Math 331-3: Abstract Algebra
Northwestern University, Lecture Notes

Written by Santiago Cañez

These are notes which provide a basic summary of each lecture for Math 331-3, the third quarter of "MENU: Abstract Algebra", taught by the author at Northwestern University. The book used as a reference is the 3rd edition of *Abstract Algebra* by Dummit and Foote. Watch out for typos! Comments and suggestions are welcome.

Contents

Lecture 1: Field Extensions

This quarter we conclude our study of abstract algebra, by first studying fields to a greater extent than we did last quarter, and then by introducing the subject of *Galois theory*, where groups will make a grand reappearance. Back in the fall we introduced groups in part as a tool for studying "symmetries/permutations", and then we introduced rings and modules in the winter as a way to provide a general notion of "number", and now we will tie these concepts together more closely and finally provide an answer to a question we posed on the first day of the fall quarter: for which polynomials is it possible to express the roots of via a "nice" algebraic formula? The theory of fields we will develop more deeply at the start will give us the language we need to talk about "constructing" such roots in an "algebraic" way, and Galois theory will allow us to turn questions about these roots—or, more precisely, about the fields to which they belong—into questions about groups, which we will then be able to answer using tools from the fall.

The fundamental object of study in Galois theory is the *Galois group* of a field extension, which essentially encodes the ways in which one field can be built up out of another. The name "Galois" (pronounced gal-WAH) comes from *Évariste Galois*, a 19th century mathematician whose work laid the foundations of modern group theory. It was he who first fully realized that questions about roots of polynomials could be recast in terms of "permutations" of those roots, an idea which nowadays has been generalized and spread far beyond the topic of "roots of polynomials" alone. Galois died young in a duel, and the story is that the night before he died he wrote a letter to a friend which summarized all of his mathematical discoveries, fearing they would be lost in the event of his death. The friend then sought to get his work recognized by the mathematical community after Galois' death, and was, apparently, successful in doing so. Of course, it is hard to tell how much of this story is true and how much is apocryphal, but certainly the importance mathematical work itself is no joke.

One type of application we will see of fields and Galois theory beyond the study roots of polynomials is to questions concerning the types of geometric constructions considered by the Ancient Greeks: given only a straightedge and compass, what geometric operations can we in fact carry out? For instance, given a circle, can we construct using straightedge and compass alone a square whose area matches that of the given circle? (This is the problem of *squaring the circle.*) Or, given a cube, can we construct another cube whose volume is double that of the first? (This is the problem of *doubling the cube.*) The answer to both of these questions is "no", and the reasons for why depend on properties of field extensions. We will also consider the problem of trisecting a given angle using straightedge and compass, and finally the problem of constructing regular polygons using straightedge and compass. Equilateral triangles, squares, regular pentagons, hexagons, and many other polygons can be constructed in this way, but not all; for instance, the regular 7-gon is not constructible using straightedge and compass alone. Determining the values of n for which the regular n-gon is constructible can be nicely approached using Galois theory, as we will do later this quarter. We will hopefully also be able to say a bit about the use of Galois theory in modern number theory, which underlies the modern proof of Fermat's Last Theorem. Good stuff lies ahead!

Extension fields. Before we can talk about any such applications, we must start with the basics of field theory. We already know what a field is and some basic properties and examples, but now our focus is on the relation *between* fields. The fundamental notion is the following: given a field F, an *extension field* of F is any field $E \supseteq F$ which contains F as a subfield. We usually say simply that E is an *extension* of F, and refer to the pair $F \subseteq E$ as a *field extension.* We commonly denote such an extension by E/F (E "over" F), taking care to not interpret this as a quotient.

Using the existing multiplication on $E \subseteq F$, we can always interpret E as a module—or more precisely a vector space—over F, simply by taking the scalar multiplication of F on E to be ordinary multiplication on E. (Recall that F is contained in E.) We then define the *degree* of the extension E/F to be the dimension of E as an F-vector space, and denote it by $[E : F]$:

$$[E : F] = \text{number of linearly independent elements of } E \text{ needed to span } E \text{ over } F.$$

(It is no coincidence that we are using the same notation here for the degree $[E : F]$ that we used in the fall for group indices $[G : H]$, as the two notions will be intimately related to one another in the context of Galois theory later.) We say that the extension E/F is *finite* if it has finite degree, and *infinite* if not.

Examples. The field $\mathbb{C}$ of complex numbers is an extension of the field $\mathbb{R}$ of real numbers of degree 2. Indeed, $1, i \in \mathbb{C}$ form a basis for $\mathbb{C}$ over $\mathbb{R}$: any element of $\mathbb{C}$ is of the form $a \cdot 1 + bi$ with $a, b \in \mathbb{R}$, so 1 and i span $\mathbb{C}$ over $\mathbb{R}$, and $a \cdot 1 + bi = 0$ implies $a = b = 0$, so 1 and i are linearly independent over $\mathbb{R}$. (In fact, $\mathbb{C}$ is the only finite extension of $\mathbb{R}$, as we will see later—this is essentially the *Fundamental Theorem of Algebra.*)

The field $\mathbb{Q}(\sqrt[3]{2})$ is a degree 3 extension of $\mathbb{Q}$. (We call such a thing a *cubic* extension; an extension of degree 2 as in the previous example is called a *quadratic* extension.) This is something we actually worked out as a Warm-Up last quarter, only we didn't use the language of extensions as the time. The fact is that an element of this field explicitly looks like

$$a + b\sqrt[3]{2} + c\sqrt[3]{4} \text{ with } a, b, c \in \mathbb{Q},$$

so that $1, \sqrt[3]{2}, \sqrt[3]{4} \in \mathbb{Q}(\sqrt[3]{2})$ span $\mathbb{Q}(\sqrt[3]{2})$ over $\mathbb{Q}$, and the only such expression which equals 0 is the one which has $a = b = c = 0$, which says that $1, \sqrt[3]{2}, \sqrt[3]{4}$ are linearly independent over $\mathbb{Q}$. Hence these three elements form a basis for $\mathbb{Q}(\sqrt[3]{2})$ over $\mathbb{Q}$, so this is indeed an extension of degree $[\mathbb{Q}(\sqrt[3]{2}) : \mathbb{Q}] = 3$ as stated. (The fact that $a + b\sqrt[3]{2} + c\sqrt[3]{4} = 0$ implies $a = b = c = 0$ is not obvious, and depends on the fact that $x^3 - 2$ is an irreducible polynomial over $\mathbb{Q}$ which has $\sqrt[3]{2}$ as a root. You can go back and check the proof of this independence we gave last quarter, but we will soon generalize this result to other fields obtained by adjoining to a base field a root of some polynomial. Of course, it is also not obvious that the inverse of such an element is again of the same form—so that these elements do in fact give a *field*—-but again we will see this in more generality soon.)

For a prime p, we denote by $\mathbb{F}_p$ the field $\mathbb{Z}/p\mathbb{Z}$. (We use this notation instead of $\mathbb{Z}/p\mathbb{Z}$ since it will fit in better with the notation $\mathbb{F}_{p^n}$ we will use for other finite fields later on. You will prove on a homework soon that any finite field must in fact have prime-power order.) The field $\mathbb{F}_p(x)$ of rational functions over $\mathbb{F}_p$ is then an infinite extension of $\mathbb{F}_p$. Recall that an element of this field is a fraction of polynomials in $\mathbb{F}_p[x]$:

$$\frac{a_0 + a_1x + \cdots + a_nx^n}{b_0 + b_1x + \cdots + b_mx^m} \text{ with } a_k, b_\ell \in \mathbb{F}_p.$$

In particular, this field contains all powers of x, which give an infinite linearly independent list, which forces the dimension of $\mathbb{F}_p(x)$ over $\mathbb{F}_p$ to be infinite. We write simply $[\mathbb{F}_p(x) : \mathbb{F}] = \infty$.

The characteristic of a field. One basic invariant that will be useful to work with is that of the *characteristic* of a field. This notion was actually introduced for rings (with unity) on the first homework last quarter, but was never used again until now. Recall that the characteristic is the minimal number of times we must add 1 to itself in order to get 0; if no such number of times exists, we say the ring has characteristic zero. So, $\mathbb{Z}/n\mathbb{Z}$ has characteristic n for instance, and $\mathbb{Z}$ has characteristic zero.

That problem on the first homework last quarter asked to show that integral domains, of which fields are examples, either have characteristic 0 or *prime* characteristic. (The proof was: suppose the characteristic (if positive) factors as $n = ab$, so that $0 = n \cdot 1 = ab \cdot 1 = (a \cdot 1)(b \cdot 1)$, then use the integral domain property to get $a \cdot 1 = 0$ or $b \cdot 1 = 0$, and then minimality of n to get $n = a$ or $n = b$.) We will denote the characteristic of a field F by char F, so that char $\mathbb{C} = 0$, char $\mathbb{Q}(\sqrt[3]{2}) = 0$, and char $\mathbb{F}_p(x) = p$ for instance.

Constructing roots of polynomials. As we saw last quarter, a basic technique for constructing fields is to take the quotient of a ring (commutative with unity) by a maximal ideal. More specifically, the case we will be interested in this quarter is the quotient of a polynomial ring over a field by the ideal generated by an irreducible polynomial: $F[x]/(p(x))$. As a first use of such quotients, we show that given an irreducible polynomial $p(x)$ over a field F, there always exists an extension of F over which $p(x)$ has a root. In some sense, this is something we already saw last quarter at times, where the point is that by forcing $p(x)$ to be 0 in the quotient, x itself becomes the root for which we are looking.

But let us be a bit pedantic about the details, to make sure they are crystal clear. Since $p(x) \in F[x]$ is irreducible, $(p(x))$ is a maximal ideal of $F[x]$, so $F[x]/(p(x))$ is a field, which we will denote by E. The claim is that E is the extension of F we want. Indeed, if we know E is in fact an extension of F (perhaps not completely obvious, so we will expand on this in a bit), then the fact that our original polynomial has a root in E is easy. To be clear, $x \in E$ now no longer denotes a "variable", but is an honest element of E. To avoid abusing notation, we should perhaps denote the variable of our polynomial by X, so that $p(X)$ is the polynomial we are considering. Setting the variable X to be the element $x \in E$ gives $p(x) = 0$ in E (by the way in which E was defined), so that $x \in E$ is indeed the root we want. (This is just a simply amazing "cheat": we force our polynomial $p(x)$ to have a root by literally declaring x to be that root. Good stuff!)

The only detail left is the claim that E as defined above is an extension of F, or in other words that F is (isomorphic to) a subfield of E. This is almost obvious, since we can consider elements of F to be constant polynomials in E, but to be precise we have to know that this identification is one-to-one, so that different elements of F give different elements of E in this way. Let us phrase this in the following way: the field homomorphism $F \to E$ that sends $a \in F$ to the constant polynomial $a \in E$ has a kernel which is an ideal of F, which must thus be either 0 or all of F since these are the only ideals a field has; this kernel cannot be F since the map in question is not the zero map, so the kernel must be trivial, meaning the map $F \to E$ is injective. It is this injectivity that guarantees F is a subfield of E as desired.

Homomorphisms of fields. The fact about the homomorphism $F \to E$ above, that it is either zero or injective, is worth singling out since it is true of any homomorphism between fields (the same fact about the kernel used above still holds), and is quite useful when studying the relation between fields. The upshot is that whenever we have a nonzero map $F \to E$ between fields, we can *always* use it to think of F as a subfield of E, or equivalently of E as an extension of F. Or, said another way, if E contains no subfield which is isomorphic to F, then the only homomorphism $F \to E$ is the zero map.

The structure of quotient extensions. The types of fields $F[x]/(p(x))$ we get above are easy enough to work with, but we already saw last quarter that in many cases these "quotient" extensions can be identified with other recognizable fields. For instance, $\mathbb{R}[x]/(x^2+1)$ is simply another way of thinking about $\mathbb{C}$, and $\mathbb{Q}[x]/(x^3-2)$ is $\mathbb{Q}(\sqrt[3]{2})$. (Apply the First Isomorphism Theorem to $\mathbb{R}[x] \to \mathbb{C}$ which sends $x \mapsto i$ and to $\mathbb{Q}[x] \to \mathbb{Q}(\sqrt[3]{2})$ which sends $x \mapsto \sqrt[3]{2}$ to get the desired isomorphisms.) After all, $\mathbb{C}$ is indeed an extension (in fact the "smallest" one) of $\mathbb{R}$ over which x^2+1 has a root, and $\mathbb{Q}(\sqrt[3]{2})$ is an extension (again "smallest") of $\mathbb{Q}$ over which x^3-2 has a root, so the fact that these isomorphisms exist should not be surprising. $\mathbb{C}$ is a degree 2 extension of $\mathbb{R}$ whose elements look like $a+bi$, and $\mathbb{Q}(\sqrt[3]{2})$ is a degree 3 extension of $\mathbb{Q}$ whose elements look like $a+b\sqrt[3]{2}+c\sqrt[3]{4}$, so one question to ask is how these degrees and corresponding bases are reflected in the structure of the quotient $F[x]/(p(x))$?

The answer is again something we already saw last quarter. Explicitly, if $p(x)$ has degree n, then elements of $F[x]/(p(x))$ are of the form

$$a_0 + a_1x + \cdots + a_{n-1}x^{n-1}$$

with $a_i \in F$, since the equality $p(x) = 0$ in the quotient gives a way to replace all larger powers of x in a random polynomial. This says that the elements

$$1, x, x^2, \ldots, x^{n-1}$$

span $F[x]/(p(x))$ over F. Moreover, we claim these elements are linearly independent over F as well. Indeed, if $a_0 + a_1x + \cdots + a_{n-1}x^{n-1} = 0$ in $F[x]/(p(x))$ for some $a_i \in F$, then $a_0 + a_1x + \cdots + a_{n-1}x^{n-1} \in (p(x))$. This means that $p(x)$ divides $a_0 + a_1x + \cdots + a_{n-1}x^{n-1}$, but since $p(x)$ has degree n, this is only possible if $a_0 + a_1x + \cdots + a_{n-1}x^{n-1}$ is the zero polynomial, which means that each $a_i = 0$. Thus $1, x, \ldots, x^{n-1}$ are indeed linearly independent over F, so they give a basis for $F[x]/(p(x))$ over F. There are n elements in this basis, so we conclude that the degree of the extension is the degree of the polynomial: $[F[x]/(p(x)) : F] = \deg p(x)$.

Thus, the structure of $F[x]/(p(x))$ as a field is pretty straightforward to understand. When we identify such a field with another well-known field, then the analysis above carries over to that second field. For instance, $\mathbb{R}[x]/(x^2+1) \cong \mathbb{C}$ with $\deg(x^2+1) = 2$ reflects the quadratic nature of the extension $\mathbb{C}/\mathbb{R}$, and $\mathbb{Q}[x]/(x^3-2) \cong \mathbb{Q}(\sqrt[3]{2})$ with $\deg(x^3-2) = 3$ reflects the cubic nature of $\mathbb{Q}(\sqrt[3]{2})/\mathbb{Q}$. In this latter example, the basis $1, x, x^2$ for $\mathbb{Q}[x]/(x^3-2)$ then becomes the basis $1, \sqrt[3]{2}, \sqrt[3]{4}$ for $\mathbb{Q}(\sqrt[3]{2})$. (Observe that we can now see why the multiplicative inverse of an element $a+b\sqrt[3]{2}+c\sqrt[3]{4}$ is indeed of the same form: this comes from the fact that the multiplicative inverse of $a+bx+cx^2$ in $\mathbb{Q}[x]/(x^3-2)$ is of the same form, which is guaranteed because we already know that $\mathbb{Q}[x]/(x^3-2)$ is in fact a field.) We will explore such observations more next time, in the context of *algebraic extensions*.

Lecture 2: Algebraic Extensions

Warm-Up. We construct a field which deserves to be called "$\mathbb{F}_{49}(\sqrt[3]{3})$", that is, an extension of $\mathbb{F}_{49}$ that contains a cube root of 3. (We will explain the use of quotations marks after the construction.) As a first step, we should construct $\mathbb{F}_{49}$—a field with 49 elements—which we do as an extension of $\mathbb{F}_7$. We actually showed as an example last quarter that there exists a field of order p^2 for any prime p, where all we needed is the existence of an irreducible quadratic polynomial we can then quotient by.

To be explicit, here let us take $x^2-3 \in \mathbb{F}_7[x]$. Since this is quadratic, being irreducible is equivalent to having no root, which we can check by hand:

$$1^2 = 1 \quad 2^2 = 4 \quad 3^2 = 2 \quad 4^2 = 2 \quad 5^2 = 4 \quad 6^2 = 1.$$

Since 3 is not a square in $\mathbb{F}_7$, x^2-3 has no root, so it is irreducible and hence $\mathbb{F}_7[x]/(x^2-3)$ is a field. Elements of this field look like

$$a+bx$$

with $a,b\in\mathbb{F}_7$, which indeed gives $7^2=49$ elements in total. So, let us take $\mathbb{F}_{49}$ to be this particular field. (We stated last quarter that any two finite fields of the same order are in fact isomorphic—still to be proved—so *any* field with 49 elements is isomorphic to this particular one, meaning that we lose nothing by using this one in our computations.)

