so we have to apply Kronecker's Theorem. Specifically, since $\mathbb{F}_2[x]$ is a PID it follows that the ideal $\langle x^2+x+1\rangle \subseteq \mathbb{F}_2[x]$ is maximal, hence we obtain a field:

$$\mathbb{E} := \frac{\mathbb{F}_2[x]}{\langle x^2+x+1\rangle}.$$

If we identify $\mathbb{F}_2$ with the subfield $\{a+\langle x^2+x+1\rangle : a \in \mathbb{F}_2\} \subseteq \mathbb{E}$ then we can think of $\mathbb{E} \supseteq \mathbb{F}_2$ as a field extension which contains an element $\alpha \in \mathbb{E}$ satisfying

$$\alpha^2+\alpha+1 = 0.^{51}$$

In fact, since $x^2+x+1 \in \mathbb{F}_2[x]$ is monic, irreducible and has $\alpha \in \mathbb{E}$ as a root, we conclude that it is the minimal polynomial for α over $\mathbb{F}_2$ and it follows from this that $\mathbb{E} = \mathbb{F}_2(\alpha)$:

$$\mathbb{E} = \frac{\mathbb{F}_2[x]}{\langle x^2+x+1\rangle} = \frac{\mathbb{F}_2[x]}{\langle m_{\alpha/\mathbb{F}_2}(x)\rangle} \cong \mathbb{F}_2(\alpha) \subseteq \mathbb{E}.$$

Furthermore, since the minimal polynomial has degree 2 we conclude that $\{1, \alpha\}$ is a basis for $\mathbb{E}$ over $\mathbb{F}_2$ and it follows that

$$\mathbb{E} = \{0+0\alpha, 1+0\alpha, 0+1\alpha, 1+1\alpha\} = \{0, 1, \alpha, 1+\alpha\}.$$

Thus we have constructed a field of size four. The addition table is just inherited from the vector space structure of $\mathbb{E}/\mathbb{F}_2$:

$+$	0	1	α	$1+\alpha$
0	0	1	α	$1+\alpha$
1	1	0	$1+\alpha$	α
α	α	$1+\alpha$	0	1
$1+\alpha$	$1+\alpha$	α	1	0

[51] Technically, $\alpha = x + \langle x^2+x+a\rangle$ is the coset generated by x but from this point on we will just call it α. Prior to Kronecker's Theorem the "imaginary" roots of polynomials over finite fields were called "Galois imaginaries" and their nature was somewhat mysterious.

But the multiplication table is more interesting because it uses the polynomial relation

$$\alpha^2 + \alpha + 1 = 0 \quad \Longrightarrow \quad \alpha^2 = -1 - \alpha = 1 + \alpha.$$

For example, we have $(1+\alpha)^2 = 1^2 + 2\alpha + \alpha^2 = 1 + 0 + (1+\alpha) = \alpha$. Here is the full table:

$\times$	0	1	α	$1+\alpha$
0	0	0	0	0
1	0	1	α	$1+\alpha$
α	0	α	$1+\alpha$	1
$1+\alpha$	0	$1+\alpha$	1	α

In fact one could use this table as the **definition of multiplication** in $\mathbb{E}$ and then check by hand that all of the field axioms are satisfied. However, that would leave the existence of $\mathbb{E}$ completely unexplained. ///

The construction of the field $\mathbb{E} = \{0, 1, \alpha, 1+\alpha\}$ above might have seemed rather arbitrary, but I claim that there were no other options.

Theorem (There is Only One Field of Size Four). If $\mathbb{E}'$ is any field of size four then we have a ring isomorphism $\mathbb{E}' \cong \mathbb{E}$.

Proof. Let $\#\mathbb{E}' = 4$. Then from previous remarks we know that $\mathbb{E}'$ is a 2-dimensional vector space over $\mathbb{F}_2$. Extend the set $\{1\}$ to a basis $\{1, \gamma\}$. Then by definition we must have

$$\mathbb{E}' = \{0, 1, \gamma, 1+\gamma\}.$$

Clearly we have a vector space isomorphism identifying $\alpha \leftrightarrow \gamma$. But does this isomorphism also preserve multiplication? For this we need to prove that $\gamma^2 = 1 + \gamma$. So consider the element $\gamma^2 \in \{0, 1, \gamma, 1+\gamma\}$. Since $\gamma \neq 0$ in a field we must have $\gamma^2 \neq 0$. Then since $\gamma \notin \{0, 1\}$ in a field we must have $\gamma^2 \neq \gamma$. Finally, assume for contradiction that we have $\gamma^2 = 1$, so that $1 - \gamma^2 = 0$. But then we have

$$0 = 1 - \gamma^2 = (1-\gamma)(1+\gamma) = (1+\gamma)(1+\gamma) = (1+\gamma)^2,$$

which contradicts the fact that $1+\gamma \neq 0$. By process of elimination we conclude that $\gamma^2 = 1+\gamma$ and hence $\mathbb{E}' \cong \mathbb{E}$ as rings. □

The next-smallest non-trivial power of a prime is $2^3 = 8$.

Example: Two Fields of Size Eight? Based on the previous example, we will be able to construct a field of size 8 if we can find an irreducible polynomial in $\mathbb{F}_2[x]$ of degree 3. In fact,

there are two such polynomials! Indeed, the polynomials x^3+x^2+1 and x^3+x+1 are both irreducible over $\mathbb{F}_2$ since they each have degree 3 and no roots in $\mathbb{F}_2$:

x	0	1
x^3+x^2+1	1	1
x^3+x+1	1	1

This guarantees that the following two vector spaces over $\mathbb{F}_2$ are fields:

$$\mathbb{E} := \{a+b\alpha+c\alpha : a,b,c \in \mathbb{F}_2, \alpha^3+\alpha^2+1=0\},$$
$$\mathbb{E}' := \{a+b\beta+c\beta^2 : a,b,c \in \mathbb{F}_2, \beta^3+\beta+1=0\}.$$

For your information, here is the multiplication table of the field $\mathbb{E}$:

$\times$	0	1	α	$1+\alpha$	α^2	$1+\alpha^2$	$\alpha+\alpha^2$	$1+\alpha+\alpha^2$
0	0	0	0	0	0	0	0	0
1	0	1	α	$1+\alpha$	α^2	$1+\alpha^2$	$\alpha+\alpha^2$	$1+\alpha+\alpha^2$
α	0	α	α^2	$\alpha+\alpha^2$	$1+\alpha^2$	$1+\alpha+\alpha^2$	1	$1+\alpha$
$1+\alpha$	0	$1+\alpha$	$\alpha+\alpha^2$	$1+\alpha^2$	1	α	$1+\alpha+\alpha^2$	α^2
α^2	0	α^2	$1+\alpha^2$	1	$1+\alpha+\alpha^2$	$1+\alpha$	α	$\alpha+\alpha^2$
$1+\alpha^2$	0	$1+\alpha^2$	$1+\alpha+\alpha^2$	α	$1+\alpha$	$\alpha+\alpha^2$	α^2	1
$\alpha+\alpha^2$	0	$\alpha+\alpha^2$	1	$1+\alpha+\alpha^2$	α	α^2	$1+\alpha$	$1+\alpha^2$
$1+\alpha+\alpha^2$	0	$1+\alpha+\alpha^2$	$1+\alpha$	α^2	$\alpha+\alpha^2$	1	$1+\alpha^2$	α

And here is the multiplication table of $\mathbb{E}'$:

$\times$	0	1	β	$1+\beta$	β^2	$1+\beta^2$	$\beta+\beta^2$	$1+\beta+\beta^2$
0	0	0	0	0	0	0	0	0
1	0	1	β	$1+\beta$	β^2	$1+\beta^2$	$\beta+\beta^2$	$1+\beta+\beta^2$
β	0	β	β^2	$\beta+\beta^2$	$1+\beta$	1	$1+\beta+\beta^2$	$1+\beta^2$
$1+\beta$	0	$1+\beta$	$\beta+\beta^2$	$1+\beta^2$	$1+\beta+\beta^2$	β^2	1	β
β^2	0	β^2	$1+\beta$	$1+\beta+\beta^2$	$\beta+\beta^2$	β	$1+\beta^2$	1
$1+\beta^2$	0	$1+\beta^2$	1	β^2	β	$1+\beta+\beta^2$	$1+\beta$	$\beta+\beta^2$
$\beta+\beta^2$	0	$\beta+\beta^2$	$1+\beta+\beta^2$	1	$1+\beta^2$	$1+\beta$	β	β^2
$1+\beta+\beta^2$	0	$1+\beta+\beta^2$	$1+\beta^2$	β	1	$\beta+\beta^2$	β^2	$1+\beta$

///

Even though these two multiplication tables look completely different I claim that

$$\mathbb{E} \cong \mathbb{E}'.$$

Proof. It is difficult to find an isomorphism by hand so we will use an indirect method. First observe that $m_{\alpha/\mathbb{F}_2}(x) = x^3+x^2+1$ is the minimal polynomial for $\alpha/\mathbb{F}_2$, so that

$$\mathbb{E} = \mathbb{F}_2(\alpha) \cong \frac{\mathbb{F}_2[x]}{\langle m_{\alpha/\mathbb{F}_2}(x)\rangle} = \frac{\mathbb{F}_2[x]}{\langle x^3+x^2+1\rangle}.$$

Sadly, β does not satisfy the same equation. However, if we can prove that there exists **some** element $\gamma \in \mathbb{E}'$ satisfying $\gamma^3 + \gamma^2 + 1 = 0$ then since $x^3 + x^2 + 1$ is irreducible over $\mathbb{F}_2$ we will conclude that $m_{\gamma/\mathbb{F}_2}(x) = x^3 + x^2 + 1$ and hence

$$\mathbb{E} = \mathbb{F}_2(\alpha) \cong \frac{\mathbb{F}_2[x]}{\langle m_{\alpha/\mathbb{F}_2}(x)\rangle} = \frac{\mathbb{F}_2[x]}{\langle m_{\gamma/\mathbb{F}_2}(x)\rangle} \cong \mathbb{F}_2(\gamma) \subseteq \mathbb{E}'.$$

Finally, since $\mathbb{F}_2(\gamma)$ and $\mathbb{E}'$ both have size 8 we will conclude that $\mathbb{E} \cong \mathbb{F}_2(\gamma) = \mathbb{E}'$.

To prove the existence of such an element we consider the group of units $(\mathbb{E}^\times, \times, 1)$. Since $\#\mathbb{E}^\times$ has size 7, Lagrange's Theorem tells us that $\upsilon^7 = 1$ for all $\upsilon \in \mathbb{E}^\times$. In particular, since $\alpha \in \mathbb{E}^\times$ we must have $\alpha^7 - 1 = 0$. Then since $x^3 + x^2 + 1$ is the minimal polynomial for $\alpha/\mathbb{F}_2$ we conclude that

$$(x^3 + x^2 + 1)f(x) = (x^7 - 1) \text{ for some } f(x) \in \mathbb{F}_2[x] \text{ of degree 4.}$$

Next consider any non-zero element $\gamma \in \mathbb{E}'$. Since the group of units of $\mathbb{E}'$ also has size 7 we conclude again from Lagrange's Theorem that $\gamma^7 = 1$ and hence

$$(\gamma^3 + \gamma^2 + 1)f(\gamma) = (\gamma^7 - 1) = 0.$$

Since this is true for all $0 \neq \gamma \in \mathbb{E}'$ and since $f(x)$ has at most 4 roots in $\mathbb{E}'$, we conclude that there exist at least three (hence exactly three) elements $\gamma \in \mathbb{E}'$ such that $\gamma^3 + \gamma^2 + 1 = 0$. $\square$

That's the best I can do by hand. To be more explicit, I used my computer to check that

$$\gamma = 1 + \beta, \quad 1 + \beta^2, \quad 1 + \beta + \beta^2$$

are the three promised roots of $x^3 + x^2 + 1$ in the field $\mathbb{E}'$. Then by sending $\alpha \mapsto \gamma$ we obtain the following three explicit isomorphisms:

0	1	α	$1+\alpha$	α^2	$1+\alpha^2$	$\alpha+\alpha^2$	$1+\alpha+\alpha^2$
0	1	$1+\beta$	β	$1+\beta^2$	β^2	$\beta+\beta^2$	$1+\beta+\beta^2$
0	1	$1+\beta^2$	β^2	$1+\beta+\beta^2$	$\beta+\beta^2$	β	$1+\beta$
0	1	$1+\beta+\beta^2$	$\beta+\beta^2$	$1+\beta$	β	β^2	$1+\beta^2$

We will prove later that there can be no other isomorphisms $\mathbb{E} \cong \mathbb{E}'$ because the Galois group $\mathrm{Gal}(\mathbb{E}/\mathbb{F}_2)$ has size $3 = [\mathbb{E}/\mathbb{F}_2]$.

Note that the above proof does not imply that **all** fields of size 8 are isomorphic, just these two **particular** fields of size 8. It happens to be true that any two finite fields of the same size are isomorphic but in order to prove this we need an extra ingredient called the Primitive Root Theorem.

The existence of finite fields beyond $\mathbb{Z}/p\mathbb{Z}$ was discovered by Galois.[52] However, the concept of isomorphism is more modern. E. H. Moore first stated and proved the uniqueness of finite fields in *A Doubly-Infinite System of Simple Groups* (1896), which was read at the International Mathematical Congress in Chicago in 1893.[53] This is the same paper in which he introduced the English term "field" for the German "Körper." Moore denoted the unique field of size p^k by GF$[p^k]$ for "Galois field," but I will use the modern notation $\mathbb{F}_{p^k}$.

In the next two lectures we will complete our discussion of finite fields by proving that for all $p, k \in \mathbb{Z}$ with p prime and $k \geq 1$, there **exists** a field of size p^k which is **unique** up to isomorphism. The full proof will require three lemmas, two of which you will prove on the homework. The first lemma shows that any finite field whatsoever has the form $\mathbb{F}_p[x]/\langle f(x)\rangle$ for some irreducible polynomial $f(x) \in \mathbb{F}_p[x]$.

Lemma (Primitive Root Theorem). If $\mathbb{E}$ is a finite field then $(\mathbb{E}^\times, \times, 1)$ is a cyclic group.

To be specific, let $\mathbb{F}_p \subseteq \mathbb{E}$ be the prime subfield and suppose that $[\mathbb{E}/\mathbb{F}_p] = k$, hence $\#\mathbb{E} = p^k$. Since $\mathbb{E}^\times$ is cyclic we can write $\mathbb{E} = \{0, 1, \alpha, \alpha^2, \ldots, \alpha^{p^k-2}\}$ for some $\alpha \in \mathbb{E}$. Then since every element of $\mathbb{E}$ can be expressed in terms of $\mathbb{F}_p \cup \{\alpha\}$ using field operations we conclude that

$$\mathbb{E} = \mathbb{F}_p(\alpha) \cong \frac{\mathbb{F}_p[x]}{\langle m_{\alpha/\mathbb{F}_p}(x)\rangle} \qquad \text{with } \deg(m_{\alpha/\mathbb{F}_p}) = k.$$

Proof. Homework. □

Remarks:

- Recall that a generator of the group $\langle e^{2\pi i/n}\rangle \subseteq \mathbb{C}^\times$ is called a *primitive n-th root of unity*. The number of generators is $\phi(n)$ and they are given by $e^{2\pi ik/n}$ for $\gcd(k, n) = 1$. Since $\langle e^{2\pi i/n}\rangle \cong \mathbb{Z}/n\mathbb{Z}$ the term "primitive roots" can also be applied to the additive generators of $\mathbb{Z}/n\mathbb{Z}$.
- In the *Disquisitiones Arithmeticae* (1801) Gauss proved for any prime p that the multiplicative group $(\mathbb{Z}/p\mathbb{Z})^\times$ is cyclic. He applied the term *primitive root* to any generator of this group and said that he was following Euler's notation.
- Next, Galois claimed without showing any details that Gauss' proof can be extended to show that the fields $GF(p^k)$ also have cyclic groups of units. He followed Gauss in calling the generators *primitive roots*.

[52] Gauss probably discovered them independently but he didn't publish the results. Gauss' approach to publication was described by his motto: *Pauca sed matura* (Few, but ripe). His extensive mathematical notebooks were published after his death and complicated many issues of priority.

[53] The main topic of the paper is the family $PSL_2(p^k)$ of finite simple groups.

- So now the term "primitive root" had three different meanings. We can extend the meaning yet again by observing that if $\alpha \in \mathbb{E}^\times$ is a primitive root (multiplicative generator for the units of a finite field) then it follows that $\mathbb{E} = \mathbb{F}_p(\alpha)$, so that α is a generator of the field extension $\mathbb{E} \supseteq \mathbb{F}_p$.
- Finally, if $\mathbb{E} = \mathbb{F}(\gamma) \supseteq \mathbb{F}$ is any field extension generated by a single element $\gamma \in \mathbb{E}$ then $\gamma \in \mathbb{E}$ is called a *primitive element* for the extension. Later we will prove the so-called Primitive Element Theorem, which says that a primitive element exists when $\text{char}(\mathbb{F}) = 0$ and $[\mathbb{E}/\mathbb{F}] < \infty$.
- In conclusion, the terms "primitive root" and "primitive element" are confusing and terrible. I prefer the term "Galois resolvent" instead of "primitive element," but the damage has already been done. The most I can do is warn you.

///

Theorem (Uniqueness of Finite Fields). Let $\mathbb{E}$ and $\mathbb{E}'$ be finite fields. Then

$$\#\mathbb{E} = \#\mathbb{E}' \quad \Longrightarrow \quad \mathbb{E} \cong \mathbb{E}'.$$

Proof. Let $\mathbb{F}_p \subseteq \mathbb{E}$ be the prime subfield and suppose that $[\mathbb{E}/\mathbb{F}_p] = k$, hence $\#\mathbb{E} = p^k$. From the Primitive Root Theorem there exists some $\alpha \in \mathbb{E}$ of multiplicative order $p^k - 1$. It follows from this that $\mathbb{E} = \{0, 1, \alpha, \alpha^2, \ldots, \alpha^{p^k-2}\} = \mathbb{F}_p(\alpha)$ and hence

$$\mathbb{E} = \mathbb{F}_p(\alpha) \cong \frac{\mathbb{F}_p[x]}{\langle m_{\alpha/\mathbb{F}_p}(x) \rangle} \quad \text{with } \deg(m_{\alpha/\mathbb{F}_p}) = k.$$

Since $\alpha^{p^k-1} = 1$ we also know that

$$m_{\alpha/\mathbb{F}_p}(x) f(x) = \left(x^{p^k-1} - 1\right) \text{ for some } f(x) \in \mathbb{F}_p[x] \text{ of degree } p^k - 1 - k.$$

Now consider the field $\mathbb{E}'$, which also has size p^k. Since the group of units of $\mathbb{E}'$ has size $p^k - 1$ we conclude from Lagrange's Theorem that $\gamma^{p^k-1} = 1$ and hence

$$m_{\alpha/\mathbb{F}_p}(\gamma) f(\gamma) = \left(\gamma^{p^k-1} - 1\right) = 0 \quad \text{for all } 0 \neq \gamma \in \mathbb{E}'.$$

Since this holds for $p^k - 1$ distinct values of γ and since $\deg(f) < p^k - 1$ there must exist some $\gamma \in \mathbb{E}'$ such that $m_{\alpha/\mathbb{F}_p}(\gamma) = 0$ and hence $m_{\alpha/\mathbb{F}_p}(x) = m_{\gamma/\mathbb{F}_p}(x) \in \mathbb{F}_p[x]$. It follows that

$$\mathbb{E} = \mathbb{F}_p(\alpha) \cong \frac{\mathbb{F}_p[x]}{\langle m_{\alpha/\mathbb{F}_p}(x) \rangle} = \frac{\mathbb{F}_p[x]}{\langle m_{\gamma/\mathbb{F}_p}(x) \rangle} \cong \mathbb{F}_p(\gamma) \subseteq \mathbb{E}'.$$

Finally, since $\#\mathbb{F}_p(\gamma) = \#\mathbb{E} = \#\mathbb{E}'$ we conclude that

$$\mathbb{E} \cong \mathbb{F}_p(\gamma) = \mathbb{E}'.$$

□

[Remark: In fact, since $x^{p^k-1}-1$ splits in $\mathbb{E}'[x]$ we conclude that $m_{\alpha/\mathbb{F}_p}(x)$ also splits in $\mathbb{E}'[x]$.[54] Then from the Repeated Root Lemma below, this implies that $m_{\alpha/\mathbb{F}_p}(x)$ has k distinct roots $\gamma \in \mathbb{E}'$, leading to k distinct isomorphisms $\mathbb{E} \cong \mathbb{E}'$. We will see later that there can be no other isomorphisms between $\mathbb{E}$ and $\mathbb{E}'$.]

So far we have proved that:

- Any finite field has size p^k for some prime p.
- Any two finite fields of the same size are isomorphic.

We have also see that irreducible polynomials in $\mathbb{F}_p[x]$ can be used to create finite fields. Indeed, if $f(x) \in \mathbb{F}_p[x]$ is irreducible of degree k then we obtain a field of size p^k:

$$\# \left(\frac{\mathbb{F}_p[x]}{\langle f(x) \rangle} \right) = p^k.$$

But it not obvious whether irreducible polynomials **exist**. Gauss gave a tricky proof for the existence of irreducible polynomials in the *Disquisitiones*. Galois was inspired by Gauss' work and he came up with an elegant direct proof for the existence of finite fields, which does not assume the existence of an irreducible polynomial, but obtains one as a corollary.

The proof requires two more lemmas, one of which you will prove on the homework.

Lemma (Repeated Roots). Let $\mathbb{F}$ be a field and let $D : \mathbb{F}[x] \to \mathbb{F}[x]$ be the "formal derivative" of polynomials. Given a polynomial $f(x) \in \mathbb{F}[x]$ we say that $\alpha \in \mathbb{F}[x]$ is a *repeated root* of $f(x)$ if $f(x) = (x-\alpha)^2 g(x)$ for some $g(x)$. Then I claim that

$$\alpha \text{ is a repeated root of } f(x) \qquad \Longleftrightarrow \qquad f(\alpha) = 0 \quad \text{and} \quad Df(\alpha) = 0.$$

Proof. Homework. □

Lemma (The Frobenius Endomorphism).[55] Let R be any ring of **prime characteristic** p. Then the map $a \mapsto a^p$ defines a ring homomorphism $R \to R$, called the *Frobenius endomorphism* of R.

[54] Here were are using the fact that $\mathbb{E}'[x]$ is a UFD.

[55] This result is sometimes called the "Freshman's Binomial Theorem," which I think is undignified.

Proof. Note that $0^p = 0$ and $1^p = 1$, and for any $a, b \in R$ note that $(ab)^p = a^p b^p$. It only remains to show that $(a+b)^p = a^p + b^p$ for all $a, b \in R$. By definition we say that R has characteristic p when the unique ring homomorphism $\iota : \mathbb{Z} \to R$ has kernel $p\mathbb{Z}$. Then for all $n \in p\mathbb{Z}$ and $a \in R$ it follows that $\iota(n)a = 0$. Now recall the binomial theorem:

$$(a+b)^p = a^p + b^p + \sum_{k=1}^{p-1} \iota\left(\frac{p!}{k!(p-k)!}\right) a^k b^{p-k} \in R.$$

Now let $1 \leq k \leq p-1$ and consider the prime factorization of the binomial coefficient $p!/\left[k!(p-k)!\right] \in \mathbb{Z}$. Clearly p divides the numerator. But the denominator $k!(p-k)!$ is a product of integers, each of which is smaller than p. Thus it follows from Euclid's lemma that p does **not** divide the denominator, and we conclude that

$$\frac{p!}{k!(p-k)!} \in p\mathbb{Z} \qquad \Longrightarrow \qquad \iota\left(\frac{p!}{k!(p-k)!}\right) = 0 \in R.$$

□

The following proof comes from Galois' paper *On the theory of numbers* (1830). This is the reason why finite fields are sometimes called "Galois fields."

Theorem (Existence of Finite Fields). For any integers $p, k \geq 1$ with p prime there exists a field of size p^k. In fact, I claim that any splitting field of $x^{p^k} - x \in \mathbb{F}_p[x]$ has size p^k.

Proof. The idea of this proof is due to Galois. Let $\mathbb{E} \supseteq \mathbb{F}_p$ be a splitting field for the polynomial $f(x) := x^{p^k} - x \in \mathbb{F}_p[x]$. From the Repeated Root Lemma we know that if $\alpha \in \mathbb{E}$ is a repeated root of $f(x)$ then we must have $f(\alpha) = 0$ and $Df(\alpha) = 0$. But the derivative is $Df(x) = p^k x^{p^k - 1} - 1 = 0 - 1 = -1 \in \mathbb{F}_p[x]$, which has no roots at all. It follows that $f(x)$ has p^k distinct roots in $\mathbb{E}$. Let $\Omega \subseteq \mathbb{E}$ be the set of roots. We will show that in fact $\Omega \subseteq \mathbb{E}$ is a subfield, hence $\Omega = \mathbb{E}$ is our desired field of size p^k.

Indeed, Ω contains 0 and 1. Furthermore, if $f(\alpha) = 0$ and $f(\beta) = 0$ with $\alpha \neq 0$ then we have

$$(\alpha\beta)^{p^k} = \alpha^{p^k}\beta^{p^k} = \alpha\beta \quad \Longrightarrow \quad f(\alpha\beta) = 0$$

and

$$\left(\alpha^{-1}\right)^{p^k} = \left(\alpha^{p^k}\right)^{-1} = \alpha^{-1} \quad \Longrightarrow \quad f(\alpha^{-1}) = 0.$$

Finally, by applying the Frobenius Automorphism k times we obtain

$$(\alpha+\beta)^{p^k} = (\alpha^p + \beta^p)^{p^{k-1}} = (\alpha^{p^2} + \beta^{p^2})^{p^{k-2}} = \cdots = \alpha^{p^k} + \beta^{p^k} = \alpha + \beta,$$

and hence $f(\alpha + \beta) = 0$. □

Notation. We have seen that there exists a unique field of size p^k for any integers $p, k \geq 1$ with p prime. We will use the following notation for this field:

$$\boxed{\mathbb{F}_{p^k} := \text{the unique field of size } p^k.}$$

Observe that this agrees with the earlier notation $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$. ///

Corollary (Existence of Irreducible Polynomials). For any integers $p, k \geq 1$ with p prime there exists at least one irreducible polynomial $f(x) \in \mathbb{F}_p[x]$ of degree k.

Proof. Consider the field $\mathbb{F}_{p^k}$. By Lagrange's Theorem, the prime subfield (being an additive subgroup) must have size dividing p^k, hence the prime subfield is $\mathbb{F}_p \subseteq \mathbb{F}_{p^k}$. Observe that $[\mathbb{F}_{p^k}/\mathbb{F}_p] = k$ since for any finite-dimensional vector space V over $\mathbb{F}_p$ we have

$$\#V = p^{\dim(V)}.$$

Next, recall recall from the Primitive Root Theorem that there exists an element $\gamma \in \mathbb{F}_{p^k}$ such that $\mathbb{F}_{p^k} = \mathbb{F}_p(\gamma)$. Finally, we have

$$[\mathbb{F}_p(\gamma)/\mathbb{F}_p] = [\mathbb{F}_{p^k}/\mathbb{F}_p] = k,$$

which implies that the minimal polynomial $m_{\gamma/\mathbb{F}_p}(x) \in \mathbb{F}_p[x]$ has degree k. (Recall that minimal polynomials are always irreducible.) □

[Remark: It is not necessarily easy to find an irreducible polynomial of a given degree.]

Remarks:

- It follows from the uniqueness of finite fields that any two splitting fields of $x^{p^k} - x \in \mathbb{F}_p[x]$ are isomorphic. Next week we will prove that the same result holds for the splitting fields of **any** polynomial over **any** field.
- Conversely, let $\mathbb{E}$ be any field of size p^k with prime subfield $\mathbb{F}_p \subseteq \mathbb{E}$. From Lagrange's Theorem applied to the group of units, one can show that **every** element of $\mathbb{E}$ is a root of $x^{p^k-1} - x \in \mathbb{F}_p[x]$ and hence $\mathbb{E}$ is a splitting field for this polynomial. Thus the uniqueness of splitting fields will give a new proof for the uniqueness of finite fields.
- As I mentioned at the beginning of this lecture, Gauss proved in the *Disquisitiones* that there exist irreducible polynomials of all degrees in $\mathbb{F}_p[x]$. The way he did this was to first **count** the polynomials. To be specific, he first showed that the number of irreducible polynomials of degree k over $\mathbb{F}_p$ is given by

$$\frac{1}{k} \sum_{d|k} \mu(k/d) p^d,$$

where $\mu : \mathbb{Z} \to \{-1, 0, 1\}$ is the number-theoretic *möbius function*. Then he gave a tricky argument that this formula can never equal zero. This result can be viewed as the proof of existence for finite fields, but Gauss never discussed this in print. It turns out that Gauss privately developed a full theory of finite fields in parallel with Galois' theory, but this was but this was only discovered after his death in 1855.[56]

Problem Set 10

1. Computing Minimal Polynomials. Define $\alpha := \sqrt[3]{2} \in \mathbb{R}$ and $\omega := e^{2\pi i/3} \in \mathbb{C}$.