Next, let us verify that $\mathbb{F}_{49}$ as constructed does not already contain a cube root of 3. A cube root of 3 would be an element $a+bx\in\mathbb{F}_{49}$ whose third power is 3. Using the fact that $x^2=3$ in $\mathbb{F}_{49}$ as constructed, we compute:

$$\begin{aligned}(a+bx)(a+bx)(a+bx) &= [(a^2+3b^2)+2abx](a+bx)\\ &= (a^3+3ab^2+6ab^2)+(a^2b+3b^3+2a^2b)x\\ &= (a^3+2ab^2)+(3a^2b+3b^3)x.\end{aligned}$$

In order for this to equal $3\in\mathbb{F}_{49}$ we must have

$$a^3+2ab^2=3 \quad\text{and}\quad 3a^2b+3b^3=0.$$

As a side computation, let us verify that $\mathbb{F}_7$ does not contain a cube root of 3:

$$1^3=1 \quad 2^3=1 \quad 3^3=6 \quad 4^3=1 \quad 5^3=6 \quad 6^3=6.$$

If b above is zero, then the first requirement becomes $a^3=3$, and there is no $a\in\mathbb{F}_7$ which satisfies this. Thus b would have to be nonzero, in which case $3a^2b+3b^3=0$ becomes $a^2+b^2=0$. Then $b^2=-a^2$, so the first requirement is

$$3=a^3+2a(-a^2)=-a^3, \text{ or } a^3=-3=4.$$

But the same side computation above also shows that there is not such a, so we conclude that no $a+bx\in\mathbb{F}_{49}$ cubes to 3. Thus x^3-3 is irreducible over $\mathbb{F}_{49}$, so $\mathbb{F}_{49}[x]/(x^3-3)$ is the extension which deserves to be called "$\mathbb{F}_{49}(\sqrt[3]{3})$". This is a field with 7^6 elements (elements look like $\alpha+\beta x+\gamma x^2$ with $\alpha,\beta,\gamma\in\mathbb{F}_{7^2}$), so in fact, taking the uniqueness of finite fields of a given order for granted, we see that *any* field of order 7^6 contains a cube root of 3.

Now, we use quotation marks in "$\mathbb{F}_{49}(\sqrt[3]{3})$ since it does not make literal sense to take $\mathbb{F}_{49}$ and adjoin $\sqrt[3]{3}$. The point is that when we write something like $\mathbb{Q}(\sqrt[3]{2})$, we already have a predetermined notion as to what $\sqrt[3]{2}$ means: in this case the real number $\sqrt[3]{2}\in\mathbb{R}$. The field $\mathbb{Q}(\sqrt[3]{2})$ is then (by definition, if you want) the smallest subfield of the already existing field $\mathbb{R}$ that contains $\mathbb{Q}$ and $\sqrt[3]{2}$. But in the case at hand, there is no already existing field that contains both $\mathbb{F}_{49}$ and the real number $\sqrt[3]{3}$, simply because $\mathbb{F}_{49}$ is *not* realizable as a subfield of $\mathbb{R}$: $\mathbb{F}_{49}$ has characteristic 7, while $\mathbb{R}$ has characteristic 0. So, we cannot construct a literal $\mathbb{F}_{49}(\sqrt[3]{3})$ as a subfield of an existing field, meaning we only have the quotient construction available. This quotient construction does adjoin an element x to $\mathbb{F}_{49}$ which behaves like $\sqrt[3]{3}$, but it is not literally the $\sqrt[3]{3}$ we already know and love. The field "$\mathbb{F}_{49}(\sqrt[3]{3})$" we get is a degree 3 extension of $\mathbb{F}_{49}$.

A similar observation can be made about our construction of $\mathbb{F}_{49}=\mathbb{F}_7[x]/(x^2-3)$: we construct $\mathbb{F}_{49}$ (an extension of $\mathbb{F}_7$ of degree 2) by adjoining to $\mathbb{F}_7$ a "square root of 3", so that $\mathbb{F}_{49}$ is *morally* (but not literally) speaking something like "$\mathbb{F}_7(\sqrt{3})$". (Consequently, any field of order 49 contains a square root of 3.) The extension of $\mathbb{F}_{49}$ we constructed above can then be thought of as "$\mathbb{F}_7(\sqrt{3},\sqrt[3]{3})$", meaning that in a sense we "adjoin" to $\mathbb{F}_7$ both a square root of 3 and a cube root of 3. Overall this is an extension of $\mathbb{F}_7$ of degree 6. (We will see next time that degrees multiply: $\mathbb{F}_{49}(\sqrt[3]{3})$ has degree 3 over $\mathbb{F}_{49}$, and $\mathbb{F}_{49}$ has degree 2 over $\mathbb{F}_7$, so $\mathbb{F}_{49}(\sqrt[3]{3})$ has degree $3 \cdot 2$ over $\mathbb{F}_7$.)

Simple extensions as polynomial quotients. Extensions such as $\mathbb{Q}(\sqrt[3]{2})$ over $\mathbb{Q}$ and $\mathbb{R}(i) = \mathbb{C}$ over $\mathbb{R}$ which are generated by adjoining single elements are known as *simple* extensions, which is somehow the field-theoretic analog of "cyclic". The element we adjoin, i.e. the generator, is called a *primitive* element for the extension. Later we will see that finite extensions of fields of characteristic zero are always simple, but proving this will require some Galois theory.

In the case where the element we adjoin is a root of an irreducible polynomial over the base field, the resulting simple extension is straightforward to describe, as we have already seen. That is, if $p(x)$ is irreducible over F and α is a root of $p(x)$ in some extension of F, then $F(\alpha)$ (i.e. the smallest subfield of that extension that contains both F and α) is isomorphic to the quotient of $F[x]$ by the ideal generated by $p(x)$:

$$F(\alpha) \cong F[x]/(p(x)).$$

(The proof, as before, comes from the First Isomorphism Theorem applied to $F[x] \to F(\alpha)$ sending x to α. The fact that α is in the image guarantees that everything is in the image.) As a consequence, anything in $F(\alpha)$ is of the form $c_0 + c_1\alpha + c_2\alpha^2 + \cdots + c_{n-1}\alpha^{n-1}$ with $c_i \in F$, where n is the degree of $p(x)$, so that in particular the multiplicative inverse of such an element is again of the same form.

Examples. Many of the examples we have seen fall into the framework above, such as:

$$\mathbb{Q}(\sqrt[3]{2}) \cong \mathbb{Q}[x]/(x^3 - 2), \quad \mathbb{R}(i) = \mathbb{R}[x]/(x^2 + 1), \quad \mathbb{F}_{49} = \text{“}\mathbb{F}_7(\sqrt{3})\text{”} \cong \mathbb{F}_7[x]/(x^2 - 3).$$

One thing to note is that the specific root we adjoint does not matter: if α, α' are both roots of the same $p(x)$, then $F(\alpha) \cong F(\alpha')$ because both are isomorphic to the same $F[x]/(p(x))$. For instance, $\sqrt[4]{2}$ (ordinary real number) and $i\sqrt[4]{2}$ are both roots of $x^4 - 2$ over $\mathbb{Q}$, so

$$\mathbb{Q}(\sqrt[4]{2}) \cong \mathbb{Q}(i\sqrt[4]{2})$$

as subfields of $\mathbb{C}$. In some sense, the extension itself (as a standalone field) cannot distinguish between the roots adjoined without more information. Bases for these degree 4 extensions are still easy to obtain: $1, \sqrt[4]{2}, \sqrt[4]{4}, \sqrt[4]{8}$ for $\mathbb{Q}(\sqrt[4]{2})$, and $1, i\sqrt[4]{2}, -\sqrt[4]{4}, -i\sqrt[4]{8}$ for $\mathbb{Q}(i\sqrt[4]{2})$. (Of course, the same elements without the negative signs in the latter case also give a basis.)

Algebraic extensions. The extension $\mathbb{Q}(\pi)$ (i.e. the smallest subfield of $\mathbb{C}$ containing $\mathbb{Q}$ and π) of $\mathbb{Q}$ is also simple, but in this case not as straightforward to describe. Elements here look like

$$\frac{a_0 + a_1\pi + \cdots + a_n\pi^n}{b_0 + b_1\pi + \cdots + b_m\pi^m}$$

with $a_k, b_\ell \in \mathbb{Q}$, which cannot be reduced to a more compact form. In particular, this extension is infinite. The reason for why this is the case as opposed to the examples $\mathbb{Q}(\sqrt[3]{2})$ and $\mathbb{R}(i)$ comes from the fact that π is *not* the root of any polynomial with rational coefficients, which leads to the following definition.

Given a field extension E/F, we say that $\alpha \in E$ is *algebraic* over F if there exists a nonzero polynomial over F having α as a root. If not, we say that α is *transcendental* over F. For instance, π is transcendental over $\mathbb{Q}$, a fact which cannot be proven using algebraic means alone, but requires some analysis, which we will not go through. If every element of E is a algebraic over F, then we say that E is an *algebraic extension* of F.

The algebraic case is the one where we get particularly nice descriptions of simple extensions which are necessarily finite. If $\alpha \in E$ is algebraic, then it is the root of some polynomial over F, so we can consider the *monic* polynomial $m_\alpha(x)$ of *smallest* degree over F which has α as a root, which we call the *minimal* polynomial of α. (Note that the minimal polynomial depends on the base field: if we consider $\sqrt{2}$ as algebraic over $\mathbb{Q}$, its minimal polynomial is $x^2 - 2$, but if we consider it as algebraic over $\mathbb{Q}(\sqrt{2})$, then its minimal polynomial is $x - \sqrt{2}$.) The minimal polynomial defined in this way is necessarily irreducible, since if not α would have to be a root of one of its factors, which would lead to a polynomial of smaller degree having α as a root.

The upshot is that when α is algebraic over F, then

$$F(\alpha) \cong F[x]/(m_\alpha(x))$$

holds, so that the simple extension on the left is of degree n equal to that of the minimal polynomial $m_\alpha(x)$, with basis $1, \alpha, \ldots, \alpha^{n-1}$. If α is transcendental over F, then $F(\alpha)$ will be an infinite extension of F, which is a consequence of the following fact.

Finite extensions are algebraic. We claim that any finite extension of a field F is algebraic. (Thus in particular, if the extension $F(\alpha)/F$ is finite, then α must be algebraic over F.) Indeed, suppose E is finite over F of degree n, and let $\alpha \in E$. Since E is an n-dimensional vector space over F, the $n+1$ elements

$$1, \alpha, \alpha^2, \ldots, \alpha^n$$

of E must in fact be linearly dependent over F. (This type of thing is true for free modules over integral domains in general, which we stated but did not prove last quarter.) Thus there exist $c_i \in F$, at least one of which is nonzero, such that

$$c_0 + c_1\alpha + \cdots + c_n\alpha^n = 0.$$

But this means that α is a root of the nonzero polynomial $c_0 + c_1x + \cdots + c_nx^n \in F[x]$, so α is algebraic over F. Since $\alpha \in E$ was arbitrary, E is algebraic over F.

One nice thing to note here is that if $\alpha \in E$ is algebraic over F, then so are α^2, α^3, $\frac{1+\alpha}{\alpha^2}$, and indeed so is any "rational function" in α. These are all elements of the finite extension $F(\alpha)$ of F, which is necessarily algebraic, which means that all of its elements are algebraic. This is not obvious, since for instance it is not at all clear how to write down a polynomial having α^2 as a root given only one which has α as a root, and this only gets harder to do for more complicated expressions involving α. Think about this: $\sqrt[3]{2}$ is a root of $x^3 - 2$, but how to you explicitly give from this alone a polynomial which has $\sqrt[3]{4}$ as a root? Of course, the answer should be $x^3 - 4$, but the problem is in constructing this polynomial solely from $x^3 - 2$. But, there is no need to do this, since the machinery of "finite $\implies$ algebraic" guarantees that this can be done. In general, if E is an extension of F, not necessarily algebraic, then the subset of E consisting of all elements of E which *are* algebraic over F will in fact be a *field* itself.

Lecture 3: More on Extensions

Warm-Up. Suppose F is a field of characteristic not equal to 2, and let E be a degree 2 extension of F. We show that E is of the form $E = F(\sqrt{D})$ for some $D \in F$ which is not a square. (If D was a square in F, $\sqrt{D}$ would be in F and hence $F(\sqrt{D}) = F$ would be of degree 1, not 2, over F.) Pick $\alpha \in E$ which is not in F. The minimal polynomial of α is then of the form

$$x^2 + bx + c \text{ with } b, c \in F.$$

(The minimal polynomial is not of degree 1, since this would require that it be $x - \alpha$ and hence that α be in F.) But we know how to express the roots of such a polynomial explicitly using the quadratic formula:

$$\alpha = \frac{-b \pm \sqrt{b^2 - 4c}}{2}.$$

(The derivation of the usual quadratic formula over $\mathbb{Q}$, or $\mathbb{R}$, or $\mathbb{C}$ works just as well over any field, *as long as* the characteristic of the field is not 2: if it were 2, then $2 = 1 + 1 = 0$ would not be invertible in the field, and so the point at which you have to divide by 2 in the derivation would not be valid. In other words, having the 2 in the denominator above would not make sense. This is the key reason why many results about fields often have a "non-characteristic 2" assumption.)

Note here that $\sqrt{b^2 - 4c}$ makes sense as an element of E (not in F) since it is equal to

$$\sqrt{b^2 - 4c} = \pm(2\alpha + b) \in E.$$

This in particular implies that $\sqrt{b^2 - 4c} \in F(\alpha)$. But also, $\alpha = \frac{-b \pm \sqrt{b^2-4c}}{2} \in F(\sqrt{b^2 - 4c})$, so we conclude that $F(\alpha) = F(\sqrt{b^2 - 4c})$. As a vector space over F, $F(\alpha)$ is a subspace of E of dimension larger than 1 (since $\alpha \notin F$), but since E has dimension 2 over F we must then have $E = F(\alpha) = F(\sqrt{b^2 - 4c})$. Thus E is of the required form $F(\sqrt{D})$ with $D = b^2 - 4c$.

To see what happens in the characteristic 2 case, consider $\mathbb{F}_4 = \mathbb{F}_2[x]/(x^2+x+1)$ as an extension of $\mathbb{F}_2$ of degree 2. (Note $x^2 + x + 1$ is irreducible over $\mathbb{F}_2$ since it has no root.) There is no $D \in F_2$ such that $\mathbb{F}_4$ as constructed is equal to $\mathbb{F}_2(\sqrt{D})$, simply because every element of $\mathbb{F}_2$ *is* a square: $0^2 = 0, 1^2 = 1$. The proof above breaks down here precisely because the quadratic formula is not valid in characteristic 2.

Algebraic number fields. Before moving on, let us touch on the subject of *algebraic number theory* a bit, where we can clarify some observations from last quarter. Consider a quadratic extension $\mathbb{Q}(\sqrt{D})$ of $\mathbb{Q}$, where $D \in \mathbb{Z}$ is not a square. Since this extension is finite, we know it is algebraic, so every element is the root of a monic polynomial over $\mathbb{Q}$. Among these algebraic elements are those whose minimal polynomials actually have *integer* coefficients, such as $\sqrt{D}$ itself with minimal polynomial $x^2 - D$. Such elements are called *algebraic integers*, and the collection of them actually forms a subring of $\mathbb{Q}(\sqrt{D})$, called the *ring of integers* of $\mathbb{Q}(\sqrt{D})$. The name comes from the observation that in the case of $\mathbb{Q}$ viewed as a degree 1 extension of itself, so that the minimal polynomial of any rational $\frac{a}{b}$ is $x - \frac{a}{b}$, the elements whose minimal polynomials have integer coefficients are precisely the ordinary integers $\mathbb{Z}$. The ring of integers of $\mathbb{Q}(\sqrt{D})$ is then the proper generalization of "integers" to this larger field. These rings are simple to describe, and showed up in various examples last quarter: in most cases, the ring of integers $\mathbb{Q}(\sqrt{D})$ is just $\mathbb{Z}[\sqrt{D}]$—specifically this is true when $D \not\equiv 1 \bmod 4$—but interestingly when $D \equiv 1 \bmod 4$ the ring of integers of $\mathbb{Q}(\sqrt{D})$ is larger than $\mathbb{Z}[\sqrt{D}]$ alone, it is actually $\mathbb{Z}[\frac{1}{2}(1 + \sqrt{D})]$. We will not verify this here, but you can if you like try to work out the minimal polynomial of $\frac{1}{2}(1 + \sqrt{D})$ in this case and see why it has integer coefficients precisely when $D \equiv 1 \bmod 4$.

In general, an *algebraic number field* is a finite extension of $\mathbb{Q}$, and provides a more general type of "rational number". As above, the elements of this extension whose minimal polynomials have integer coefficients form the ring of algebraic integers of this number field. For instance, $\mathbb{Q}(\sqrt[3]{2})$ is a number field (of degree 3 over $\mathbb{Q}$), and its ring of algebraic integers is $\mathbb{Z}[\sqrt[3]{2}]$. (This is actually quite non-trivial to show!) One important fact about rings of integers inside number fields is that they are always *Dedekind domains*, which are a type of ring we briefly mentioned last quarter as one where unique factorization of *ideals* into prime ideals always holds. Much of modern number theory is devoted to studying such fields and rings.

Tower law. It will be important for us to study successive extensions $F \subseteq E_1 \subseteq E_2 \subseteq \cdots \subseteq E_n$ of a base field F. A first key tool for studying such extensions is known as the *Tower law* (our book does not use this name), whose name comes from drawing the given fields as a "tower", such as

$$\begin{tikzcd}
E \arrow[d, dash] \\
K \arrow[d, dash] \\
F
\end{tikzcd}$$

in the case of two successive extensions. The Tower law says that degrees multiply along such a tower: if $F \subseteq K \subseteq E$, then $[E : F] = [E : K][K : F]$. Induction then gives

$$[E_n : F] = [E_n : E_{n-1}][E_{n-2} : E_{n-3}] \cdots [E_1 : F]$$

when we have more successive extensions. Before proving this, note the similarity with a result we saw in the fall on the multiplicative property of group indices: if $A \leq H \leq G$ is a chain of groups, then $[G : A] = [G : H][H : A]$. Of course, this is no coincidence, as Galois theory will make clear.

Suppose then that $F \subseteq K \subseteq E$ are two field extensions. First note that if either E/K or K/F are of infinite degree, then so is E/F: if E/K is infinite, then an infinite collection of linearly independent elements of E over K will remain linearly independent over F since restricting the coefficients preserves independence, while if K/F is infinite then an infinite collection of linearly independent elements of K over F will also be an infinite linearly collection in E over F since $K \subseteq E$. Thus we are left with the case where E/K and K/F are both finite, in which case we show that E/F is necessarily finite and determine the degree.

Say that $\alpha_1, \dots, \alpha_n$ is a basis for E over K (so $[E : K] = n$) and $\beta_1, \dots, \beta_m$ is a basis for K over F (so $[K : F] = m$). Any element x of E is then of the form

$$x = k_1\alpha_1 + \cdots + k_n\alpha_n$$

for some $k_i \in K$ since the α_i span E over K. Each k_i here can be written as

$$k_i = f_{i1}\beta_1 + \cdots + f_{im}\beta_m$$

for some $f_{ij} \in F$ since the β_j span K over F, and making these substitutions into x above will express x as a linear combination of the products $\alpha_i\beta_j$ over F:

$$x = \sum_{i,j} (\text{coefficient in } F)\, \alpha_i\beta_j.$$

This shows that the $\alpha_i\beta_j$ span E over F. We claim that these products are also linearly independent over F, in which case we will be done: the $\alpha_i\beta_j$ will give a basis for E over F, and there are nm such elements, giving $[E : F] = nm = [E : K][K : F]$ as desired.

To check independence, suppose

$$\begin{aligned}
&f_{11}\alpha_1\beta_1 + \cdots + f_{1m}\alpha_1\beta_m \\
&+f_{21}\alpha_2\beta_1 + \cdots + f_{2m}\alpha_2\beta_m \\
&\;\vdots \\
&+f_{n1}\alpha_n\beta_1 + \cdots + f_{nm}\alpha_n\beta_m = 0
\end{aligned}$$

is a linear combination of all the $\alpha_i\beta_j$ over F. (We have written it this way so that all the terms involving α_1 are in the first two, all the terms with α_2 are in the second, and so on.) Factoring α_i out of each term in the i-th row gives

$$(f_{11}\beta_1 + \cdots + f_{1m}\beta_m)\alpha_1 + \cdots + (f_{n1}\beta_1 + \cdots + f_{nm}\beta_m)\alpha_n = 0.$$

Each coefficient (i.e. term in front of α_i) here is an element of K, so since the α_i are linearly independent over K we have

$$f_{i1}\beta_1 + \cdots + f_{im}\beta_m = 0$$

for all i. But the β_j are independent over F, so we get $f_{ij} = 0$ for all i, j, which shows that the $\alpha_i\beta_j$ are linearly independent over F as claimed. (If you go back and look at the proof of the multiplicative index property for groups, note that it is "morally" but not *literally* similar to this one, since it too involves multiplying different representatives together and using the "intermediate" object to move one step at a time.)

Example. As a quick example, we use the Tower law to show that the real number $\sqrt[3]{2}$ cannot be obtained from $\mathbb{Q}$ by the basic algebraic operations (add, substract, mulitply, divide) and (repeated) square root extractions alone. That is, $\sqrt[3]{2}$ cannot be written as something like

$$\frac{\sqrt{\sqrt{3} + 5 - \sqrt{\sqrt{7}}}}{4 - \sqrt{4 + \sqrt{5 + \sqrt{11}}}}.$$

(Not this exact number necessarily, but something *like* this.) The point is to interpret the construction of such a number in terms of (successive) field extensions, and then to use to degrees to study these extensions.

To express a number in this form requires constructing extensions of $\mathbb{Q}$ where we adjoin a square root at each step. For instance, the specific number written above, we must first introduce $\sqrt{3}$ in the denominator by extending to $\mathbb{Q}(\sqrt{3})$, then introduce $\sqrt{7}$ in the denominator by extending further to $\mathbb{Q}(\sqrt{3}, \sqrt{7})$, then introduce $\sqrt{\sqrt{7}}$ by extending to $\mathbb{Q}(\sqrt{3}, \sqrt{7}, \sqrt{\sqrt{7}})$, and so on. That is, we build up all the required terms by considering a quadratic extension of what we had before at each step. The specific number above thus lies in an extension like

$$\frac{\sqrt{\sqrt{3} + 5 - \sqrt{\sqrt{7}}}}{4 - \sqrt{4 + \sqrt{5 + \sqrt{11}}}} \in \mathbb{Q}\left(\sqrt{3}, \sqrt{7}, \sqrt{\sqrt{7}}, \sqrt{\sqrt{3} + 5 - \sqrt{\sqrt{7}}}, \sqrt{11}, \sqrt{5 + \sqrt{11}}, \sqrt{4 + \sqrt{5 + \sqrt{11}}}\right),$$

and a general number expressible in terms of repeated square roots in this way will lie in a similar extension E of $\mathbb{Q}$. Since at each step the extension is quadratic–because we adjoin a single square root—the degree of each intermediate extension is 2, so that the degree of the final extension is

$$[E : \mathbb{Q}] = 2 \cdot 2 \cdot 2 \cdots 2 = 2^n$$

for some n by the Tower law. If $\sqrt[3]{2}$ were expressible in this way, we would have $\sqrt[3]{2} \in E$ in such an E, and then the Tower law applied to $\mathbb{Q} \subseteq \mathbb{Q}(\sqrt[3]{2}) \subseteq E$ would give

$$[E : \mathbb{Q}] = [E : \mathbb{Q}(\sqrt[3]{2})][\mathbb{Q}(\sqrt[3]{2}) : \mathbb{Q}] = [E : \mathbb{Q}(\sqrt[3]{2})]3,$$

which is not possible because the left side is a power of 2. Thus $\sqrt[3]{2}$ is not expressible in this way as claimed. (This result is precisely the reason why the Ancient Greek problem of "doubling the cube" by straightedge and compass alone is impossible, but we'll clarify this later.)

Lecture 4: Splitting Fields

Warm-Up. Suppose $a \in \mathbb{Q}$ is positive and that its n^{th} root $a^{1/n}$ is not rational for all $n \geq 2$. (This guarantees that the minimal polynomial of $a^{1/n}$ over $\mathbb{Q}$ is $x^n - a$. For instance, the minimal polynomial of $4^{1/8}$ over $\mathbb{Q}$ is not $x^8 - 4$—it is $x^4 - 2$.) If $gcd(m, n) = 1$, we show that $\mathbb{Q}(a^{1/m}, a^{1/n}) = \mathbb{Q}(a^{1/mn})$. To give some context here, we will see later that any finite extension of a field of characteristic zero is in fact simple, so we know that $\mathbb{Q}(a^{1/m}, a^{1/n})$—although written with two generators here—can in fact be generated by a simple element alone, and the claim is that $a^{1/mn}$ does the job. We will consider this from two perspectives: one more brute-force, and another more "field-theoretic".

First, since $a^{1/m} = (a^{1/mn})^n$ and $a^{1/n} = (a^{1/mn})^m$, both the m^{th} and n^{th} roots of a lie in the field generated by the mn^{th} root, so

$$\mathbb{Q}(a^{1/m}, a^{1/n}) \subseteq \mathbb{Q}(a^{1/mn}).$$

Conversely, since m and n are relatively prime, there exist $p, q \in \mathbb{Z}$ such that $mp + nq = 1$, so that

$$(a^{1/m})^q (a^{1/n})^p = a^{q/m+p/n} = a^{(mp+nq)/mn} = a^{1/mn}.$$

This shows that $a^{1/mn} \in \mathbb{Q}(a^{1/m}, a^{1/n})$, so $\mathbb{Q}(a^{1/mn}) \subseteq \mathbb{Q}(a^{1/m}, a^{1/n})$ and we get equality as claimed.

Now, for a second approach, let us still begin with the observation that

$$\mathbb{Q}(a^{1/m}, a^{1/n}) \subseteq \mathbb{Q}(a^{1/mn})$$

as argued above. The point is that we can verify the opposite containment, not via a brute-force computation in terms of generators, but perhaps more elegantly by considering degrees. Note that $[\mathbb{Q}(a^{1/mn}) : \mathbb{Q}] = mn$ since the minimal polynomial of the mn^{th} root of a is $x^{mn} - a$. Thus for sure $\mathbb{Q}(a^{1/m}, a^{1/n})$ has degree no larger than mn over $\mathbb{Q}$. We claim that it is exactly mn, which will verify the equality we want. Consider the extensions

$$\mathbb{Q} \subseteq \mathbb{Q}(a^{1/m}) \subseteq \mathbb{Q}(a^{1/m}, a^{1/n}) \quad \text{and} \quad \mathbb{Q} \subseteq \mathbb{Q}(a^{1/n}) \subseteq \mathbb{Q}(a^{1/n}, a^{1/m}).$$

Since $\mathbb{Q}(a^{1/m})$ has degree m over $\mathbb{Q}$ and $\mathbb{Q}(a^{1/n})$ has degree n over $\mathbb{Q}$, we get

$$[\mathbb{Q}(a^{1/m}, a^{1/n}) : \mathbb{Q}] = (\text{something})m \quad \text{and} \quad [\mathbb{Q}(a^{1/m}, a^{1/n}) : \mathbb{Q}] = (\text{something})n$$

respectively. Thus $[\mathbb{Q}(a^{1/m}, a^{1/n}) : \mathbb{Q}]$ is divisible by both m and n, so it is divisible by their least common multiple, which is mn since $gcd(m, n) = 1$. Hence mn divides $[\mathbb{Q}(a^{1/m}, a^{1/n}) : \mathbb{Q}]$, but at the same time this degree is at most mn, so we conclude that $[\mathbb{Q}(a^{1/m}, a^{1/n}) : \mathbb{Q}] = mn$ as desired.

Composite fields. The field $\mathbb{Q}(a^{1/m}, a^{1/n})$ in the Warm-Up is by definition the smallest field (in $\mathbb{C}$) containing $\mathbb{Q}$, $a^{1/m}$, and $a^{1/n}$. Thus we can also characterize it as the smallest field containing the two fields $\mathbb{Q}(a^{1/m})$ and $\mathbb{Q}(a^{1/n})$, so it is the *composite* $\mathbb{Q}(a^{1/m})\mathbb{Q}(a^{1/n})$ of these two fields. In general, the composite K_1K_2 of two subfields $K_1, K_2 \subseteq E$ of a given field is the smallest subfield of E containing both K_1, K_2. Composites will be useful tools for constructing fields from given ones.

The result about the degree of $\mathbb{Q}(a^{1/m}, a^{1/n})$ we derived in the second approach to the Warm-Up is a special case of the following results of degrees of composite fields. Suppose K_1, K_2 are both finite extensions of F contained in a larger extension. Then

$$[K_1K_2 : F] \leq [K_1 : F][K_2 : F],$$

so that the degree of the composite is bounded by the product of the individual degrees. The book has a nice visualization of this as a tower: if $m = [K_1 : F]$ and $n = [K_2 : F]$, then we have

$$\begin{tikzcd}
 & K_1K_2 \arrow[dl, dash, "\leq n"'] \arrow[dr, dash, "\leq m"] & \\
K_1 \arrow[dr, dash, "m"'] & & K_2 \arrow[dl, dash, "n"] \\
 & F &
\end{tikzcd}$$

where the key point is that $[K_1K_2 : K_1]$ is *at most* n and $[K_1K_2 : K_2]$ at most m. The proof is as follows. Consider for instance K_1K_2 over K_1 and let $\alpha_1, \ldots, \alpha_n$ be a basis for K_2 over F, so that we can write K_2 as $K_2 = F(\alpha_1, \ldots, \alpha_n)$. Then K_1K_2 over K_1 is

$$K_1K_2 = K_1(\alpha_1, \ldots, \alpha_n).$$

The elements $\alpha_1, \ldots, \alpha_n$ still span K_1K_2 over K_1, so that $[K_1K_2 : K_1] \leq n$. Swapping the roles of K_1 and K_2 shows that $[K_1K_2 : K_2] \leq m$ by a similar argument, and we have our result. Note that we cannot guarantee equality, since it could be that $\alpha_1, \ldots, \alpha_n$, although linearly independent over F, might become linearly *dependent* over K_1, and similarly with K_1, K_2 switched. In fact, equality in $[K_1K_2 : F] \leq [K_1 : F][K_2 : F]$ holds if and only if the basis for K_2 over F remains linearly independent over K_1, or if the basis for K_1 over F remains linearly independent over K_2. This is what happens in the Warm-Up for instance.

Two more facts about extensions. Let us state two more facts about extensions which are useful to know. The book gives more thorough (perhaps too thorough) justifications for these than what we'll say. First, we have seen that $F(\alpha)$ is a finite extension of F if and only if α is algebraic over F, and more generally we can characterize *all* finite extensions of a given field: E is a finite extension of F if and only if E is generated by a finite number of algebraic elements, i.e. $E = F(\alpha_1, \ldots, \alpha_n)$ where each α_i is algebraic over F. The forward direction follows from the fact that all finite extensions are algebraic, so that the basis elements are the finitely many algebraic generators we need. The backwards direction comes from applying the case of a simple extension $F(\alpha)$ repeatedly along with the tower law: if $E = F(\alpha_1, \ldots, \alpha_n)$ with each α_i algebraic, then

$$F \subseteq F(\alpha_1) \subseteq F(\alpha_1)(\alpha_2) \subseteq \ldots \subseteq E$$

is a sequence of simple extensions, which are each finite since the generator α_i at each step is algebraic over the previous field, so that the tower law implies $[E : F]$ is finite.

The second fact we'll state is that "algebraic extensions of algebraic extensions are algebraic": if E is algebraic over K and K is algebraic over F, then E is algebraic over F. Indeed, let $\alpha \in E$. Then α is algebraic over K, so α is the root of some

$$k_0 + k_1x + \cdots + k_nx^n \text{ with } k_i \in K.$$

In particular α is actually algebraic over $F(k_1, \ldots, k_n)$, since this field already contains the coefficients needed for the minimal polynomial of α. Then $F(k_1, \ldots, k_n, \alpha) = F(k_1, \ldots, k_n)(\alpha)$ is a simple extension of $F(k_1, \ldots, k_n)$ by an algebraic generator, so it is a finite extension of $F(k_1, \ldots, k_n)$. Each k_i is algebraic over F, so $F(k_1, \ldots, k_n)$ is a finite extension of F by the first fact above, and thus

$$[F(k_1, \ldots, k_n, \alpha) : F] = [F(k_1, \ldots, k_n, \alpha) : F(k_1, \ldots, k_n)][F(k_1, \ldots, k_n) : F]$$

is finite as well. Since α belongs to a finite extension of F, it is algebraic over F, so E is algebraic over F as claimed since $\alpha \in E$ was arbitrary.

Splitting fields. Given a polynomial $p(x)$ over a field F, we say that $p(x)$ *splits* in an extension $E \supseteq F$ if it factors into a product of linear terms:

$$p(x) = (x - \alpha_1) \cdots (x - \alpha_n) \text{ in } E[x].$$

We call E the *splitting field* of $p(x)$ if $p(x)$ splits over E and if E is the smallest extension of F with this property, meaning that $p(x)$ does not split over any proper subfield of E. Splitting fields will be essential tools for studying roots of polynomials. Note that we called E here *the* splitting field of F, rather than *a* splitting field, and indeed splitting fields are in fact unique up to isomorphism: any two splitting fields of $p(x)$ are isomorphic by an isomorphism which fixes elements of the base field F. We will prove this later. We will also prove that splitting fields always exist, so that there will be no confusion when talking about the splitting field of a given polynomial.