(a) Prove that $x^3 - 2$ is the minimal polynomial for α over $\mathbb{Q}(\omega)$.

(b) Prove that $x^2 + x + 1$ is the minimal polynomial for ω over $\mathbb{Q}(\alpha\omega)$.

(c) Prove that $x^2 + (\alpha\omega)x + (\alpha\omega)^2$ is the minimal polynomial for α over $\mathbb{Q}(\alpha\omega)$.

[Hint: Consider any $\beta \in \mathbb{E} \supseteq \mathbb{F}$ and let $f(x) \in \mathbb{F}[x]$ be a polynomial satisfying $\deg(f) = [\mathbb{E}/\mathbb{F}]$. Suppose also that $f(x)$ is monic and satisfies $f(\beta) = 0$, hence $m_{\beta/\mathbb{F}}(x)|f(x)$. Then since $m_{\beta/\mathbb{F}}(x)$ and $f(x)$ are monic of the same degree we conclude that $m_{\beta/\mathbb{F}}(x) = f(x)$.]

(a) Let $f(x) = x^3 - 2 \in \mathbb{Q}(\omega)[x]$ and observe that $f(\alpha) = 0$ and $\deg(f) = 3$. From the hint, it suffices to show that $[\mathbb{Q}(\omega)(\alpha)/\mathbb{Q}(\omega)] = 3$. To see this, first note that $\mathbb{Q}(\omega)(\alpha) = \mathbb{Q}(\alpha, \omega)$ and recall from class that $[\mathbb{Q}(\alpha, \omega)/\mathbb{Q}] = 6$. One can see that $[\mathbb{Q}(\omega)/\mathbb{Q}] = 2$ because ω is a root of the polynomial $x^2 + x + 1 \in \mathbb{Q}[x]$, which is irreducible over $\mathbb{Q}$ because it has no rational root. Then from Dedekind's Law we conclude that

$$[\mathbb{Q}(\alpha, \omega)/\mathbb{Q}(\omega)] = [\mathbb{Q}(\alpha, \omega)/\mathbb{Q}] \,/\, [\mathbb{Q}(\omega)/\mathbb{Q}] = 6/2 = 3.$$

(b) Let $g(x) = x^2 + x + 1 \in \mathbb{Q}(\alpha\omega)$ and observe that $g(\omega) = 0$ and $\deg(g) = 2$. From the hint, it suffices to show that $[\mathbb{Q}(\alpha\omega)(\omega)/\mathbb{Q}(\alpha\omega)] = 2$. To see this, first note that $\mathbb{Q}(\alpha\omega)(\omega) = \mathbb{Q}(\alpha, \omega)$. Next, observe that $\alpha\omega$ is a root of the irreducible polynomial $x^3 - 2 \in \mathbb{Q}[x]$, hence $[\mathbb{Q}(\alpha\omega)/\mathbb{Q}] = 3$. Finally, we conclude from Dedekind's Law that

$$[\mathbb{Q}(\alpha, \omega)/\mathbb{Q}(\alpha\omega)] = [\mathbb{Q}(\alpha, \omega)/\mathbb{Q}] \,/\, [\mathbb{Q}(\alpha\omega)/\mathbb{Q}] = 6/3 = 2.$$

(c) Let $h(x) = x^2 + (\alpha\omega)x + (\alpha\omega)^2 \in \mathbb{Q}(\alpha\omega)[x]$ and observe that $\deg(h) = 2$. Since $\mathbb{Q}(\alpha\omega)(\alpha) = \mathbb{Q}(\alpha, \omega)$, we already know from part (b) that $[\mathbb{Q}(\alpha\omega)(\alpha)/\mathbb{Q}(\alpha\omega)] = 2$. Thus it only remains to show that $g(\alpha) = 0$. And, indeed, we have

$$g(\alpha) = \alpha^2 + (\alpha\omega)\alpha + (\alpha\omega)^2$$

[56] According to Günther Frei (2005) these results appear in an early unpublished section eight of the *Disquisitiones*, written by Gauss in 1797. Moreover, it seems that Gauss' treatment was more rigorous than that of Galois, since he treated the "Galois imaginagies" as cosets of polynomials. The "unpublished section eight" was first published by Dedekind in 1863 with a second printing in 1876.

$$
\begin{aligned}
&= \alpha^2 + \alpha^2\omega + \alpha^2\omega^2 \\
&= \alpha^2(1 + \omega + \omega^2) \\
&= \alpha \cdot 0 \\
&= 0.
\end{aligned}
$$

2. Formal Derivation and Repeated Roots. If $\mathbb{F}$ is a field then we can think of the ring of polynomials $\mathbb{F}[x]$ as an infinite dimensional $\mathbb{F}$-vector space with basis $\{1, x, x^2, \ldots\}$. Let $D : \mathbb{F}[x] \to \mathbb{F}[x]$ be the unique $\mathbb{F}$-linear function defined by

$$D(x^n) = nx^{n-1} \text{ for all } n \geq 0.$$

(a) For all polynomials $f(x), g(x) \in \mathbb{F}[x]$ prove that the *product rule* holds:

$$D(fg) = f \cdot Dg + Df \cdot g.$$

[Hint: Show that each side is an $\mathbb{F}$-bilinear function of f and g. Thus it suffices to check the case when $f = x^m$ and $g = x^n$ are basis elements.]

(b) For all polynomials $f(x) \in \mathbb{F}[x]$ use part (a) and induction to prove the *power rule*:

$$D(f^n) = nf^{n-1} \cdot Df \text{ for all } n \geq 0.$$

(c) Consider a polynomial $f(x) \in \mathbb{F}[x]$ and a field extension $\mathbb{E} \supseteq \mathbb{F}$. We say that $\alpha \in \mathbb{E}$ is a *repeated root* of f when $f(x) = (x - \alpha)^2 g(x)$ for some polynomial $g(x) \in \mathbb{E}[x]$. Use part (a) to prove that

$$\alpha \text{ is a repeated root of } f \quad \iff \quad f(\alpha) = 0 \text{ and } Df(\alpha) = 0.$$

(a) For all $f(x), g(x) \in \mathbb{F}[x]$ we define

$$\Phi(f, g) := D(fg) \qquad \text{and} \qquad \Psi(f, g) := f \cdot Dg + Df \cdot g.$$

Since each of these functions is symmetric in f and g, it suffices to show that each function is linear in the second coordinate. So consider any polynomials $f(x), g(x), h(x) \in \mathbb{F}[x]$ and scalar $\lambda \in \mathbb{F}[x]$. Then since D is a linear function, we have

$$
\begin{aligned}
\Phi(f, g + \lambda h) &= D(f(g + \lambda h)) \\
&= D(fg + \lambda fh) \\
&= D(fg) + \lambda D(fh) \\
&= \Phi(f, g) + \lambda \Phi(f, h)
\end{aligned}
$$

and

$$\Psi(f, g + \lambda h) = f \cdot D(g + \lambda h) + Df \cdot (g + \lambda h)$$

$$
\begin{aligned}
&= f \cdot (Dg + \lambda Dh) + Df \cdot (g + \lambda h) \\
&= (f \cdot Dg + Df \cdot g) + \lambda(f \cdot Dh + Df \cdot h) \\
&= \Psi(f, g) + \lambda\Psi(f, h).
\end{aligned}
$$

Now since Φ and Ψ are each bilinear, in order to prove that $\Phi(f, g) = \Psi(f, g)$ for all $f(x), g(x) \in \mathbb{F}[x]$, it suffices to show that this identity holds for ordered pairs of basis vectors. So let $f(x) = x^m$ and $g(x) = x^n$ and observe that

$$
\begin{aligned}
\Psi(x^m, x^n) &= x^m \cdot D(x^n) + D(x^m) \cdot x^n \\
&= x^m \cdot nx^{n-1} + mx^{m-1} \cdot x^n \\
&= (m + n)x^{m+n-1} \\
&= D(x^{m+n}) \\
&= D(x^m x^n) \\
&= \Phi(x^m, x^n).
\end{aligned}
$$

(b) If we define $f(x)^0 = 1 \in \mathbb{F}[x]$ then the statement holds for $n = 0$. Now fix some integer $n \geq 0$ and let us assume for induction that

$$D(f^n) = nf^{n-1} \cdot Df.$$

Then the product rule from part (a) gives

$$
\begin{aligned}
D(f^{n+1}) &= D(f \cdot f^n) \\
&= f \cdot D(f^n) + Df \cdot f^n \\
&= f \cdot nf^{n-1} \cdot Df + Df \cdot f^n \\
&= (n + 1)f^n \cdot Df
\end{aligned}
$$

as desired.

(c) Consider a polynomial $f(x) \in \mathbb{F}[x]$ and an element of a field extension, $\alpha \in \mathbb{E} \supseteq \mathbb{F}$. First suppose that $f(x) = (x - \alpha)^2 g(x)$ for some $g(x) \in \mathbb{F}[x]$. We clearly have $f(\alpha) = 0$. Now apply the product and power rules to get

$$Df(x) = 2(x - \alpha)g(x) + (x - \alpha)^2 Dg(x) = (x - \alpha)\left[2g(x) + (x - \alpha)Dg(x)\right].$$

It follows that $Df(\alpha) = 0$. Conversely, suppose that $f(\alpha) = 0$ and $Df(\alpha) = 0$. Since $f(\alpha) = 0$, Descartes' Theorem says that $f(x) = (x-\alpha)g(x)$ for some $g(x) \in \mathbb{F}[x]$. Now apply the product and power rules to get

$$Df(x) = g(x) + (x - \alpha)Dg(x).$$

Then since $0 = Df(\alpha) = g(\alpha)$, Descartes' Theorem says that $g(x) = (x - \alpha)h(x)$ for some $h(x) \in \mathbb{F}[x]$ and it follows that $f(x) = (x - \alpha)^2 h(x)$ as desired.

3. Cyclotomic Polynomials. Fix an integer n and consider the polynomial $x^n - 1 \in \mathbb{Z}[x]$.

(a) Factor $x^n - 1$ into irreducible polynomials over $\mathbb{C}$. [Hint: Let $\omega := e^{2\pi/n}$.]

(b) Factor $x^n - 1$ into irreducible polynomials over $\mathbb{R}$. [Hint: For all integers $k \in \mathbb{Z}$ we have $\omega^k + \omega^{-k} = 2\cos(2\pi k/n)$.]

(c) We define the n-th *cyclotomic polynomial* $\Phi_n(x) \in \mathbb{C}[x]$ as follows:

$$\Phi_n(x) := \prod_{\omega \in \Omega'_n} (x - \omega) \quad \text{where} \quad \Omega'_n := \{e^{2\pi i k/n} : 0 \le k < n, \gcd(k, n) = 1\}.$$

Prove that

$$x^n - 1 = \prod_{d|n} \Phi_d(x) = \prod_{d|n} \Phi_{n/d}(x).$$

[Hint: The elements of Ω'_n are called the *primitive* nth roots of unity. Prove that the set of **all** nth roots of unity can be expressed as a disjoint union $\coprod_{d|n} \Omega'_d$.]

(d) Use part (c) and induction to prove that actually $\Phi_n(x) \in \mathbb{Z}[x]$. [Hint: For any $f(x), g(x) \in \mathbb{Z}[x]$ with $g(x)$ monic there exist **unique** polynomials $q(x), r(x) \in \mathbb{Z}[x]$ such that $f(x) = q(x)g(x) + r(x)$ and $\deg(r) < \deg(g)$.]

(a) If $\omega = 2^{2\pi i/n}$ then the n-th roots of unity are $1, \omega, \ldots, \omega^{n-1}$, hence

$$x^n - 1 = \prod_{k=0}^{n-1} (x - \omega^k).$$

Note that each linear factor is irreducible.

(b) For any $k \in \mathbb{Z}$ we have $\omega^k + \omega^{-k} = 2\cos(2\pi k/n)$ and $\omega^k \omega^{-k} = 1$, hence

$$(x - \omega^k)(x - \omega^{-k}) = x^2 - 2\cos(2\pi k/n)x + 1 \in \mathbb{R}[x].$$

If $2k \notin n\mathbb{Z}$ then this degree 2 polynomial is irreducible over $\mathbb{R}$ because it has no real roots. Now there are two cases. If n is odd then we have

$$\begin{aligned} x^n - 1 &= (x - 1) \prod_{k=1}^{(n-1)/2} (x - \omega^k)(x - \omega^{-k}) \\ &= (x - 1) \prod_{k=1}^{(n-1)/2} (x^2 - 2\cos(2\pi k/n)x + 1) \end{aligned}$$

and if n is even then we have

$$x^n - 1 = (x - 1)(x + 1) \prod_{k=1}^{(n-2)/2} (x - \omega^k)(x - \omega^{-k})$$

$$= (x-1)(x+1)\prod_{k=1}^{(n-2)/2}(x^2 - 2\cos(2\pi k/n)x + 1).$$

In either case we observe that each factor is irreducible over $\mathbb{R}$.

(c) Given a fraction $\alpha \in \mathbb{Q}$ we observe that $e^{2\pi i\alpha} \in \mathbb{C}$ is a primitive n-th root of unity if and only if α has denominator n when written in lowest terms. So let us define the following sets:

$$F_n := \{k/n : 0 \le k < n\},$$
$$F'_n := \{k/n : 0 \le k < n, \gcd(k,n) = 1\} \subseteq F_n.$$

Then the map $\alpha \mapsto e^{2\pi i\alpha}$ defines a bijection $F_n \to \Omega_n$ with the set of **all** n-th roots of unity, and restricts to a bijection $F'_n \to \Omega'_n$ with the **primitive** n-th roots of unity. I claim that the set of fractions F_n decomposes as a disjoint union as follows:

$$F_n = \coprod_{d|n} F'_d.$$

There are three things to show:

- $\left[F_n \subseteq \cup_{d|n} F'_d\right]$ Consider any fraction $k/n \in F_n$ with $e := \gcd(k,n)$. Then we have $\gcd(k/e, n/e) = 1$ and hence
$$\frac{k}{n} = \frac{k/e}{n/e} \in F'_{n/e} \subseteq \cup_{d|n} F'_d.$$

- $\left[\cup_{d|n} F'_d \subseteq F_n\right]$ Suppose that $k/d \in F'_d$ for some $d|n$ with $d\ell = n$. Then since $0 \le k < d$ implies that $0 \le k\ell < d\ell = n$ we conclude that
$$\frac{k}{d} = \frac{k\ell}{d\ell} = \frac{k\ell}{n} \in F_n.$$

- $[d \ne e \Rightarrow F'_d \cap F'_e = \emptyset]$ Assume that $F'_d \cap F'_e \ne \emptyset$. Thus there exist some reduced fractions $a/d \in F'_d$ and $b/e \in F'_e$ such that $a/d = b/e$, and hence $ae = bd$. Since $a|bd$ and $\gcd(a,d) = 1$ we conclude that $a|b$ and, similarly, since $b|ae$ and $\gcd(b,e) = 1$ we conclude that $b|a$. Since a and b are both positive this implies that $a = b$ and hence $d = e$.

Finally, we conclude that

$$\begin{aligned} x^n - 1 &= \prod_{\alpha \in F_n} (x - e^{2\pi i\alpha}) \\ &= \prod_{d|n} \prod_{\alpha \in F'_d} (x - e^{2\pi i\alpha}) \\ &= \prod_{d|n} \Phi_d(x). \end{aligned}$$

(d) Observe that $\Phi_1(x) = x - 1$ has coefficients in $\mathbb{Z}$. Now fix $n \geq 1$ and assume for induction that $\Phi_k(x) \in \mathbb{Z}[x]$ for all $k < n$. From part (c) we have

$$x^n - 1 = \Phi_n(x)f(x) + 0.$$

where $f(x)$ is the product of $\Phi_d(x)$ over all divisors $d|n$ such that $d \neq n$. In particular, we see that $\Phi_n(x)$ is the **unique** quotient of $x^n - 1 \bmod f(x)$ in the ring $\mathbb{C}[x]$. On the other hand, we know by induction that $f(x) \in \mathbb{Z}[x]$. Then since $f(x)$ is monic, the Division Theorem tells us that there exist some **integer** polynomials $q(x), r(x) \in \mathbb{Z}[x]$ such that

$$x^n - 1 = q(x)f(x) + r(x) \qquad \text{with } \deg(r) < \deg(f).$$

By viewing both of these equations in the ring $\mathbb{C}[x]$ we conclude by uniqueness that

$$\Phi_n(x) = q(x) \in \mathbb{Z}[x].$$

[Remark: We will prove later that in fact $\Phi_n(x)$ is **irreducible** over $\mathbb{Q}$, hence is the minimal polynomial over $\mathbb{Q}$ for any primitive n-th root of unity ω. It will follow that the cyclotomic field $\mathbb{Q}(\omega)/\mathbb{Q}$ has dimension equal to Euler's totient $\phi(n)$.]

4. Quadratic Field Extensions, Part III Prove that $[\mathbb{E}/\mathbb{F}] = 2$ implies $\mathbb{E} = \mathbb{F}(\iota)$ for some $\iota \in \mathbb{E} - \mathbb{F}$ with $\iota^2 \in \mathbb{F}$. [Hint: For any $\alpha \in \mathbb{E} - \mathbb{F}$ note that the set $1, \alpha, \alpha^2$ is linearly dependent, hence we have $f(\alpha) = 0$ for some polynomial $f(x) \in \mathbb{F}[x]$ of degree 2. Let $\beta \in \mathbb{E}$ be the other root of $f(x)$ and define $\iota := \alpha - \beta \in \mathbb{E}$.]

Proof. Let $\mathbb{E} \supseteq \mathbb{F}$ be any field extension with $[\mathbb{E}/\mathbb{F}] = 2$. For any element $\alpha \in \mathbb{E} - \mathbb{F}$ we know that the set $\{1, \alpha, \alpha^2\}$ is linearly dependent over $\mathbb{F}$, hence we have $f(\alpha) = 0$ for some nonzero polynomial $f(x) = ax^2 + bx + c \in \mathbb{F}[x]$. Furthermore, we must have $a \neq 0$ because $\alpha \notin \mathbb{F}$. Then dividing $f(x)$ by a and applying Descartes' Theorem gives

$$x^2 + (b/a)x + (c/a) = (x - \alpha)(x - \beta).$$

for some $\beta \in \mathbb{E}$. Comparing coefficients gives $\alpha + \beta = -b/a \in \mathbb{F}$ and $\alpha\beta = c/a \in \mathbb{F}$. If $\beta \in \mathbb{F}$ then we obtain the contradiction that $\alpha = -b/a - \beta \in \mathbb{F}$. Then since $f(x)$ has degree 2 and no roots in $\mathbb{F}$ we conclude that $f(x)/a$ is the minimal polynomial for $\alpha/\mathbb{F}$. It follows from this that $[\mathbb{F}(\alpha)/\mathbb{F}] = 2$ and hence $\mathbb{E} = \mathbb{F}(\alpha)$.

Finally, let $\iota := \alpha - \beta$ and observe that $\mathbb{F}(\alpha) = \mathbb{F}(\iota)$ because $2\alpha = \iota + (\alpha + \beta) = \iota - b/a$. In particular, this implies that $\iota \neq \mathbb{F}$ and it only remains to show that $\iota^2 \in \mathbb{F}$. For this we use Newton's Theorem:

$$\iota^2 = (\alpha - \beta)^2 = (\alpha + \beta)^2 - 4\alpha\beta = (-b/a)^2 - 4(c/a) \in \mathbb{F}.$$

Indeed, ι^2 is just the discriminant of the minimal polynomial for $\alpha/\mathbb{F}$. $\square$

5. Impossible Constructions. We say that a number $\alpha \in \mathbb{R}$ is *constructible over* $\mathbb{Q}$ if there exists a chain of field extensions

$$\alpha \in \mathbb{F}_k \supseteq \mathbb{F}_{k-1} \supseteq \cdots \supseteq \mathbb{F}_1 \supseteq \mathbb{F}_0 = \mathbb{Q}$$

such that $[\mathbb{F}_{i+1}/\mathbb{F}_i] = 2$ for all i. [Reason: A point of $\mathbb{R}^2$ is "constructible with straightedge and compass" if and only if both of its coordinates are constructible in the above sense.]

(a) Let $f(x) \in \mathbb{Q}[x]$ be any polynomial of degree 3. Prove that

$$f \text{ has a constructible root } \alpha \in \mathbb{R} \quad \Longrightarrow \quad f \text{ has a root in } \mathbb{Q}.$$

[Hint: You proved the induction step on the previous homework.]

(b) Prove that the real numbers $\sqrt[3]{2}$, $\cos(2\pi/18)$ and $\cos(2\pi/7)$ are not constructible. It follows from this that the classical problems of "doubling the cube," "trisecting the angle," and "constructing the regular heptagon" are impossible. [Hint: Show that each is a root of some irreducible polynomial $f(x) \in \mathbb{Q}[x]$ of degree 3.]

(a) Let $f(x) \in \mathbb{Q}[x]$ have degree 3 and assume that $f(\alpha) = 0$ for some constructible $\alpha \in \mathbb{R}$. In other words, assume that there exists a chain of field extensions

$$\alpha \in \mathbb{F}_k \supseteq \mathbb{F}_{k-1} \supseteq \cdots \supseteq \mathbb{F}_1 \supseteq \mathbb{F}_0 = \mathbb{Q}$$

with $[\mathbb{F}_{i+1}/\mathbb{F}_i] = 2$ for all i. From part (a) it follows that for each i we have $\mathbb{F}_{i+1} = \mathbb{F}_i(\gamma_i)$ for some $\gamma_i \in \mathbb{F}_{i+1} - \mathbb{F}_i$ such that $\gamma_i^2 \in \mathbb{F}_i$.

If we think of $f(x)$ as an element of F_{k-1} then since $f(x)$ has degree 3 and has a root in a quadratic field extension, we conclude from the previous homework that $f(x)$ has a root in $\mathbb{F}_{k-1}$. Finally, we use induction on k to conclude that $f(x)$ has a root in $\mathbb{Q}$.

[Remark: This is a surprisingly powerful tool for proving that certain classical "straightedge and compass" constructions are impossible.]

(b) First note that $\sqrt[3]{2} \in \mathbb{R}$ is a root of the polynomial $x^3 - 2 \in \mathbb{Q}[x]$. We have already seen that this polynomial has no roots in $\mathbb{Q}$, hence the number $\sqrt[3]{2}$ is not constructible [i.e., it is impossible to "double the cube" using straightedge and compass.]

Next consider the number $\alpha = \cos(2\pi/18) = \cos(\pi/9) \in \mathbb{R}$ and recall the "triple angle identity"

$$\cos(3\theta) = 4\cos^3\theta - 3\cos\theta$$

which can be proved with de Moivre's Theorem. Then putting $\theta = \pi/9$ gives

$$4\alpha^3 - 3\alpha = \cos(\pi/3)$$
$$4\alpha^3 - 3\alpha = 1/2$$

$$4\alpha^3 - 3\alpha - 1/2 = 0$$
$$8\alpha^3 - 6\alpha - 1 = 0.$$

One can check using the Rational Root Test that this cubic polynomial has no roots in $\mathbb{Q}$, hence α is not constructible. [But the number $\cos(\pi/3) = 1/2$ is constructible, hence one can construct the angle $\pi/3$ but one cannot trisect this angle using straightedge and compass.]

Finally, consider the number $\beta = \cos(2\pi/7) \in \mathbb{R}$ and define $\omega = e^{2\pi i/7} \in \mathbb{C}$ so that

$$\omega + \omega^{-1} = 2\cos(2\pi/7) = 2\beta.$$

Here are the first few powers of 2β:

$$\begin{array}{lllllll} 2\beta & = 0\omega^3 & +0\omega^2 & +1\omega & +0 & +\omega^{-1} & +0\omega^{-2} & +0\omega^{-3}, \\ (2\beta)^2 & = 0\omega^3 & +1\omega^2 & +0\omega & +2 & +0\omega^{-1} & +1\omega^{-2} & +0\omega^{-3}, \\ (2\beta)^3 & = 1\omega^3 & +0\omega^2 & +3\omega & +0 & +3\omega^{-1} & +0\omega^{-2} & +1\omega^{-3}. \end{array}$$

But recall that the sum of all 7-th roots of unity equals zero, hence

$$(2\beta)^3 + (2\beta)^2 - 2(2\beta) - 1 = \omega^3 + \omega^2 + \omega + 1 + \omega^{-1} + \omega^{-2} + \omega^{-3}$$
$$8\beta^3 + 4\beta^2 - 4\beta - 1 = 0.$$

One can check using the Rational Root Test that this rational cubic has no roots in $\mathbb{Q}$, hence β is not constructible [i.e., it is impossible to construct a regular heptagon using straightedge and compass.]

[Remark: Here is an indirect proof that does not require us to find the polynomials.

Let $\omega = e^{2\pi i/n}$ (with $n \geq 3$) and observe that the following polynomial is irreducible over $\mathbb{R}$ because it has degree 2 and no real roots:

$$(x - \omega)(x - \omega^{-1}) = x^2 - 2\cos(2\pi/n)x + 1.$$

In particular, since $\mathbb{Q}(\cos(2\pi/n)) \subseteq \mathbb{R}$ we conclude that this is the minimal polynomial for ω over the field $\mathbb{Q}(\cos(2\pi/n))$, and hence

$$[\mathbb{Q}(\omega)/\mathbb{Q}(\cos(2\pi/n))] = 2.$$

Then applying Dedekind's Law gives

$$[\mathbb{Q}(\cos(2\pi/n))/\mathbb{Q}] = [\mathbb{Q}(\omega)/\mathbb{Q}] \,/\, [\mathbb{Q}(\omega)/\mathbb{Q}(\cos(2\pi/n))] = \frac{1}{2}[\mathbb{Q}(\omega)/\mathbb{Q}].$$

If we assume that the cyclotomic polynomial $\Phi_n(x) \in \mathbb{Q}[x]$ is irreducible then it follows that

$$[\mathbb{Q}(\cos(2\pi/n))/\mathbb{Q}] = \frac{1}{2}[\mathbb{Q}(\omega)/\mathbb{Q}] = \frac{1}{2}\phi(n).$$

In particular, since $\phi(7) = \phi(18) = 6$ we conclude that the numbers $\cos(2\pi/18)$ and $\cos(2\pi/7)$ each satisfy a minimal polynomial of degree 3 over $\mathbb{Q}$. Incidentally, the numbers $\sin(2\pi/n)$ are a bit more complicated. If $n \neq 2$ and $n = 2^e m$ with m odd, then one can show that[57]

$$[\mathbb{Q}(\sin(2\pi/n))/\mathbb{Q}] = \begin{cases} \phi(n) & e = 0, 1 \\ \phi(n)/4 & e = 2, \\ \phi(n)/2 & e \geq 3.] \end{cases}$$

6. Primitive Root Theorem. If $\mathbb{F}$ is a finite field then the group of units $\mathbb{F}^\times$ is cyclic.

(a) Consider $m, n \in \mathbb{Z}$ with $\gcd(m, n) = 1$. If $m|nk$ prove that $m|k$. If $m|k$ and $n|k$ prove that $mn|k$. [Hint: Since $\gcd(m, n) = 1$ there exist $x, y \in \mathbb{Z}$ with $mx + ny = 1$.]

(b) Let A be an abelian group. If elements $a, b \in A$ have orders m, n with $\gcd(m, n) = 1$, prove that ab has order mn. [Hint: Show that $(ab)^k = \varepsilon$ implies $m|k$ and $n|k$.]

(c) Let A be an abelian group. If m is the **maximal order** of an element, prove that every element has order dividing m. [Hint: Let $a, b \in A$ have orders ℓ, m with $\ell \nmid m$. Then for some prime p we have $\ell = p^i \ell'$ and $m = p^j m'$ with $p \nmid \ell', m'$ and $i > j$. Use (b) to show that $a^{\ell'} b^{p^j}$ has order greater than m.]

(d) If $\alpha \in \mathbb{F}^\times$ is an element of **maximal order** m, prove that $\mathbb{F}^\times = \{1, \alpha, \ldots, \alpha^{m-1}\}$. [Hint: If not then the polynomial $x^m - 1 \in \mathbb{F}[x]$ has too many roots. Use (c).]

(a) If $m, n \in \mathbb{Z}$ with $\gcd(m, n) = 1$ then there exist integers $x, y \in \mathbb{Z}$ such that $mx + ny = 1$.

First suppose that $m|nk$ with $mm' = nk$. Then we have

$$\begin{aligned} mx + by &= 1 \\ mkx + bky &= k \\ mkx + mm'y &= k \\ m(kx + m'y) &= k, \end{aligned}$$

hence $m|k$. Next, suppose that $m|k$ and $n|k$ with $k = mm' = nn'$. Then we have

$$\begin{aligned} mx + ny &= 1 \\ mkx + nky &= k \\ m(nn')x + n(mm')y &= k \\ (mn)(n'x + m'y) &= k, \end{aligned}$$

hence $mn|k$.

[57] See Beslin and De Angelis (2004), Mathematics Magazine.

[Remark: One could give shorter proofs of these facts by applying unique prime factorization.]

(b) Let A be an abelian group and suppose that elements $a, b \in A$ have orders m, n (respectively) with $\gcd(m,n) = 1$. Then I claim that $ab \in A$ has order mn. Indeed, since $a^m = b^n = \varepsilon$ we have $(ab)^{mn} = a^m b^n = \varepsilon$. Now suppose that $(ab)^k = \varepsilon$ for some integer $k \in \mathbb{Z}$. Then raising to the power of m gives

$$(ab)^{km} = \varepsilon \quad \Longrightarrow \quad (a^m)^k b^{km} = \varepsilon \quad \Longrightarrow \quad b^{km} = \varepsilon \quad \Longrightarrow \quad n | km \quad \stackrel{\text{(a)}}{\Longrightarrow} \quad n | k.$$

And raising to the power of n gives

$$(ab)^{kn} = \varepsilon \quad \Longrightarrow \quad a^{kn}(b^n)^k = \varepsilon \quad \Longrightarrow \quad a^{kn} = \varepsilon \quad \Longrightarrow \quad m | kn \quad \stackrel{\text{(a)}}{\Longrightarrow} \quad m | k.$$

It follows from (a) that $mn|k$, hence ab has order mn.

(c) Let m be the maximal order of an element in an abelian group group A. Now consider elements $a, b \in A$ with orders ℓ, m (respectively). In order to prove that $\ell | m$, let us assume for contradiction that $\ell \nmid m$. By definition there exists some prime p such that $\ell = p^i \ell'$ and $m = p^j m'$ where $i > j$ and where m' and ℓ' are not divisible by p. Now consider the element $a^{\ell'} b^{p^j} \in A$. Since a has order $\ell = p^i \ell'$ we observe observe that $(a^{\ell'})^{p^i} = \varepsilon$ and for any $q \in \mathbb{Z}$ we have

$$(a^{\ell'})^q \quad \Longrightarrow \quad a^{q\ell'} = \varepsilon \quad \Longrightarrow \quad q\ell' | p^i \ell' \quad \Longrightarrow \quad q | p^i,$$

hence $a^{\ell'}$ has order p^i. Similarly, one can show that $b^{p^j} \in A$ has order m'. Finally, since $\gcd(p^i, m') = 1$ we conclude from part (b) that $a^{\ell'} b^{p^j}$ has order $p^i m'$, which contradicts the minimality of m because

$$p^i m' > p^j m' = m.$$

(d) Let $\mathbb{F}$ be a finite field and let $\alpha \in \mathbb{F}^\times$ be an nonzero element of maximal multiplicative order m. Since $\mathbb{F}^\times$ is an abelian group we know from part (c) that $\beta^m = 1$ for **all** $\beta \in \mathbb{F}^\times$. In particular, we see that $\{1, \alpha, \ldots, \alpha^{m-1}\}$ is a set of m distinct roots of the polynomial $x^m - 1 \in \mathbb{F}[x]$. Since this polynomial can have **no more than** m **roots**, it follows that any $\beta \in \mathbb{F}^\times$ is equal to a power of α. In other words, $\mathbb{F}^\times$ is a cyclic group.

7. Laplace's Proof of the FTA. The FTA is easily proved with complex analysis. However, it is still nice to have an elementary proof that is mostly algebraic. The following proof from Laplace (1795) builds on earlier ideas of Euler (1749) and Lagrange (1770). A logical gap in the proof was later filled by Kronecker's Theorem (1887). Specifically, we will prove that

every non-constant polynomial $f(x) \in \mathbb{R}[x]$ has a root in $\mathbb{C}$.

(a) Observe that every polynomial $f(x) \in \mathbb{R}[x]$ of odd degree has a root in $\mathbb{R}$.

(b) Now let $f(x) \in \mathbb{R}[x]$ have degree $n = 2^e m$ with $e \geq 1$ and m odd. Consider $f(x)$ as an element of $\mathbb{C}[x]$ and let $\mathbb{E} \supseteq \mathbb{C}$ be a splitting field:

$$f(x) = (x - \alpha_1)(x - \alpha_2) \cdots (x - \alpha_n) \in \mathbb{E}[x].$$

Now for any real number $\lambda \in \mathbb{R}$ we define the polynomial

$$g_\lambda(x) := \prod_{1 \leq i < j \leq n} (x - \beta_{ij\lambda}) \in \mathbb{E}[x] \quad \text{with} \quad \beta_{ij\lambda} := \alpha_i + \alpha_j + \lambda \alpha_i \alpha_j \in \mathbb{E}.$$

Prove that $g_\lambda(x) \in \mathbb{R}[x]$ and $\deg(g_\lambda) = 2^{e-1} m'$ with m' odd. [Hint: Newton.]

(c) By induction on e we can assume that $g_\lambda(x)$ has a complex root $\beta_{ij\lambda} \in \mathbb{C}$. If we apply this argument for more than $\binom{n}{2}$ different values of $\lambda \in \mathbb{R}$ then we will find specific indices $i < j$ and real numbers $\lambda \neq \mu$ such that $\beta_{ij\lambda}$ and $\beta_{ij\mu}$ are **both** in $\mathbb{C}$. In this case prove that α_i and α_j are in $\mathbb{C}$, hence $f(x)$ has a complex root.

(a) Consider any polynomial $f(x) \in \mathbb{R}[x]$ of odd degree. Since $f(x)$ is a continuous function whose graph goes from $\pm\infty$ to $\mp\infty$ we conclude that the graph must cross the x-axis at some point. In other words, $f(x)$ has a real root.

[Remark: This is just plain common sense. It was basically assumed as an axiom until the 1800s. Bolzano (1817) and Cauchy (1821) were the first to write down "proofs," however their arguments were not very convincing. Weierstrass was probably the first person to write a convincing proof. You can decide for yourself if a proof is necessary.]

(b) Now let $f(x) \in \mathbb{R}[x]$ have degree $n = 2^e m$ with m odd and and assume for induction that every real polynomial of degree $2^{e'} m'$ with $e' < e$ and m' odd has a complex root. Consider $f(x)$ as an element of $\mathbb{C}[x]$ and let $\mathbb{E} \supseteq \mathbb{C}$ be a splitting field, which exists by Kronecker's Theorem. To be explicit, suppose that

$$f(x) = (x - \alpha_1)(x - \alpha_2) \cdots (x - \alpha_n) \in \mathbb{E}[x].$$

For any integers $1 \leq i < j \leq n$ and any for any real number $\lambda \in \mathbb{R}$ we define the elements

$$\beta_{ij\lambda} := \alpha_i + \alpha_j + \lambda \alpha_i \alpha_j \in \mathbb{E}$$

and the polynomial

$$g_\lambda(x) := \prod_{1 \leq i < j \leq n} (x - \beta_{ij\lambda}) \in \mathbb{E}[x].$$

But note that each coefficient of $g_\lambda(x)$ is a symmetric polynomial expression of the roots α_i with coefficients in $\mathbb{R}$. It follows from Newton's Theorem that every coefficient of $g_\lambda(x)$ is a real polynomial function of the coefficients of $f(x)$. Since the coefficients of $f(x)$ are in $\mathbb{R}$ we conclude that $g_\lambda(x) \in \mathbb{R}[x]$. Furthermore, we observe that the degree of $g_\lambda(x)$ is

$$\deg(g_\lambda) = \binom{n}{2} = \frac{n(n-1)}{2} = \frac{2^e m(2^e m - 1)}{2} = 2^{e-1} \left[m(2^e m - 1) \right] = 2^{e-1} \left[\text{odd} \right].$$

(c) By induction we conclude that for each $\lambda \in \mathbb{R}$ the polynomial $g_\lambda(x)$ has a complex root, which must have the form $\beta_{ij\lambda}$. By fixing the indices $i < j$ and letting λ run over more than $\binom{n}{2}$ different values we will find two distinct real numbers $\lambda \neq \mu$ such that $\beta_{ij\lambda}$ and $\beta_{ij\mu}$ are both in $\mathbb{C}$. It follows that

$$(\lambda - \mu)\alpha_i\alpha_j = \beta_{ij\lambda} - \beta_{ij\mu} \in \mathbb{C} \quad \Longrightarrow \quad \alpha_i\alpha_j \in \mathbb{C} \quad \Longrightarrow \quad \alpha_i + \alpha_j = \beta_{ij\lambda} - \lambda\alpha_i\alpha_j \in \mathbb{C}.$$

Finally, since the field $\mathbb{C}$ is closed under taking square roots,[58] we can use the quadratic formula to show that the following polynomial has complex roots:

$$(x - \alpha_i)(x - \alpha_j) = x^2 - (\alpha_i + \alpha_j)x + \alpha_i\alpha_j \in \mathbb{C}[x].$$

In other words, we conclude that $\alpha_i \in \mathbb{C}$ and $\alpha_j \in \mathbb{C}$. This shows that the polynomial $f(x)$ has at least one complex root. □

Week 21

After our detour through ring and field theory, we are finally ready to resume our discussion of Galois groups. The following definition is due to Dedekind, although he was only interested in the case when $\mathbb{E}$ is a subfield of $\mathbb{C}$.

Dedekind's Definition of Galois Groups. For any field extension $\mathbb{E} \supseteq \mathbb{F}$ we define

$$\mathrm{Gal}(\mathbb{E}/\mathbb{F}) := \{ \text{ field automorphisms } \sigma : \mathbb{E} \to \mathbb{E} \text{ such that } \sigma(a) = a \text{ for all } a \in \mathbb{F} \ \}.$$

///

Our first goal is to show that the Galois group $\mathrm{Gal}(\mathbb{E}/\mathbb{F})$ is **finite** whenever $\mathbb{E}$ is a **finite-dimensional** vector space over $\mathbb{F}$. The proof will involve the notion of multi-variable polynomials. We skirted around this concept before, but now I will give you the official definition.

Definition of Multi-Variable Polynomials. Let R be a field and let $\{x_1, \ldots, x_n\}$ be a set of formal symbols, called "variables." We define the ring of polynomials by induction:

$$R[x_1, \ldots, x_n] := R[x_1, \ldots, x_{n-1}][x_n].$$

Explicitly, each element of this ring has the form

$$f(x_1, \ldots, x_n) = \sum a_{k_1,\ldots,k_n} x_1^{k_1} \cdots x_n^{k_n},$$

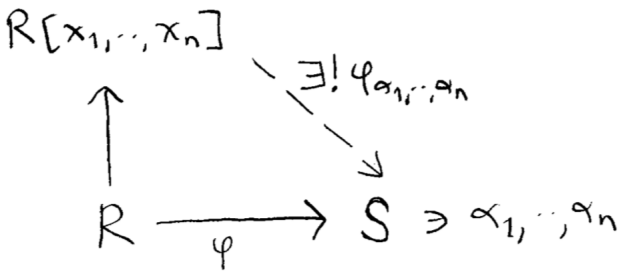
[58] This assumes the existence of a real square root $\sqrt{r} \in \mathbb{R}$ for each positive real number $r > 0$, which also follows from the Intermediate Value Theorem. Then any complex number $re^{i\theta} \in \mathbb{C}$ in polar form has square roots $\pm\sqrt{r}e^{i\theta/2} \in \mathbb{C}$.

where the sum is over all n-tuples of natural numbers $(k_1, \ldots, k_n) \in \mathbb{N}^n$ and all but finitely many of the coefficients $a_{k_1,\ldots,k_n} \in R$ are equal to zero. This ring also satisfies a universal property, which is inherited from the one variable case.

Let $\varphi : R \to S$ be any ring homomorphism and let $\alpha_1, \ldots, \alpha_n \in S$ be any elements, not necessarily distinct. Then there exists a unique ring homomorphism $\varphi_{\alpha_1,\ldots,\alpha_n} : R[x_1, \ldots, x_n] \to S$ sending $x_i \mapsto \alpha_i$ for all i and acting on the coefficients by φ. Here is the explicit definition:

$$\varphi_{\alpha_1,\ldots,\alpha_n}\left(\sum a_{k_1,\ldots,k_n} x_1^{k_1} \cdots x_n^{k_n}\right) = \sum \varphi(a_{k_1,\ldots,k_n}) \alpha_1^{k_1} \cdots \alpha_n^{k_n}.$$

And here is the commutative diagram:



In modern terms we say that

$$R[x_1, \ldots, x_n] \text{ is the } \textit{free } R\textit{-algebra generated by the set } \{x_1, \ldots, x_n\}.$$

///

The definition of multi-variable polynomials is motivated by the following fact, which generalizes the case of one variable. For any ring extension $E \supseteq R$ and for any elements $\alpha_1, \ldots, \alpha_n \in E$, the smallest subring containing the set $R \cup \{\alpha_1, \ldots, \alpha_n\}$ is equal to the image of the evaluation:

$$R[\alpha_1, \ldots, \alpha_n] = \mathrm{im}(\mathrm{id}_{\alpha_1,\ldots,\alpha_n}).$$

Proof. Since $\mathrm{im}(\mathrm{id}_{\alpha_1,\ldots,\alpha_n}) \subseteq E$ is a subring containing the set $R \cup \{\alpha_1, \ldots, \alpha_n\}$, it must contain the smallest such subring. Conversely, since $R[\alpha_1, \ldots, \alpha_n] \subseteq E$ is a subring containing the set $R \cup \{\alpha_1, \ldots, \alpha_n\}$, it must contain every polynomial expression $f(\alpha_1, \ldots, \alpha_n)$. □

We are ready to prove our first theorem about Galois groups. Before reading the proof you may want to go back and remind yourself about the Orbit-Stabilizer Theorem for group actions.

The Finiteness Theorem. Let $[\mathbb{E}/\mathbb{F}] < \infty$ and $G = \mathrm{Gal}(\mathbb{E}/\mathbb{F})$. Then

(1) There exist elements $\alpha_1, \ldots, \alpha_n \in \mathbb{E}$ such that $\mathbb{E} = \mathbb{F}(\alpha_1, \ldots, \alpha_n)$.

(2) Every element of $\mathbb{E}$ is algebraic over $\mathbb{F}$, hence

$$\mathbb{F}(\alpha_1, \ldots, \alpha_n) = \mathbb{F}[\alpha_1, \ldots, \alpha_n].$$

(3) If $\sigma \in G$ satisfies $\sigma(\alpha_i) = \alpha_i$ for all i then we have $\sigma = \mathrm{id}$.

(4) Let $m_i(x) \in \mathbb{F}[x]$ be the minimal polynomial of $\alpha_i/\mathbb{F}$. Then we have

$$\#G \leq \deg(m_1)\deg(m_2)\cdots\deg(m_n)$$

and hence G is finite.

///

Proof. (1) If $\mathbb{E} = \mathbb{F}$ then we are done. Otherwise, let $\alpha_1 \in \mathbb{E} - \mathbb{F}$ and consider the extension $\mathbb{F}(\alpha_1) \supseteq \mathbb{F}$. Since $[\mathbb{F}(\alpha_1)/\mathbb{F}] > 1$ we have $[\mathbb{E}/\mathbb{F}(\alpha_1)] < [\mathbb{E}/\mathbb{F}]$ and it follows by induction that there exist elements $\alpha_2, \ldots, \alpha_n \in \mathbb{E}$ such that

$$\mathbb{E} = \mathbb{F}(\alpha_1)(\alpha_2, \ldots, \alpha_n) = \mathbb{F}(\alpha_1, \alpha_2, \ldots, \alpha_n).$$

(2) Consider any element $\alpha \in \mathbb{E}$. Since $[\mathbb{E}/\mathbb{F}] < \infty$ we know that the set $\{1, \alpha, \alpha^2, \ldots\}$ is linearly dependent over $\mathbb{F}$, hence there exist some coefficients $a_0, \ldots, a_k \in \mathbb{F}$, not all zero, such that $a_0 + a_1\alpha + \cdots + a_k\alpha^k = 0$. In other words, α is algebraic over $\mathbb{E}$ over $\mathbb{F}$.

Now consider the elements $\alpha_1, \ldots, \alpha_n$ from part (1). Since α_1 is algebraic over $\mathbb{F}$ the Minimal Polynomial Theorem tells us that $\mathbb{F}[\alpha_1]$ is a field and hence $\mathbb{F}[\alpha_1] = \mathbb{F}(\alpha_1)$. Then since α_2 is algebraic over $\mathbb{F}$, hence also over $\mathbb{F}(\alpha_1)$, the Minimal Polynomial Theorem tells us that $\mathbb{F}[\alpha_1](\alpha_2) = \mathbb{F}[\alpha_1][\alpha_2] = \mathbb{F}[\alpha_1, \alpha_2]$. Continuing in this way gives the result.

(3) Consider any element $\sigma \in G$ such that $\sigma(\alpha_i) = \alpha_i$ for all i. From part (2) and the remarks before the theorem we know that every element of $\mathbb{E}$ can be expressed as $f(\alpha_1, \ldots, \alpha_n)$ for some polynomial $f(x_1, \ldots, x_n) \in \mathbb{F}[x_1, \ldots, x_n]$. But then since σ fixes $\mathbb{F}$ and preserves ring operations we have

$$\sigma\left(f(\alpha_1, \ldots, \alpha_n)\right) = f\left(\sigma(\alpha_1), \ldots, \sigma(\alpha_n)\right) = f(\alpha_1, \ldots, \alpha_n),$$

and hence $\sigma = \mathrm{id}$.

(4) From (2) we know that each generator $\alpha_i \in \mathbb{E}$ has a minimal polynomial $m_i(x) \in \mathbb{F}[x]$. For any $\sigma \in G$ we observe that $\sigma(\alpha_i)$ is a root of $m_i(x)$ because

$$m_i(\sigma(\alpha_i)) = \sigma(m_i(\alpha)) = \sigma(0) = 0.$$

In other words, the group G acts on the set $\Omega_i \subseteq \mathbb{E}$ of roots of $m_i(x)$ in the field $\mathbb{E}$. Moreover, G acts on the Cartesian product of sets:

$$G \curvearrowright (\Omega_1 \times \cdots \times \Omega_n).$$

Let $\vec{\alpha} = (\alpha_1, \ldots, \alpha_n)$ be the only element of this set that we know,[59] and consider the G-orbit:

$$\mathrm{Orb}(\vec{\alpha}) = \{(\sigma(\alpha_1), \ldots, \sigma(\alpha_n)) : \sigma \in G\} \subseteq \Omega_1 \times \cdots \times \Omega_n.$$

In part (3) we proved that the stabilizer is trivial: $\mathrm{Stab}(\vec{\alpha}) = \{\mathrm{id}\}$. Hence from the Orbit-Stabilizer Theorem we obtain a bijection

$$G \leftrightarrow \frac{G}{\{\mathrm{id}\}} = \frac{G}{\mathrm{Stab}(\vec{\alpha})} \leftrightarrow \mathrm{Orb}(\vec{\alpha}) \subseteq \Omega_1 \times \cdots \times \Omega_n.$$

Since $m_i(x)$ has at most $\deg(m_i)$ roots in the field $\mathbb{E}$ we have $\#\Omega_i \leq \deg(m_i)$ and hence

$$\#G = \#\mathrm{Orb}(\vec{\alpha}) \leq \#(\Omega_1 \times \cdots \times \Omega_n) = \#\Omega_1 \times \cdots \times \#\Omega_n \leq \prod \deg(m_i) < \infty.$$

□

In fact, Dedekind proved a sharper bound:

$$\#\mathrm{Gal}(\mathbb{E}/\mathbb{F}) \leq [\mathbb{E}/\mathbb{F}].$$

However, his proof used very different methods.[60] For pedagogical reasons I chose to prove a weaker result using more relevant methods.

Below we will see that Galois theory is concerned with certain "nice" field extensions $\mathbb{E} \supseteq \mathbb{F}$ that achieve the upper bound: $\#\mathrm{Gal}(\mathbb{E}/\mathbb{F}) = [\mathbb{E}/\mathbb{F}]$. From the proof of the Finiteness Theorem we can already see how this might happen. Suppose that an extension $\mathbb{E} \supseteq \mathbb{F}$ satisfies:

- $\mathbb{E} = \mathbb{F}(\gamma)$ for some element $\gamma \in \mathbb{E}$ with minimal polynomial $m(x) \in \mathbb{F}[x]$,
- the polynomial $m(x)$ splits in $\mathbb{E}[x]$,
- the polynomial $m(x)$ has no multiple roots in $\mathbb{E}$,
- the Galois group $G = \mathrm{Gal}(\mathbb{E}/\mathbb{F})$ acts transitively on the roots of $m(x)$. In other words, for any two roots α, β there exists a group element $\sigma \in G$ with $\sigma(\alpha) = \beta$.

In this case let $\Omega \subseteq \mathbb{E}$ be the set of roots of $m(x)$. From the assumptions we have

$$\#\Omega = \deg(m) = [\mathbb{F}(\gamma)/\mathbb{F}] = [\mathbb{E}/\mathbb{F}].$$

Now if $\sigma \in G$ fixes γ then for every element $f(\gamma) \in \mathbb{F}[\gamma] = \mathbb{F}(\gamma) = \mathbb{E}$ we have $\sigma(f(\gamma)) = f(\sigma(\gamma)) = f(\gamma)$ and hence $\sigma = \mathrm{id}$. Finally, since G acts transitively on Ω we obtain bijections

$$G \leftrightarrow \frac{G}{\{\mathrm{id}\}} = \frac{G}{\mathrm{Stab}(\gamma)} \leftrightarrow \mathrm{Orb}(\gamma) = \Omega$$

and it follows that $\#G = \#\Omega = [\mathbb{E}/\mathbb{F}]$. ///

It may seem to you that the four properties above are rather special, but we will soon prove that these properties hold for a large and natural class of field extensions. To be specific, we will show that these four properties hold whenever:

[59] In fact, there may be no other elements.

[60] It uses the "linear independence of characters."

- $\mathbb{F}$ is finite or has characteristic zero,[61]
- $\mathbb{E}$ is a splitting field for some polynomial $f(x) \in \mathbb{F}[x]$.

From now on we will only consider finite-dimensional field extensions. Last time we proved that if $[\mathbb{E}/\mathbb{F}] < \infty$ then the group $\mathrm{Gal}(\mathbb{E}/\mathbb{F})$ is finite. I also mentioned (but did not prove) Dedekind's theorem, which says that

$$\#\mathrm{Gal}(\mathbb{E}/\mathbb{F}) \leq [\mathbb{E}/\mathbb{F}].$$

We have a special name for field extensions that achieve this bound.

Definition of Galois Extensions. Let $\mathbb{E} \supseteq \mathbb{F}$ be a finite-dimensional field extension. We say that $\mathbb{E}/\mathbb{F}$ is a *Galois extension*[62] if the following equality holds:

$$\#\mathrm{Gal}(\mathbb{E}/\mathbb{F}) = [\mathbb{E}/\mathbb{F}].$$

///

Galois extensions are the field-theoretic version of "normal subgroups." (The Fundamental Theorem of Galois Theory will make this analogy precise.) And, just as with normal subgroups, there are several equivalent ways to state the definition. Before investigating this, let me show you some small examples.

Example: If $[\mathbb{E}/\mathbb{F}] = 1$ or 2 then $\mathbb{E}/\mathbb{F}$ is Galois.

Proof. If $[\mathbb{E}/\mathbb{F}] = 1$ then we have $\mathbb{E} = \mathbb{F}$ and it follows that $\mathrm{Gal}(\mathbb{E}/\mathbb{F}) = \{\mathrm{id}\}$. Now suppose that $[\mathbb{E}/\mathbb{F}] = 2$. On a previous homework you showed that this implies $\mathbb{E} = \mathbb{F}(\iota)$ for some element $\iota \in \mathbb{E}$ with $\iota \notin \mathbb{F}$ and $\iota^2 \in \mathbb{F}$. Let me briefly recall the proof.

Choose any $\alpha \in \mathbb{E} - \mathbb{F}$. Since $[\mathbb{E}/\mathbb{F}] = 2$ we know that the set $\{1, \alpha, \alpha^2\}$ is linearly dependent over $\mathbb{Q}$. Since $\alpha \notin \mathbb{F}$ it follows that $f(\alpha) = 0$ for some $f(x) \in \mathbb{Q}[x]$ of degree 2. Let $\beta \in \mathbb{E}$ be the other root of $f(x)$ and define $\iota := \alpha - \beta$. Then $\iota^2 = (\alpha - \beta)^2 \in \mathbb{F}$ is the discriminant of $f(x)$ and one can show that $\mathbb{E} = \mathbb{F}(\alpha) = \mathbb{F}(\iota)$.

[61] This includes every field that you have ever seen, so it is barely a restriction.

[62] I don't know the origin of this terminology but it seems reasonable. In the literature you will see a "Galois extension" defined as "finite-dimensional, normal and separable." These last two terms have technical meanings that are only relevant for infinite fields of positive characteristic. I think it is appropriate to ignore such fields in a first course on the subject. (Also, I plan never to teach a second course.)

If follows that $\{1, \iota\}$ is a basis for the vector space $\mathbb{E}/\mathbb{F}$. To compute the Galois group, let $\sigma \in \mathrm{Gal}(\mathbb{E}/\mathbb{F})$. Then for any element $a + b\iota \in \mathbb{E}$ we have

$$\sigma(a + b\iota) = a + b\sigma(\iota).$$

Furthermore, since $\iota^2 = a$ for some $a \in \mathbb{F}$ we must also have

$$\sigma(\iota)^2 = \sigma(\iota^2) = \sigma(a) = a.$$

Since the polynomial $x^2 - a \in \mathbb{F}[x]$ has at most two roots in $\mathbb{E}$, this implies that $\sigma(\iota) = \iota$ or $\sigma(\iota) = -\iota$. The first choice corresponds to the identity element and the second choice yields the following function:

$$\tau(a + b\iota) := a - b\iota.$$

One can check by hand that this function is, indeed, a field automorphism and hence

$$\#\mathrm{Gal}(\mathbb{E}/\mathbb{F}) = \#\{\mathrm{id}, \tau\} = 2 = [\mathbb{E}/\mathbb{F}].$$

□

[Remark: This result is analogous to the fact that a subgroup $H \subseteq G$ satisfying $\#(G/H) = 2$ is necessarily normal. Again, The Fundamental Theorem of Galois Theory will make this analogy precise.]

Non-Example: The field extension $\mathbb{Q}(\sqrt[3]{2}) \supseteq \mathbb{Q}$ is not Galois.

Proof. Let $\alpha = \sqrt[3]{2} \in \mathbb{R}$ be the real cube root of 2. Since α is a root of the irreducible polynomial $x^3 - 2 \in \mathbb{Q}[x]$ we know from the Minimal Polynomial Theorem that

$$[\mathbb{Q}(\alpha)/\mathbb{Q}] = \deg(x^3 - 2) = 3.$$

On the other hand, for any $\sigma \in \mathrm{Gal}(\mathbb{Q}(\alpha)/\mathbb{Q})$ we must have

$$0 = \sigma(0) = \sigma(\alpha^3 - 2) = \sigma(\alpha)^3 - 2.$$

Then since $\mathbb{Q}(\alpha) \subseteq \mathbb{R}$ and since α is the only real root of $x^3 - 2$ we must have $\sigma(\alpha) = \alpha$. Finally, since $\alpha/\mathbb{Q}$ is algebraic we know that every element of $\mathbb{Q}(\alpha) = \mathbb{Q}[\alpha]$ has the form $f(\alpha)$ for some polynomial $f(x) \in \mathbb{Q}[x]$ and hence

$$\sigma\left(f(\alpha)\right) = f\left(\sigma(\alpha)\right) = f(\alpha).$$

If follows that $\#\mathrm{Gal}(\mathbb{Q}(\alpha)/\mathbb{Q}) = \#\{\mathrm{id}\} = 1 < 3 = [\mathbb{Q}(\alpha)/\mathbb{Q}]$. □

[Remark: The ultimate problem with this example is that the field $\mathbb{Q}(\sqrt[3]{2})$ contains only one root of the irreducible polynomial $x^3 - 2 \in \mathbb{Q}[x]$. We can fix this problem by passing to the splitting field.]

Example: The splitting field of $x^3 - 2 \in \mathbb{Q}[x]$ is Galois.