Let us look at a few examples. The splitting field of $x^2 - 2 \in \mathbb{Q}[x]$ is $\mathbb{Q}(\sqrt{2})$, which is probably simple enough to see without much justification: $x^2 - 2$ splits as $(x - \sqrt{2})(x + \sqrt{2})$ over this extension, and it does not split in any subextension since the only proper subfield of $\mathbb{Q}(\sqrt{2})$ is $\mathbb{Q}$. This latter result is due to the fact that $\mathbb{Q}(\sqrt{2})$ has degree 2 over $\mathbb{Q}$, so that there can be no nontrivial proper intermediate extension by the tower law.

Second, the splitting field of $(x^2 - 2)(x^2 - 3)$ is $\mathbb{Q}(\sqrt{2}, \sqrt{3})$, which again is probably simple enough to believe. The only nontrivial part of this is arguing that no proper subfield will do the trick, but this can shown by verifying that $\sqrt{3}$ does not belong to $\mathbb{Q}(\sqrt{2})$, nor does $\sqrt{2}$ belong to $\mathbb{Q}(\sqrt{3})$. Now, we mentioned before that any finite extension of $\mathbb{Q}$ is simple, so $\mathbb{Q}(\sqrt{2}, \sqrt{3})$ must be simple. Indeed, we claim that $\sqrt{2} + \sqrt{3}$ is a primitive element, i.e. a generator. It is clear that $\mathbb{Q}(\sqrt{2} + \sqrt{3}) \subseteq \mathbb{Q}(\sqrt{2}, \sqrt{3})$, and for the reverse containment we can argue as follows. We have

$$(\sqrt{2} + \sqrt{3})^2 = 5 + 2\sqrt{6},$$

which implies that $\sqrt{6} \in \mathbb{Q}(\sqrt{2} + \sqrt{3})$. Then $\sqrt{6}(\sqrt{2} + \sqrt{3}) = 2\sqrt{3} + 3\sqrt{2}$ is also in this field, and hence so are

$$3(\sqrt{2} + \sqrt{3}) - (2\sqrt{3} + 3\sqrt{2}) = \sqrt{3} \text{ and } -2(\sqrt{2} + \sqrt{3}) + (2\sqrt{3} + 3\sqrt{2}) = \sqrt{2},$$

which proves the claim.

More elegantly, we can argue by degrees by finding the minimal polynomial of $\sqrt{2} + \sqrt{3}$. If we set $\alpha = \sqrt{2} + \sqrt{3}$, then

$$\alpha^2 = 5 + 2\sqrt{6}, \text{ so } \frac{1}{2}(\alpha^2 - 5) = \sqrt{6}.$$

Squaring this gives a degree 4 polynomial expression satisfied by α, so the minimal polynomial of $\sqrt{2} + \sqrt{3}$ is of degree 4 and thus $[\mathbb{Q}(\sqrt{2} + \sqrt{3}) : \mathbb{Q}] = 4$. Since composite $\mathbb{Q}(\sqrt{2}, \sqrt{3}) = \mathbb{Q}(\sqrt{2})\mathbb{Q}(\sqrt{3})$ has degree at most $2 \cdot 2 = 4$, it must be exactly 4 since the subfield $\mathbb{Q}(\sqrt{2} + \sqrt{3})$ has degree 4 over $\mathbb{Q}$, so these two fields are the same.

Roots of unity. Let us look at a final example, which we also use as an opportunity to introduce the roots of unity. The polynomial $x^3 - 2 \in \mathbb{Q}[x]$ has a root in $\mathbb{Q}(\sqrt[3]{2})$, but it does not split in this extension since it does not contain the other two roots of $x^3 - 2$. More precisely, we have

$$x^3 - 2 = (x - \sqrt[3]{2})(x^2 + x\sqrt[3]{2} + \sqrt[3]{4})$$

and $x^2 + x\sqrt[3]{2} + \sqrt[3]{4}$ is actually irreducible over $\mathbb{Q}(\sqrt[3]{2})$, as can be checked by, say, the rational root test from last quarter applied to the ring $\mathbb{Z}[\sqrt[3]{2}]$, of which $\mathbb{Q}(\sqrt[3]{2})$ is the fraction field. The details of this are a bit tedious to check, so we will skip the verification here in favor of discussing the other two roots of $x^3 - 2$ instead.

The other two roots of $x^3 - 2$ are not real, and can be easily described using the complex roots of unity. The complex n^{th} *roots of unity* are the complex numbers whose n^{th} power is 1, or in other words the roots of $x^n - 1$ over $\mathbb{C}$. These form a group under multiplication. For $n \geq 1$, set

$$\zeta_n = e^{2\pi i/n} := \cos(\tfrac{2\pi}{n}) + i\sin(\tfrac{2\pi}{n}).$$

Then $\zeta_n^n = 1$, and moreover $(\zeta_n)^k$ also has n^{th} power equal to 1. Thus

$$1,\ \zeta,\ \zeta^2, \ldots,\ \zeta^{n-1}$$

gives all the n^{th} roots of unity. Hence the group of n^{th} roots of unity is cyclic with ζ_n a generator, which we call a *primitive* n^{th} root of unity. In general, $(\zeta_n w)^n = w^n$ for any $w \in \mathbb{C}$, so the point is that the other two non-real roots of $x^3 - 2$ are $\zeta_3\sqrt[3]{2}$ and $\zeta_3^2\sqrt[3]{2}$, where ζ_3 is a primitive third root of unity. (The ζ_n we wrote down above is not the only root of unity which generates all the others, it is just the most common one to use.)

The splitting field of $x^3 - 2$ over $\mathbb{Q}$ is thus $\mathbb{Q}(\sqrt[3]{2}, \zeta_3\sqrt[3]{2}, \zeta_3^2\sqrt[3]{2})$. If we use

$$\zeta_3 = \cos(\tfrac{2\pi}{3}) + i\sin(\tfrac{2\pi}{3}) = -\tfrac{1}{2} + i\tfrac{\sqrt{3}}{2},$$

then $\zeta_3^2 = \cos(4\pi/3) + i\sin(4\pi/3) = -\frac{1}{2} - i\frac{\sqrt{3}}{2}$. Thus both ζ_3 and ζ_3^2 can be obtained by adjoining $i\sqrt{3}$ alone to $\mathbb{Q}$, so the splitting field of $x^3 - 2$ can more simply be described as $\mathbb{Q}(\sqrt[3]{2}, i\sqrt{3})$. In general, the splitting field of $x^n - a$ over $\mathbb{Q}$ will be $\mathbb{Q}(\sqrt[n]{a}, \zeta_n)$, provided that $\sqrt[n]{a}$ makes sense in $\mathbb{R}$.

Lecture 5: Algebraic Closures

Warm-Up. We determine the degree of the splitting field of $x^5 - 3$ over $\mathbb{Q}$. The roots of $x^5 - 3$ in $\mathbb{C}$ are

$$\sqrt[5]{3},\ \zeta_5\sqrt[5]{3},\ \zeta_5^2\sqrt[5]{3},\ \zeta_5^3\sqrt[5]{3},\ \zeta_5^4\sqrt[5]{3}$$

where ζ_5 is a primitive fifth root of unity, say

$$\zeta_5 = e^{2\pi i/5} = \cos(\tfrac{2\pi}{n}) + i\sin(\tfrac{2\pi}{5}).$$

The splitting field is thus $\mathbb{Q}(\sqrt[5]{3}, \zeta_5)$, which is the composite of $\mathbb{Q}(\sqrt[5]{3})$ and $\mathbb{Q}(\zeta_5)$. The extension $\mathbb{Q}(\sqrt[5]{3})$ has degree 5 over $\mathbb{Q}$, and $\mathbb{Q}(\zeta_5)$ has degree 4 over $\mathbb{Q}$ since the minimal polynomial of ζ_5 over $\mathbb{Q}$ is the 5^{th} *cyclotomic polynomial*:

$$\phi_5(x) = x^4 + x^3 + x^2 + x + 1.$$

This polynomial was shown to be irreducible last quarter as a consequence of Eisenstein's criterion (this was the "replace x by $x + 1$" trick), and it does have ζ_5 as a root since

$$\phi_5(x) = x^4 + x^3 + x^2 + x + 1 = \frac{x^5 - 1}{x - 1}$$

and the numerator is zero at $x = \zeta_5$.

The composite $\mathbb{Q}(\sqrt[3]{5}, \zeta_5)$ thus has degree at most $5 \cdot 4 = 20$ over $\mathbb{Q}$. But upon considering the intermediate extensions $\mathbb{Q} \subseteq \mathbb{Q}(\sqrt[5]{3}) \subseteq \mathbb{Q}(\sqrt[5]{3}, \zeta_5)$ and $\mathbb{Q} \subseteq \mathbb{Q}(\zeta_5) \subseteq \mathbb{Q}(\sqrt[5]{3}, \zeta_5)$, we see that the degree of the composite should be divisible by both 5 and 4, and hence by their least common multiple, which is 20. We thus conclude that $[\mathbb{Q}(\sqrt[5]{3}, \zeta_5) : \mathbb{Q}] = 20$.

Normal extensions. The types of fields which can arise as splitting fields of polynomials over F are called the *normal* extensions of F. To be precise, E/F is a normal extension if it is algebraic and if E is the splitting field of a collection of polynomials over F. (A splitting field for a *collection* of polynomials just means the smallest field over which all polynomials in that collection split simultaneously. The splitting field of one polynomial in the collection will be contained in the splitting field of the entire collection.) The book tends to avoid the term "normal" and simply calls E a "splitting field" over F.

But, it is more common to state the definition of normal in the following equivalent way: E is normal over F if it is algebraic and whenever an irreducible polynomial $p(x)$ over F has a root in E, then it splits completely over E. For instance, $\mathbb{Q}(\sqrt[5]{3})$ is not normal over $\mathbb{Q}$ since $x^5 - 3$ has a root in this field but does not split in this field. The full splitting field $\mathbb{Q}(\sqrt[5]{3}, \zeta_5)$ of $x^5 - 3$ over $\mathbb{Q}$ *is* normal over $\mathbb{Q}$. We will tend to use "E is normal over F" instead of "E is a splitting field over F" in these notes, and use whichever of the two equivalent definitions of "normal" is appropriate for the problem at hand.

Existence of splitting fields. We now prove that splitting fields always exist. Suppose F is a field and $p(x) \in F[x]$. To get an idea for what to do, suppose $p(x)$ facts into irreducibles as

$$p(x) = (x - \alpha_1) \cdots (x - \alpha_k) q_1(x) \cdots q_m(x)$$

where we have written the linear factors first (for $\alpha_i \in F$) and each $q_i(x)$ has degree at least 2. The roots of the linear factors are already in F, so we need only adjoin roots of the irreducible factors. But this we know we can do by taking a quotient: for instance,

$$E_1 := F[x]/(q_1(x))$$

contains a root of $q_1(x)$; let us call this root β (which is really the element $x \in E_1$), so that we can continue to refer to the variable of the polynomials we are working with as x. In this extension we can then "split off" a linear factor $x - \beta$ from $q_1(x) = (x - \beta)h(x)$, for some $h(x) \in E_1[x]$, and then factor $h(x)$ further into irreducibles. Then we do the same thing: take one of these irreducible factors and take another quotient

$$E_2 := E_1[x]/(\text{irreducible})$$

to get a further extension of F in which $q_1(x)$ now has an additional root, so that we can split off another linear factor. And so on we keep going, taking more and more quotients, and going back and doing the same with the other original irreducible factors $q_2(x), \ldots, q_m(x)$, until we get a final extension E of F over which $p(x)$ has split completely.

To make this process "cleaner" we can phrase it as an induction on the degree of $p(x)$. If $\deg p(x) = 1$ then $p(x)$ is already split in F itself, so F serves as the splitting field. Suppose now $n > 1$. There exists an extension K/F in which $p(x)$ has a root α: if $p(x)$ already has a root in F, this extension is just F, but if not then the quotient by some irreducible factor of $p(x)$ will do. Over K we have

$$p(x) = (x - \alpha)q(x)$$

for some $q(x) \in K[x]$. Since $\deg q(x) = n - 1 < n$, by induction we may assume that there exists a splitting field E for $q(x)$ over K, and since $\alpha \in K \subseteq E$, this E also serves as a splitting field for $p(x)$ over F as desired. (Note $p(x)$ does not split in a proper subfield of E because $q(x)$ does not.)

Uniqueness and degree. As a consequence of the same ideas as those used in proof of the existence of splitting fields, we also get a proof of uniqueness and a bound on the degree of the splitting field. For uniqueness, we use the fact that $F(\alpha) \cong F(\alpha')$ whenever α, α' are roots of the same irreducible polynomial over F, which if you recall is true simply because both of these extensions are isomorphic to the same quotient of $F[x]$. Intuitively, adjoining roots of $p(x) \in F[x]$ one at a time to extend F into the splitting field should thus produce only one possible field in the end. To be more precise we can again phrase this as an induction on the degree. Suppose $p(x) \in F[x]$ has splitting fields E and E'. If $p(x)$ is of degree 1, then $F \cong E \cong E'$ since E, E' cannot be proper extensions of F because already splits in F. If $p(x)$ has degree at least 2, we pass to an extension containing a root α of $p(x)$, where it does not matter which root we use by the $F(\alpha) \cong F(\alpha')$ observation above. In this extension we have

$$p(x) = (x - \alpha)q(x)$$

for some $q(x)$. Then E and E' both serve as splitting fields for $q(x)$, so since $\deg q(x) < \deg p(x)$, we have that $E \cong E'$ by induction as desired.

As for the degree, note that if $p(x)$ was irreducible to begin with, then an extension (constructed as a quotient) containing a first root α will have degree $n = \deg p(x)$. Then we factor $p(x) = (x - \alpha)q(x)$: if $q(x)$ is now irreducible, the next extension we get (adjoining a root of $q(x)$) has degree $\deg q(x) = n - 1$. And so on, if we get a new irreducible factor at each step, the degrees of the resulting extensions decrease to $n - 2$, $n - 3$, and all the way down to 3 and finally 2 once we reach the splitting field. The tower law then says that the degree of the splitting field is $n!$. Now, there is no guarantee that we actually get irreducible factors at each step, and for instance perhaps $p(x)$ was not irreducible to begin with. The point is that the degrees in the scenario above are only *upper* bounds in general, in that the degree at each step in the construction of the splitting fields are *at most* n, then $n-1$, then $n-2$, etc. Thus we get that in general the splitting field has degree at most $n!$ over the base field.

Now, this bound on the degree will be quite large in general. For example, in the Warm-Up we see that the splitting field of $x^5 - 3$ over $\mathbb{Q}$ has degree 20, which is significantly smaller than $5! = 120$. But, this will be a useful bound nonetheless. One observation to make now is that we have seen the number $n!$ in another context previously, as the order of the symmetric group S_n. This is a reflection of the fact that the Galois group (whatever that means) of the splitting field of a polynomial of degree n will always be a subgroup of S_n, as we'll see.

Algebraically closed fields. We say that a field F is *algebraically closed* if any polynomial over F has a root in F. Equivalently, by factoring out more and more linear terms, which is always possible since any remaining factor still has a root in F, we see that this definition is equivalent to the statement that any polynomial over F splits in F. Thus, an algebraically closed field serves as the splitting field for any polynomial over it. Another way of saying all this is that an algebraically closed field has no algebraic (and hence no finite) extensions: if E is algebraic over F and F is algebraically closed, then any $\alpha \in E$ is the root of a polynomial over F, which means that $\alpha \in F$ since F is algebraically closed, so that in fact $E = F$.

The most basic example of an algebraically closed field is $\mathbb{C}$, where the statement that $\mathbb{C}$ is algebraically closed is known as the *Fundamental Theorem of Algebra*, which we will prove later using Galois theory. Our proof will be almost purely algebraic, except for one fact from analysis (or calculus, even) we will need. For now we just comment that this is theorem has tons of other non-algebraic proof as well, which highlights its importance across mathematics broadly: if you take a course in complex analysis (MATH 325), you will see a proof there using what's called *Liouville's Theorem*, and if you take a course in algebraic topology (MATH 344-2), you will see a proof using the notion of the *fundamental group* of the circle. Good stuff!

Fields of algebraic elements. $\mathbb{C}$ is thus an algebraically closed field containing $\mathbb{Q}$, but is in fact not the smallest such field. ($\mathbb{C}$ *is* the smallest algebraically closed field containing $\mathbb{R}$ however.) Denote by $\overline{\mathbb{Q}}$ the set of elements of $\mathbb{C}$ that are algebraic over $\mathbb{Q}$:

$$\overline{\mathbb{Q}} = \{z \in \mathbb{C} \mid z \text{ is algebraic over } \mathbb{Q}\}.$$

We have argued before that sums, products, and quotients (nonzero denominator) of algebraic elements are algebraic as well, so that $\overline{\mathbb{Q}}$ is a field, called the field of *complex algebraic numbers*. We now claim that $\overline{\mathbb{Q}}$ is algebraically closed, so that it is the smallest algebraically closed field containing $\mathbb{Q}$. (No proper subfield will be algebraically closed since for $\alpha \in \overline{\mathbb{Q}}$ which is not in that subfield, the minimal polynomial of α over $\mathbb{Q}$ does not split in the subfield.)