Proof. The roots of $x^3 - 2 \in \mathbb{Q}[x]$ are the complex numbers $\alpha, \omega\alpha, \omega^2\alpha \in \mathbb{C}$ where $\alpha = \sqrt[3]{2} \in \mathbb{R}$ and $\omega = e^{2\pi i/3}$. It follows that the splitting field is

$$\mathbb{E} = \mathbb{Q}(\alpha, \omega\alpha, \omega^2\alpha) = \mathbb{Q}(\alpha, \omega).$$

Since x^3-2 is the minimal polynomial for $\alpha/\mathbb{Q}$ and since x^2+x+1 is the minimal polynomial for $\omega/\mathbb{Q}(\alpha)$ we conclude from the Minimal Polynomial Theorem and Dedekind's Tower Law that $[\mathbb{E}/\mathbb{Q}] = 6$ with basis $\{1, \alpha, \alpha^2, \omega, \omega\alpha, \omega\alpha^2\}$. To compute the Galois group, let $\sigma \in \mathrm{Gal}(\mathbb{E}/\mathbb{Q})$. Then for any $a, b, c, d, e, f \in \mathbb{Q}$ we have

$$\begin{aligned}\sigma(a + b\alpha + c\alpha^2 + d\omega + e\alpha\omega + f\alpha^2\omega)& \\ = a + b\sigma(\alpha) + c\sigma(\alpha)^2 + c\sigma(\omega) + e\sigma(\alpha)\sigma(\omega) + f\sigma(\alpha)^2\sigma(\omega).&\end{aligned}$$

It follows that the function σ is determined by the two numbers $\sigma(\alpha)$ and $\sigma(\omega)$. Furthermore, since $\sigma(\alpha)$ is a root of $x^3 - 2$ and since $\sigma(\omega)$ is a root of $x^2 + x + 1$ we must have

$$\sigma(\alpha) \in \{\alpha, \omega\alpha, \omega^2\alpha\} \qquad \text{and} \qquad \sigma(\omega) \in \{\omega, \omega^2\}.$$

Since all of these roots exist in $\mathbb{E}$ we obtain six different functions $\sigma : \mathbb{E} \to \mathbb{E}$. These functions are necessarily $\mathbb{F}$-linear, hence they fix $\mathbb{F}$ and preserve addition.

But how do we know that these functions preserve multiplication?

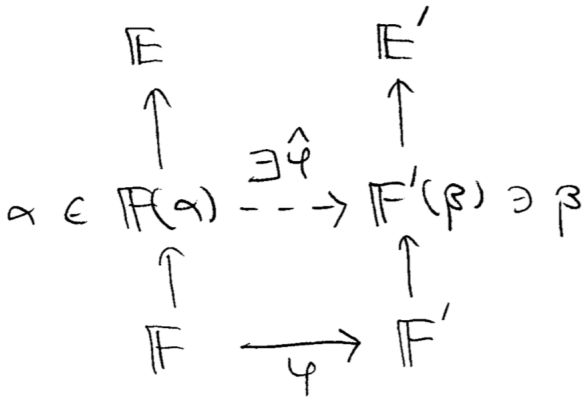
For the moment let me just say that one can check this by hand, or, better, with a computer. It follows that $\#\mathrm{Gal}(\mathbb{E}/\mathbb{Q}) = 6 = [\mathbb{E}/\mathbb{Q}]$ and hence $\mathbb{E}/\mathbb{Q}$ is Galois. □

This last example illustrates two points:

- Splitting fields are likely to be Galois. In fact, we will prove below that any splitting field $\mathbb{E}/\mathbb{F}$ is Galois as long as $\mathbb{F}$ is finite or has characteristic zero.
- The hard part of the proof is to show that certain functions defined on the roots can be **lifted** to automorphisms of the splitting field. Clearly the brute force method is not good enough. We will need a general theorem about this.

The Finiteness Theorem showed that Galois groups are small. Now we want to prove that the Galois group of a splitting field is big. The Splitting Field Theorem below is probably the most subtle theorem in this course. It is good to prepare for this theorem with a lemma.

The Lifting Lemma. Let $\mathbb{E} \supseteq \mathbb{F}$ and $\mathbb{E}' \supseteq \mathbb{F}'$ be field extensions and let $\varphi : \mathbb{F} \to \mathbb{F}'$ be an isomorphism. Let $\alpha \in \mathbb{E}$ be a root of an **irreducible** polynomial $f(x) \in \mathbb{F}[x]$ and let $\beta \in \mathbb{E}'$ be any root of $f^\varphi(x) \in \mathbb{E}'$. Then there exists a field isomorphism $\hat{\varphi} : \mathbb{F}(\alpha) \to \mathbb{F}'(\beta)$ lifting $\varphi : \mathbb{F} \to \mathbb{F}'$ and sending $\alpha \mapsto \beta$. Here is a picture:



It is worth highlighting the special case when $\mathbb{F} = \mathbb{F}'$, $\mathbb{E} = \mathbb{E}'$ and $\varphi = \mathrm{id}$. In this case if $\alpha, \beta \in \mathbb{E}$ are any two roots of an irreducible polynomial $f(x) \in \mathbb{F}$ then there exists an isomorphism $\mathbb{F}(\alpha) \to \mathbb{F}(\beta)$ sending $\alpha \mapsto \beta$ and fixing the elements of $\mathbb{F}$. ///

Proof. The proof is easy, but only because we have developed the right technology. Let $\varphi : \mathbb{F} \to \mathbb{F}'$ be a field isomorphism and let $\alpha \in \mathbb{E} \supseteq \mathbb{F}$ be a root of an **irreducible** polynomial $f(x) \in \mathbb{F}[x]$. Since $\varphi : \mathbb{F} \to \mathbb{F}'$ is a ring isomorphism we obtain an isomorphism of polynomial rings $\mathbb{F}[x] \to \mathbb{F}'[x]$ by letting φ act on the coefficients:

$$\begin{array}{ccc} \mathbb{F}[x] & \xrightarrow{\sim} & \mathbb{F}'[x] \\ f(x) & \mapsto & f^\varphi(x). \end{array}$$

Suppose that there exist a root $\beta \in \mathbb{E}'$ of the image polynomial $f^\varphi(x) \in \mathbb{F}[x]$. Since $f(x)$ and $f^\varphi(x)$ are both irreducible, it follows that these polynomial are (up to non-zero scalar multiples) the minimal polynomials for $\alpha/\mathbb{F}$ and $\beta/\mathbb{F}'$, respectively. Then from the isomorphism $\mathbb{F}[x] \cong \mathbb{F}'[x]$ and the Minimal Polynomial Theorem we obtain a sequence of three isomorphisms:

$$\begin{array}{ccccccc} \mathbb{F}(\alpha) & \cong & \mathbb{F}[x]/\langle f(x)\rangle & \cong & \mathbb{F}'[x]/\langle f^\varphi(x)\rangle & \cong & \mathbb{F}'(\beta) \\ \alpha & \leftrightarrow & x + \langle f(x)\rangle & \leftrightarrow & x + \langle f^\varphi(x)\rangle & \leftrightarrow & \beta. \end{array}$$

Composing these gives the desired isomorphism $\mathbb{F}(\alpha) \cong \mathbb{F}'(\beta)$. □

Application: Complex Conjugation. Let $\mathbb{E} = \mathbb{F}(\iota)$ with $\iota^2 = a \in \mathbb{F}$ and $\iota \notin \mathbb{F}$. Then $\pm\iota$ are roots of the irreducible polynomial $x^2 - a \in \mathbb{F}[x]$. It follows from the Lifting Lemma that there exists an isomorphism $\mathbb{E} = \mathbb{F}(\iota) \to \mathbb{F}(-\iota) = \mathbb{E}$ sending $\iota \mapsto -\iota$ and fixing $\mathbb{F}$. This proves that the "conjugation" map $a + b\iota \mapsto a - b\iota$ is a field automorphism, without doing any calculations. ///

That was a small time savings, but the next one is substantial.

Application: The Splitting Field of $x^3 - 2 \in \mathbb{Q}[x]$. Recall that there exist six functions $\sigma : \mathbb{E} \to \mathbb{E}$ defined by letting $\sigma(\alpha)$ and $\sigma(\omega)$ be any roots of $x^3 - 2$ and $x^2 + x + 1$, respectively. Let's prove that these functions are field automorphisms.

First, let $\alpha' \in \{\alpha, \omega\alpha, \omega^2\alpha\}$ be any root of $x^3 - 2$. Since $x^3 - 2$ is irreducible there exists a field isomorphism $\varphi : \mathbb{Q}(\alpha) \to \mathbb{Q}(\alpha')$ sending $\alpha \mapsto \alpha'$ and fixing $\mathbb{Q}$. Next, observe that the polynomial $x^2 + x + 1$ is still irreducible over $\mathbb{Q}(\alpha)$ because $\mathbb{Q}(\alpha) \subseteq \mathbb{R}$ and $x^2 + x + 1$ has no real roots. Thus if $\omega' \in \{\omega, \omega^2\}$ is any root of $x^2 + x + 1$ then there exists an isomorphism $\hat{\varphi} : \mathbb{Q}(\alpha)(\omega) \to \mathbb{Q}(\alpha')(\omega')$ lifting φ and sending $\omega \mapsto \omega'$. In particular, this $\hat{\varphi}$ also sends $\alpha \mapsto \alpha'$ and fixes $\mathbb{Q}$. Finally, since

$$\mathbb{E} = \mathbb{Q}(\alpha, \omega) = \mathbb{Q}(\alpha)(\omega) \cong \mathbb{Q}(\alpha')(\omega') \subseteq \mathbb{E}$$

we conclude that $\hat{\varphi} : \mathbb{E} \to \mathbb{E}$ is a field automorphism. Thus we have proved the **existence** of the six desired elements of the Galois group $\mathrm{Gal}(\mathbb{E}/\mathbb{Q})$. ///

The following theorem simply generalizes this procedure. The theorem is strangely worded, but this is only for the purposes of the induction proof. Our real interest is the special case when $\varphi = \mathrm{id} : \mathbb{F} \to \mathbb{F}$ is the identity. Before stating the theorem it is worth restating the definition of a splitting field.

Let $f(x) \in \mathbb{F}[x]$ be a polynomial. We say that $\mathbb{E} \supseteq \mathbb{F}$ is a *splitting field* for $f(x)$ if

- The polynomial $f(x)$ splits in $\mathbb{E}[x]$. That is, we have

$$f(x) = (x - \alpha_1)(x - \alpha_2) \cdots (x - \alpha_n) \text{ for some } \alpha_1, \ldots, \alpha_n \in \mathbb{E}.$$

- If $\mathbb{E} \supseteq \mathbb{K} \supseteq \mathbb{F}$ and if $f(x)$ splits in $\mathbb{K}[x]$ then $\mathbb{K} = \mathbb{E}$. Equivalently, we have

$$\mathbb{E} = \mathbb{F}(\alpha_1, \ldots, \alpha_n).$$

The Splitting Field Theorem (Existence of Automorphisms). Consider the following:

- Let $\varphi : \mathbb{F} \to \mathbb{F}'$ be an isomorphism of fields and let $f(x) \in \mathbb{F}[x]$ be any polynomial.
- Let $\mathbb{E} \supseteq \mathbb{F}$ and $\mathbb{E}' \supseteq \mathbb{F}'$ be splitting fields for $f(x) \in \mathbb{F}[x]$ and $f^\varphi(x) \in \mathbb{F}'[x]$.

- Let $p_i(x)|f(x)$ be a list of distinct[63] irreducible factors in $\mathbb{F}[x]$.
- For each i let $\alpha_i \in \mathbb{E}$ be a root of $p_i(x)$ and let $\beta_i \in \mathbb{E}'$ be a root of $p^\varphi(x)$. Note that such roots always exist because $\mathbb{E}$ and $\mathbb{E}'$ are splitting fields.[64]

Then there exists an isomorphism $\Phi : \mathbb{E} \to \mathbb{E}'$ lifting φ and sending $\alpha_i \mapsto \beta_i$ for all i. ///

[Remark: There are two small issues in this proof that you will check on the homework. Namely, (a) any divisor of a split polynomial is also split, and (b) non-associate irreducible polynomials have no roots in common.]

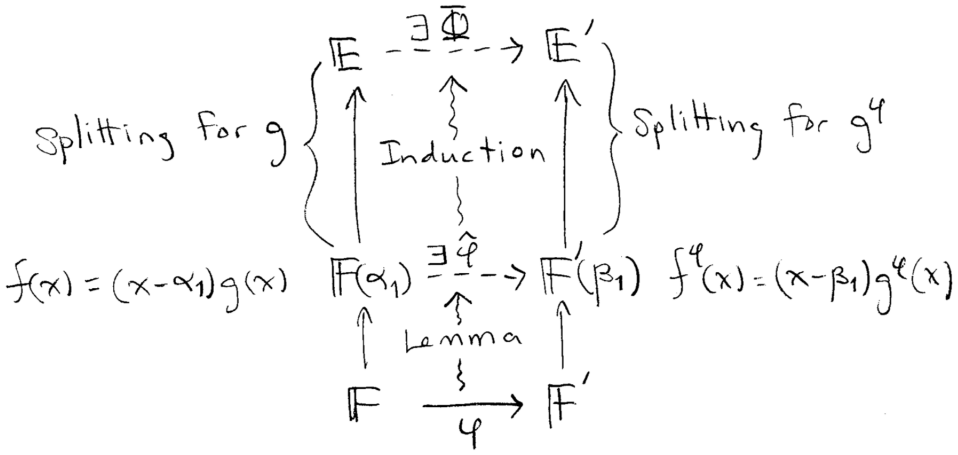
Proof. We will use induction on $\deg(f)$. The result is vacuously true when $\deg(f) = 1$ so let $\deg(f) = n \geq 2$ and let $\mathbb{E} \supseteq \mathbb{F}$ be a splitting field for $f(x) \in \mathbb{F}[x]$. If $p_1(x)|f(x)$ is any irreducible factor then since $p_1(x)$ splits in $\mathbb{E}$ we know that $p_1(x)$ has a root, say $\alpha_1 \in \mathbb{E}$. Next, observe that $p_1^\varphi(x)|f^\varphi(x)$ in $\mathbb{F}'[x]$. Since $\mathbb{E}'$ is a splitting field for $f^\varphi(x)$ this implies that $p_1^\varphi(x)$ has some root, say $\beta_1 \in \mathbb{E}'$. Thus from the Lifting Lemma we obtain an isomorphism $\hat{\varphi} : \mathbb{F}(\alpha_1) \to \mathbb{F}'(\beta_1)$ lifting φ and sending $\alpha_1 \mapsto \beta_1$.

Next, by Descartes' Theorem there exists $g(x) \in \mathbb{F}(\alpha_1)[x]$ of degree $n-1$ such that $f(x) = (x-\alpha_1)g(x)$ and by applying the isomorphism $\hat{\varphi}$ we obtain $f^\varphi(x) = (x-\beta_1)g^\varphi(x)$ for some $g^\varphi(x) \in \mathbb{F}'[x]$. Observe that $\mathbb{E} \supseteq \mathbb{F}(\alpha)$ is a splitting field for $g(x)$ since if $g(x)$ splits over an intermediate field $\mathbb{E} \supseteq \mathbb{K} \supseteq \mathbb{F}(\alpha)$ then $f(x)$ also splits over $\mathbb{K}$. Since $\mathbb{E}$ is a splitting field for $f(x)$ this implies that $\mathbb{K} = \mathbb{E}$. Similarly, $\mathbb{E}'$ is a splitting field for $g^\varphi(x)$.

Furthermore, if $p_2(x)|f(x)$ is irreducible and not a scalar multiple of $p_1(x)$ then since $p_1(x), p_2(x)$ have no common root we must have $p_2(x) \nmid (x-\alpha_1)$ and hence $p_2(x)|g(x)$. Finally, let $\alpha_i \in \mathbb{E}$ and $\beta_i \in \mathbb{E}'$ be any roots of the polynomials $p_i(x) \in \mathbb{F}[x]$ and $p^\varphi(x) \in \mathbb{F}'[x]$ for $i \geq 2$. Since $\deg(g) < \deg(f)$ we conclude by induction that there exists an isomorphism $\Phi : \mathbb{E} \to \mathbb{E}'$ lifting $\hat{\varphi} : \mathbb{F}(\alpha_1) \to \mathbb{F}'(\beta_1)$ and sending $\alpha_i \mapsto \beta_i$ for all $i \geq 2$, hence Φ also lifts φ and sends $\alpha_1 \mapsto \beta_1$. Here is a picture:

[63] Technically: We assume that the polynomials a pairwise non-associate. That is, for any $i \neq j$ and $\lambda \in \mathbb{F}$ we have $p_i(x) \neq \lambda p_j(x)$.

[64] This is fairly obvious but it still needs a proof. You will provide a proof on the homework, using the fact that $\mathbb{E}[x]$ is a UFD.



□

To see how this applies to the previous example, let $\mathbb{E} \supseteq \mathbb{Q}$ be a splitting field for $x^3 - 2$ and observe that $\mathbb{E}$ is also a splitting field for $f(x) = (x^3 - 2)(x^2 + x + 1)$. Let $p_1(x) = x^3 - 2$ and $p_2(x) = x^2 + x + 1$. Then for any roots α, α' of $x^3 - 2$ and roots ω, ω' of $x^2 + x + 1$ there exists an isomorphism $\mathbb{E} \to \mathbb{E}$ sending $(\alpha, \omega) \mapsto (\alpha', \omega')$ and fixing $\mathbb{Q}$. This is a powerful theorem.

In addition to helping us compute Galois groups, the Splitting Field Theorem has an important theoretical corollary.

Corollary (Uniqueness of Splitting Fields). Let $\mathbb{E}, \mathbb{E}' \supseteq \mathbb{F}$ be splitting fields for the same polynomial $f(x) \in \mathbb{F}[x]$. Then there exists a (non-unique) isomorphism $\Phi : \mathbb{E} \to \mathbb{E}'$ fixing $\mathbb{F}$.

Proof. Take $\mathbb{F} = \mathbb{F}'$ and $\varphi = \mathrm{id}$ in the theorem. Ignore the roots of $f(x)$. □

So far we have only defined Galois groups for extensions. This corollary allows us to define the Galois group of a polynomial. Note that this is the reverse of the historical development.

The Galois Group of a Polynomial. Let $f(x) \in \mathbb{F}[x]$ be any polynomial and let $\mathbb{E} \supseteq \mathbb{F}$ be any splitting field for $f(x)$. We define the *Galois group of f over $\mathbb{F}$* as follows:

$$\mathrm{Gal}(f/\mathbb{F}) := \mathrm{Gal}(\mathbb{E}/\mathbb{F}).$$

I claim that this group is well-defined up to isomorphism.

Proof. Let $\mathbb{E}, \mathbb{E}' \supseteq \mathbb{F}$ be any two splitting fields and let $\Phi : \mathbb{E} \to \mathbb{E}'$ be an isomorphism fixing $\mathbb{F}$, which exists by the corollary. Then claim that the map $\sigma \mapsto \Phi \circ \sigma \circ \Phi^{-1}$ is a group

isomorphism $\mathrm{Gal}(\mathbb{E}/\mathbb{F}) \to \mathrm{Gal}(\mathbb{E}'/\mathbb{F})$. Indeed, for any $\sigma \in \mathrm{Gal}(\mathbb{E}/\mathbb{F})$ we observe that the function $\Phi \circ \sigma \circ \Phi^{-1} : \mathbb{E}' \to \mathbb{E}'$ is a field automorphism that fixes $\mathbb{F}$. Then we observe that the function $\sigma \mapsto \Phi \circ \sigma \circ \Phi^{-1}$ is invertible and preserves composition. $\square$

Remarks:

- The uniqueness of splitting fields is not news for finite extensions of $\mathbb{Q}$. Indeed, if $[\mathbb{F}/\mathbb{Q}] < \infty$ then since every element of $\mathbb{F}$ is algebraic over $\mathbb{Q}$ we know from the FTA that $\mathbb{F} \subseteq \mathbb{C}$. Then for any polynomial $f(x) \in \mathbb{F}[x]$ we may view the splitting field as the intersection of all subfields of $\mathbb{C}$ that contain the roots of $f(x)$.
- However, for fields of characteristic p we get new information. For example, let $\mathbb{E}$ be finite of characteristic p. Then we have previously shown that $\mathbb{E}$ is a splitting field for the polynomial $x^{p^k} - x \in \mathbb{F}_p[x]$ for some k. It follows from the uniqueness of splitting fields that any two fields of size p^k are isomorphic. This new proof is a bit more elegant than our old proof because it avoids the Primitive Root Theorem.
- It is worth emphasizing one more consequence of the Splitting Field Theorem. If $\mathbb{E} \supseteq \mathbb{F}$ is a splitting field for $f(x) \in \mathbb{F}[x]$ and if $p(x)|f(x)$ is any irreducible factor, then for any two roots $\alpha, \beta \in \mathbb{E}$ of $p(x)$ there exists an automorphism $\sigma : \mathbb{E} \to \mathbb{E}$ fixing $\mathbb{F}$ and sending $\alpha \mapsto \beta$. In other words:

  The Galois group $\mathrm{Gal}(\mathbb{E}/\mathbb{F})$ acts **transitively** on the roots of $p(x)$.

  Note that this fact does **not** apply to reducible polynomials. For example, let $\mathbb{E} \supseteq \mathbb{Q}$ be a splitting field for $f(x) = (x^2 - 2)(x^2 - 3)$. Then we have $f(\sqrt{2}) = f(\sqrt{3}) = 0$, but there does **not** exist any group element $\sigma \in \mathrm{Gal}(\mathbb{E}/\mathbb{F})$ sending $\sqrt{2} \mapsto \sqrt{3}$. (Why not?)

Week 22

We will see that that the basic theorems of Galois theory hold for finite fields and for fields of characteristic zero. However, it is a sad fact that there exist certain infinite fields of characteristic p for which the theorems break down. Ernst Steinitz (1910) was the first person to deal uniformly with the good cases, while excluding the pathological cases. He did this with the following definition.

Steinitz' Definition of Perfect Fields. We say that a field $\mathbb{F}$ is *perfect* [vollkommene] if

- $\mathrm{char}(\mathbb{F}) = 0$, or
- $\mathrm{char}(\mathbb{F}) = p$ and the function $\mathbb{F} \to \mathbb{F}$ defined by $a \mapsto a^p$ is surjective. You will show on the homework that this case includes **all finite fields**.

///

Unfortunately, this definition is just a notational device because any theorem about perfect fields requires two separate proofs for the two cases.[65] Here are the properties that we need for Galois theory.

Nice Properties of Perfect Fields. Let $\mathbb{F}$ be a perfect field. Then:

(1) **Irreducible Polynomials are Separable.**

If $f(x) \in \mathbb{F}[x]$ is irreducible then $f(x)$ has no repeated roots in any field extension.

(2) **Primitive Elements Exist.**

If $\mathbb{E} \supseteq \mathbb{F}$ is finite-dimensional then there exists an element $\gamma \in \mathbb{E}$ such that $\mathbb{E} = \mathbb{F}(\gamma)$.

///

As I mentioned, this theorem requires separate proofs for the cases $\text{char}(\mathbb{F}) = 0$ and $\#\mathbb{F} < \infty$. I hope you don't mind that I relegated some of the steps to the homework. Also, we will ignore the case of infinite perfect fields of characteristic p.

Proof. (1) Let $f(x) \in \mathbb{F}[x]$ be irreducible and assume for contradiction that $f(x)$ has a repeated root in some field extension. On the homework you will show that this implies $g(x) = \gcd(f, Df)$ has degree ≥ 1, where $Df(x) \in \mathbb{F}[x]$ is the formal derivative. Since $f(x)$ is irreducible this implies that $g(x) = \lambda f(x)$ for some $\lambda \in \mathbb{F}$. But we also know that $g(x)|Df(x)$. If $\text{char}(\mathbb{F}) = 0$ then this is a contradiction because $\deg(Df) = \deg(f) - 1$. If $\#\mathbb{F} < \infty$ then it could possibly be the case that $Df(x)$ is identically zero, but you will show on the homework that this also leads to a contradiction.

(2) If $\#\mathbb{F} < \infty$ and $[\mathbb{E}/\mathbb{F}] < \infty$ then we also have $\#\mathbb{E} < \infty$. You showed on a previous homework that the group $(\mathbb{E}^\times, \times, 1)$ is cyclic (we called this the **Primitive Root Theorem**). Say $\mathbb{E}^\times = \{\gamma^n : n \in \mathbb{Z}\}$ for some $\gamma \in \mathbb{E}$. Then clearly every element of $\mathbb{E}$ can be expressed in the form $f(\gamma)$ for some $f(x) \in \mathbb{F}[x]$, hence $\mathbb{E} = \mathbb{F}(\gamma)$.

(2) Next suppose that $\text{char}(\mathbb{F}) = 0$ and $[\mathbb{E}/\mathbb{F}] < \infty$. This case is sometimes called the **Primitive Element Theorem**. The proof goes back to Galois.

By the Finiteness Theorem we know that $\mathbb{E} = \mathbb{F}(\alpha_1, \ldots, \alpha_n)$ for some algebraic elements $\alpha_1, \ldots, \alpha_n \in \mathbb{E}$. Thus by induction it suffices to prove for all algebraic $\alpha, \beta \in \mathbb{E}$ that

$$\mathbb{F}(\alpha, \beta) = \mathbb{F}(\gamma) \quad \text{for some } \gamma \in \mathbb{E}.$$

Let $m_\alpha(x), m_\beta(x) \in \mathbb{F}[x]$ be the minimal polynomials of α, β over $\mathbb{F}$. Then since $\mathbb{F}$ is **infinite** we may choose a non-zero element $c \in \mathbb{F}$ such that

$$c \neq \frac{\alpha' - \alpha}{\beta - \beta'} \quad \text{for all roots } \alpha' \neq \alpha \text{ of } m_\alpha \text{ and } \beta' \neq \beta \text{ of } m_\beta.$$

[65] Maybe there is a deep connection between the two cases that I don't know about?

In this case I claim that $\gamma := \alpha + c\beta$ is a primitive element. Indeed, since $\gamma \in \mathbb{F}(\alpha, \beta)$ we have $\mathbb{F}(\gamma) \subseteq \mathbb{F}(\alpha, \beta)$. Conversely, we want to show that $\alpha, \beta \in \mathbb{F}(\gamma)$ and hence $\mathbb{F}(\alpha, \beta) \subseteq \mathbb{F}(\gamma)$, and for this it suffices to prove $\beta \in \mathbb{F}(\gamma)$ since then we also have $\alpha = \gamma - c\beta \in \mathbb{F}(\gamma)$.

We will show that $\beta \in \mathbb{F}(\gamma)$ by an indirect argument. That is, let $m'_\beta(x) \in \mathbb{F}(\gamma)[x]$ be the minimal polynomial of β over $\mathbb{F}(\gamma)$. We will prove that $\deg(m'_\beta) = 1$ and hence $\beta \in \mathbb{F}(\gamma)$. By thinking of $m_\beta(x)$ as an element of $\mathbb{F}(\gamma)[x]$ we clearly have $m'_\beta(x) | m_\beta(x)$. Now we need to get α in on the action. So (TRICK) define the polynomial

$$f(x) := m_\alpha(\gamma - cx) \in \mathbb{F}(\gamma)[x].$$

By construction we have $f(\beta) = m_\alpha(\gamma - c\beta) = m_\alpha(\alpha) = 0$ which implies that $m'_\beta(x) | f(x)$. It follows that any root of $m'_\beta(x)$ is a common root of $m_\beta(x)$ and $f(x)$.

Finally, let $\mathbb{E}' \supseteq \mathbb{E}$ be a splitting field for the polynomial $m_\alpha(x)m_\beta(x)$. We know from part (1) that each of the polynomials $m_\alpha(x), f(x), m_\beta(x), m'_\beta(x)$ splits and has no repeated roots in $\mathbb{E}'$. It follows that the number of common roots of $m_\beta(x)$ and $f(x)$ in $\mathbb{E}'$ is equal to $\deg(m'_\beta)$. We will be done if we can show that β **is the only common root**. So assume for contradiction that we have $m_\beta(\beta') = f(\beta') = 0$ for some $\beta' \neq \beta$. By definition of $f(x)$ this means that $\alpha' := \gamma - c\beta'$ is a root of $m_\alpha(x)$. But then we have

$$\begin{aligned} \alpha' &= \gamma - c\beta' \\ \alpha' &= (\alpha + c\beta) - c\beta' \\ c &= (\alpha' - \alpha)/(\beta - \beta'), \end{aligned}$$

which contradicts the definition c. □

Here is an application to our favorite example.

Example: A Primitive Element for the Splitting Field of $x^3 - 2$. Recall that the splitting field is $\mathbb{Q}(\alpha, \omega)$ where $\alpha = \sqrt[3]{2} \in \mathbb{R}$ and $\omega = e^{2\pi i/3}$. We are looking for an element of the form $\gamma = \alpha + c\omega$ with nonzero $c \in \mathbb{Q}$ such that $\mathbb{Q}(\alpha, \omega) = \mathbb{Q}(\gamma)$. By the proof of the previous theorem it suffices to choose c such that

$$c \neq \frac{\alpha' - \alpha}{\omega - \omega'} \quad \text{for all } \alpha' \in \{\omega\alpha, \omega^2\alpha\} \text{ and } \omega' \in \{\omega^2\}.$$

But note that $\omega - \omega^2$ is purely imaginary and $\alpha' - \alpha$ never is. Thus we may take **any nonzero element** $c \in \mathbb{Q}$. For example, $c = 1$. ///

Remarks:

- The proof of (2) for finite fields goes back to Gauss and the proof of (2) for characteristic zero fields goes back to Galois. In fact, John Stillwell[66] says that this was the first

[66] *Elements of Algebra*, page 160

substantial result in Galois' 1831 memoir. In modern language, Galois' version says that for any algebraic imaginaries $\alpha, \beta \in \mathbb{C}$ there exists an integer $c \in \mathbb{Z}$ such that $\mathbb{Q}(\alpha, \beta) = \mathbb{Q}(\alpha + c\beta)$. By induction it follows that any finite extension satisfies

$$\mathbb{Q}(\alpha_1, \ldots, \alpha_n) = \mathbb{Q}(c_1\alpha_1 + \cdots + c_n\alpha_n) \text{ for some integers } c_1, \ldots, c_n \in \mathbb{Z}.$$

Such an element $\gamma = c_1\alpha_1 + \cdots + c_n\alpha_n$ was called a *Galois resolvent*, but today it is usually called a *primitive element*.

- Sadly, there exist pathological examples of finite-dimensional field extensions which do not have a primitive element. For example, consider the field $\mathbb{F} = \mathbb{F}_p(x, y)$ consisting of fractions $f(x, y)/g(x, y)$ where $f, g \in \mathbb{F}_p[x, y]$ and $g \neq 0$. Let $\mathbb{E} = \mathbb{F}(\alpha, \beta)$ where $\alpha^p = x$ and $\beta^p = y$. Then one can show that $[\mathbb{E}/\mathbb{F}] = p^2 < \infty$ but has no primitive element.
- When Dedekind modernized Galois theory he continued to use primitive elements because his main concern was with finite extensions of $\mathbb{Q}$. However, after Steinitz included characteristic p in his 1910 memoir, other mathematicians such as Emmy Noether began to reject the use of primitive elements because they are not completely general. In a 1935 memorial address[67] after Emmy Noether's death, Hermann Weyl praised her "drive toward axiomatic purity," but he thought that it was not always appropriate:

  > *This can be carried too far, however, as when she disdained to employ a primitive element in the development of Galois theory.*

- I agree with with Weyl that the use of primitive elements leads to the cleanest development of Galois theory, at least for beginners. You will see this in the next lecture.

Dedekind proved for any field extension $\mathbb{E} \supseteq \mathbb{F}$ that $\#\mathrm{Gal}(\mathbb{E}/\mathbb{F}) \leq [\mathbb{E}/\mathbb{F}]$.[68] Recall that finite-dimensional extensions satisfying $\#\mathrm{Gal}(\mathbb{E}/\mathbb{F}) = [\mathbb{E}/\mathbb{F}]$ are called *Galois extensions*. As with normal subgroups, there are several equivalent ways to state the definition. Today we will prove a big characterization theorem for Galois extensions over perfect fields.

We isolate the following lemma for pedagogical reasons. Emil Artin proved this lemma for general fields,[69] using linear algebraic techniques inspired by Dedekind. We will only prove it for finite-dimensional extensions over perfect fields.

Artin's Fixed Field Lemma. Let $\mathbb{E}$ be any field and let $G \subseteq \mathrm{Aut}(\mathbb{E})$ be any **finite** group of automorphisms with fixed field $\mathrm{Fix}_{\mathbb{E}}(G) \subseteq \mathbb{E}$. Then we have

$$[\mathbb{E}/\mathrm{Fix}_{\mathbb{E}}(G)] = \#G.$$

[67] Reprinted as an appendix in *Emmy Noether: 1882–1935* by Auguste Dick (1981).
[68] We didn't prove this, but we did prove a weaker version called the Finiteness Theorem.
[69] It is Theorem 14 in his *Galois Theory* (1942), reprinted by Dover (1998).

///

Proof. As I said, we will only prove a special case of this. Let $\mathbb{F}$ be perfect and let $\mathbb{E} \supseteq \mathbb{F}$ be a finite-dimensional extension, so there exists a primitive element $\mathbb{E} = \mathbb{F}(\gamma)$. Now consider any (finite) subgroup $G \subseteq \mathrm{Gal}(\mathbb{E}/\mathbb{F})$ and let $\mathrm{Orb}_G(\gamma) = \{\gamma_1, \ldots, \gamma_n\}$ be the G-orbit of γ. Since $\mathbb{E}$ is generated by γ over $\mathbb{F}$ we see that $\mathrm{Stab}_G(\gamma) = \{\mathrm{id}\}$ and hence

$$n = \#\mathrm{Orb}_G(\gamma) = \#G/\#\,\mathrm{Stab}_G(\gamma) = \#G.$$

Now consider the following polynomial with degree n and no repeated roots:

$$f(x) = (x - \gamma_1) \cdots (x - \gamma_n) \in \mathbb{E}[x].$$

Since every element of G permutes the roots of $f(x)$ it also fixes the coefficients, hence $f(x) \in \mathrm{Fix}_{\mathbb{E}}(G)[x]$. I claim in fact that $f(x)$ is the minimal polynomial for γ over $\mathrm{Fix}_{\mathbb{E}}(G)$. Indeed, let $m(x) \in \mathrm{Fix}_{\mathbb{E}}(G)[x]$ be the minimal polynomial. Then since $f(\gamma) = 0$ we have $m(x)|f(x)$. Conversely, since every $\gamma_i \in \mathrm{Orb}_G(\gamma)$ has the form $\sigma(\gamma)$ for some $\sigma \in G$ we must have

$$m(\gamma_i) = m(\sigma(\gamma)) = \sigma(m(\gamma)) = \sigma(0) = 0.$$

Then it follows from Descartes' Theorem that $f(x)|m(x)$ and hence $f(x) = m(x)$. Finally, since $\mathrm{Fix}_{\mathbb{E}}(\gamma) = \mathbb{F}(\gamma) = \mathbb{E}$ we conclude from the Minimal Polynomial Theorem that

$$[\mathbb{E}/\mathrm{Fix}_{\mathbb{E}}(G)] = [\mathrm{Fix}_{\mathbb{E}}(G)(\gamma)/\mathrm{Fix}_{\mathbb{E}}(G)] = \deg(m) = \deg(f) = n = \#G.$$

□

Before stating today's theorem let me note that if $\mathbb{F}$ is perfect and if $\mathbb{E} \supseteq \mathbb{F}$ is finite-dimensional then $\mathbb{E}$ is also perfect. Indeed, for all $\mathbb{E} \supseteq \mathbb{F}$ we have $\mathrm{char}(\mathbb{F}) = 0 \Rightarrow \mathrm{char}(\mathbb{E}) = 0$ and for all $[\mathbb{E}/\mathbb{F}] < \infty$ we have $\#\mathbb{F} < \infty \Rightarrow \#\mathbb{E} < \infty$. Again, we don't care about the other cases.

I find the following theorem amazing. Galois is lucky to have this concept named after him.

Characterization Theorem for Galois Extensions (of Perfect Fields). Let $\mathbb{E} \supseteq \mathbb{F}$ be a finite-dimensional extension of perfect fields. Then the following five conditions are equivalent:

(GE1) $\#\mathrm{Gal}(\mathbb{E}/\mathbb{F}) = [\mathbb{E}/\mathbb{F}]$

(GE2) $\mathrm{Fix}_{\mathbb{E}}(\mathrm{Gal}(\mathbb{E}/\mathbb{F})) = \mathbb{F}$

(GE3) $\mathbb{E}$ is a splitting field for some polynomial $f(x) \in \mathbb{F}[x]$.

(GE4) For any $\mathbb{E}' \supseteq \mathbb{E}$ and $\sigma \in \mathrm{Gal}(\mathbb{E}'/\mathbb{F})$ we have $\sigma(\mathbb{E}) \subseteq \mathbb{E}$.[70]

(GE5) If $m(x) \in \mathbb{F}[x]$ is irreducible and has a root in $\mathbb{E}$, then $m(x)$ splits in $\mathbb{E}[x]$.

[70] And hence $\sigma(\mathbb{E}) = \mathbb{E}$. Indeed, since $\sigma(1) = 1$ we know that $\ker\sigma \neq \mathbb{E}$. Then since a field has no non-trivial ideals we must have $\ker\sigma = \{0\}$. Finally, since $\sigma : \mathbb{E}/\mathbb{F} \to \mathbb{E}/\mathbb{F}$ is an injective endomorphism of a finite-dimensional $\mathbb{F}$-vector space we conclude from the Rank-Nullity Theorem that σ is also surjective.

///

Proof. To save space we will write $G = \mathrm{Gal}(\mathbb{E}/\mathbb{F})$.

(GE1)⇔(GE2) Since $\mathbb{F}$ is perfect and $[\mathbb{E}/\mathbb{F}] < \infty$ there exists a primitive element $\gamma \in \mathbb{E}$ with $\mathbb{E} = \mathbb{F}(\gamma)$. Since $\mathbb{E} \supseteq \mathrm{Fix}_{\mathbb{E}}(G) \supseteq \mathbb{F}$ this also implies $\mathbb{E} = \mathrm{Fix}_{\mathbb{E}}(G)(\gamma)$. Then from the Fixed Field Lemma and Dedekind's Tower Law we have

$$[\mathbb{E}/\mathbb{F}] = [\mathbb{E}/\mathrm{Fix}_{\mathbb{E}}(G)] \cdot [\mathrm{Fix}_{\mathbb{E}}(G)/\mathbb{F}] = \#G \cdot [\mathrm{Fix}_{\mathbb{E}}(G)/\mathbb{F}].$$

It follows that

$$\#G = [\mathbb{E}/\mathbb{F}] \iff [\mathrm{Fix}_{\mathbb{E}}(G)/\mathbb{F}] = 1 \iff \mathrm{Fix}_{\mathbb{E}}(G) = \mathbb{F}.$$

(GE2)⇒(GE3) Assume that $\mathrm{Fix}_{\mathbb{E}}(G) = \mathbb{F}$. Since $\mathbb{F}$ is perfect there exists a primitive element $\mathbb{E} = \mathbb{F}(\gamma)$ with minimal polynomial $m(x) \in \mathbb{F}[x]$ satisfying $\deg(m) = [\mathbb{F}(\gamma)/\mathbb{F}] = [\mathbb{E}/\mathrm{Fix}_{\mathbb{E}}(G)] = \#G$. From the proof of the Fixed Field Lemma we also know that $m(x)$ has $\#G$ distinct roots in $\mathbb{E}$. It follows that $m(x)$ splits in $\mathbb{E}[x]$ and hence $\mathbb{E} = \mathbb{F}(\gamma)$ is a splitting field for $m(x) \in \mathbb{F}[x]$.

(GE3)⇒(GE4) Assume that there exists some $f(x) \in \mathbb{F}[x]$ with $f(x) = (x-\alpha_1)\cdots(x-\alpha_n) \in \mathbb{E}[x]$ and $\mathbb{E} = \mathbb{F}(\alpha_1, \ldots, \alpha_n)$. From the Finiteness Theorem we know that $\mathbb{F}(\alpha_1, \ldots, \alpha_n) = \mathbb{F}[\alpha_1, \ldots, \alpha_n]$. In other words, $\mathbb{E}$ is the set of evaluations $g(\alpha_1, \ldots, \alpha_n)$ of polynomials $g \in \mathbb{F}[x_1, \ldots, x_n]$. Now consider any field extension $\mathbb{E}' \supseteq \mathbb{E}$ and any automorphism $\sigma \in \mathrm{Gal}(\mathbb{E}'/\mathbb{F})$. Since σ fixes $\mathbb{F}$ it necessarily permutes the roots of $f(x)$. Then for any element $g(\alpha_1, \ldots, \alpha_n) \in \mathbb{E}$ we have

$$\sigma\left(g(\alpha_1, \ldots, \alpha_n)\right) = g\left(\sigma(\alpha_1), \ldots, \sigma(\alpha_n)\right) \in \mathbb{E},$$

since this last expression is also a polynomial evaluated at the roots of $f(x)$.

(GE4)⇒(GE5) Let $m(x) \in \mathbb{F}[x]$ be irreducible and let $\mathbb{E}' \supseteq \mathbb{E}$ be a splitting field for $m(x)$. Let $\Omega \subseteq \mathbb{E}'$ be the roots of $m(x)$ and assume that this set contains an element of $\mathbb{E}$, say $\alpha \in \Omega \cap \mathbb{E}$. Now consider the group $G' = \mathrm{Gal}(\mathbb{E}'/\mathbb{F})$. Since $\mathbb{E}'$ is a splitting field for the irreducible polynomial $m(x) \in \mathbb{F}[x]$ we know from the Splitting Field Theorem that G' acts transitively on Ω. In other words, we have $\mathrm{Orb}_{G'}(\alpha) = \Omega$. But by assumption we also know that G' sends $\mathbb{E}$ to $\mathbb{E}$. It follows that

$$\Omega = \mathrm{Orb}_{G'}(\alpha) \subseteq \mathbb{E},$$

and hence $m(x)$ splits in $\mathbb{E}[x]$.

(GE5)⇒(GE1) Since $\mathbb{F}$ is perfect there exists a primitive element $\mathbb{E} = \mathbb{F}(\gamma)$. If $m(x) \in \mathbb{F}[x]$ is the minimal polynomial for $\gamma/\mathbb{F}$ then by assumption we know that $m(x)$ splits in $\mathbb{E}[x]$. Let $\Omega \subseteq \mathbb{E}$ be the set of roots of $m(x)$ and consider the action of G on Ω. Since γ generates $\mathbb{E}$ over $\mathbb{F}$ we have $\mathrm{Stab}_G(\gamma) = \{\mathrm{id}\}$ and since $\mathbb{E} = \mathbb{F}(\gamma)$ is a splitting field for $m(x)$ we know from the Splitting Field Theorem that $\mathrm{Orb}_G(\gamma) = \Omega$. Finally, since $\mathbb{F}$ is perfect[71] we know that $m(x)$

[71] Alternatively, we could argue again that $m(x) = \prod_i (x - \gamma_i)$ where $\mathrm{Orb}_G(\gamma) = \{\gamma_i\}_i$.

has no repeated roots in $\mathbb{E}$ and it follows that

$$\#G = \frac{\#G}{\#\operatorname{Stab}_G(\gamma)} = \#\mathrm{Orb}_G(\gamma) = \#\Omega = \deg(m_\gamma) = [\mathbb{F}(\gamma)/\mathbb{F}] = [\mathbb{E}/\mathbb{F}].$$

[This is the argument that I previewed after the proof of the Finiteness Theorem.] □

Remarks:

- Normally I don't like TFAE[72] theorems, but I can't think of any pedagogically better way to state these results.
- Observe that this theorem contains $5 \cdot 4 = 20$ implications. I tried to make the whole proof as short as possible, which has the drawback that your favorite implication might not be optimized.
- Many of these equivalences break for extensions of non-perfect fields. If you want to know the details about that then you are reading the wrong book.

The previous lecture was the most difficult one of the course. It's all downhill from here.

Consider any field extension $\mathbb{E} \supseteq \mathbb{F}$ with Galois group $G = \mathrm{Gal}(\mathbb{E}/\mathbb{F})$. Let $\mathscr{L}(\mathbb{E}, \mathbb{F})$ be the lattice of intermediate fields and let $\mathscr{L}(G)$ be the lattice of subgroups. Now recall from the introduction of Part II that we have an abstract Galois connection:

$$\mathrm{Gal}(\mathbb{E}/-) : \mathscr{L}(\mathbb{E}, \mathbb{F}) \rightleftarrows \mathscr{L}(G)^{\mathrm{op}} : \mathrm{Fix}_{\mathbb{E}}(-).$$

Technically, this means that for all subfields $\mathbb{E} \supseteq \mathbb{K} \supseteq \mathbb{F}$ and for all subgroups $H \subseteq G$ we have

$$\mathbb{K} \subseteq \mathrm{Fix}_{\mathbb{E}}(H) \iff \mathrm{Gal}(\mathbb{E}/\mathbb{K}) \supseteq H.$$

Recall that an abstract Galois connection always restricts to an isomorphism between certain subposets of "closed elements." In general, it follows from Artin's Fixed Field Lemma that every **finite** subgroup of G is "closed." If $\mathbb{E} \supseteq \mathbb{F}$ is a Galois extension of perfect fields then it turns out that every intermediate field is also "closed," and in this case we have an isomorphism of lattices $\mathscr{L}(\mathbb{E}, \mathbb{F}) \cong \mathscr{L}(G)^{\mathrm{op}}$.[73] Here is the full statement.

The Fundamental Theorem of Galois Theory. Let $\mathbb{E} \supseteq \mathbb{F}$ be a Galois extension of perfect fields and let $G = \mathrm{Gal}(\mathbb{E}/\mathbb{F})$ be the Galois group. Then:

[72] "The following are equivalent."

[73] This result can be extended to certain "infinite Galois extensions" by replacing the lattice of subgroups with the lattice of "profinite subgroups." Never mind.

(1) The Galois connection $\mathrm{Gal}(\mathbb{E}/-) : \mathscr{L}(\mathbb{E},\mathbb{F}) \leftrightarrows \mathscr{L}(G)^{\mathrm{op}} : \mathrm{Fix}_{\mathbb{E}}(-)$ is actually a **bijection**. That is, for all intermediate fields $\mathbb{E} \supseteq \mathbb{K} \supseteq \mathbb{F}$ and for all subgroups $H \subseteq G$ we have

$$\mathrm{Fix}_{\mathbb{E}}(\mathrm{Gal}(\mathbb{E}/\mathbb{K})) = \mathbb{K} \qquad \text{and} \qquad \mathrm{Gal}(\mathbb{E}/\mathrm{Fix}_{\mathbb{E}}(H)) = H.^{74}$$

(2) For any pair $\mathbb{K} = \mathrm{Fix}_{\mathbb{E}}(H)$ and $H = \mathrm{Gal}(\mathbb{E}/\mathbb{K})$ we have

$$\#\{\text{cosets of } H \text{ in } G\} = \#(G/H) = [\mathbb{K}/\mathbb{F}] = \dim(\mathbb{K} \text{ as a vector space over } \mathbb{F}).$$

(3) Furthermore, we have

$$\mathbb{K} \supseteq \mathbb{F} \text{ is a Galois field extension} \quad \Longleftrightarrow \quad H \trianglelefteq G \text{ is a normal subgroup,}$$

in which case the quotient group is isomorphic to the Galois group:

$$\frac{G}{H} = \frac{\mathrm{Gal}(\mathbb{E}/\mathbb{F})}{\mathrm{Gal}(\mathbb{E}/\mathbb{K})} \cong \mathrm{Gal}(\mathbb{K}/\mathbb{F}).$$

///

Proof. The proof will refer to the Characterization Theorem for Galois extensions.

(1) Consider any intermediate field $\mathbb{E} \supseteq \mathbb{K} \supseteq \mathbb{F}$. Since $\mathbb{E}/\mathbb{F}$ is Galois we know from (GE3) that $\mathbb{E}$ is a splitting field for some polynomial $f(x) \in \mathbb{F}[x]$. But then $\mathbb{E}$ is also a splitting field for $f(x) \in \mathbb{K}[x]$ which implies that $\mathbb{E}/\mathbb{K}$ is Galois. We conclude from (GE2) that $\mathrm{Fix}_{\mathbb{E}}(\mathrm{Gal}(\mathbb{E}/\mathbb{K})) = \mathbb{K}$.

Now consider any subgroup $H \subseteq G$ and let $\mathbb{K} = \mathrm{Fix}_{\mathbb{E}}(G)$, so that $H \subseteq \mathrm{Gal}(\mathbb{E}/\mathbb{K})$. As above we know that $\mathbb{E}/\mathbb{K}$ is Galois, hence from (GE1) we have $\#\mathrm{Gal}(\mathbb{E}/\mathbb{K}) = [\mathbb{E}/\mathbb{K}]$. On the other hand, we know from the Fixed Field Lemma that $\#H = [\mathbb{E}/\mathbb{K}]$ and it follows that $H = \mathrm{Gal}(\mathbb{E}/\mathbb{K})$.

(2) Consider any pair $\mathbb{K} = \mathrm{Fix}_{\mathbb{E}}(H)$ and $H = \mathrm{Gal}(\mathbb{E}/\mathbb{K})$. Since $\mathbb{E}/\mathbb{F}$ and $\mathbb{E}/\mathbb{K}$ are both Galois, we know from Lagrange's Theorem, (GE1) and Dedekind's Tower Law that

$$\#(G/H) = \frac{\#G}{\#H} = \frac{\#\mathrm{Gal}(\mathbb{E}/\mathbb{F})}{\#\mathrm{Gal}(\mathbb{E}/\mathbb{K})} = \frac{[\mathbb{E}/\mathbb{F}]}{[\mathbb{E}/\mathbb{K}]} = [\mathbb{K}/\mathbb{F}].$$

(3) Furthermore, I claim that

$$\mathrm{Gal}(\mathbb{E}/\sigma(\mathbb{K})) = \sigma\,\mathrm{Gal}(\mathbb{E}/\mathbb{K})\,\sigma^{-1} = \sigma H \sigma^{-1} \qquad \text{for all } \sigma \in G.$$

Indeed, this follows immediately from the definitions:

$$\mu \in \mathrm{Gal}(\mathbb{E}/\sigma(\mathbb{K})) \Longleftrightarrow \mu(\sigma(a)) = \sigma(a) \text{ for all } a \in \mathbb{K}.$$

[74] As I mentioned above, the equation $\mathrm{Gal}(\mathbb{E}/\mathrm{Fix}_{\mathbb{E}}(H)) = H$ holds for any field $\mathbb{E}$ and for any finite group of automorphisms $H \subseteq \mathrm{Aut}(\mathbb{E})$. The proof only depends on Artin's Fixed Field Lemma (which, however, we did not prove in full generality). The other equation $\mathrm{Fix}_{\mathbb{E}}(\mathrm{Gal}(\mathbb{E}/\mathbb{K})) = \mathbb{K}$ is more interesting.

$$\iff (\sigma^{-1}\mu\sigma)(a) = a \text{ for all } a \in \mathbb{K}.$$
$$\iff \sigma^{-1}\mu\sigma \in H$$
$$\iff \mu \in \sigma H \sigma^{-1}.$$

Now suppose that $\mathbb{K}/\mathbb{F}$ is Galois. Then from (GE4) we have $\sigma(\mathbb{K}) = \mathbb{K}$ and hence $\sigma H \sigma^{-1}$ for all $\sigma \in G$. In other words, $H \trianglelefteq G$ is normal. Conversely, suppose that $H \trianglelefteq G$ is normal. Then we have $\sigma H \sigma^{-1} = H$ and hence $\mathrm{Gal}(\mathbb{E}/\sigma(\mathbb{K})) = \mathrm{Gal}(\mathbb{E}/\mathbb{K})$ for all $\sigma \in G$. We conclude from the bijection in part **(1)** that $\sigma(\mathbb{K}) = \mathbb{K}$ for all $\sigma \in G$.

Now since each $\sigma \in G$ restricts to an element of $\mathrm{Gal}(\mathbb{K}/\mathbb{F})$ we obtain a "restriction homomorphism" $\varphi : \mathrm{Gal}(\mathbb{E}/\mathbb{F}) \to \mathrm{Gal}(\mathbb{K}/\mathbb{F})$ with kernel $\mathrm{Gal}(\mathbb{K}/\mathbb{F}) = H$. Furthermore, since $\mathbb{E}/\mathbb{K}$ is a splitting field we know from the Splitting Field Theorem that each automorphism $\sigma : \mathbb{K} \to \mathbb{K}$ fixing $\mathbb{F}$ lifts to an automorphism $\hat{\sigma} : \mathbb{E} \to \mathbb{E}$. Hence the restriction homomorphism is **surjective** and we conclude from the First Isomorphism Theorem that

$$\frac{\mathrm{Gal}(\mathbb{E}/\mathbb{F})}{\mathrm{Gal}(\mathbb{E}/\mathbb{K})} = \frac{G}{H} = \frac{G}{\ker\varphi} \cong \mathrm{im}\,\varphi = \mathrm{Gal}(\mathbb{K}/\mathbb{F}).$$

Finally, from part **(2)** we have $\#\mathrm{Gal}(\mathbb{K}/\mathbb{F}) = \#(G/H) = [\mathbb{K}/\mathbb{F}]$ and it follows from (GE1) that $\mathbb{K} \supseteq \mathbb{F}$ is a Galois extension. □

Mathematical Remarks:

- Note that this proof was quite short because already did the hard work. The details are spread over three previous results: the Splitting Field Theorem, the Fixed Field Lemma and the Characterization Theorem for Galois Extensions.
- One surprising corollary of this theorem is that any finite-dimensional extension $\mathbb{E} \supseteq \mathbb{F}$ of perfect fields has **finitely many intermediate fields**. Indeed, if $\mathbb{E} \supseteq \mathbb{F}$ is not Galois then let $\mathbb{E}' \supseteq \mathbb{E}$ be a splitting field for some polynomial $f(x) \in \mathbb{F}[x]$. Then it follows from the Fundamental Theorem that that $\mathscr{L}(\mathbb{E}', \mathbb{F}) \cong \mathscr{L}(G)^{\mathrm{op}}$ where $G = \mathrm{Gal}(\mathbb{E}'/\mathbb{F})$. Since G is a finite group this implies that the lattice $\mathscr{L}(\mathbb{E}', \mathbb{F})$ is finite. Finally, since $\mathscr{L}(\mathbb{E}, \mathbb{F})$ is a subposet of $\mathscr{L}(\mathbb{E}', \mathbb{F})$ we conclude that $\mathscr{L}(\mathbb{E}, \mathbb{F})$ is also finite.
- If the field $\mathbb{E}$ is infinite then one can prove from the finiteness of $\mathscr{L}(\mathbb{E}, \mathbb{F})$ that there exists a primitive element $\mathbb{E} = \mathbb{F}(\gamma)$. Indeed, suppose that $\mathbb{E}$ has finitely many maximal subfields over $\mathbb{F}$. Since each of these is a proper $\mathbb{F}$-subspace of $\mathbb{E}$ we conclude that there exists some $\gamma \in \mathbb{E}$ that is not in any maximal subfield.[75] Then since $\mathbb{F}(\gamma)$ is not contained in any maximal subfield we must have $\mathbb{F}(\gamma) = \mathbb{E}$.
- In fact, Steinitz (1910) proved that the existence of a primitive element is **equivalent** to the existence of only finitely many intermediate fields. But this equivalence is useless for us because it doesn't help us to prove either statement.

[75] We are using intuitively obvious fact that the complement of finitely many proper subspaces is not empty. I prefer not to prove this.

Historical Remarks:

- The Fundamental Theorem is a theorem of *Galois Theory*, but it is not *Galois' Theorem*. The original version of the theorem appears in Dedekind's 11th supplement (1894) to Dirichlet's *Vorlesungen über Zahlentheorie* (Lectures on Number Theory). According to Walther Purkert (1976), Dedekind had lectured on this material at Göttingen as early as 1856. See Dean (2009) for more details on Dedekind's version of the Fundamental Theorem.
- The modern statement of the theorem for abstract fields (i.e., not just subfields of $\mathbb{C}$) is generally attributed to Emil Artin in his Notre Dame lectures (1942).
- So what did Galois actually do? Recall from the introduction that his main concern was the solvability of polynomial equations with rational or integer coefficients. Next week we will return to this subject and we will apply the Fundamental Theorem to finally prove Galois' Solvability Theorem (in modern language).

///

For now let me show you a "toy example" of the Fundamental Theorem.