To see that $\overline{\mathbb{Q}}$ is algebraically closed, let $p(x) \in \overline{\mathbb{Q}}[x]$ and let α be a root of $p(x)$ in $\mathbb{C}$. Then $\overline{\mathbb{Q}}(\alpha)$ is an algebraic extension of $\overline{\mathbb{Q}}$, which in turn is an algebraic extension of $\mathbb{Q}$. Since algebraic extensions of algebraic extensions are algebraic, $\overline{\mathbb{Q}}(\alpha)$ is algebraic over $\mathbb{Q}$, so in particular α is algebraic over $\mathbb{Q}$. Thus $\alpha \in \overline{\mathbb{Q}}$, so $\overline{\mathbb{Q}}$ is algebraically closed. More generally, if $F \subseteq K$ and K is algebraically closed, the same argument shows that the set $\overline{F}$ of elements of K which are algebraic over F is an algebraically closed field, and indeed the smallest such one containing F.

Algebraic closures. The field $\overline{F}$ constructed above, as the smallest algebraically closed field containing F, is called the *algebraic closure* of F. But, the construction of $\overline{F}$ here depends on the existence of *some* algebraically closed field containing F, since we define $\overline{F}$ to be the set of algebraic elements of that field. We seek to make sense of this notion of "algebraic closure" without having to make reference to a larger algebraically closed field to begin with.

Here is a definition we can give without such a reference: an *algebraic closure* of a field F is an algebraic extension $\overline{F}$ of F in which every polynomial over F splits. Another way of saying this that $\overline{F}$ is the simultaneous splitting field for the collection of *all* polynomials over F. Yet another way of saying this is that $\overline{F}$ is the maximal algebraic extension of F, but justifying this will take a bit of effort, which we'll look at next time. It turns out that an algebraic closure, if it exists (which it always does, as we'll see), is unique up to isomorphism, as we'll prove next time. Because of this, it makes sense to talk about *the* algebraic closure of F.

The same proof we gave above showing that $\overline{\mathbb{Q}}$ is algebraically closed also shows that $\overline{F}$ (assuming it exists) is algebraically closed, with a slight modification: $p(x) \in \overline{F}[x]$ has a root α in some algebraic extension of $\overline{F}$, and then the "algebraic over algebraic is algebraic" reasoning will show that α is algebraic over F; the minimal polynomial of α over F then splits in $\overline{F}$ by the definition of algebraic closure, which means that α must actually lie in $\overline{F}$, so that $\overline{F}$ is algebraically closed. We will show next time that the two uses of the notation $\overline{F}$ we have given—as an algebraic closure of F and as the set of elements in a larger algebraically closed field that are algebraic over F—are the same, and use this to justify the fact that any field has an algebraic closure.

Lecture 6: More on Closures

Warm-Up. For p prime, we show that $\bigcup_n \mathbb{F}_{p^n}$ serves as an algebraic closure of $\mathbb{F}_p$. (We will take for granted the fact that a finite field of order p^n always exists and is unique, which is something we will be able to prove very shortly.) Now, there is some ambiguity here, in that at first it is not clear how to define the required union: taking the union of sets ordinarily requires that all of those sets be subsets of a common larger superset. For instance, it technically does not make sense to take the union of $\mathbb{R}$ and $\{$people in this class$\}$, unless these two were already subsets of some larger S, in which case the union is defined as the set of all $x \in S$ that belong to at least one of the subsets in question. Given the sets $\mathbb{F}_{p^n}$, it is not clear that there is a set which will contain all, in order for the required union to make sense. But, we will appeal to the fact we will prove in a bit, that an algebraic closure $\overline{\mathbb{F}_p}$ always exists. This algebraic closure will contain all $\mathbb{F}_{p^n}$, simply because it contains all algebraic (hence finite) extensions of $\mathbb{F}_p$. Thus, this problem is not really asking to show that $\overline{\mathbb{F}_p}$ exists in the first place, but rather that once we know it does, to verify that it is the union of all $\mathbb{F}_{p^n}$.

We verify that $\bigcup_n \mathbb{F}_{p^n}$ satisfies the second definition of algebraic closure we gave last time: it is an algebraic extension of $\mathbb{F}_p$ in which every polynomial over $\mathbb{F}_p$ splits. First, observe that this union is in fact a field, basically because every algebraic operation we need to consider (addition, multiplication, etc) when verifying the field axioms will take place within a specific $\mathbb{F}_{p^n}$: if $\alpha, \beta \in \bigcup_n \mathbb{F}_{p^n}$, with say $\alpha \in \mathbb{F}_{p^k}$ and $\beta \in \mathbb{F}_{p^\ell}$, then α and β both belong to the composite $\mathbb{F}_{p^k}\mathbb{F}_{p^\ell}$, which is a finite field itself since it has degree at most $[\mathbb{F}_{p^k} : \mathbb{F}_p][\mathbb{F}_{p^\ell} : \mathbb{F}_p] = p^k p^\ell$ over $\mathbb{F}_p$; thus $\alpha, \beta \in \mathbb{F}_{p^i}$ for some i, so the the field axioms (closure under addition, closure under multiplication, etc) can be checked in $\mathbb{F}_{p^i}$ if nothing else. Next, $\bigcup_n \mathbb{F}_{p^n}$ is algebraic over $\mathbb{F}_p$, since if α is in the union, then α lies in some finite (hence algebraic) extension $\mathbb{F}_{p^k}$ of $\mathbb{F}_p$, so that α is algebraic over $\mathbb{F}_p$.

Now, let $p(x) \in \mathbb{F}_p[x]$. If $p(x)$ is of degree n, then the splitting field of $p(x)$ is of degree at most $n!$ over $\mathbb{F}_p$, so in particular this splitting field is a finite extension of $\mathbb{F}_p$, and is thus $\mathbb{F}_{p^k}$ for some k. Hence $p(x)$ splits in $\mathbb{F}_{p^k} \subseteq \bigcup_n \mathbb{F}_{p^n}$, so $\bigcup_n \mathbb{F}_{p^n} = \overline{\mathbb{F}_p}$ is an algebraic closure of $\mathbb{F}_p$ as claimed.

Algebraic elements form a closure. We now verify the statement that $\overline{F}$, constructed as the set of elements in an algebraically closed extension K of F which are algebraic over F, is an algebraic closure of F in the sense of having every polynomial over F split in $\overline{F}$, so that the various definitions of "algebraic closure" we gave are equivalent. To this end, let $p(x) \in F[x]$. Since $p(x) \in K[x]$ as well and K is algebraically closed, $p(x)$ splits over K:

$$p(x) = (x - \alpha_1) \cdots (x - \alpha_n) \text{ for some } \alpha_i \in K.$$

By construction, each α_i is the root of a polynomial over F—namely $p(x)$—so each α_i is algebraic over F and hence belongs to $\overline{F}$. Thus the splitting of $p(x)$ above is actually a valid splitting in $\overline{F}[x]$, so $\overline{F}$ is an algebraic closure of F.

Closures exist. So, algebraic closures exist, at least for fields which are contained in algebraically closed fields. But, for F without an obvious previously-defined algebraically closed extension, how do we actually check for the existence of an algebraic closure?

One approach comes from thinking of an algebraic closure as a *maximal* algebraic extension: if we can show that F has a maximal (meaning not contained in any larger) algebraic extension, then that *should* be the algebraic closure. Showing that there is in fact a maximal algebraic extension seems like the type of thing which Zorn's Lemma should apply to: naively, we take the set of all algebraic extensions of F, show that it has a maximal element by verifying the hypotheses of Zorn's Lemma, and then show that this maximal element is an algebraic closure. But, this approach runs into a similar issue as the "take the union of all $\mathbb{F}_{p^n}$" ambiguity we had in the Warm-Up, in that taking the "set of all algebraic extensions" only makes sense for extensions which sit inside of an already-constructed larger set. In other words,

$$\{E \mid E \text{ is an algebraic extension of } F\}$$

is not actually a well-defined set since it does not specify what types of objects the E's should be in the first place; only something like

$$\{E \subseteq S \mid E \text{ is an algebraic extension of } F\}$$

is well-defined, where we say definitively that we only consider E's that come from some set S which already exists. (This distinction is the same reason why "the set of all sets" is not actually a well-defined set.) In the language of set theory, $\{E \mid E \text{ is an algebraic extension of } F\}$ is what's called a *proper class*, which is no good for our purposes since Zorn's Lemma only applies to things which are actually *sets*. (This was not an issue in past quarters, since for instance when using Zorn's Lemma to construct maximal ideals, the ideals used *did* in fact all belong to an already-existing set, namely the ring in question.) To make such a "maximal algebraic extension" argument work, we would need to know that there in fact exists a (very) large set S which *does* in fact contain every possible algebraic extension of F. It turns out that this is actually true, but constructing the required set S would take us too far into the realm of advanced set theory and the study of cardinality to be worth it here.

Instead, we give an alternative approach, which is the one the book gives and is based on the idea of using quotients to construct extensions over which polynomials have roots. Our goal is to construct an algebraically closed field K containing F in this way, after which the set of elements of K which are algebraic over F will be an algebraic closure of F, as we have already shown. So, given F, we first construct an extension K_1 over which every polynomial in $F[x]$ has *at least one* root. For a single polynomial, say irreducible, we can simply quotient out by the ideal generated by that polynomial, so we aim to do something similar here, only where we quotient out by *all* polynomials at once. To this end, for each monic nonconstant $p(x) \in F[x]$, introduce a new "variable" x_p. (We need only consider monic polynomials since the roots of a non-monic polynomial are the roots of the monic polynomial obtained by dividing through by the leading coefficient.) Then we consider the polynomial ring over F generated by *all* of these (infinitely many) variables:

$$R := F[x_p : p(x) \in F[x] \text{ is nonconstnat and monic}].$$

We then consider $p(x_p) \in R[x]$, which is the same polynomial as $p(x)$ only with x_p as the variable instead of x. The point is that we use these new variables to keep track of which polynomial we are looking at: $x_p^2 + 2x_p + 3$ corresponds to the original $p(x) = x^2 + 2x + 3$, $x_q^4 - 2x_q^3 + 4x_q - 3$ corresponds to the original $q(x) = x^4 - 2x^3 + 4x - 3$, and so on.

We would like to quotient out by all the $p(x_p)$, which will force every polynomial in $F[x]$ to now have a root: x_p in the quotient is a root of $p(x_p)$. But, the issue is that there is no guarantee we get a field when quotienting out by the $p(x_p)$, or more precisely by the ideal I they generate, since we do not know that I will be maximal. (This is avoided in the previous $F[x]/(p(x))$ constructions we've used since $p(x)$ was irreducible in those cases, but the construction we are attempting here is much more general.) We have a quick fix however: by Zorn's Lemma, I is contained in a maximal ideal of R, *as long as* we know that I is a *proper* ideal of R. If so, then we use $K_1 := R/M$ where M is such a maximal ideal containing I; since M still contains all $p(x_p)$, every $p(x) \in F[x]$ will have a root over the extension $K_1 \supseteq F$. To show that I is proper, we show that $1 \in R$ is not in I. If it was, we would have

$$1 = q_1 p_1(x_{p_1}) + \cdots + q_n p_n(x_{p_n})$$

for some $q_i \in R$ and $p_i(x) \in F[x]$. (Recall that I is the deal generated by the $p(x_p)$.) There is an extension of F over which each $p_i(x)$ has a root, since we can first take such an extension for $p_1(x)$ alone, then enlarger further if need be to ensure $p_2(x)$ has a root, then enlarger further for $p_3(x)$, and so on until we do so for $p_n(x)$. If α_i is a root of $p_i(x)$ in this common extension, then setting $x_{p_i} = \alpha_i$ in the equation above gives $1 = 0$, which is not possible. Thus $1 \notin I$, so I is proper.

We thus get a field K_1 in which each $p(x) \in F[x]$ has at least one root. We want to get an algebraically closed field which contains F, but we do not know anything now about polynomials in $K_1[x]$ in terms of whether they have roots in K_1. But no matter, we do the exact same thing as above only with K_1 instead of F: we get an extension K_2 of K_1 over which every polynomial in $K_1[x]$ has a root! And so on we keep going, to obtain a chain of extensions

$$F \subseteq K_1 \subseteq K_2 \subseteq \dots$$

with the property that every $p(x) \in K_i[x]$ (set $K_0 := F$) has a root in $K_{i+1}[x]$. We claim that the union $\bigcup_n K_n$ is in fact an algebraically closed field containing F, which is what we want to obtain. As stated before, an algebraic closure of F is then obtained by taking the set of elements of K that are algebraic over F.

The union $\bigcup_n K_n$ clearly contains F, and the proof that is a field is the same as the argument for why $\bigcup_n \mathbb{F}_{p^n}$ was a field in the Warm-Up: given two elements, all the necessary algebraic operations needed to verify the field axioms will take place within a specific K_i, which we already know to be a field. Thus all that remains is to check that the union is algebraically closed. But if $p(x)$ is a polynomial over $\bigcup_n K_n$, the finitely many coefficients of $p(x)$ will lie in finitely many of the K_n, so there is one K_m that contains all coefficients. Then $p(x) \in K_m[x]$, so $p(x)$ has a root in K_{m+1}, which belongs to $\bigcup_n K_n$ as well. Hence every polynomial over $\bigcup_n K_n$ has a root in $\bigcup_n K_n$, so $\bigcup_n K_n$ is algebraically closed.

Closures contain all algebraic extensions. Our final goal is to show that algebraic closures are unique. But for this we need to know first that a given algebraic closure of F will contain all algebraic extensions of F. Now, there is some ambiguity as to what we actually mean by this: until we know that algebraic closures are unique, it will not be true that a given one *literally* contains every possible algebraic extension. That is, if, say, $\overline{F}$ and $\overline{F}'$ are two algebraic closures of F without any a prior relation to one another, an algebraic extension $F \subseteq K \subseteq \overline{F}$ contained in $\overline{F}$ does not have to be a literal subset of $\overline{F}'$, so we have to be careful about what "$\overline{F}'$ contains K" actually means in this context. What we really mean is that a given algebraic closure will contain an *isomorphic copy* of any algebraic extension: more precisely, if $\overline{F}$ is an algebraic closure of F and K is an algebraic extension of F, then there exists a injective field homomorphism $K \to \overline{F}$. (The image of this map is then the isomorphic copy of K inside $\overline{F}$ we want.)

To show this, we again exploit the wonders of Zorn's Lemma. Let K be an algebraic extension of F, and consider the set $\mathcal{S}$ of all injective mappings of subextensions $F \subseteq L \subseteq K$ into $\overline{F}$, where $\overline{F}$ is a fixed algebraic closure of F:

$$\mathcal{S} := \{\text{injective } L \to \overline{F} \mid L \text{ is a subextension of } K\}.$$

(We should actually require that these injective maps fix elements of the base field F. Also, this is a well-defined set, since we only consider things which sit inside of the already-existing K.) Our goal is to show that K is actually in this set, so that we do have an injective mapping $K \to \overline{F}$. (We're being a bit sloppy here conflating the fields L with the injective maps $L \to \overline{F}$, but that's not a big deal and is easy to avoid. We only do this to not get bogged down in notation.) A straightforward application of Zorn's Lemma shows that $\mathcal{S}$ has a maximal element: $\mathcal{S}$ is not empty since $F \in \mathcal{S}$, and the usual "union of elements in a chain" argument will give an upper bound for the chain; we'll omit the specific details, but there is nothing new we haven't seen before. Say that M is the maximal element. If $M \neq K$, then there exists $\alpha \in K - M$. This element is algebraic over F, since

K is algebraic over F, so it has a minimal polynomial over F. But this minimal polynomial splits and thus has a root in the algebraic closure $\overline{F}$; if β is this root, then $F(\alpha) \cong F(\beta)$ since the specific root of the minimal polynomial we adjoin does not matter. This shows that $M(\alpha) \in \mathcal{S}$, since the injective map $M \to \overline{F}$ can be extended to α using the isomorphism $F(\alpha) \cong F(\beta) \subseteq \overline{F}$ (i.e. send α to β). This contradicts maximality of M, so we must in fact have $M = K$, so that any algebraic extension K of F does map injectively into the specified algebraic closure $\overline{F}$.

Closures are unique. Thus, suppose $\overline{F}$ and $\overline{F}'$ are two algebraic closures of F. Since $\overline{F}$ is algebraic over F, there exists an injective map $\overline{F} \to \overline{F}'$ by the fact above, so that $\overline{F}$ is isomorphic to a subfield E of $\overline{F}'$. If $\alpha \in \overline{F}'$, then α has some minimal polynomial $m_\alpha(x)$ over F. But this minimal polynomial splits in $\overline{F} \cong E$ since $\overline{F}$ is also an algebraic closure of F, which means that the root α must belong to E. Hence $\overline{F}' = E$, so $\overline{F} \cong \overline{F}'$ as desired. We conclude that algebraic closures are unique (up to isomorphism), so that we can speak of *the* algebraic closure of F. (There is a lot of machinery that goes into the existence and uniqueness of algebraic closures in general!)

p-adic complex numbers. As a fun aside, let us introduce the *p-adic complex numbers*, who construction depends on the notion of an algebraic closure. For p prime, last quarter we saw the example of the $\mathbb{Q}_p$, the field of p-adic numbers, whose elements are most easily described as Laurent series in the "variable" p. Those elements whose Laurent series expansions are actually power series (no negative exponents) make up $\mathbb{Z}_p$, the ring of p-adic integers, and $\mathbb{Q}_p$ is the fraction field of $\mathbb{Z}_p$.