Example: Galois Theory of Finite Fields. We have seen that any finite field has the form $\mathbb{E} = \mathbb{F}_{p^k}$ where $\mathbb{F}_{p^k}$ is the splitting field of the polynomial $x^{p^k} - x \in \mathbb{F}_p[x]$. It follows that $\mathbb{F}_{p^k} \supseteq \mathbb{F}_p$ is a Galois extension of perfect fields. Furthermore, you will show on the homework that the Galois group is **cyclic** and generated by the **Frobenius automorphism**:

$$\begin{array}{rccc} \varphi: & \mathbb{F}_{p^k} & \to & \mathbb{F}_{p^k} \\ & \alpha & \mapsto & \alpha^p. \end{array}$$

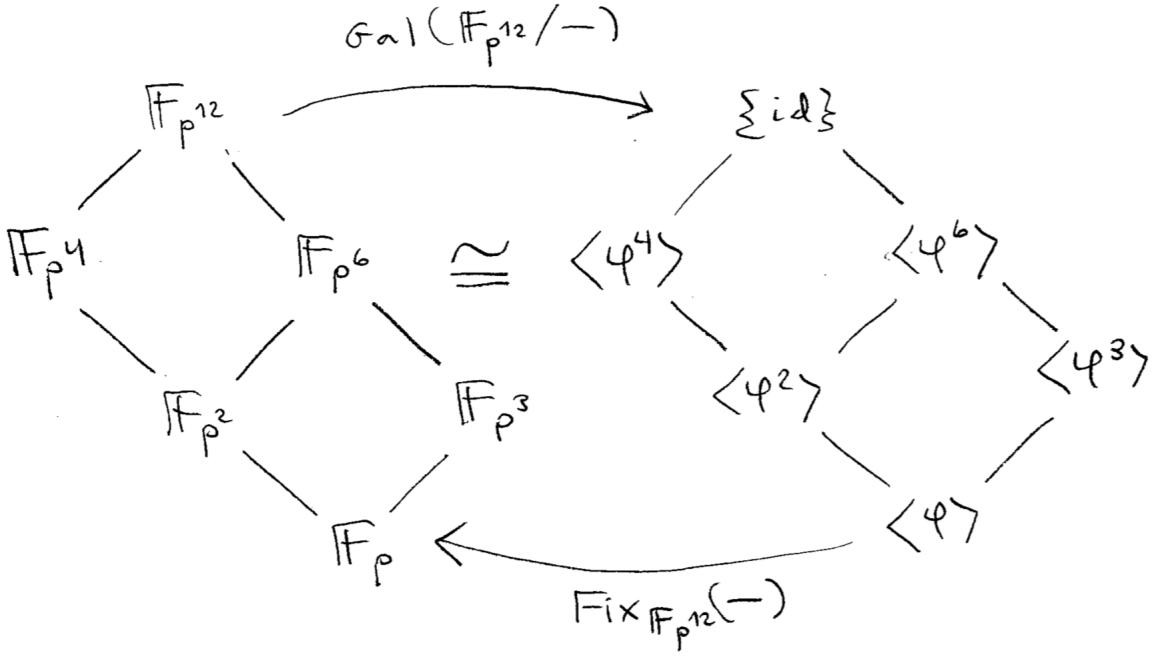
In other words, you will show that

$$\mathrm{Gal}(\mathbb{F}_{p^k}/\mathbb{F}_p) = \langle\varphi\rangle = \{\mathrm{id}, \varphi, \varphi^2, \ldots, \varphi^{k-1}\}.$$

Then it follows from the Fundamental Theorems of Galois Theory and Cyclic Groups (which we proved early last semester) that the lattice of intermediate fields $\mathscr{L}(\mathbb{F}_{p^k}, \mathbb{F}_p)$ is isomorphic to the lattice of positive divisors $d|k$ of the integer k:

$$\begin{array}{ccccc} \mathscr{L}(\mathbb{F}_{p^k}, \mathbb{F}_p) & \cong & \mathscr{L}\langle\varphi\rangle^{\mathrm{op}} & \cong & \mathrm{Div}(k) \\ \mathbb{F}_{p^d} & \leftrightarrow & \langle\varphi^d\rangle & \leftrightarrow & d. \end{array}$$

Here is a picture for $k = 12$:



///

In hindsight, we see that the theory of finite fields is roughly as complicated as the theory of cyclic groups. That is, not very. Galois studied finite fields in his paper *On the Theory of Numbers*, and I am sure that this directly inspired his later work on extensions of $\mathbb{Q}$. The passage from finite fields to fields of characteristic zero is analogous to the passage from **cyclic groups** to **all finite groups**.[76] We should not expect it to be easy.

Problem Set 11

1. Two Small Issues. Let $\mathbb{E} \supseteq \mathbb{F}$ be any field extension.

(a) If $f(x) \in \mathbb{F}[x]$ splits in $\mathbb{E}[x]$ and $g(x)|f(x)$ in $\mathbb{F}[x]$, prove that $g(x)$ also splits in $\mathbb{E}[x]$.

(b) Let $p(x), q(x) \in \mathbb{F}[x]$ be irreducible polynomials that are not associate. Prove that $p(x)$ and $q(x)$ have no common root in $\mathbb{E}$. [Hint: Since $p(x), q(x)$ are coprime in $\mathbb{F}[x]$ we have $p(x)f(x) + q(x)g(x) = 1$ for some $f(x), g(x) \in \mathbb{F}[x]$.]

(a) Suppose that $f(x) = g(x)h(x)$ for some $h(x) \in \mathbb{F}[x]$ and suppose that $f(x)$ splits over the

[76] The major open problem of Galois theory is to establish whether or not every finite group G can be expressed in the form $G = \mathrm{Gal}(\mathbb{E}/\mathbb{Q})$. This is called the "inverse Galois problem." Shafarevich (1954) proved that every **solvable** group can be expressed in this way.

field extension $\mathbb{E} \supseteq \mathbb{F}$. That is, suppose that there exist $\alpha_1, \ldots, \alpha_n \in \mathbb{E}$ such that

$$g(x)h(x) = f(x) = (x - \alpha_1) \cdots (x - \alpha_n) \text{ in the ring } \mathbb{E}[x].$$

Now let $p(x)|g(x)$ be any monic irreducible factor in the ring $\mathbb{E}[x]$. Since $\mathbb{E}[x]$ is a UFD and since $p(x)$ divides the product $(x - \alpha_1) \cdots (x - \alpha_n)$, we must have $p(x)|(x - \alpha_i)$ and hence $p(x) = x - \alpha_i$ for some i. It follows that $g(x)$ splits in $\mathbb{E}[x]$.

(b) Let $p(x), q(x) \in \mathbb{F}[x]$ be irreducible polynomials that are not associate, i.e., such that one is not a scalar multiple of the other. Then since $\mathbb{F}[x]$ is a PID we can write $p(x)f(x)+q(x)g(x) = 1$ for some polynomials $f(x), g(x) \in \mathbb{F}[x]$. Now let $\mathbb{E} \supseteq \mathbb{F}$ be any field extension and assume for contradiction that $\alpha \in \mathbb{E}$ is a common root of $p(x)$ and $q(x)$. From Descartes' Theorem it follows that $(x-\alpha)|p(x)$ and $(x-\alpha)|q(x)$ in $\mathbb{E}[x]$. Since the equation $p(x)f(x)+q(x)g(x) = 1$ also holds in $\mathbb{E}[x]$, we conclude that $x - \alpha$ divides 1. Contradiction.

2. The Galois Group of a Finite Field. Let $\mathbb{E}$ be a field of size p^k and recall that the *Frobenius endomorphism* $\varphi : \mathbb{E} \to \mathbb{E}$ is defined by $\varphi(\alpha) = \alpha^p$.

(a) Use the fact that $\mathbb{E}$ is finite to prove that $\varphi \in \mathrm{Gal}(\mathbb{E}/\mathbb{F}_p)$.

(b) Prove that φ has order k as an element of $\mathrm{Gal}(\mathbb{E}/\mathbb{F}_p)$.

(c) Conclude that $\mathrm{Gal}(\mathbb{E}/\mathbb{F}_p) = \langle\varphi\rangle$ is cyclic of size k.

(a) We know from Fermat's Little Theorem that the Frobenius endomorphism $\varphi : \mathbb{E} \to \mathbb{E}$ fixes the prime subfield $\mathbb{F}_p \subseteq \mathbb{E}$. Furthermore, since $\ker\varphi \subsetneq \mathbb{E}$ is a proper ideal and since $\mathbb{E}$ is a field we must have $\ker\varphi = \{0\}$, hence φ is injective. Finally, since $\varphi : \mathbb{E} \to \mathbb{E}$ is an injective function from a **finite set** to itself, we conclude that φ is also surjective,[77] hence $\varphi \in \mathrm{Gal}(\mathbb{E}/\mathbb{F}_p)$.

(b) From Lagrange's Theorem applied to the group of units $(\mathbb{E}^\times, \times, 1)$ we know that $\alpha^{p^k-1} = 1$ for all $\alpha \in \mathbb{E}^\times$ and hence $\alpha^{p^k} = \alpha$ for all $\alpha \in \mathbb{E}$. In other words, we have $\varphi^k = \mathrm{id}$. Now assume for contradiction that we have $\varphi^\ell = \mathrm{id}$ for some $1 \leq \ell < k$. This would imply that $\alpha^{p^\ell} = \alpha$ and hence $\alpha^{p^\ell} - \alpha = 0$ for all $\alpha \in \mathbb{E}$. But then the polynomial $x^{p^\ell} - x \in \mathbb{F}_p[x]$ of degree p^ℓ has $p^k > p^\ell$ distinct roots in the field extension $\mathbb{E}$, which contradicts Descartes' Factor Theorem.

(c) We have shown that the cyclic subgroup $\langle\varphi\rangle \subseteq \mathrm{Gal}(\mathbb{E}/\mathbb{F}_p)$ generated by the Frobenius automorphism φ has size k. On the other hand, since $\mathbb{E}/\mathbb{F}_p$ is a Galois extension (indeed, it is a splitting field for the polynomial $x^{p_k} - x$) we know from the Characterization Theorem for Galois Extensions of Perfect Fields[78] that

$$\#\mathrm{Gal}(\mathbb{E}/\mathbb{F}_p) = [\mathbb{E}/\mathbb{F}_p] = k,$$

[77] Actually we don't even need to use the fact that $\mathbb{E}$ is a finite set. It suffices to know that $\varphi : \mathbb{E}/\mathbb{F}_p \to \mathbb{E}/\mathbb{F}_p$ is an injective endomorphism of a finite-dimensional vector space. Then from the Rank-Nullity Theorem we conclude that φ is also surjective. There are levels of "finiteness" in algebra.

[78] It follows from 2(a) that $\mathbb{E}$ and $\mathbb{F}_p$ are perfect fields.

and hence $\langle\varphi\rangle = \mathrm{Gal}(\mathbb{E}/\mathbb{F}_p)$.

3. Repeated Roots, Part II We say that a polynomial $f(x) \in \mathbb{F}[x]$ is *inseparable* if it has a repeated root in some field extension. Otherwise we say that $f(x)$ is *separable*. Prove that

$$f(x) \text{ is separable} \quad \Longleftrightarrow \quad \gcd(f, Df) = 1.$$

Proof. Consider a polynomial $f(x) \in \mathbb{F}[x]$ and let $\mathbb{E} \supseteq \mathbb{F}$ be any field extension. On a previous homework you showed that $\alpha \in \mathbb{E}$ is a multiple root of $f(x)$ if and only if $f(\alpha) = 0$ and $Df(\alpha) = 0$. By Descartes' Theorem this means that $(x - \alpha)$ is a common divisor of $f(x)$ and $Df(x)$ in the ring $\mathbb{E}[x]$.

To prove the equivalence, first assume that $f(x) \in \mathbb{F}[x]$ is *inseparable*, i.e., has some multiple root $\alpha \in \mathbb{E} \supseteq \mathbb{F}$. From the above remark this implies that $(x - \alpha)$ is a common factor of $f(x)$ and $Df(x)$ in $\mathbb{E}[x]$. Now assume for contradiction that $f(x)$ and $Df(x)$ are coprime in $\mathbb{F}[x]$. Since $\mathbb{F}[x]$ is a PID this means that $f(x)a(x)+Df(x)b(x) = 1$ for some $a(x), b(x) \in \mathbb{F}[x]$. Finally, since this equation also holds in $\mathbb{E}[x]$ we conclude that $(x-\alpha)$ divides 1, contradiction.

Conversely, suppose that $\gcd(f, Df) = 1$ in $\mathbb{F}[x]$. Again, this means that $f(x)a(x)+Df(x)b(x) = 1$ for some $a(x), b(x) \in \mathbb{F}[x]$. Now suppose for contradiction that $f(x)$ has a multiple root α in some field extension $\mathbb{E} \supseteq \mathbb{F}$. As above, this implies that $(x - \alpha)$ is a common divisor of $f(x)$ and $Df(x)$ in $\mathbb{E}[x]$. Since the equation $f(x)a(x) + Df(x)b(x) = 1$ still holds in $\mathbb{E}[x]$ we conclude that $x - \alpha$ divides 1. Contradiction. □

4. Finite Fields are Separable. Let $\mathbb{E}$ be finite field of characteristic p. For all polynomials $f(x) \in \mathbb{E}[x]$ we will show that

$$f(x) \text{ is irreducible} \quad \Longrightarrow \quad f(x) \text{ is separable.}$$

(a) Let $f(x) \in \mathbb{E}[x]$ be irreducible and assume for contradiction that $f(x)$ is inseparable. Prove that the derivative $Df(x) \in \mathbb{E}[x]$ is the zero polynomial.

(b) Use part (a) to show that $f(x) = g(x^p)$ for some polynomial $g(x) \in \mathbb{E}[x]$.

(c) Finally, show that $g(x^p) = h(x)^p$ for some polynomial $h(x) \in \mathbb{E}[x]$. Contradiction. [Hint: You showed in a previous problem that the Frobenius map $\alpha \mapsto \alpha^p$ is surjective.]

(a) Let $f(x) \in \mathbb{E}[x]$ be irreducible and assume for contradiction that $f(x)$ is *inseparable*. By the previous problem this means that there exists a non-constant common divisor $d(x)|f(x)$ and $d(x)|Df(x)$. But then since $f(x)$ is irreducible we must have $d(x) = \lambda f(x)$ for some $\lambda \in \mathbb{E}$, and hence $f(x)|Df(x)$. If $Df(x) \neq 0$ then since $\deg(Df) < \deg(f)$ this is a contradiction.

(b) Thus we must have $Df(x) = 0$. To be specific, suppose that $f(x) = \sum a_k x^k$ so that $Df(x) = \sum k \cdot a_k x^k = 0$. To say that this polynomial is zero means that each coefficient $k \cdot a_k$

is zero. There are two ways this can happen: (1) If $a_k = 0$ then we also have $k \cdot a_k = 0$. (2) If $a_k \neq 0$ then $k \cdot a_k$ implies $k = 0$, which is equivalent to $p|k$ because p is the characteristic of the field. In any case, we conclude that $a_k = 0$ unless $p|k$. It follows that $f(x) = \sum a_{ip}x^{ip} = \sum a_{ip}(x^p)^i = g(x^p)$ where $g(x) = \sum a_{ip}x^i$.

(c) Finally, we know from 2(a) that the Frobenius endomorphism $\varphi : \mathbb{E} \to \mathbb{E}$ defined by $\alpha \mapsto \alpha^p$ is surjective. It follows that each coefficient $a_{ip} \in \mathbb{E}$ of $g(x)$ has the form $a_{ip} = \alpha_i^p$ for some $\alpha_i \in \mathbb{E}$. Now define $h(x) = \sum \alpha_i x^i$ and extend φ to a ring homomorphism $\mathbb{E}[x] \to \mathbb{E}[x]$ by sending $x \mapsto x^p$. Then we conclude that

$$\begin{aligned} h(x)^p &= \varphi(h(x)) \\ &= \varphi(\sum \alpha_i x^i) \\ &= \sum \varphi(\alpha_i)\varphi(x)^i \\ &= \sum a_{ip}x^{ip} = g(x^p). \end{aligned}$$

Since $f(x) = g(x^p) = h(x)^p$ is not constant we observe that $h(x)$ is not constant. But then since $p \geq 2$ this contradicts the fact that $f(x)$ is irreducible. □

5. Cyclotomic Extensions are Abelian. Let $\mathbb{F} \supseteq \mathbb{Q}$ and let $\omega \in \mathbb{C}$ be a primitive n-th root of unity. That is, assume that we have

$$x^n - 1 = (x-1)(x-\omega)\cdots(x-\omega^{n-1}) \text{ in } \mathbb{C}[x].$$

(a) For all $\sigma \in \mathrm{Gal}(\mathbb{F}(\omega)/\mathbb{F})$ prove that $\sigma(\omega) = \omega^{k_\sigma}$ for some $\gcd(k_\sigma, n) = 1$.

(b) Prove that the map $\sigma \mapsto k_\sigma$ defines an injective group homomorphism

$$\mathrm{Gal}(\mathbb{F}(\omega)/\mathbb{F}) \hookrightarrow (\mathbb{Z}/n\mathbb{Z})^\times,$$

hence $\mathrm{Gal}(\mathbb{F}(\omega)/\mathbb{F})$ is abelian.

(c) Let $\Phi_n(x) \in \mathbb{Q}[x]$ be the cyclotomic polynomial. Prove that

$$\mathrm{Gal}(\mathbb{F}(\omega)/\mathbb{F}) \cong (\mathbb{Z}/n\mathbb{Z})^\times \quad \Longleftrightarrow \quad \Phi_n(x) \text{ is irreducible in } \mathbb{F}[x].$$

(a) Recall that the cyclotomic polynomial is defined as $\Phi_n(x) = \prod(x - \omega^k)$, where the product is taken over $0 \leq k < n$ such that $\gcd(k, n) = 1$. You showed on a previous homework that $\Phi_n(x) \in \mathbb{Z}[x]$. If $\sigma \in \mathrm{Gal}(\mathbb{F}(\omega)/\mathbb{F})$ then since $\mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{F}$ we must have

$$\Phi_n(\sigma(\omega)) = \sigma\left(\Phi_n(\omega)\right) = \sigma(0) = 0,$$

and hence $\sigma(\omega) = \omega^k$ for some $\gcd(k, n) = 1$. Since k depends on σ we will call it $k\sigma$.

(b) Thus we obtain a function $\mathrm{Gal}(\mathbb{F}(\omega)/\mathbb{F}) \to (\mathbb{Z}/n\mathbb{Z})^\times$ defined by $\sigma \mapsto k_\sigma$. This function is a group homomorphism because

$$(\sigma \circ \mu)(\omega) = \sigma(\mu(\omega)) = \sigma(\omega^{k_\mu}) = \sigma(\omega)^{k_\mu} = (\omega^{k_\sigma})^{k_\mu} = \omega^{k_\sigma k_\mu},$$

and hence $k_{\sigma\circ\mu} = k_\sigma k_\mu$. Furthermore, I claim that the kernel is trivial. Indeed, suppose that $k_\sigma = 1$, so that $\sigma(\omega) = \omega$. Then since every element of $\mathbb{F}(\omega)$ has the form $f(\omega)$ for some $f(x) \in \mathbb{F}[x]$ we conclude that $\sigma(f(\omega)) = f(\sigma(\omega)) = f(\omega)$, and hence $\sigma = \mathrm{id}$. Finally, since $\mathrm{Gal}(\mathbb{F}(\omega)/\mathbb{F})$ is isomorphic to the image of the map $\sigma \mapsto k_\sigma$, which is a subgroup of the abelian group $(\mathbb{Z}/n\mathbb{Z})^\times$, we conclude that $\mathrm{Gal}(\mathbb{F}(\omega)/\mathbb{F})$ is abelian.

(c) Consider the injective homomorphism $\varphi : \mathrm{Gal}(\mathbb{F}(\omega)/\mathbb{F}) \hookrightarrow (\mathbb{Z}/n\mathbb{Z})^\times$ from part (b), and let $m(x) \in \mathbb{F}[x]$ be the minimal polynomial for ω over $\mathbb{F}$, so that $m(x)|\Phi_n(x)$. Since $\mathbb{F}(\omega)/\mathbb{F}$ is a Galois extension — indeed, it is a splitting field for $\Phi_n(x)$ — we know from the Characterization Theorem that $\#\mathrm{Gal}(\mathbb{F}(\omega)/\mathbb{F}) = [\mathbb{F}(\omega)/\mathbb{F}]$.

First assume that $\mathrm{im}\,\varphi = (\mathbb{Z}/n\mathbb{Z})^\times$. Then we have

$$\deg(\Phi_n) = \phi(n) = \#\left(\mathbb{Z}/n\mathbb{Z}\right)^\times = \#\mathrm{im}\,\varphi = \#\mathrm{Gal}(\mathbb{F}(\omega)/\mathbb{F}) = [\mathbb{F}(\omega)/\mathbb{F}] = \deg(m).$$

Since $m(x)|\Phi_n(x)$ it follows that $\Phi_n(x) = m(x)$ is irreducible in $\mathbb{F}[x]$. Conversely, suppose that $\Phi_n(x)$ is irreducible in $\mathbb{F}[x]$, so that $m(x) = \Phi_n(x)$. Then we have

$$\#\mathrm{im}\,\varphi = \#\mathrm{Gal}(\mathbb{F}(\omega)/\mathbb{F}) = [\mathbb{F}(\omega)/\mathbb{F}] = \deg(m) = \deg(\Phi_n) = \phi(n) = \#\left(\mathbb{Z}/n\mathbb{Z}\right)^\times,$$

and hence $\mathrm{im}\,\varphi = (\mathbb{Z}/n\mathbb{Z})^\times$. □

[Remark: Part (c) is not necessary for the proof of Galois' Solvability Theorem. However, it is necessary if one wants to fill in all of the details of the Gauss-Wantzel Theorem.]

6. Radical Extensions are Abelian. Consider field extensions $\mathbb{E} \supseteq \mathbb{F}(\alpha) \supseteq \mathbb{F} \supseteq \mathbb{Q}$ where $\alpha^n \in \mathbb{F}$ for some $n \geq 2$ and suppose that $\mathbb{F}$ contains a primitive n-th root of unity $\omega \in \mathbb{F}$.

(a) For any $\sigma \in \mathrm{Gal}(\mathbb{E}/\mathbb{F})$ and $\beta \in \mathbb{F}(\alpha)$ prove that $\sigma(\beta) \in \mathbb{F}(\alpha)$.

(b) Prove that $\mathrm{Gal}(\mathbb{E}/\mathbb{F}(\alpha)) \subseteq \mathrm{Gal}(\mathbb{E}/\mathbb{F})$ is a normal subgroup. [Hint: Use part (a) to define a group homomorphism $\mathrm{Gal}(\mathbb{E}/\mathbb{F}) \to \mathrm{Gal}(\mathbb{F}(\alpha)/\mathbb{F})$ with kernel $\mathrm{Gal}(\mathbb{E}/\mathbb{F}(\alpha))$.]

(c) Prove that the quotient group is abelian.

(a) Suppose that $\alpha^n = a \in \mathbb{F}$. By assumption the polynomial $x^n - a \in \mathbb{F}[x]$ has n distinct roots in the field $\mathbb{F}(\alpha)$; namely, $\alpha, \omega\alpha, \ldots, \omega^{n-1}\alpha$. Now consider any automorphism $\sigma \in \mathrm{Gal}(\mathbb{E}/\mathbb{F})$ and any element $\beta \in \mathbb{F}(\alpha)$. Since α is algebraic we know that β has the form $f(\alpha)$ for some $f(x) = \sum a_i x^i \in \mathbb{F}[x]$. Furthermore, since $\sigma(\alpha)$ is also a root of $x^n - a$ we know that $\sigma(\alpha) = \omega^k\alpha$ for some k. It follows that

$$\sigma(f(\alpha)) = \sigma\left(\sum a_i\alpha^i\right) = \sum a_i\sigma(\alpha)^i = \sum a_i(\omega^k\alpha)^i = \sum (a_i\omega^{ik})\alpha_i = g(\alpha),$$

where $g(x) = \sum a_i \omega^{ik} x^i \in \mathbb{F}[x]$. We conclude that $\sigma(\beta) = g(\alpha) \in \mathbb{F}(\alpha)$.

(b) In part (a) we showed that any element $\sigma \in \mathrm{Gal}(\mathbb{E}/\mathbb{F})$ restricts to a function $\sigma : \mathbb{F}(\alpha) \to \mathbb{F}(\alpha)$. Since this function is a non-zero ring homomorphism that fixes $\mathbb{F}$ we conclude from the usual Rank-Nullity argument[79] that $\sigma : \mathbb{F}(\alpha) \to \mathbb{F}(\alpha)$ is surjective, hence $\sigma \in \mathrm{Gal}(\mathbb{F}(\alpha)/\mathbb{F})$. Then since restriction preserves composition we obtain a group homomorphism:

$$\begin{array}{rcl} \varphi : \mathrm{Gal}(\mathbb{E}/\mathbb{F}) & \to & \mathrm{Gal}(\mathbb{F}(\alpha)/\mathbb{F}) \\ \sigma & \mapsto & \sigma. \end{array}$$

The kernel of this homomorphism is $\mathrm{Gal}(\mathbb{E}/\mathbb{F}(\alpha))$ **by definition**.

(c) From the First Isomorphism Theorem we obtain

$$\frac{\mathrm{Gal}(\mathbb{E}/\mathbb{F})}{\mathrm{Gal}(\mathbb{E}/\mathbb{F}(\alpha))} = \frac{G}{\ker \varphi} \cong \mathrm{im}\, \varphi \subseteq \mathrm{Gal}(\mathbb{F}(\alpha)/\mathbb{F}).$$

We will be done if we can show that $\mathrm{Gal}(\mathbb{F}(\alpha)/\mathbb{F})$ is abelian. So consider any element $\sigma \in \mathrm{Gal}(\mathbb{F}(\alpha)/\mathbb{F})$. From part (a) we know that $\sigma(\alpha) = \omega^k \alpha$ for some k. Furthermore, since σ is determined by its action on α we might as well write $\sigma = \sigma_k$. Observe that $\sigma_k^{-1} = \sigma_{-k}$. Then for all $k, \ell \in \mathbb{Z}$ we have

$$\sigma_k \circ \sigma_\ell \circ \sigma_k^{-1} \circ \sigma_\ell^{-1}(\alpha) = \omega^k \omega^\ell \omega^{-k} \omega^{-\ell} \alpha = \alpha,$$

which implies that

$$\begin{aligned} \sigma_k \circ \sigma_\ell \circ \sigma_k^{-1} \circ \sigma_\ell^{-1} &= \mathrm{id} \\ \sigma_k \circ \sigma_\ell &= \sigma_\ell \circ \sigma_k. \end{aligned}$$

□

[Remark: In the special case that $\mathbb{E}/\mathbb{F}$ is a Galois extension we can use the Splitting Field Theorem to show that φ is also surjective, as in part **(3)** of the Fundamental Theorem.]

Epilogue: Galois' Solvability Theorem

We have come full circle. At the very beginning of this course I told you that Galois established a relationship between the "solvability of polynomial equations by radicals" and a certain structural property of abstract groups (which for this reason is called "solvability of groups"). Now we have (almost) all of the tools that we need to prove Galois' theorem.

However, let me warn you that you might find some of the details unsatisfying. To illustrate this, let's consider the case of Emil Artin, who — more than anyone — is responsible for the modern form of the subject. Here is a quote from a lecture he gave in 1950:

[79] You explored this argument on the very first homework.

Since my mathematical youth I have been under the spell of the classical theory of Galois. This charm has forced me to return to it again and again, and to try to find new ways to prove its fundamental theorems.[80]

However, in Artin's Notre Dame lectures (1942) which are considered his definitive statement on the subject, **he did not include a proof of the solvability theorem!** Instead, this theorem appears in an appendix[81] on "Applications," written by the American mathematician Arthur Milgram. It seems that in the preceding hundred years, the core of Galois theory had shifted from the "solvability theorem" to the "fundamental theorem," and that Milgram's appendix was included only as an accommodation to tradition.

It often happens in mathematics that the original motivation for a subject is discarded after we have discovered "what is really going on." But I believe that tradition still has pedagogical value.

So on to the Solvability Theorem. Let me recall the important definitions.

Definition of Solvable Groups. We say that a finite group G is *solvable* if there exists a chain of subgroups

$$G = G_0 \supseteq G_1 \supseteq G_2 \supseteq \cdots \supseteq G_r = \{\mathrm{id}\}$$

in which $G_i \trianglelefteq G_{i-1}$ is normal for all i and the quotient group G_{i-1}/G_i is abelian.

By inserting extra groups into the chain as necessary, we may assume without loss of generality that the that there does not exist any subgroup $G_{i-1} \supsetneq H \supsetneq G_i$ with $H \trianglelefteq G_{i-1}$ normal, which, by the Correspondence Theorem, is equivalent to assuming that the quotient groups G_{i-1}/G_i have no non-trivial normal subgroups. Finally, since each G_{i-1}/G_i is abelian, we may assume without loss of generality that $G_{i-1}/G_i \cong \mathbb{Z}/p_i\mathbb{Z}$ for some prime numbers $p_i \in \mathbb{Z}$. ///

Next let me recall Dedekind's algebraic version of "solvable by radicals."

Definition of Solvable Field Extensions. We say that a field extension $\mathbb{E} \supseteq \mathbb{F}$ is *solvable* if there exists a chain of field extensions

$$\mathbb{F} = \mathbb{F}_0 \subseteq \mathbb{F}_1 \subseteq \mathbb{F}_2 \subseteq \cdots \subseteq \mathbb{F}_r \supseteq \mathbb{E}$$

in which for all i we have $\mathbb{F}_i = \mathbb{F}_{i-1}(\alpha_i)$ for some element with $\alpha_i^{n_i} \in \mathbb{F}_{i-1}$. In the special case that $\mathbb{E}$ is the splitting field for a polynomial $f(x) \in \mathbb{F}[x]$ we say that that the equation $f(x) = 0$ is *solvable by radicals.* ///

Galois' Solvability Theorem. Consider a polynomial $f(x) \in \mathbb{F}[x]$ over a field $\mathbb{F}$ of characteristic zero, and let $\mathbb{E} \supseteq \mathbb{F}$ be a the splitting field. Then

$$f(x) = 0 \text{ is solvable by radicals} \quad \Longleftrightarrow \quad \mathrm{Gal}(\mathbb{E}/\mathbb{F}) \text{ is a solvable group.}$$

[80] Quoted in *The development of Galois Theory from Lagrange to Artin* (1971) by B. Melvin Kiernan.
[81] Technically, it is Part III.

///

Even though we have some rather powerful theorems at our disposal, the proof of this result is still trickier than one might guess. It is surprising how much effort is required to appreciate the the insights of an 18 year old who lived almost 200 years ago! Today we will prove that

$$f(x) = 0 \text{ is solvable by radicals} \quad \Longrightarrow \quad \mathrm{Gal}(\mathbb{E}/\mathbb{F}) \text{ is a solvable group,}$$

and for this we still need a few lemmas.

Lemma (Quotient of a Solvable Group is Solvable). Let G be a solvable group and let $\varphi : G \to G'$ be a surjective group homomorphism. Then I claim that G' is solvable. It follows that any quotient group G/N is solvable since it is the image of the projection $G \to G/N$.

Proof. By assumption we have a chain of subgroups

$$G = G_0 \supseteq G_1 \supseteq G_2 \supseteq \cdots \supseteq G_r = \{\mathrm{id}\}$$

where each quotient G_{i-1}/G_i exists and is abelian. Now apply φ to obtain a chain of subgroups

$$G' = G'_0 \supseteq G'_1 \supseteq G'_2 \supseteq \cdots \supseteq G'_r = \{\mathrm{id}\},$$

where $G'_i := \varphi[G_i]$ for all i. It remains to prove that each quotient G'_{i-1}/G'_i exists an is abelian. So consider any elements $\varphi(h) \in G'_i$ and $\varphi(g) \in G'_{i-1}$. Then since $G_i \trianglelefteq G_{i-1}$ is normal we have

$$\varphi(g)\varphi(h)\varphi(g)^{-1} = \varphi(ghg^{-1}) \in \varphi[G_i] = G'_i,$$

which implies that $G'_i \trianglelefteq G'_{i-1}$ is normal. Furthermore, I claim that the rule $\Phi(gG_i) := \varphi(g)G'_i$ defines a (surjective) group homomorphism $\Phi : G_{i-1}/G_i \to G'_{i-1}/G'_i$. Indeed, we only need to check that this function is well-defined:

$$\begin{aligned} gG_i = hG_i &\Longrightarrow h^{-1}g \in G_i \\ &\Longrightarrow \varphi(h^{-1}g) \in G'_i \\ &\Longrightarrow \varphi(h)^{-1}\varphi(g) \in G'_i \\ &\Longrightarrow \varphi(g)G'_i = \varphi(h)G'_i. \end{aligned}$$

Finally, consider any two elements $\Phi(a), \Phi(b) \in G'_{i-1}/G'_i$. Since G_{i-1}/G_i is abelian we have

$$\Phi(a)\Phi(b) = \Phi(ab) = \Phi(ba) = \Phi(b)\Phi(a),$$

and hence G'_{i-1}/G'_i is abelian. □

The next two lemmas were proved by you on the previous homework. I will state them in exactly the form that we will use them.

Abelian Lemmas. Let $\mathbb{F}$ be a field of chacteristic zero.

(1) For any root of unity $\omega \in \mathbb{C}$ the extension $\mathbb{F}(\omega)/\mathbb{F}$ is Galois with abelian Galois group.

(2) If $\mathbb{F}$ contains a primitive n-th root of unity and if $\alpha \in \mathbb{C}$ satisfies $\alpha^n \in \mathbb{F}$ then the extension $\mathbb{F}(\alpha)/\mathbb{F}$ is Galois with abelian Galois group.

Proof. Homework. □

Proof That Radical Implies Solvable. Let $\mathbb{E} \supseteq \mathbb{F} \supseteq \mathbb{Q}$ be the splitting field of a polynomial $f(x) \in \mathbb{F}[x]$ and assume that there exists a chain of radical extensions

$$\mathbb{F} = \mathbb{F}_0 \subseteq \mathbb{F}_1 \subseteq \mathbb{F}_2 \subseteq \cdots \subseteq \mathbb{F}_r \supseteq \mathbb{E}$$

where for each i we have $\mathbb{F}_i = \mathbb{F}_{i-1}(\alpha_i)$ for some element $\alpha_i \in \mathbb{F}_i$ with $\alpha_i^{n_i} = a_i \in \mathbb{F}_{i-1}$. Our goal is to construct a field $\mathbb{F}'_r \supseteq \mathbb{F}_r$ such that $\mathbb{F}'_r/\mathbb{F}$ is Galois and $\mathrm{Gal}(\mathbb{F}'_r/\mathbb{F})$ is a solvable group.[82] Then since $\mathbb{E} \supseteq \mathbb{F}$ (being a splitting field) is Galois we will conclude from the Fundamental Theorem and the Lemma on quotient groups that

$$\mathrm{Gal}(\mathbb{E}/\mathbb{F}) \cong \frac{\mathrm{Gal}(\mathbb{F}'_r/\mathbb{F})}{\mathrm{Gal}(\mathbb{F}'_r/\mathbb{E})} \quad \text{is also solvable.}$$

The difficulty has to do with the existence of enough roots of unity. I will follow Milgram's proof from the appendix of Artin's Notre Dame lectures. First let $\mathbb{F}'_0 := \mathbb{F}_0 = \mathbb{F}$. Then let $\mathbb{F}'_1$ be the splitting field of the polynomial $f_1(x) := x^{n_1} - a_1 \in \mathbb{F}[x]$ and observe that

- $\mathbb{F}'_1/\mathbb{F}$ is Galois,
- $\mathbb{F}'_1 \supseteq \mathbb{F}_1 = \mathbb{F}(\alpha_1)$,
- If ω_{n_1} is a primitive n_1-th root of unity then we observe that the splitting field contains α_1 and $\omega_{n_1}\alpha_1$, hence it also contains ω_{n_1}. Furthermore, we can get from $\mathbb{F} = \mathbb{F}'_0$ to $\mathbb{F}'_1$ by **first** adjoining ω_{n_1} and **then** adjoining α_1. From the Abelian Lemma we know that each of these extensions is Galois with abelian Galois group.

Next let $\mathbb{F}'_2$ be a splitting field for the following polynomial:

$$f_2(x) := f_1(x) \cdot \prod_{\sigma \in \mathrm{Gal}(\mathbb{F}'_1/\mathbb{F})} (x^{n_2} - \sigma(a_2)) \in \mathbb{F}[x].^{83}$$

Observe that

- $\mathbb{F}'_2/\mathbb{F}$ is Galois,
- $\mathbb{F}'_2 \supseteq \mathbb{F}_2 = \mathbb{F}(\alpha_1, \alpha_2)$,

[82] If $\mathbb{F}'_r/\mathbb{F}$ is Galois then to prove that $\mathrm{Gal}(\mathbb{F}'_r/\mathbb{F})$ is solvable it suffices by the Fundamental Theorem to show that we can get from $\mathbb{F}$ to $\mathbb{F}'_r$ by a sequence of Galois extensions, each of which has an abelian Galois group.

[83] This polynomial has coefficients in $\mathbb{F}$ because each coefficient is a symmetric polynomial in the elements $\{\sigma(a_2) : \sigma \in \mathrm{Gal}(\mathbb{F}'_1/\mathbb{F})\}$. But the elements of this set are permuted by the action of $\mathrm{Gal}(\mathbb{F}'_1/\mathbb{F})$, hence every coefficient is in the fixed field. Finally, since $\mathbb{F}'_1$ is a Galois extension we know that the fixed field is $\mathbb{F}$.

- Again we note that the splitting field contains a primitive n_2-th root of unity: $\omega_{n_2} \in \mathbb{F}'_2$. Then we can get from $\mathbb{F}'_1$ to $\mathbb{F}'_2$ by **first** adjoining ω_{n_2} and **then** adjoining (in any order) a primitive n_2-th root of each element $\sigma(a_2)$. Again we know from the Abelian Lemma that each of these extensions is Galois with abelian Galois group.

One more time. Let $\mathbb{F}'_3$ be the splitting field of

$$f_3(x) := f_2(x) \cdot \prod_{\sigma \in \mathrm{Gal}(\mathbb{F}_3/\mathbb{F})} (x^{n_3} - \sigma(a_3)) \in \mathbb{F}[x].$$

For the same reasons as above we see that

- $\mathbb{F}'_3/\mathbb{F}$ is Galois,
- $\mathbb{F}'_3 \supseteq \mathbb{F}_3 = \mathbb{F}(\alpha_1, \alpha_2, \alpha_3)$,
- We can get from $\mathbb{F}'_2$ to $\mathbb{F}'_3$ by **first** adjoining a primitive root ω_{n_3} and **then** adjoining (in any order) a primitive n_3-th root of each element $\sigma(a_3) \in \mathbb{F}'_2$. We know that each of these extensions is Galois with abelian Galois group.

By continuing in this way we will obtain a field extension $\mathbb{F}'_r \supseteq \mathbb{F}_r$ such that $\mathbb{F}'_r/\mathbb{F}$ is Galois and such that we can get from $\mathbb{F}$ to $\mathbb{F}'_r$ by a sequence of Galois extensions with abelian groups, hence the Galois group $\mathrm{Gal}(\mathbb{F}'_r/\mathbb{F})$ is solvable. □

Corollary. For $n \geq 5$ the general polynomial equation of degree n is not solvable by radicals.

Proof. We will prove below that the "general polynomial equation of degree n" has Galois group S_n. We proved last semester that this group is not solvable when $n \geq 5$. □

Of course, the unsolvabiity of the quintic was not an original discovery of Galois. It is generally attributed to Abel (1824) and Ruffini (1799), so is called the Abel-Ruffini Theorem. The original contribution of Galois was to explain precisely which equations are solvable and to provide a method by which one could (in principle, but not usually in practice) solve these equations. We will prove this next time.

Today we will prove that any polynomial with a solvable Galois group is (in principle) solvable by radicals. For this we will need two more lemmas. The first is a straightforward translation of the Second Isomorphism Theorem for Groups into the language of field extensions. I will prove this at the maximum level of generality.

Lemma (The Second Isomorphism Theorem). Let $\mathbb{E} \supseteq \mathbb{F}$ be a finite-dimensional extension of perfect fields and consider any two intermediate fields $\mathbb{K}, \mathbb{L} \in \mathscr{L}(\mathbb{E}, \mathbb{F})$. If $\mathbb{L} \supseteq \mathbb{F}$ is Galois then $(\mathbb{K}\mathbb{L}) \supseteq \mathbb{K}$ and $\mathbb{L} \supseteq (\mathbb{K} \cap \mathbb{L})$ are both Galois and we have

$$\mathrm{Gal}(\mathbb{K}\mathbb{L}/\mathbb{K}) \cong \mathrm{Gal}(\mathbb{L}/\mathbb{K} \cap \mathbb{L}).$$

///

For the purpose of the proof we may assume that $\mathbb{E} \supseteq \mathbb{F}$ is a Galois extension, since otherwise we can enlarge $\mathbb{E}$ to a splitting field for some polynomial over $\mathbb{F}$. The proof will use the Fundamental Theorem of Galois Theory.

Proof. Let $G = \mathrm{Gal}(\mathbb{E}/\mathbb{F})$. Since $\mathbb{E} \supseteq \mathbb{F}$ is Galois we have the Galois correspondence:

$$\mathscr{L}(\mathbb{E}, \mathbb{F}) \cong \mathscr{L}(G)^{\mathrm{op}}.$$

Now define $H = \mathrm{Gal}(\mathbb{E}/\mathbb{K})$ and $N = \mathrm{Gal}(\mathbb{E}/\mathbb{L})$. By assumption we know that $N \trianglelefteq G$ is normal. Since any isomorphism of posets preserves[84] meet and join we also have

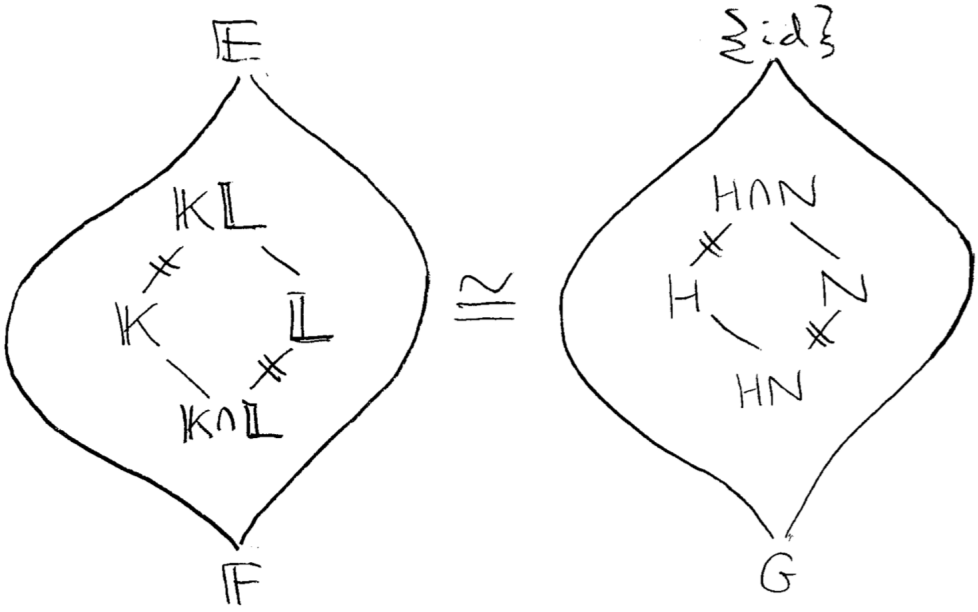
$$H \cap N = \mathrm{Gal}(\mathbb{E}/\mathbb{K}\mathbb{L}) \quad \text{and} \quad HN = \mathrm{Gal}(\mathbb{E}/\mathbb{K} \cap \mathbb{L}).$$

Then since $(H \cap N) \trianglelefteq H$ and $N \trianglelefteq HN$ are normal subgroups we see that $\mathbb{K}\mathbb{L}/\mathbb{K}$ and $\mathbb{L}/(\mathbb{K} \cap \mathbb{L})$ are Galois extensions and it follows from the Second Isomorphism Theorem that

$$\mathrm{Gal}(\mathbb{K}\mathbb{L}/\mathbb{K}) \cong \frac{\mathrm{Gal}(\mathbb{E}/\mathbb{K})}{\mathrm{Gal}(\mathbb{E}/\mathbb{K}\mathbb{L})} = \frac{H}{H \cap N} \cong \frac{HN}{N} = \frac{\mathrm{Gal}(\mathbb{E}/\mathbb{K} \cap \mathbb{L})}{\mathrm{Gal}(\mathbb{E}/\mathbb{L})} \cong \mathrm{Gal}(\mathbb{L}/\mathbb{K} \cap \mathbb{L}).$$

Here is a picture:

[84] Note that the meet and join in $\mathscr{L}(G)$ are flipped because we are using the opposite partial order.



□

The second lemma is similar in spirit to the Primitive Root Theorem and the Primitive Element Theorem. Today this result is regarded as part of "Kummer Theory," so we will call it "Kummer's Lemma."[85] However, the key idea of the proof goes all the way back to Lagrange's 1770 work on algebraic equations.

Kummer's Lemma (Existence of Lagrange Resolvents). Let $\mathbb{E} \supseteq \mathbb{F} \supseteq \mathbb{Q}$ be a Galois extension and let $[\mathbb{E}/\mathbb{F}] = p$ be prime. If $\mathbb{F}$ contains a primitive p-th root of unity $\omega \in \mathbb{F}$ then we can find some element $\alpha \in \mathbb{E} - \mathbb{F}$ such that $\alpha^p \in \mathbb{F}$ and $\mathbb{E} = \mathbb{F}(\alpha)$. We will call this element α a *Lagrange resolvent* for the extension $\mathbb{E}/\mathbb{F}$.

Proof. Let $G = \mathrm{Gal}(\mathbb{E}/\mathbb{F})$. Since $\#G = [\mathbb{E}/\mathbb{F}] = p$ is prime we know that $G = \{\mathrm{id}, \sigma, \ldots, \sigma^{p-1}\}$ is cyclic. Furthermore, we know from Dedekind's Tower Law that $\mathbb{E} \supseteq \mathbb{F}$ has no nontrivial intermediate field. Our goal is to find some $\alpha \in \mathbb{E} - \mathbb{F}$ with $\sigma(\alpha^p) = \alpha^p$. Then $\alpha \notin \mathbb{F}$ implies that $\mathbb{E} = \mathbb{F}(\alpha)$ because there are no intermediate fields, and $\sigma(\alpha^p) = \alpha^p$ implies that $\alpha^p \in \mathbb{F}$ because σ generates G and because $\mathbb{F}$ is the fixed field of G. For fun, I will give two proofs: (1) an easy existence proof, (2) a tricky constructive proof.

(1) We have assumed that there exists a primitive p-th root of unity $\omega \in \mathbb{F}$. Thus we have $x^p - 1 = \prod_{k=0}^{p-1}(x - \omega^k)$ in $\mathbb{F}[x]$. Since powers of σ commute under composition we have

[85] Ernst Eduard Kummer developed these ideas in the 1840s as part of his work on Fermat's Last Theorem.

an "evaluation homomorphism" from $\mathbb{F}[x]$ into the endomorphism ring $\mathrm{End}(\mathbb{E}/\mathbb{F})$[86] sending $x \mapsto \sigma$ and $1 \mapsto \mathrm{id}$. Applying this to $x^p - 1$ gives $\prod_k^{p-1}(\sigma - \omega^k \cdot \mathrm{id}) = \sigma^p - \mathrm{id} = \mathbf{0}$, where the product on the left denotes composition of functions and $\mathbf{0}$ denotes the zero function. Since $\sigma \neq \mathrm{id}$ there exists some $\beta \in \mathbb{E}$ with $\sigma(\beta) \neq \beta$ and hence $(\sigma - \mathrm{id})(\beta) \neq 0$. But note that

$$(\sigma - \omega^{p-1} \cdot \mathrm{id}) \cdots (\sigma - \omega^2 \cdot \mathrm{id})(\sigma - \omega \cdot \mathrm{id})(\sigma - \mathrm{id})(\beta) = \mathbf{0}(\beta) = 0.$$

Let k be minimal such that $0 \neq (\sigma - \omega^k \cdot \mathrm{id}) \cdots (\sigma - \omega \cdot \mathrm{id})(\sigma - \mathrm{id})(\beta)$ and call this nonzero element $\alpha \in \mathbb{E}$. By definition of k we have $(\sigma - \omega^{k+1} \cdot \mathrm{id})(\alpha) = 0$ and hence $\sigma(\alpha) = \omega^{k+1}\alpha \neq \alpha$. Since $\mathbb{F}$ is the fixed field of G this implies that $\alpha \notin \mathbb{F}$. Finally, note that

$$\sigma(\alpha^p) = \sigma(\alpha)^p = (\omega^{k+1})^p \alpha^p = (\omega^p)^{k+1}\alpha^p = \alpha^p.$$

(2) **Lagrange's Proof.** Choose any $\alpha \in \mathbb{E} - \mathbb{F}$ and for each $0 \leq j \leq p-1$ define the element

$$\alpha_j := \sum_{i=0}^{p-1} \omega^{ij}\sigma^i(\alpha) \in \mathbb{E}.$$

Since $\omega \in \mathbb{F}$ we have for all $\sigma \in G$ that

$$\sigma(\alpha_j) = \sum_{i=0}^{p-1} \omega^{ij}\sigma^{i+1}(\alpha) = \omega^{-j} \sum_{i=0}^{p-1} \omega^{(i+1)j}\sigma^{i+1}(\alpha) = \omega^{-j}\alpha_j$$

which implies that $\sigma(\alpha_j^p) = \sigma(\alpha_j)^p = (\omega^{-j})^p \alpha_j^p = (\omega^p)^{-j}\alpha_j^p = \alpha_j^p$. It only remains to show that $\alpha_j \notin \mathbb{F}$ for some j. To prove this, we observe for all $1 \leq i \leq p-1$ that ω^i is a primitive p-th root of unity and hence $1 + \omega^i + (\omega^i)^2 + \cdots + (\omega^i)^{p-1} = 0$. Then we have

$$\sum_{j=0}^{p-1} \alpha_j = \sum_{i,j=0}^{p-1} \omega^{ij}\sigma^i(\alpha) = \sum_{i=0}^{p-1} \sigma^i(\alpha) \sum_{j=0}^{p-1} (\omega^i)^j = \alpha^0(\alpha) \cdot p = p\alpha \notin \mathbb{F},$$

which implies that $\alpha_j \notin \mathbb{F}$ for some j. □

Proof that Solvable Implies Radical. Let $\mathbb{E} \supseteq \mathbb{F} \supseteq \mathbb{Q}$ be the splitting field for some polynomial $f(x) \in \mathbb{F}[x]$. Suppose that the Galois group $G = \mathrm{Gal}(\mathbb{E}/\mathbb{F})$ is solvable. From the above definition this means that we have a chain of subgroups

$$G = G_0 \supseteq G_1 \supseteq G_2 \supseteq \cdots \supseteq G_r = \{\mathrm{id}\}$$

where $G_i \trianglelefteq G_{i-1}$ is normal for all i and each quotient G_i/G_{i+1} is isomorphic to $\mathbb{Z}/p_i\mathbb{Z}$ for some prime number $p_i \in \mathbb{Z}$. Since $\mathbb{E} \supseteq \mathbb{F}$ is a Galois extension we can apply the Galois correspondence to obtain a chain of subfields

$$\mathbb{F} = \mathbb{F}_0 \subseteq \mathbb{F}_1 \subseteq \mathbb{F}_2 \subseteq \cdots \subseteq \mathbb{F}_r = \mathbb{E}$$

[86] This is the **noncommutative** ring of $\mathbb{F}$-linear functions $\mathbb{E}/\mathbb{F} \to \mathbb{E}/\mathbb{F}$ under pointwise addition and composition. Note that we have a natural inclusion $\mathbb{F} \to \mathrm{End}(\mathbb{E}/\mathbb{F})$ defined by $a \mapsto a \cdot \mathrm{id}$. Since the subring generated over $\mathbb{F}$ by a single element σ is **commutative**, the evaluation at σ is still a ring homomorphism.

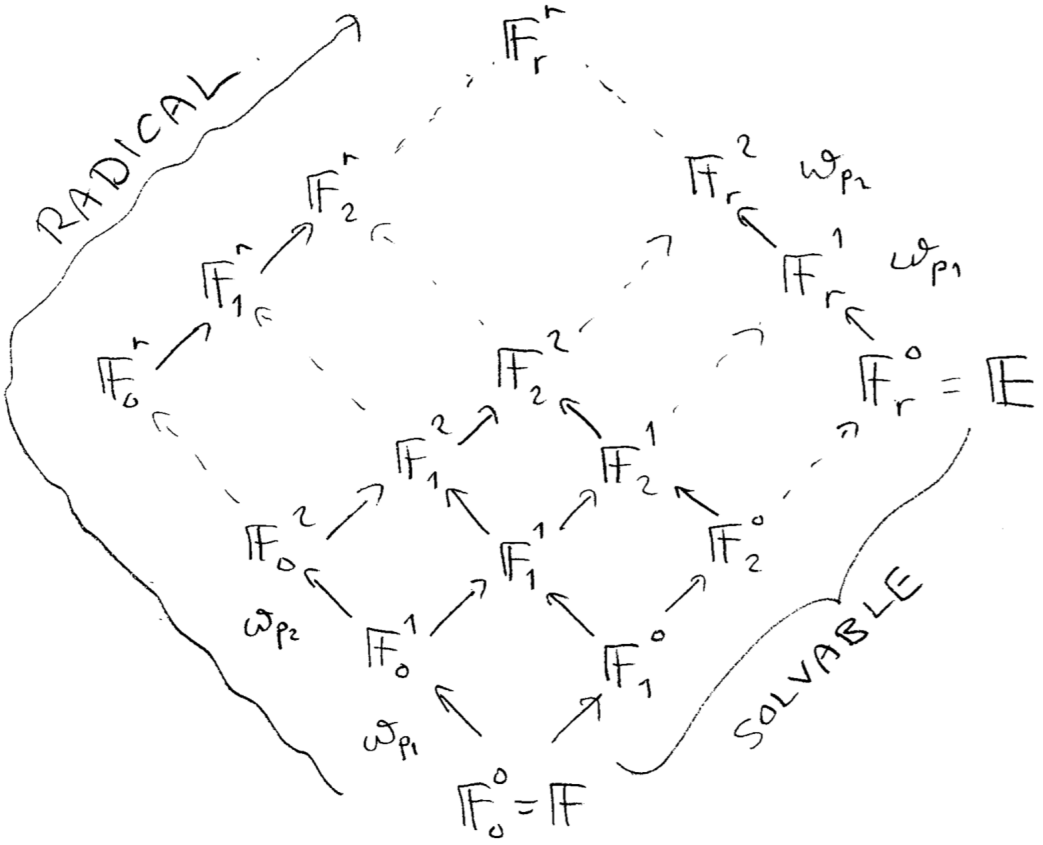
where $G_i = \mathrm{Gal}(\mathbb{E}/\mathbb{F}_i)$ for all i. Furthermore, since $G_i \trianglelefteq G_{i-1}$ is normal we know from the Fundamental Theorem that each $\mathbb{F}_i \supseteq \mathbb{F}_{i-1}$ is a Galois extension with

$$\mathrm{Gal}(\mathbb{F}_i/\mathbb{F}_{i-1}) \cong G_{i-1}/G_i \cong \mathbb{Z}/p_i\mathbb{Z}.$$

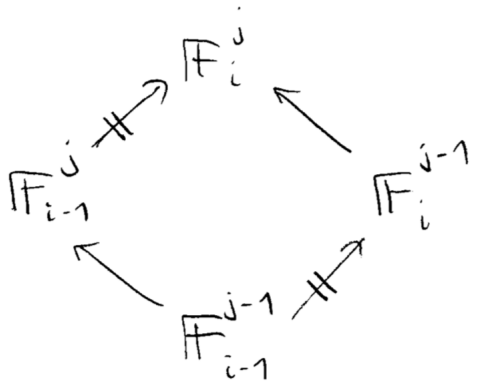
This chain of fields might not be radical, due to the fact that it might not contain enough roots of unity. To fix this situation I will describe a method to construct a chain of radical field extensions containing $\mathbb{E}$:

$$\mathbb{F} = \mathbb{F}_0^0 \subseteq \mathbb{F}_0^1 \subseteq \cdots \subseteq \mathbb{F}_0^r \subseteq \mathbb{F}_1^r \subseteq \cdots \subseteq \mathbb{F}_r^r \supseteq \mathbb{E}.$$

First let $\mathbb{F}_i^0 := \mathbb{F}_i$ for all i. Then for all j let $\mathbb{F}_i^j := \mathbb{F}_i^{j-1}(\omega_{p_j})$ where ω_{p_j} is a primitive p_j-th root of unity. Here is a picture:



By definition, each extension $\mathbb{F}_0^j \supseteq \mathbb{F}_0^{j-1}$ is either trivial or a radical extension generated by a root of unity $\omega_{p_j} \in \mathbb{F}_0^j - \mathbb{F}_0^{j-1}$ with $\omega_{p_j}^{p_j} = 1 \in \mathbb{F}_0^{j-1}$. Furthermore, for all indices i, j we observe that the following "diamond" satisfies the hypotheses of the Second Isomorphism Theorem:



Therefore we have $\mathrm{Gal}(\mathbb{F}_i^j/\mathbb{F}_{i-1}^j) \cong \mathrm{Gal}(\mathbb{F}_i^{j-1}/\mathbb{F}_{i-1}^{j-1})$ and by induction it follows that

$$\mathrm{Gal}(\mathbb{F}_i^r/\mathbb{F}_{i-1}^r) \cong \mathrm{Gal}(\mathbb{F}_i^0/\mathbb{F}_{i-1}^0) = \mathrm{Gal}(\mathbb{F}_i/\mathbb{F}_{i-1}) \cong \mathbb{Z}/p_i\mathbb{Z} \quad \text{for all } i.$$

Finally, since the field $\mathbb{F}_{i-1}^r$ contains ω_{p_i} by construction, we conclude from Kummer's Lemma that the extension $\mathbb{F}_i^r \supseteq \mathbb{F}_{i-1}^r$ is radical. □

[Remark: In fact it is sufficient to take the zig-zag chain of field extensions

$$\mathbb{F} = \mathbb{F}_0^0 \subseteq \mathbb{F}_0^1 \subseteq \mathbb{F}_1^1 \subseteq \mathbb{F}_1^2 \subseteq \mathbb{F}_2^2 \subseteq \cdots \subseteq \mathbb{F}_{r-1}^r \subseteq \mathbb{F}_r^r \supseteq \mathbb{E}$$

since each pair $\mathbb{F}_{i-1}^i \subseteq \mathbb{F}_i^i$ satisfies the hypotheses of Kummer's Lemma for the prime p_i.]

That was it. To end the course I will show you how to apply Galois' theorem to the solution of the general polynomial equations of degrees $2, 3, 4$. But first, what is a "general polynomial equation"?