Now, we briefly (in the notes, at least) discussed how $\mathbb{Q}_p$ can be obtained analytically from $\mathbb{Q}$, as the *completion* of $\mathbb{Q}$ with respect to the *p-adic metric*, just as $\mathbb{R}$ is the completion of $\mathbb{Q}$ with respect to the standard Euclidean metric. (No worries if you haven't seen enough analysis to know what of this means, we are just mentioning it for fun! Check the notes from last quarter to see what the p-adic metric is.) In the case of $\mathbb{R}$, we can then take the algebraic closure to obtain $\mathbb{C}$, which also inherits a metric and is in fact complete with respect to that metric. So, the process of beginning with $\mathbb{Q}$ and alternating between taking completions and closures ends with $\mathbb{C}$:

$$\mathbb{Q} \rightsquigarrow \mathbb{R} \rightsquigarrow \mathbb{C}.$$

In the p-adic case, $\mathbb{Q}_p$ is complete, but not algebraically closed, so we must take an algebraic closure in order to obtain an algebraically closed field $\overline{\mathbb{Q}_p}$. As opposed to the "standard" case above, however, this algebraic closure is *not* complete with respect to the inherited p-adic metric, and so we must take another completion in order to obtain a complete metric space.

This completion, it turns out, is in fact still an algebraically closed field, and is commonly denoted by $\mathbb{C}_p$, the field of "p-adic complex numbers":

$$\mathbb{Q} \rightsquigarrow \mathbb{Q}_p \rightsquigarrow \overline{\mathbb{Q}_p} \rightsquigarrow \mathbb{C}_p.$$

This field $\mathbb{C}_p$ is the ultimate field in which "p-adic analysis" takes place, and is used through number theory and algebraic geometry. In fact, as a *field*, $\mathbb{C}_p$ is actually isomorphic to $\mathbb{C}$, but the point is that the analytic metric structures (i.e. the way in which you measure distance) is very different.

Separable polynomials. We finish by moving away from the main topic of the day, to quickly introduce the notion of a *separable* polynomial. This will be very brief, and it probably makes sense to just postpone the definition until next time, but we introduce it now just so that we can hit the ground running with a Warm-Up which deals with this concept next time.

Our eventual goal is to the define the notion of a *separable* extension. We are working towards Galois theory, and want to understand the types of extensions over which Galois theory works as nicely as possible, and it turns out that separability is one key ingredient towards this; the other is normality, which we introduced last time. (An extension which is both normal and separable will be called a *Galois* extension.) Now, the definition of a separable extension (which we will give next time) will possibly seem counterintuitive at first, or more precisely it will seem counterintuitive that such a definition is needed, since at first glance it will appear as if *every* extension should be separable. Indeed, pretty much every field extension you have seen so far in your life is in fact separable, in particular because in characteristic zero every extension *is* indeed separable. To have any hope of seeing extensions which are not separable we must move to prime characteristic, but even that is not enough since, as we'll see, all finite extensions of finite fields are also separable. The point is that separability can possibly only fail for infinite fields of prime characteristic, and we just have not seen many examples of such things apart from $\overline{\mathbb{F}_p}$. (But actually, $\overline{\mathbb{F}_p}$ is separable over $\mathbb{F}_p$, so this doesn't cut it either.) Even though our main focus will be on field extensions where separability is always guaranteed to hold, it will nevertheless be useful to come to terms with this concept in order to understand why it is so important to Galois theory. (It will also be crucial to the construction and uniqueness of all possible finite fields.)

Here is the key definition: we say that a polynomial $p(x) \in F[x]$ is *separable* if it has no repeated roots in any extension, or equivalently if it has distinct roots in its splitting field; otherwise, we say that $p(x)$ is *inseparable*. (The name "separable" comes from the idea that the roots can be "separated" from one another, because they are distinct.) For example, $(x-1)^2 = x^2+2x+1$ is inseparable over $\mathbb{Q}$ since it has a repeated root of 1 (of multiplicity 2), and x^2-2 is separable over $\mathbb{Q}$ since its roots $\pm\sqrt{2}$ in its splitting field are distinct. The polynomial x^3-2 is also separable over $\mathbb{Q}$ since its distinct roots are $\sqrt[3]{2}$, $\zeta_3\sqrt[3]{2}$, and $\zeta_3^2\sqrt[3]{2}$ (where ζ_3 is a primitive third root of unity), each of multiplicity 1.

The property that will characterize separable extensions comes from asking whether there exist any irreducible polynomials which are not separable. As alluded to before, the answer is "no" over a field of characteristic zero, and "no" for polynomials over finite fields, but can be "yes" in other settings, with a basic example given below. We will look at all this more closely next time.

Example. Consider the field $\mathbb{F}_3(x)$ of rational functions (fractions of polynomials) over $\mathbb{F}_3$. This field has characteristic 3, and is an infinite extension of $\mathbb{F}_3$. Now, take the polynomial

$$X^3 - x$$

over this field, meaning in the polynomial ring $(\mathbb{F}_3(x))[X]$. (So, x here is no longer the variable of polynomials, but rather an element of the coefficient field $\mathbb{F}_3(x)$. Capital X is the polynomial variable.) We claim first that this polynomial is irreducible over $\mathbb{F}_3(x)$. Indeed, if we recognize $\mathbb{F}_3(x)$ as the fraction field of $\mathbb{F}_3[x]$, we see that x is a prime in the ring $\mathbb{F}_3[x]$ since the quotient $\mathbb{F}_3[x]/(x) \cong \mathbb{F}_3$ is an integral domain. Eisenstein's criterion with this prime then applies to show that X^3-x is irreducible over $\mathbb{F}_3[x]$, and hence over $\mathbb{F}_3(x)$ by Gauss's Lemma.

But, in the extension $\mathbb{F}_3(x)(\sqrt[3]{x}) \cong (\mathbb{F}_3(x))[X]/(X^3-x)$, where we adjoin a cube root of $x \in \mathbb{F}_3(x)$, we have:

$$X^3 - x = (X - \sqrt[3]{x})^3.$$

Indeed, multiplying out the right side gives

$$(X-\sqrt[3]{x})^3 = X^3 - 3\sqrt[3]{x}X^2 + 3\sqrt[3]{x^2}X - x,$$

which simplifies to X^3-x since $3=0$ in characteristic 3 so that the terms in the middle vanish. This shows that X^3-x has only one root in its splitting field, namely $\sqrt[3]{x}$, but that this root is repeated with multiplicity 3. Hence $X^3 - x$ is an irreducible polynomial over $\mathbb{F}_3(x)$ which is inseparable. (The same is true of $X^p - x$ over $\mathbb{F}_p(x)$ for a similar reason for any prime p.) We will see next time what inseparability of irreducibles can fail to give us, and see hints of why it is something we want to avoid.

Lecture 7: Separable Extensions

Warm-Up. We show that irreducible polynomials over $\mathbb{R}$ are always separable. (In fact, the same argument will apply to irreducible polynomials over $\mathbb{Q}$.) This is a special case of the fact we mentioned last time that irreducible polynomials over a field of characteristic zero are always separable, which we will prove in full in a bit. Here, however, we can give a more brute-force proof since we know where roots of polynomials over $\mathbb{R}$ (or $\mathbb{Q}$) lie, namely in $\mathbb{C}$.

A key observation is that if $\alpha \in \mathbb{C}$ is a root of a polynomial with real coefficients, then so is the complex conjugate $\overline{\alpha}$. Indeed, if α satisfies

$$c_0 + c_1\alpha + \cdots + c_n\alpha^n = 0$$

for some $c_i \in \mathbb{R}$, then conjugating both sides yields

$$c_0 + c_1\overline{\alpha} + \cdots + c_n\overline{\alpha}^n = 0,$$

where we use the fact that conjugation preserves addition and multiplication of complex numbers, and that $\overline{c_i} = c_i$ since c_i is real. So, suppose $p(x) \in \mathbb{R}[x]$ is irreducible. Then $p(x)$ cannot have a repeated real root $a \in \mathbb{R}$, since such a repeated root would lead to a factorization like

$$p(x) = (x-a)^2 q(x)$$

for some $q(x) \in \mathbb{R}[x]$, but then $p(x)$ is reducible since it factors as $x - a$ times $(x-a)q(x)$. (For a similar reason, $p(x)$ has at most one real root, since otherwise we would have at least two distinct real factors $x - a$ and $x - b$.)

Now, suppose $\alpha \in \mathbb{C}$ is non-real root of $p(x)$, so that $\overline{\alpha}$ is also a root. Note that

$$(x-\alpha)(x-\overline{\alpha}) = x^2 - (\alpha + \overline{\alpha})x + \alpha\overline{\alpha}$$

is a polynomial over $\mathbb{R}$ since $\alpha + \overline{\alpha}$ and $\alpha\overline{\alpha}$ are both real. Over $\mathbb{C}$ then, $p(x)$ is divisible by $x - \alpha$ and $x - \overline{\alpha}$, so it is divisible by the *real* polynomial $(x-\alpha)(x-\overline{\alpha})$, so $p(x)$ factors over $\mathbb{R}$ as

$$p(x) = (x-\alpha)(x-\overline{\alpha})q(x)$$

for some $q(x) \in \mathbb{R}[x]$. Now, any other roots of $p(x)$ have to come from roots of $q(x)$, so if α is to be a repeated root of $p(x)$ it must be a root of $q(x)$. But for the same reason as above, this would give a factorization like

$$q(x) = (x-\alpha)(x-\overline{\alpha})h(x)$$

for some $h(x) \in \mathbb{R}[x]$. But then we have

$$p(x) = [x^2 - (\alpha + \overline{\alpha}) + \alpha\overline{\alpha}]^2 h(x),$$

so $p(x)$ would be reducible over $\mathbb{R}$. Thus $p(x)$ has no repeated roots in $\mathbb{C}$, so it is separable. (This also essentially classifies irreducible polynomials over $\mathbb{R}$: there are the linear ones $x - a$ for $a \in \mathbb{R}$, and the quadratic ones $(x-\alpha)(x-\overline{\alpha})$ for α a non-real complex number.)

Separability via derivatives. For irreducible polynomials over other fields, an argument such as the one above will not work since we do not know how to describe the roots explicitly; in particular, there is no analog of "if $\alpha \in \mathbb{C}$ is a root, then so is $\overline{\alpha}$". So, we need a way to characterize separability in a way which does not make explicit reference to the unknown roots.

But, if a is a repeated root of $p(x) \in F[x]$ in some extension, so that we can write $p(x)$ as

$$p(x) = (x-a)^2 q(x)$$

over that extension (note this allows for the possibility that a has multiplicity greater than 2, since $q(x)$ could still have a as a root), taking *derivatives* gives

$$p'(x) = 2(x-a)q(x) + (x-a)^2 q'(x).$$

To be clear, by "derivative" here we mean the usual algebraic formula for the derivative of a polynomial, which makes sense over any field; there is no "limit" definition being used, which would *not* make sense over arbitrary fields. (One can show using induction that the product rule still holds for such derivatives, which we used above.) We see that a is then still a root of $p'(x)$ due to the $x-a$ factors present in the expression for $p'(x)$ above. Conversely, if a is a root of $p(x)$, write $p(x)$ as

$$p(x) = (x-a)h(x)$$

for some $h(x)$, so that

$$p'(x) = h(x) + (x-a)h'(x).$$

If a is also a root of $p'(x)$, then this equality shows that a is a root of $h(x)$ as well, so $h(x) = (x-a)g(x)$ for some $g(x)$. This then gives

$$p(x) = (x-a)h(x) = (x-a)^2 g(x),$$

so that a is a repeated root of $p(x)$. We conclude that a is a repeated root of $p(x)$ (over any field) if and only if a is a root of both $p(x)$ and $p'(x)$.

If a is a root of both $p(x)$ and $p'(x)$, then $x-a$ is a divisor of both $p(x)$ and $p'(x)$, so that $p(x)$ and $p'(x)$ are not relatively prime, assuming neither is constant Conversely, if $p(x)$ and $p'(x)$ are not relatively prime, then their greatest common divisor is a polynomial $g(x)$ of degree at least 1; this $g(x)$ then has a root in some extension of F, which will be a root of both $p(x)$ and $p'(x)$ since $x-a$ divides $g(x)$ implies $x-a$ divides $p(x)$ and $p'(x)$. The final upshot is the following result:

> $p(x)$ is inseparable (i.e. has a repeated root) if and only if $p(x)$ and $p'(x)$ (assuming $p'(x) \neq 0$) are not relatively prime; or equivalently, $p(x)$ is separable if and only if $p(x)$ and $p'(x)$ are relatively prime.

We thus have our desired characterization of separability which does not depend on being able to describe, or even mention, the roots of a polynomial.

Examples. The polynomial $p(x) = x^2 - 2x + 1 = (x-1)^2$ has derivative $p'(x) = 2(x-1)$, so $p(x)$ and $p'(x)$ are both divisible by $x-1$ and are thus not relatively prime. Hence we recover the fact that $p(x) = x^2 - 2x + 1$ is inseparable over $\mathbb{Q}$. For $q(x) = x^3 - 2$ over $\mathbb{Q}$, we have $q'(x) = 3x^2$, which is relatively prime to $x^3 - 2$ since only the only non-units dividing $q'(x)$ are x and x^2 (or these times units), and neither divide $q(x)$. (Also note that $q(x)$ and $q'(x)$ do not share a root.) Thus $q(x) = x^3 - 2$ is, as expected, separable. The polynomial $X^p - x \in \mathbb{F}_p(x)$ (we looked at the special case $p = 3$ last time) has derivative pX^{p-1}, which is 0 since $\mathbb{F}_p(x)$ has characteristic p. Thus we cannot say anything definitive using derivatives in this case, but of course we argued that $X^p - x$ was inseparable over $\mathbb{F}_p(x)$ last time.

The polynomial $x^n - 1$, whose roots are the n^{th} roots of unity, has derivative nx^{n-1}. Over a field of characteristic zero, this derivative is nonzero and relatively prime to $x^n - 1$, so $x^n - 1$ is separable over such a field. This means that there are always n distinct n^{th} roots of unity lying in some extension (say in the algebraic closure if nothing else) of a field of characteristic 0. (In the case of $\mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}$, we know this already since we can describe the roots of unity explicitly as ζ_n^k where $\zeta_n = e^{2\pi i/n}$.) Over a field of characteristic not dividing n, nx^{n-1} is still nonzero and relatively prime to $x^n - 1$, so the algebraic closure of such a field also contains n distinct n^{th} roots of unity. If the characteristic of $\mathbb{F}$, however, divides n, then $nx^{n-1} = 0$ so no information about the number of roots of unity can be derived from this alone. (In fact, it turns out there are strictly fewer than n distinct n^{th} roots of unity in this case, each with multiplicity greater than 1.)

Irreducible over characteristic zero. We can now prove that irreducible polynomials over fields of characteristic zero are always separable, generalizing the Warm-Up. Suppose $q(x)$ is irreducible, say of degree $n \geq 1$. Then $q'(x)$ is nonzero (since the characteristic is not zero) of degree $n - 1$. Since $q(x)$ is irreducible, its only divisors are 1 and $q(x)$ (or these times units), and $q(x)$ does not divide $q'(x)$ since $q'(x)$ has smaller degree. Thus $q(x)$ and $q'(x)$ are relatively prime, so $q(x)$ is separable over a field of characteristic zero.

The reason why this does not work over a field of characteristic p is that $q'(x)$ could in fact be zero, which happens when all the coefficients of $q'(x)$, or equivalently all exponents in $q(x)$, are divisible by p. (In which case $q(x)$ does divide $q'(x) = 0$.) If $q'(x)$ is in fact nonzero, then the proof above *does* work, and we can conclude that irreducible polynomials with nonzero derivatives *are* separable over fields of prime characteristic.

Separable extensions. We say that E is a *separable* extension of F if it is algebraic over F and the minimal polynomial of every element of E is separable over F. (Equivalently, all irreducible polynomials over F are separable.) The claim proved above thus says that any finite (or more generally algebraic) extension of a field of characteristic zero is separable.

Freshman's dream and Frobenius. Now, let us focus on the characteristic p case, where, as discussed above, it is not immediate that irreducible polynomials are separable. (Indeed, this is not true over $\mathbb{F}_p(x)$ for example.) We claim that this is in fact true over *finite* fields at least, and to prove this we need what's called the *freshman's dream*: $(x + y)^n = x^n + y^n$.