Definition/Theorem (The General Polynomial Equation). Let $\{x_1, \ldots, x_n\}$ be a set of variables representing the unknown roots of a general degree n polynomial over $\mathbb{Q}$. We will denote by $\mathbb{Q}(x_1, \ldots, x_n)$ the field of fractions of the ring of polynomials $\mathbb{Q}[x_1, \ldots, x_n]$ (which is an integral domain). To be explicit, we consider the set of formal fractions

$$\mathbb{E} = \mathbb{Q}(x_1, \ldots, x_n) := \left\{ \frac{f(x_1, \ldots, x_n)}{g(x_1, \ldots, x_n)} : f, g \in \mathbb{Q}[x_1, \ldots, x_n] \text{ and } g \neq 0 \right\}.$$

with respect to the equivalence relation $f/g = f'/g' \Leftrightarrow fg' = f'g$. We know from a previous homework that this set is a field with respect to the obvious operations. Now consisider the *elementary symmetric polynomials* $e_1, e_2, \ldots, e_n \in \mathbb{Q}[x_1, x_2, \ldots, x_n]$ defined by

$$f(x) = x^n - e_1 x^{n-1} + e_2 x^{n-2} - \cdots + (-1)^n e_n = (x - x_1)(x - x_2) \cdots (x - x_n)$$

and let $\mathbb{F} := \mathbb{Q}(e_1, \ldots, e_n) \subseteq \mathbb{E}$ be the smallest subfield containing these polynomials. Then clearly $\mathbb{E} \supseteq \mathbb{F}$ is a splitting field of $f(x) \in \mathbb{F}[x]$ and hence $\mathbb{E}/\mathbb{F}$ is a finite-dimensional Galois extension. Furthermore, since any element of the group $\mathrm{Gal}(\mathbb{E}/\mathbb{F})$

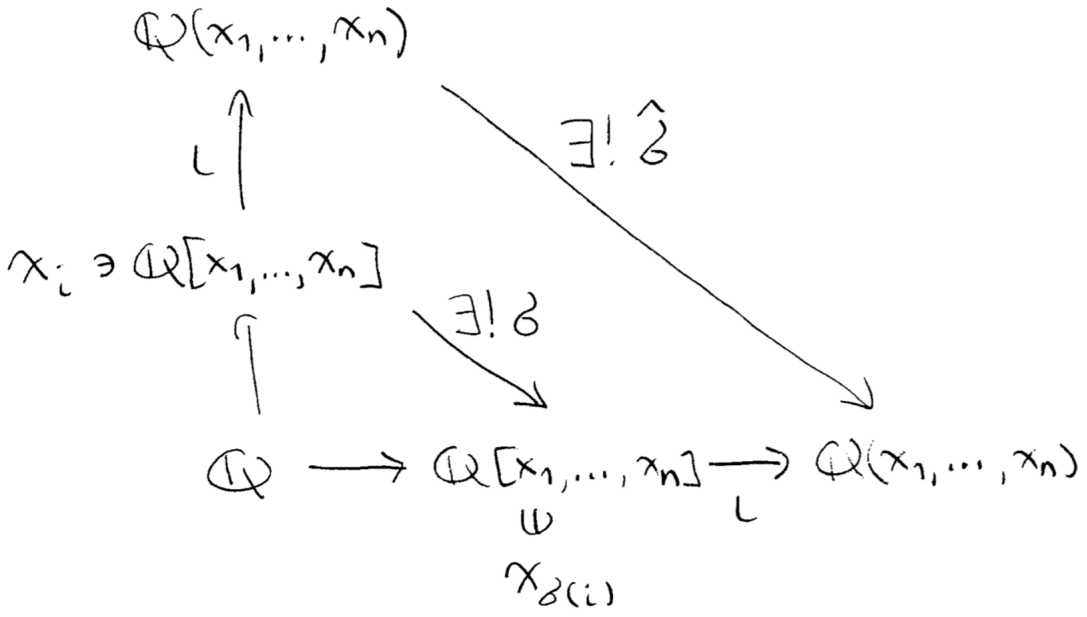
- permutes the variables $x_1, \ldots, x_n$ (i.e., the roots of $f(x)$), and
- is determined by its action on the variables $x_1, \ldots, x_n$ (i.e., the generators of $\mathbb{E}/\mathbb{F}$),

we obtain an injective group homomorphism $\mathrm{Gal}(\mathbb{E}/\mathbb{F}) \hookrightarrow S_n$ into the group of permutations of the variables. I claim that this homomorphism is also **surjective**, and hence

$$\mathrm{Gal}(\mathbb{E}/\mathbb{F}) = \mathrm{Gal}(\mathbb{Q}(x_1, \ldots, x_n)/\mathbb{Q}(e_1, \ldots, e_n)) \cong S_n.$$

Proof. We need to show that every permutation $\sigma \in S_n$ of the variables $\{x_1, \ldots, x_n\}$ extends to a field automorphism $\hat{\sigma} : \mathbb{Q}(x_1, \ldots, x_n) \to \mathbb{Q}(x_1, \ldots, x_n)$ that fixes the subfield $\mathbb{Q}(e_1, \ldots, e_n)$.

First, we will prove the existence of $\hat{\sigma}$ by messing around with universal properties. For any permutation $\sigma \in S_n$ we know from the universal property of polynomials that the inclusion $\mathbb{Q} \hookrightarrow \mathbb{Q}[x_1, \ldots, x_n]$ extends to a unique ring homomorphism $\sigma : \mathbb{Q}[x_1, \ldots, x_n] \to \mathbb{Q}[x_1, \ldots, x_n]$ fixing $\mathbb{Q}$ and sending $x_i \mapsto x_{\sigma(i)}$ for all i. Furthermore, this homomorphism is **injective**. Next, consider the inclusion $\iota : \mathbb{Q}[x_1, \ldots, x_n] \hookrightarrow \mathbb{Q}(x_1, \ldots, x_n)$ of the domain $\mathbb{Q}[x_1, \ldots, x_n]$ into its field of fractions. Then since $\iota \circ \sigma$ is an **injective** homomorphism from a domain to a field, we know from the universal property of fractions that there exists a unique extension $\hat{\sigma} : \mathbb{Q}(x_1, \ldots, x_n) \to \mathbb{Q}(x_1, \ldots, x_n)$ satisfying $\hat{\sigma} \circ \iota = \iota \circ \sigma$. Here is a picture:



We only need to show that the endomorphism $\hat{\sigma} : \mathbb{Q}(x_1, \ldots, x_n) \to \mathbb{Q}(x_1, \ldots, x_n)$ is invertible. To see this we will prove that if $\sigma, \mu \in S_n$ are inverse permutations then $\hat{\sigma}, \hat{\mu}$ are inverse

endomorphisms, hence automorphisms of $\mathbb{Q}(x_1, \ldots, x_n)$. Indeed, we have $\hat{\sigma} \circ \iota = \iota \circ \sigma$ and $\hat{\mu} \circ \iota = \iota \circ \mu$ by definition. But then

$$(\hat{\sigma} \circ \hat{\mu}) \circ \iota = \hat{\sigma} \circ (\hat{\mu} \circ \iota) = \hat{\sigma} \circ (\iota \circ \mu) = (\hat{\sigma} \circ \iota) \circ \mu = (\iota \circ \sigma) \circ \mu = \iota \circ (\sigma \circ \mu) = \iota \circ \mathrm{id}$$

implies by uniqueness that $\hat{\sigma} \circ \hat{\mu} = \hat{\mathrm{id}} = \mathrm{id}$. For the same reason we have $\hat{\mu} \circ \hat{\sigma} = \mathrm{id}$.

Next we need to show that each group element $\hat{\sigma}$ fixes the subfield $\mathbb{Q}(e_1, \ldots, e_n)$. Clearly we have $\hat{\sigma}(e_i) = e_i$ for each elementary symmetric polynomial, and hence $\hat{\sigma}(f(e_1, \ldots, e_n))$ for each polynomial $f(x_1, \ldots, x_n) \in \mathbb{Q}[x_1, \ldots, x_n]$. If the polynomials e_i were algebraic over $\mathbb{Q}$ then we would be done. Since they are not, we need one more step. We observe that

$$\mathbb{Q}(e_1, \ldots, e_n) = \left\{ \frac{f(e_1, \ldots, e_n)}{g(e_1, \ldots, e_n)} : f, g \in \mathbb{Q}[x_1, \ldots, x_n] \text{ and } g(e_1, \ldots, e_n) \neq 0 \right\}.^{87}$$

Indeed, the set on the right is a subfield of $\mathbb{Q}(x_1, \ldots, x_n)$ containing the elements $e_1, \ldots, e_n$, hence it contains the smallest such subfield. Conversely, since every element of the set on the right can be formed from the set $\mathbb{Q} \cup \{e_1, \ldots, e_n\}$ using field operations, we see that this set is contained in $\mathbb{Q}(e_1, \ldots, e_n)$. Finally, we conclude that every element of this field is fixed:

$$\hat{\sigma}\left(\frac{f(e_1, \ldots, e_n)}{g(e_1, \ldots, e_n)} \right) = \frac{\sigma(f(e_1, \ldots, e_n))}{\sigma(g(e_1, \ldots, e_n))} = \frac{f(e_1, \ldots, e_n)}{g(e_1, \ldots, e_n)}.$$

□

Remarks:

- This finally completes our proof that the general polynomial equation of degree $n \geq 5$ is not solvable by radicals.
- With a bit of extra work, one can use this result to give a non-constructive proof of Newton's Theorem. Here's a sketch. Since $\mathbb{E}/\mathbb{F}$ is a Galois extension with Galois group $S_n = \mathrm{Gal}(\mathbb{E}/\mathbb{F})$ we know from condition (GE2) of the Characterization Theorem that $\mathbb{F} = \mathrm{Fix}_{\mathbb{E}}(S_n)$. Thus for any symmetric polynomial $f(x_1, \ldots, x_n) \in \mathrm{Fix}_{\mathbb{E}}(S_n)$ we must have $f(x_1, \ldots, x_n) \in \mathbb{F} = \mathbb{Q}(e_1, \ldots, e_n)$ and hence

  $$f(x_1, \ldots, x_n) = \frac{g(e_1, \ldots, e_n)}{h(e_1, \ldots, e_n)} \text{ for some } g, h \in \mathbb{Q}[x_1, \ldots, x_n].$$

  Finally, one can argue[88] that the denominator must be constant, hence $f(x_1, \ldots, x_n)$ can be expressed as a polynomial in the elementary symmetric polynomials.

[87] In his second proof of the Fundamental Theorem of Algebra, Gauss proved that $g(x_1, \ldots, x_n) \neq 0$ implies $g(e_1, \ldots, e_n) \neq 0$. In other words, the elementary symmetric polynomials are *algebraically independent* over $\mathbb{Q}$.

[88] This is the hardest part. It involves the development of a ring-theoretic generalization of the field-theoretic concept of an "algebraic element," called an "integral element." This is more suitable for a graduate course.

///

Example: The General Quadratic. Let $\mathbb{E} = \mathbb{Q}(x_1, x_2)$ and $\mathbb{F} = \mathbb{Q}(e_1, e_2)$, so $\mathbb{E} \supseteq \mathbb{F}$ is the splitting field of the general quadratic polynomial

$$f(x) = x^2 - e_1 x + e_2 = (x - x_1)(x - x_2) \in \mathbb{F}[x].$$

Since $[\mathbb{E}/\mathbb{F}] = 2$ and since $\mathbb{F}$ contains a primitive 2-nd root of unity (namely, $-1 \in \mathbb{F}$) then we know from Kummer's Lemma that there exists an element $\gamma \in \mathbb{E} - \mathbb{F}$ with $\gamma^2 \in \mathbb{F}$ and $\mathbb{E} = \mathbb{F}(\gamma)$. Furthermore, note that $\sigma = (12)$ is a generator of $\mathrm{Gal}(\mathbb{E}/\mathbb{F}) = S_2 = \{\mathrm{id}, (12)\}$. Thus for any $\alpha \in \mathbb{E} - \mathbb{F}$ we know from Lagrange's proof that at least one of the following two elements is a resolvent:

$$\alpha_1 = \alpha + \sigma(\alpha),$$
$$\alpha_2 = \alpha - \sigma(\alpha).$$

In fact, we know that α_1 is **not** a resolvent because $\alpha(\alpha_1) = \alpha_1$ implies that α_1 is in the fixed field $\mathbb{F}$. Thus α_2 is **always** a resolvent. For simplicity, let's take $\alpha = x_1$ so that $\alpha_1 = x_1 + x_2$ and $\alpha_2 = x_1 - x_2$ is a resolvent. To be specific, we have

$$\alpha_2^2 = (x_1 - x_2)^2 = e_1^2 - 4e_2 \in \mathbb{F},$$

and then each of x_1 and x_2 is guaranteed to have the form $a + b\alpha_2 = a + b\sqrt{e_1^2 - 4e_2}$ for some $a, b \in \mathbb{F}$. With a bit of thought we find that

$$x_1 = (\alpha_1 + \alpha_2)/2 = (e_1 + \sqrt{e_1^2 - 4e_2})/2,$$
$$x_2 = (\alpha_1 - \alpha_2)/2 = (e_2 - \sqrt{e_1^2 - 4e_2})/2.$$

///

Example: The General Cubic. Let $\mathbb{E} = \mathbb{Q}(x_1, x_2, x_3)$ and $\mathbb{F} = \mathbb{Q}(e_1, e_2, e_3)$ so that $\mathbb{E} \supseteq \mathbb{F}$ is the splitting field of the general cubic polynomial

$$f(x) = x^3 - e_1 x^2 + e_2 x - e_3 = (x - x_1)(x - x_2)(x - x_3) \in \mathbb{F}[x].$$

From the above theorem we also have $\mathrm{Gal}(\mathbb{E}/\mathbb{F}) = S_3$. Recall that S_3 is a **solvable** group with composition series

$$S_3 \supsetneq A_3 \supsetneq \{\mathrm{id}\}.$$

Explicitly, $A_3 \subseteq S_3$ is the cyclic subgroup generated by the 3-cycle (123). Now apply the Galois correspondence to obtain a chain of field extensions

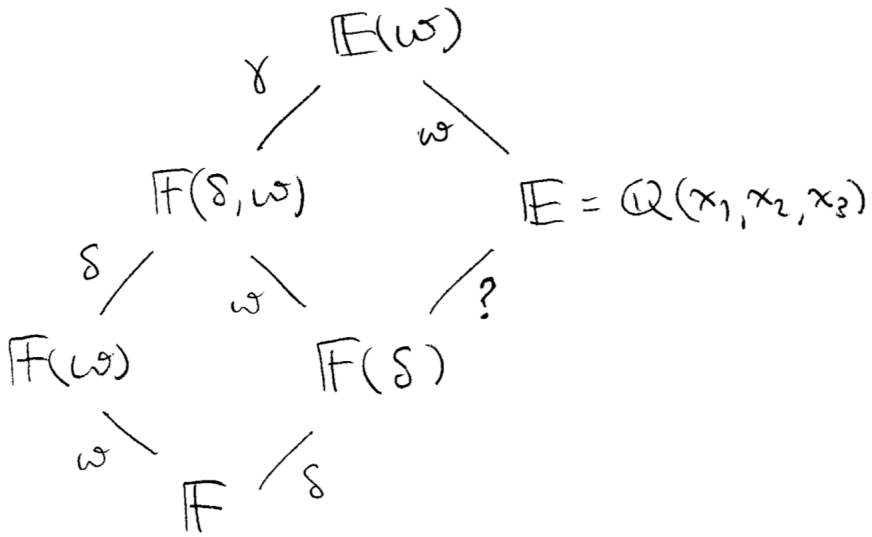
$$\mathbb{F} \subsetneq \mathbb{K} \subsetneq \mathbb{E},$$

with $[\mathbb{K}/\mathbb{F}] = 2$, $[\mathbb{E}/\mathbb{K}] = 3$ and $\mathrm{Gal}(\mathbb{E}/\mathbb{K}) = A_3$. Next I claim that $\mathbb{K} = \mathbb{F}(\delta)$, where

$$\delta := (x_1 - x_2)(x_1 - x_3)(x_2 - x_3).$$

Indeed, we have $\delta \in \mathbb{K} - \mathbb{F}$ because δ is fixed by the alternating group A_3 but not by the full symmetric group S_3. And we have $\delta^2 \in \mathbb{F}$ because δ^2 is fixed by S_3.[89]

Unfortunately, the extension $\mathbb{E} \supseteq \mathbb{K} = \mathbb{F}(\delta)$ is **not** radical. To fix this, let $\omega^2 + \omega + 1 = 0$ be a primitive third root of unity and adjoin ω to every field in the chain:



Now since $\mathbb{E}(\omega) \supseteq \mathbb{F}(\delta, \omega)$ is a Galois extension of (prime) degree 3 which contains a primitive 3-rd root of unity, Kummer's Lemma guarantees that there exists a Lagrange resolvent $\gamma \in \mathbb{E}(\omega)$ with $\mathbb{E}(\omega) = \mathbb{F}(\delta, \omega, \gamma)$ and $\gamma^3 \in \mathbb{F}(\delta, \omega)$. To be explicit, consider the generator $\sigma = (123)$ of the Galois group $A_3 = \mathrm{Gal}(\mathbb{E}(\omega)/\mathbb{F}(\delta, \omega))$. Then for any element $\alpha \in \mathbb{E}(\omega) - \mathbb{F}(\delta, \omega)$ we know that at least one of the following elements[90] is a Lagrange resolvent:

$$\alpha_2 = \alpha + \omega\sigma(\alpha) + \omega^2\sigma^2(\alpha),$$
$$\alpha_3 = \alpha + \omega^2\sigma(\alpha) + \omega\sigma^2(\alpha).$$

To simplify things, let's take $\alpha = x_1$ so the two potential Lagrange resolvents become

$$\alpha_2 = x_1 + \omega x_2 + \omega^2 x_3 \quad \text{and} \quad \alpha_3 = x_1 + \omega^2 x_2 + \omega x_3.$$

Since α_2^3 and α_3^3 are elements of $\mathbb{F}(\delta, \omega)$ and since $\mathbb{F}(\delta, \omega) \supseteq \mathbb{F}(\omega)$ has degree 2, we are guaranteed that each of α_2^3 and α_3^3 is a root of a quadratic equation with coefficients in $\mathbb{F}(\omega)$. In fact, the choice $\alpha = x_1$ is particularly nice because it turns out that α_2^3 and α_3^3 are both roots

[89] Recall that δ^2 is called the *discriminant* of the polynomial $f(x)$. On a previous homework you showed that

$$\delta^2 = e_1^2e_2^2 - 4e_2^3 - 4e_1^3e_3 + 18e_1e_2e_3 - 27e_3^2.$$

[90] Again, the element $\alpha_1 = \alpha + \sigma(\alpha) + \sigma^2(\alpha)$ is **not** a resolvent because $\sigma(\alpha_1) = \alpha_1$ implies that α_1 is in the fixed field $\mathbb{F}(\delta, \omega)$.

of a certain quadratic polynomial with coefficients in $\mathbb{F}$. The rest of the details are called "Cardano's Formula," which we discussed at the beginning of last semester. ///

Example: The General Quartic. Let $\mathbb{E} = \mathbb{Q}(x_1, x_2, x_3, x_4)$ and $\mathbb{F} = \mathbb{Q}(e_1, e_2, e_3, e_4)$ so that $\mathbb{E} \supseteq \mathbb{F}$ is the splitting field of the general quartic polynomial

$$f(x) = x^4 - e_1x^3 + e_2x^2 - e_3x + e_4 = (x - x_1)(x - x_2)(x - x_3)(x - x_4) \in \mathbb{F}[x]$$

with Galois group $S_4 = \mathrm{Gal}(\mathbb{E}/\mathbb{F})$. Since $4! = 24$ is still a small number, it is a lucky accident that the group S_4 is solvable. To be explicit, we have the following composition series:

$$S_4 \supsetneq A_4 \supsetneq V_4 \supsetneq \langle (12)(34) \rangle \supsetneq \{\mathrm{id}\}.$$

Here V_4 is the *Kleinsche Vierergruppe*:[91]

$$V_4 = \{\mathrm{id}, (12)(34), (13)(24), (14)(23)\}.$$

For the same reason as above, the fixed field of the subgroup A_4 is $\mathrm{Fix}_{\mathbb{E}}(A_4) = \mathbb{F}(\delta)$, where

$$\delta := (x_1 - x_2)(x_1 - x_3)(x_1 - x_4)(x_2 - x_3)(x_2 - x_4)(x_2 - x_4)(x_3 - x_4)$$

is "the square root of the discriminant" $\delta^2 \in \mathbb{F}$.[92] Now apply the Galois correspondence to obtain a chain of fields

$$\mathbb{F} \subsetneq \mathbb{F}(\delta) \subsetneq \mathbb{K} \subsetneq \mathbb{L} \subsetneq \mathbb{E},$$

where $\mathbb{K} = \mathrm{Fix}_{\mathbb{E}}(V_4)$ and $\mathbb{L} = \mathrm{Fix}_{\mathbb{E}}(\langle (12)(34) \rangle)$. With a bit of thought, one can show that

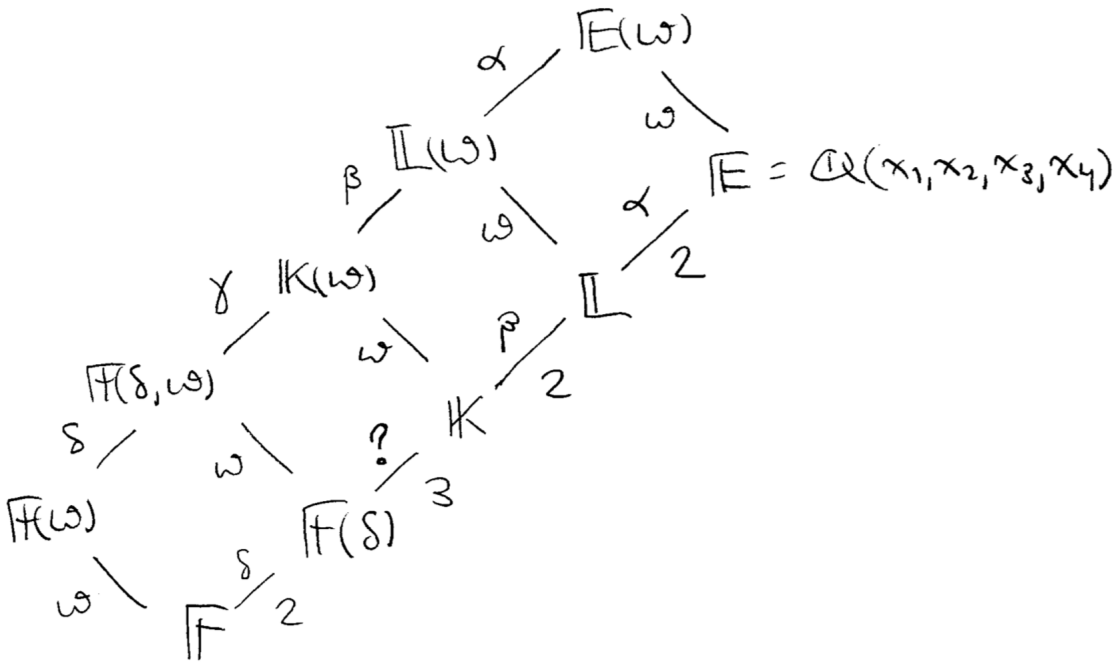
$$\mathbb{L} = \mathbb{Q}(x_1 + x_2, x_1x_2, x_3x_4, x_3 + x_4),$$
$$\mathbb{K} = \mathbb{Q}(x_1x_2 + x_3x_4, x_1x_3 + x_2x_4, x_1x_4 + x_2x_3).$$

Since $[\mathbb{E}/\mathbb{L}] = 2$ and $[\mathbb{L}/\mathbb{K}] = 2$, we are guaranteed that each of these extensions is radical. However, since $[\mathbb{K}/\mathbb{F}(\delta)] = 3$ and since $\mathbb{F}(\delta)$ does not contain a primitive 3-rd root of unity,[93] this extension is **not** radical. Thus we should adjoin a primitive root $\omega^2 + \omega + 1 = 0$ to obtain the following diagram:

[91] *Klein's four-group* is isomorphic to $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$, and hence is the smallest non-cyclic group. G. A. Miller tells us but according to Miller (*Group theory in the history of mathematics*, 1938) the term was bestowed by "various German writers" in honor of Felix Klein, who used the term "Vierergruppe" in his work. I think *Ruffini's four-group* is a better name, since Paolo Ruffini (1799) was the first to apply the group to the solvability of the quartic.

[92] Believe me, you do not want to see the explicit formula for δ^2 in terms of the coefficients e_1, e_2, e_3, e_4.

[93] for reasons of degree



The rest of the solution follows from the quadratic and cubic cases. First choose any element of $\mathbb{E}(\omega)-\mathbb{L}(\omega)$; for example x_1. Then since $\sigma = (12)(34)$ generates the group $\mathrm{Gal}(\mathbb{E}(\omega)/\mathbb{L}(\omega))$ we know that $\alpha := x_1 - \sigma(x_1) = x_1 - x_2$ is a resolvent. Next choose any element of $\mathbb{L}(\omega)-\mathbb{K}(\omega)$; for example x_1+x_2. Then since (the coset of) $\sigma = (13)(24)$ generates the group $\mathrm{Gal}(\mathbb{L}(\omega)/\mathbb{K}(\omega))$ we know that $\beta := x_1 + x_2 - \sigma(x_1+x_2) = x_1 + x_2 - x_3 - x_4$ is a resolvent. Finally, we need to choose an element of $\mathbb{K}(\omega) - \mathbb{F}(\delta,\omega)$; for example $x_1x_2 + x_3x_4$. Then since (the coset of) $\sigma = (123)$ generates the group $\mathrm{Gal}(\mathbb{K}(\omega)/\mathbb{F}(\delta,\omega) = A_4/V_4$ we know that

$$\begin{aligned}\gamma &= (x_1x_2 + x_3x_4) + \omega\sigma(x_1x_2 + x_3x_4) + \omega^2\sigma^2(x_1x_2 + x_3x_4)\\ &= (x_1x_2 + x_3x_4) + \omega(x_1x_4 + x_2x_3) + \omega^2(x_1x_3 + x_2x_3)\end{aligned}$$

is a resolvent. From this recipe it is possible to find explicit radical formulas for the roots x_1, x_2, x_3, x_4 in terms of the coefficients e_1, e_2, e_3, e_4, but what would be the point? The full solution will certainly not fit on a page.[94] ///

Galois knew that he had achieved a complete conceptual understanding of the solvability of polynomial equations. But he also knew that this understanding was mostly useless because the solutions are too complicated to write down. I will end this course by quoting Galois on this issue. The following excerpt is from the preface to a planned pair of manuscripts. Galois

[94]Here I chose only the most obvious resolvents. The history of the quartic equation is filled with more elegant choices. But even the most beautiful version of the "quartic formula" will still not fit on a page.

wrote this in prison in December 1832. He was released in April 1832, and died in May. The corrections and modifications are copied from the handwritten original:[95]

> *Long algebraic calculations were at first hardly necessary for progress in Mathematics; the very simple theorems hardly gained from being translated into the language of analysis. It is only since Euler that this briefer language has become indispensable to the new extensions which this great geometer has given to science. Since Euler calculations have become more and more necessary but more and more ~~complicated~~ difficult, at least insofar as they are applied to the most advanced objects of science. Since the beginning of this century algorithmics had attained such a degree of complication that any progress had become impossible by these means, ~~except~~ without the elegance with which new modern geometers have believed they should imprint their research, and by means of which the mind promptly and with a single glance grasps a large number of operations.*
>
> *It is clear that such vaunted elegance, and so properly claimed, has no other goal. From the well established fact that the efforts of the most advanced geometers have elegance as their object, ~~it follows that~~ ~~we have come to~~ ~~science has come to~~ ~~on this point~~ one may therefore ~~deduce~~ conclude with certainty ~~that the further the research of one advances, the more it is~~ that it becomes more and more necessary to embrace several operations ~~at a single glance~~ at once ~~in other words~~ because ~~the less~~ the mind does not have the time any more to stop ~~at each~~ at details.*
>
> *Thus I believe that the simplifications produced by elegance of calculations (intellectual simplifications, of course; there are no material ones) have their limits; I believe that the time will come when the ~~calculations~~ algebraic transformations foreseen by the speculations of analysts will find neither the time nor the place for their realisation; at which point one will have to be content with having foreseen them.*
>
> ~~*That is, according to me, the mission of future geometers; that is the path that I have entered.*~~ *I would not wish to say that there is nothing new for analysis without this rescue; but I believe that without this one day all will run out.* ~~*Embrace*~~ *Jump with both feet on calculations.* ~~*embrace*~~ *put operations into groups,* ~~*distinguish*~~ *class them according to their difficulty and not according to their form; that is, according to me, the mission of future geometers, that is the path that I have entered in this work.*

[95] Quoted from Dossier 11 in *The mathematical writings of Évartiste Galois* (2011) by Peter M. Neumann.