Now, of course, this seems like wishful thinking (and it often is for those attempting to use this in a high school algebra or calculus course), but the fact is that there is *some* truth to this, in that it actually holds in characteristic p for $n = p$! (That's an exclamation mark, not a factorial.) Indeed, suppose $\mathbb{F}$ has characteristic p. Then for $x, y \in \mathbb{F}$ we have:

$$(x + y)^p = x^p + px^{p-1}y + \tfrac{1}{2}p(p-1)x^{p-2}y^2 + \cdots + pxy^{p-1} + y^p$$

by the binomial theorem, where the coefficient of $x^{p-k}y^k$ is $\binom{p}{k} = \frac{p!}{(p-k)!k!}$. For p prime, this binomial coefficient is in fact divisible by p for $0 < k < p$ (we actually proved this in last quarter's notes when discussing the irreducibility of the p^{th} cyclotomic polynomial using Eisenstein's criterion), so all of these intermediate terms vanish and we are left with:

$$(x + y)^p = x^p + y^p,$$

i.e. the freshman's dream. (Perhaps all of those freshmen are actually on to something!)

As a consequence of this, the p^{th} power map $\mathbb{F} \to \mathbb{F}$ sending x to x^p is a field homomorphism, since it preserves addition (freshman's dream) and multiplication. This homomorphism is known as the *Frobenius* map, or often simply "Frobenius" on its own. As we'll see, it plays an important role in Galois theory in prime characteristic. The Frobenius map $\mathbb{F} \to \mathbb{F}$ is always injective (the kernel is trivial), and when $\mathbb{F}$ is finite it is actually surjective too, simply because any injective map between sets of equal finite size is automatically surjective. In the finite case then, Frobenius is actually an automorphism.

Perfect fields. If $\mathbb{F}$ is finite of characteristic p, the fact that Frobenius is surjective says that every element of $\mathbb{F}$ is a p^{th} power: for any $a \in \mathbb{F}$, there exists $b \in \mathbb{F}$ such that $b^p = a$. Or in other words, every element of $\mathbb{F}$ has a p^{th} root in $\mathbb{F}$. We say that a field F is *perfect* if it has characteristic zero, or characteristic p and every element of F is a p^{th} power, which we symbolically denote as $K = K^p$ where K^p denotes the set of p^{th} powers. Thus, finite fields are perfect, as are our usual $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, and subfields thereof. (I do not know where the name "perfect" came from historically, but it's a good name in my mind since these are the "perfect" fields to use in Galois theory.)

Irreducible over perfect field. We can now extend the fact that irreducible polynomials over fields of characteristic zero are separable to hold over perfect fields more generally. So, suppose $\mathbb{F}$ is perfect of characteristic p (which thus encompasses at least all finite fields) and that $p(x)$ is inseparable. As discussed earlier, this forces $p'(x)$ to be the zero polynomial, which requires that all exponents showing up in $p(x)$ be multiples of p:

$$p(x) = a_0 + a_1 x^{m_1 p} + a_2 x^{m_2 p} + \cdots + a_k x^{m_k p}$$

for some $a_i \in \mathbb{F}$ and $m_i \in \mathbb{N}$. Since $\mathbb{F}$ is perfect, each a_i is a p^{th} power, say $a_i = b_i^p$ for some $b_i \in \mathbb{F}$. Then we have:

$$\begin{aligned} p(x) &= a_0 + a_1 x^{m_1 p} + a_2 x^{m_2 p} + \cdots + a_k x^{m_k p} \\ &= b_0^p + b_1^p x^{m_1 p} + b_2^p x^{m_2 p} + \cdots + b_k^p x^{m_k p} \\ &= b_0^p + (b_1 x^{m_1})^p + (b_2 x^{m_2})^p + \cdots + (b_k x^{m_k})^p \\ &= (b_0 + b_1 x^{m_1} + b_2 x^{m_2} + \cdots + b_k x^{m_k})^p, \end{aligned}$$

where the last equality is the freshman's dream. But this shows that $p(x)$ is reducible (the term in parentheses above is a p-fold factor), so we conclude that irreducible polynomials over a perfect field are separable.

Finite over perfect is separable. As a consequence, any finite (or algebraic) extension of a perfect field is separable, which, in particular, means that finite extensions of finite fields are separable. This is why we have to move to more exotic infinite prime characteristic examples, such as extensions of $\mathbb{F}_p(x)$, if we want to see settings where separability can fail. As we develop Galois theory, we will see what use separability truly has.

Lecture 8: Cyclotomic Extensions

Warm-Up. Suppose $\mathbb{F}$ is a finite field of characteristic p, so that it is an extension of $\mathbb{F}_p$. We show that that the fixed points of the Frobenius automorphism on $\mathbb{F}$ are precisely the elements of $\mathbb{F}_p$: $a^p = a$ for $a \in \mathbb{F}$ if and only if $a \in \mathbb{F}_p \subseteq \mathbb{F}$. First, if $a \in \mathbb{F}_p^\times$, then since $\mathbb{F}_p^\times$ is a (multiplicative) group of order $p - 1$ we have $a^{p-1} = 1$. (In other words, all nonzero elements of $\mathbb{F}_p$ are $(p-1)$-st roots of unity.) Multiplying by a gives $a^p = a$, so $a \in \mathbb{F}_p^\times$ is a fixed point of Frobenius. Since $0^p = 0$ is also true, all elements of $\mathbb{F}_p$ are fixed points.

Now, $a^p = a$ is equivalent to $a^p - a = 0$, which says that a is a root of the polynomial $x^p - x \in \mathbb{F}_p[x]$. This polynomial has at most p roots in an extension of $\mathbb{F}_p$ (in fact exactly p roots since it is separable and hence has no repeated roots: the derivative of $x^p - x$ is -1, which is relatively prime to $x^p - x$.) But of course, we know from above that elements $\mathbb{F}_p$ already give p such roots, so these must be all the roots. Hence if $\beta \in \mathbb{F}$ satisfies $\beta^p = \beta$, so that it is a root of $x^p - x$, β must actually be in $\mathbb{F}_p \subseteq \mathbb{F}$. Thus the fixed field of Frobenius on $\mathbb{F}$ is precisely $\mathbb{F}_p$.

As a quick application, we show that if $f(x)$ is a polynomial over $\mathbb{F}_p$, then Frobenius sends roots to roots. Let α be a root of

$$f(x) = c_0 + c_1 x + c_2 x^2 + \cdots + c_n x^n, \ c_i \in \mathbb{F}_p$$

in some extension. Then

$$c_0 + c_1\alpha + c_2\alpha^2 + \cdots + c_n\alpha^n = 0.$$

Taking the p^{th} power of both sides and applying the freshman's dream to the result gives:

$$c_0^p + c_1^p\alpha^p + c_2^p(\alpha^p)^2 + \cdots + c_n^p(\alpha^p)^n = 0.$$

But Frobenius fixes $c_i \in \mathbb{F}_p$, so this becomes

$$c_0 + c_1\alpha^p + c_2(\alpha^p)^2 + \cdots + c_n(\alpha^p)^n = 0,$$

which says that α^p is a root of $f(x)$. Iterating this map then gives α^{p^2} as a root, then α^{p^3} as a root, and so on. In fact, this process gives *all* roots of $f(x) \in \mathbb{F}_p[x]$ (to be proved later), so that the Frobenius map has the effect of permuting the roots of a polynomial over $\mathbb{F}_p$. We will come back to see what this says about the Galois theory of extensions of $\mathbb{F}_p$ later.

Existence and uniqueness of finite fields. We can now finally achieve the long sought-after understanding of finite fields: for any prime power p^n, there exists a unique field of order p^n. Let us first demonstrate uniqueness. To this end, suppose $\mathbb{F}$ is a field of order p^n. Since the group of units $\mathbb{F}^\times$ has order $p^n - 1$, any nonzero $a \in \mathbb{F}$ satisfies $a^{p^n-1} = 1$, which gives $a^{p^n} = a$. The element $a = 0$ also satisfies this, so we see that all elements of $\mathbb{F}$ are roots of $x^{p^n} - x$. But this polynomial has at most p^n roots since its degree is p^n, so the elements of $\mathbb{F}$ must give all the roots and hence $\mathbb{F}$ is the splitting field of $x^{p^n} - x$. (A proper subfield would exclude some $a \in \mathbb{F}$, and hence would not contain all roots of $x^{p^n} - x$.) Since splitting fields are unique, this shows that a field of order p^n—if it exists—is unique, up to isomorphism of course.

Now for the existence. The algebraic closure $\overline{\mathbb{F}_p}$ of $\mathbb{F}_p$ contains the set S of roots of $x^{p^n} - x \in \mathbb{F}_[x]$:

$$S := \{\alpha \in \overline{\mathbb{F}_p} \mid \alpha^{p^n} = \alpha\}.$$

Since the derivative of $x^{p^n} - x$ is $p^n x^{p^n-1} - 1 = -1$, $x^{p^n} - x$ is relatively prime to its derivative and hence is separable, so the roots of $x^{p^n} - x$ are all distinct, which means that S contains exactly p^n elements. We claim that S is actually a field, which is the sought-after field of order p^n we want. If $\alpha, \beta \in S$, then

$$(\alpha\beta)^{p^n} = \alpha^{p^n}\beta^{p^n} = \alpha\beta$$

and

$$(\alpha + \beta)^{p^n} = \alpha^{p^n} + \beta^{p^n} = \alpha + \beta,$$

so $\alpha\beta, \alpha + \beta \in S$, which means S is closed under addition and multiplication. (The first step in the second line above comes from repeated application of the freshman's dream: for example, $(\alpha+\beta)^{p^2} = ((\alpha+\beta)^p)^p = (\alpha^p + \beta^p)^p = \alpha^{p^2} + \beta^{p^2}$, and similarly for larger powers of p.) If $\alpha^{p^n} = \alpha$, then taking reciprocals gives $(\alpha^{-1})^{p^n} = \alpha^{-1}$, so that S is also closed under taking inverses, and hence we conclude that S is a field as claimed.

We make one final note that doing computations with $\mathbb{F}_{p^n}$ (now unambiguously defined) described as the roots of $x^{p^n} - x$ in the algebraic closure of $\mathbb{F}_p$, or as the splitting field of $x^{p^n} - x$, is not so easy, since these descriptions alone do not give an obvious way to manipulate elements directly. In practice, if one is seeking to do concrete computations, one should instead first describe $\mathbb{F}_{p^n}$ as $\mathbb{F}_p[x]/(f(x))$ by finding an irreducible polynomial $f(x)$ of degree n over $\mathbb{F}_p$, and then work with elements in this quotient instead. Finding such a polynomial, or rather proving that your candidate is actually irreducible, is not easy, but nowadays various pieces of computing software can easily handle such computations.

Cyclotomic extensions. We have seen that the group of complex n^{th} roots of unity, often denoted by μ_n, is cyclic with generator

$$\zeta_n = e^{2\pi i/n} = \cos(\tfrac{2\pi}{n}) + i \sin(\tfrac{2\pi}{n}).$$

The other *primitive* roots of unity, i.e. those which can also be taken as generators of μ_n, can be described as ζ_n^a where a is relatively prime to n, since we saw back in the fall that for a cyclic group $G = \langle x \rangle$ of order n, the order of x^k in general is $n/(n,k)$, which is n if and only if $(n,k) = 1$. The field $\mathbb{Q}(\zeta_n)$ is the splitting field of $x^n - 1$ over $\mathbb{Q}$, and we refer to this field as the n^{th} *cyclotomic extension* of $\mathbb{Q}$. (Any other primitive n^{th} root of unity will generate the same field.)

The degree of this extension is the degree of the minimal polynomial $\phi_n(x)$ of ζ_n over $\mathbb{Q}$, and we call this $\phi_n(x)$ the n^{th} cyclotomic polynomial over $\mathbb{Q}$. We saw one special case of this last quarter, namely when $n = p$ is prime:

$$\phi_p(x) = x^{p-1} + x^{p-2} + \cdots + x + 1.$$

We showed this was irreducible using Eisenstein's Criterion on the shifted polynomial $\phi_p(x+1)$, and it has ζ_p as a root since

$$x^{p-1} + x^{p-2} + \cdots + x + 1 = \frac{x^p - 1}{x - 1}.$$

(Note that it is not at all obvious to see without this identity that $1 + \zeta_p + \zeta_p^2 + \cdots + \zeta_p^{p-1}$ is zero!) Thus $[\mathbb{Q}(\zeta_p) : \mathbb{Q}] = p - 1$ when p is prime. When n is not prime, we still have that ζ_n is a root of $x^{n-1} + \cdots + x^2 + x + 1$, but this polynomial is no longer irreducible.

Cyclotomic polynomials. We now give an explicit description of $\phi_n(x)$ in general. We claim that:

$$\phi_n(x) = \prod_{\substack{\text{primitive } n\text{-th} \\ \text{roots of unity } \zeta}} (x - \zeta) = \prod_{(a,n)=1} (x - \zeta_n^a),$$

where ζ_n in third expression is our usual primitive n^{th} root of unity. Note that ζ_n is a root of this polynomial, since $(\zeta_n - \zeta_n) = 0$ is one of the factors which occurs. What remains to be seen is that this actually has rational—in fact integer—coefficients (not obvious, since for now the coefficients appear to lie only in $\mathbb{Q}(\zeta_n)$) and that it is irreducible over $\mathbb{Q}$. It is these properties that will guarantee this definition of $\phi_n(x)$ matches the one we gave before, as the minimal polynomial of

ζ_n. We will prove irreducibility next time, but for now note that as a consequence we can say definitively what the degree of $\mathbb{Q}(\zeta_n)$ over $\mathbb{Q}$ is: it is the number of positive integers smaller than n that are relatively prime to n, or in other words $\varphi(n)$ where φ is the Euler phi-function.

Here are the first few examples. The only first root of unity is 1, so

$$\phi_1(x) = x - 1.$$

There are two second roots of unity, ± 1, but only -1 is primitive. (Indeed, 1 is a second root of unity which is also a first root of unity. In general, a d^{th} root of unity is also an n^{th} root of unity when $d \mid n$, but not a primitive one. The primitive ones are those which do not occur as roots of unity for any smaller exponent.) Thus

$$\phi_2 = x - (-1) = x + 1.$$

Since 3 is prime, $\phi_3(x) = x^2+x+1$. To see this via the new characterization above, take $\zeta_3 = e^{2\pi i/3}$, so that ζ_3 and ζ_3^2 are the primitive third roots of unity, and compute:

$$\phi_3(x) = (x - \zeta_3)(x - \zeta_3^2) = x^2 - (\zeta_3 + \zeta_3^2)x + \zeta_3^3 = x^2 + x + 1$$

where we use $\zeta_3^3 = 1$ and $1+\zeta_3+\zeta_3^2 = 0$. (Again recall that ζ_n always satisfies $1+\zeta_n+\cdots+\zeta^{n-1} = 0$.)

The fourth roots of unity are $\pm 1, \pm i$, and of these $\pm i$ are primitive, so

$$\phi_4(x) = (x - i)(x + i) = x^2 + 1.$$

The $n = 5$ case is prime, so $\phi_5(x) = x^4 + x^3 + x^2 + x + 1$, and for $n = 6$ we have two primitive roots of unity, so:

$$\phi_6(x) = (x - \zeta_6)(x - \zeta_6^5) = x^2 - (\zeta_6 + \zeta_6^5)x + \zeta_6^6.$$

Since

$$\zeta_6 = \cos(\tfrac{2\pi}{6}) + i\sin(\tfrac{2\pi}{6}) = \tfrac{1}{2} + i\tfrac{\sqrt{3}}{2} \text{ and } \zeta_6^5 = \cos(\tfrac{5\cdot 2\pi}{6}) + i\sin(\tfrac{5\cdot 2\pi}{6}) = \tfrac{1}{2} - i\tfrac{\sqrt{3}}{2},$$

we have $\zeta_6 + \zeta_6^5 = 1$, so $\phi_6(x) = x^2 - x + 1$.

Recursion and integer coefficients. Computing $\phi_n(x)$ using the factorization via primitive roots of unity above gets tedious quickly, but we can use these factorization to obtain a recursive method for computing these polynomials. As an example, consider $n = 6$, where the roots of unity, including the non-primitive ones, are $1, \zeta_6, \zeta_6^2, \zeta_6^3, \zeta_6^4, \zeta_6^5$. These are all the roots of $x^6 - 1$, so this factors as:

$$x^6 - 1 = (x - 1)(x - \zeta_6)(x - \zeta_6^2)(x - \zeta_6^3)(x - \zeta_6^4)(x - \zeta_6^5).$$

But among here we can find the cyclotomic polynomials for $n = 1, 2, 3,$ and 6 by grouping roots according to their status as primitive roots for possibly smaller exponents: 1 is the primitive first root of unity, $\zeta_6^3 = \zeta_2 - 1$ is the primitive second root of unity, $\zeta_6^2 = \zeta_3$ and $\zeta_6^4 = \zeta_3^2$ are the primitive third roots of unity, and ζ_6 and ζ_6^5 are the primitive sixth roots of unity, which means

$$x^6 - 1 = \phi_1(x)\phi_2(x)\phi_3(x)\phi_6(x).$$

If we know $\phi_1(x), \phi_2(x)$, and $\phi_3(x)$, then from this we can determine $\phi_6(x)$. More generally, by grouping roots in the same way for any n, we get:

$$x^n - 1 = \prod_{\substack{n\text{-th roots} \\ \text{of unity } \zeta}} (x - \zeta) = \prod_{d \mid n} \prod_{\substack{\text{primitive } d\text{-th} \\ \text{roots of unity } \zeta}} (x - \zeta) = \prod_{d \mid n} \phi_d(x)$$

which thus gives a recursive way to find $\phi_n(x)$ from the previous cyclotomic polynomials.

As a consequence of this recursion, we can now see inductively that $\phi_n(x)$ will always have integer coefficients. Suppose we know already that $\phi_d(x)$ has integer coefficients for $d \mid n, d \neq n$. Then

$$\phi_n(x) = \frac{x^n - 1}{\prod_{d|n, d\neq n} \phi_d(x)},$$

considered as a polynomial over $\mathbb{Q}(\zeta_n)$, is a quotient of two integer polynomials. If the denominator above did not divide $x^n - 1$ over $\mathbb{Q}$, then we would have

$$x^n - 1 = (\text{denominator})q(x) + r(x)$$

for some $q(x), r(x) \in \mathbb{Q}[x]$ with $r(x) \neq 0$ by the division algorithm, but this equality would then also hold in $\mathbb{Q}(\zeta_n)[x]$, so that the denominator would not divide $x^n - 1$ in $\mathbb{Q}(\zeta_n)[x]$ either. Since it does, the denominator does divide $x^n - 1$ in $\mathbb{Q}[x]$. By Gauss's Lemma (note all polynomials here are monic), this implies the denominator also divides $x^n - 1$ in $\mathbb{Z}[x]$, so $\phi_n(x) \in \mathbb{Z}[x]$ as claimed.

Lecture 9: Geometric Constructions

Warm-Up. We determine which finite field serves as the 9^{th} cyclotomic extension of $\mathbb{F}_5$. Now, the notion of a cyclotomic extension of a finite field is analogous to the one we gave last time for $\mathbb{Q}$: in this case, it is the smallest extension of $\mathbb{F}_5$ which contains a primitive 9^{th} root of unity, or equivalently the splitting field of $x^9 - 1$ over $\mathbb{F}_5$. A "primitive" 9^{th} root of unity in this context is an element of multiplicative order 9.

Any root of unity is necessarily nonzero, and so belongs to $\mathbb{F}_{5^n}^\times$, which is a cyclic group of order $5^n - 1$. (We proved last quarter that the group of units of any finite field is cyclic.) In order for this to contain an element of order 9 requires that 9 divide $5^n - 1$, and our work from the fall shows that this is condition sufficient as well. Thus we are looking for the smallest n such that $9 \mid 5^n - 1$. Checking increasing n by brute force shows that $n = 6$ is the first value that works, so we conclude that $\mathbb{F}_{5^6}$ is the smallest extension of $\mathbb{F}_5$ that contains a primitive 9^{th} root of unity. This is thus the 9^{th} cyclotomic extension of $\mathbb{F}_5$.

Irreducibility of cyclotomic polynomials. We now prove that $\phi_n(x)$, defined as the monic polynomial whose roots are precisely the primitive n^{th} roots of unity, is irreducible over $\mathbb{Q}$. (We argued last time using the recursive expression for these polynomials that $\phi_n(x) \in \mathbb{Z}[x]$ for all n.) This then justifies the claim that $\phi_n(x)$ is the minimal polynomial of ζ_n, so that $\mathbb{Q}(\zeta_n)$ has degree $\varphi(n)$ over $\mathbb{Q}$, where $\varphi(n)$ is the number of relatively prime positive integers less than n. The proof give here is essentially the same as the book's proof, only written in a (I think) clearer way.

Suppose $\phi_n(x) = f(x)g(x)$ for some $f(x), g(x) \in \mathbb{Q}[x]$ (in fact, since $\phi_n(x) \in \mathbb{Z}[x]$ we can assume by Gauss's Lemma that $f(x), g(x) \in \mathbb{Z}[x]$) and let ζ be a primitive n^{th} root of unity. (Note that ζ is not necessarily $\zeta_n = e^{2\pi i/n}$, although it is a power of ζ_n.) Then ζ is a root of $\phi_n(x)$, so it must be a root of either $f(x)$ or $g(x)$, so let us say that it is a root of $f(x)$. We claim that in fact then *all other* primitive n^{th} roots of unity are roots of $f(x)$ as well, which implies that $\phi_n(x) = f(x)$ (since $\phi_n(x)$ and $f(x)$ will have exactly the same roots), which implies that $\phi_n(x)$ is irreducible.

Suppose $p \nmid n$ (p prime) and consider the primitive n^{th} root of unity ζ^p. This is a root of $\phi_n(x)$, so it is a root of either $f(x)$ or $g(x)$, and by way of contradiction let us assume it is a root of $g(x)$. Then $g(\zeta^p) = 0$, which can be instead interpreted as saying that ζ is a root of the polynomial $g(x^p)$; i.e. plugging ζ^p into $g(x)$ is the same as plugging ζ into $g(x^p)$. Since ζ is now a root of both $f(x)$ and $g(x^p)$, both are divisible by $x - \zeta$, and hence these are not relatively prime over $\mathbb{Q}(\zeta)$. But we claim this implies they are also not relatively prime over $\mathbb{Q}$ (note that $x-\zeta$ would not work as a common factor over $\mathbb{Q}$ since $\zeta \notin \mathbb{Q}$, I guess unless $n=1,2$): if $f(x), g(x^p)$ *are* relatively prime over $\mathbb{Q}$, then

$$f(x)h(x)+g(x)\ell(x)=1$$

for some $h(x), \ell(x) \in \mathbb{Q}[x]$ by the Euclidean algorithm, but this same identity holds over $\mathbb{Q}(\zeta)$ as well, so that $f(x)$ and $g(x^p)$ would be relatively prime over $\mathbb{Q}(\zeta)$ too, which they are not.

Hence there is a common divisor $d(x) \in \mathbb{Z}[x]$ (not just in $\mathbb{Q}[x]$ but in $\mathbb{Z}[x]$ as well by Gauss's Lemma) of both $f(x)$ and $g(x^p)$ of positive degree, which we can take to be irreducible. (We can just take an irreducible factor of the gcd of $f(x)$ and $g(x^p)$.) Reducing coefficients mod p then shows that $\overline{d(x)}$ divides both $\overline{f(x)}$ and $\overline{g(x^p)}$ over $\mathbb{F}_p$. (Bars indicate the reductions.) But $\overline{g(x^p)} = \overline{g(x)}^p$ by the freshman's dream and the fact that any $a \in \mathbb{F}_p$ is a fixed point of Frobenius:

$$c_0 + c_1x^p + \cdots + c_n(x^p)^n = c_0^p + c_1^p x^p + \cdots + c_n^p (x^n)^p = (c_0 + c_1 x + \cdots + c_n x^n)^p.$$

Since $\overline{d(x)}$ is irreducible (hence a prime element of $\mathbb{F}_p[x]$) and divides $\overline{g(x)}^p$, it divides $\overline{g(x)}$ as well. But then, a root α of $\overline{d(x)}$ in an extension of $\mathbb{F}_p$ will be a root of both $\overline{f(x)}$ and $\overline{g(x)}$, hence a repeated root of $\overline{f(x)}\,\overline{g(x)} = \overline{\phi_n(x)}$, and hence a repeated root of x^n-1 over $\mathbb{F}_p$ because $\phi_n(x)$ divides x^n-1. However, x^n-1 is separable over $\mathbb{F}_p$ since its derivative nx^{n-1} is nonzero (because $p \nmid n$) and relatively prime to x^n-1, so it cannot have repeated roots. We thus conclude that the original ζ^p must have been a root of $f(x)$ and not $g(x)$ all along.

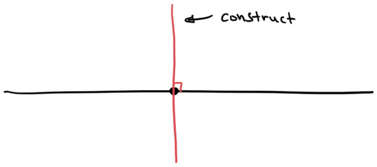
Thus if ζ is any primitive n^{th} root of unity which is a root of $f(x)$, then ζ^p is also a root of $f(x)$ for any prime $p \nmid n$. If p_1, p_2 are two primes not dividing n, then ζ^{p_1} is a root of $f(x)$, so $(\zeta^{p_1})^{p_2} = \zeta^{p_1p_2}$ is also a root of $f(x)$ by this reasoning since ζ^{p_1} is itself a primitive n^{th} root of unity. This new root is still primitive, so if p_3 is a third prime that does not divide n then

$$((\zeta^{p_1})^{p_2})^{p_3} = \zeta^{p_1p_2p_3}$$

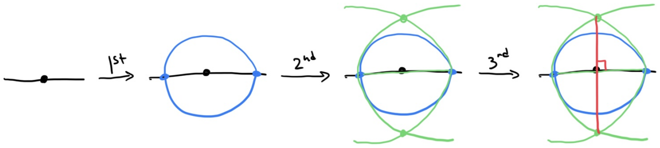
is also a root of $f(x)$. And so on, we conclude that for any primes p_i not dividing n, $\zeta^{p_1\cdots p_n}$ is a root of $f(x)$. All $a \in \mathbb{N}$ relatively prime to n are products of such primes, so we get that ζ^a is a root of $f(x)$ whenever $(a,n)=1$, which means that all primitive n^{th} roots of unity are roots of $f(x)$. As stated earlier, this means that $\phi_n(x) = f(x)$, so whenever $\phi_n(x) = f(x)g(x)$ we have that $\phi_n(x)$ equals one of the factors, so it is irreducible as claimed. (A lot of good stuff went into this argument!)

Geometric constructions. Before moving on to Galois theory, we discuss the topic of *straightedge and compass* constructions, focusing on how they can be formulated in a field-theoretic manner. The starting point is a straightedge, which is a tool we can use to draw straight lines, and a compass, which we can use to draw circles. (Note that the straightedge is not assumed to have any markings, so that it is not a ruler, just a literal straight edge.) The fundamental problem is that of determining what types of geometric objects can be constructed using these tools alone.

We will not delve too heavily into the process of actually *constructing* said objects, since our goal is to interpret them field-theoretically, so for the most part we will assume that the constructions we claim are possible are in fact possible. But, to get a sense for what we mean, we will show that given a line and point on it, we can construct a line passing through that point that is *perpendicular* to the given line:



Take the compass and draw any circle (radius doesn't matter) centered at the given point, and mark off the two points where our given line intersects the circle. Extend the legs of the compass to match the diameter of the circle, with endpoints to the points we just drew, and draw two circles with radius equal to this diameter, each centered at one of our two points:



Finally, mark off the two points where these two circles intersect each other, and then draw a line connecting them using the straightedge. This final line is that one we want: it is perpendicular to the original line, passes through the original point (see picture above), and was constructed using straightedge and compass alone.

Given a line and a point not on that line, by constructing one line perpendicular to the first line and then a third line perpendicular to the second line, we can also construct parallel lines using straightedge and compass. Other geometric objects which can be constructed in similar ways include equilateral triangles, squares, and bisected angles. The four historical problems considered in this subject are:

- *squaring the circle*: given a circle, can a square of area equal to that of the circle be constructed using straightedge and compass?
- *doubling the cube*: given a cube, can a cube of double the volume be constructed using straightedge and compass?
- *trisecting the angle*: given an angle, can an angle of a third the measure of the first be constructed using straightedge and compass? and
- *constructing polygons*: for which n can a regular n-gon be constructed using straightedge and compass?

The answers to these depends on the field theory we have developed, and, in the fourth case, the Galois theory we will soon develop. (We should note that these problems might seem like merely "cute" problems by today's standards, in that they might seem all that important as stated. But, to the ancient Greeks, this was all that mattered: they had no concept of "number" as the abstract notion we interpret it as today, and to them numbers only made sense as *lengths*, so that constructing things via straightedge and compass was how they actually interpreted "arithmetic".)

Constructible numbers. Given a line segment that we interpret as having length "1", we can determine which "numbers"—i.e. lengths—can be constructed from this using straightedge and compass. (From now, we will simply say "construct" instead of the full "construct using straightedge and compass".) For example, by putting this line segment next to itself (set the legs of the compass to be at the endpoints of this line segment, and use this as the radius of circle we can use to "copy" the first line segment) we get a segment of length "2". Then in a similar way we get 3, 4, and in fact any positive integer. It turns out that given line segments of lengths α and β, it is possible to construct a line segment of length α/β, so that we get all positive rational numbers as well.

We say that a positive real number α is *constructible* if it a line segment of length α can be obtained from the starting length 1 through a sequence of intermediate constructions. So, any positive rational is constructible. In general, it turns out (check the book or take MATH 340 to see how to actually do these things) that the set of constructible numbers is closed under addition, subtraction (as long as the result is positive), multiplication, division, and square root extractions, where for the last one we mean that if α is constructible, then $\sqrt{\alpha}$ is constructible. So, the set of constructible numbers is *almost* a field with no degree 2 extensions (such things required square roots not in the base field), except for the lack of negatives.

Now, to determine precisely which numbers are constructible requires thinking about the algebra that underlies straightedge and compass constructions. The possible line segments—and hence constructible numbers—we can obtain via these operations are those connecting points obtained by intersecting constructible lines with constructible lines, constructible lines with constructible circles, and constructible circles with constructible circles. (See the construction of perpendicular lines we gave before, for example. The constructible lines and circles are those whose defining data—slopes, radii, centers—are constructible.) Such lines and circles have equations of the form

$$ax + bx = c \quad \text{and} \quad (x-p)^2 + (y-q)^2 = r^2$$

where a, b, c, p, q, r are all constructible, and the point is that determining the intersections of these by substituting one thing to another leads to equations which are at most quadratic. Solving these equations produces new constructible numbers that are obtained from previous ones (a, b, c, d, p, q, and r) using the operations of addition, subtraction, multiplication, division, and square root extraction alone, so the conclusion is that α is constructible if and only if α can be obtained from 1 using a sequence of operations involving only addition, subtraction, multiplication, division, and square root extractions alone. For example,

$$\frac{\sqrt{\sqrt{3}+5-\sqrt{\sqrt{7}}}}{4+\sqrt{4+\sqrt{5+\sqrt{11}}}}$$

is constructible, and to actually construct it we would construct 7, then $\sqrt{7}$, then $\sqrt{\sqrt{7}}$, then $5-\sqrt{\sqrt{7}}$, and so on, working our way "outward" in both the numerator and denominator and then constructing the quotient.

Constructibility in terms of fields. The upshot is that if $\alpha \in \mathbb{R}$ is constructible, then in fact $\mathbb{Q}(\alpha)$ must be an extension of $\mathbb{Q}$ with degree a power of 2. This is something we saw in a previous example which asked about the impossibility of expressing $\sqrt[3]{2}$ in terms rational numbers and addition, subtraction, multiplication, division, and square root extractions. The point was that each additional square root (not already in the set we have so far) introduced in the construction increases the degree by exactly 2, so by the tower law the final result has degree 2^n where n denotes the number of root extractions which were necessary. For example,

$$\frac{\sqrt{\sqrt{3}+5-\sqrt{\sqrt{7}}}}{4+\sqrt{4+\sqrt{5+\sqrt{11}}}}$$

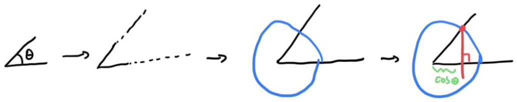
lies in a degree 2^7 extension of $\mathbb{Q}$. (The numerator needs four root extractions, and the denominator three, which are all different from the ones needed in the numerator.)

With this in mind, we can now answer some of our Greek construction questions. First, to "square" a circle of area πr^2 requires constructing of square of side length $r\sqrt{\pi}$. But $\sqrt{\pi}$ is transcendental over $\mathbb{Q}$ (if it was algebraic, π would be as well), so $\mathbb{Q}(\sqrt{\pi})$ has infinite degree over $\mathbb{Q}$ and hence $[\mathbb{Q}(\sqrt{\pi}) : \mathbb{Q}]$ is not a power of 2. Thus $\sqrt{\pi}$ is not constructible, so circles cannot be squared. Second, to double a cube of side length a and volume a^3 requires constructing a cube of side length $a\sqrt[3]{2}$. But $\sqrt[3]{2}$ is not constructible since $[\mathbb{Q}(\sqrt[3]{2}) : \mathbb{Q}] = 3$, so cubes cannot be doubled.

Angle trisection in general is not possible either (it is for certain angles), as can be shown by considering the trisection of a 60° angle: this angle is constructible (take an angle in a constructed equilateral triangle), but trisecting requires constructing a 20° angle, which is not possible. We will come back to this next time to see why, and will wrap things up by relating the construction of regular n-gons to the construction of primitive roots of unity and hence to cyclotomic extensions.

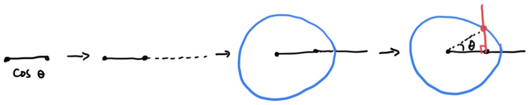
Lecture 10: Galois Groups

Warm-Up. We show that an angle θ is constructible using straightedge and compass—meaning that line segments intersecting at and angle θ can be constructed—if and only if $\cos\theta$ is a constructible real number. Take a construction of θ. Extend the line segments used in this construction so that they are longer than the base length "1", and draw a circle of radius 1 with center as the angle vertex:



Mark where this circle intersects one leg from the angle, and construct the perpendicular from this intersection to the other leg. (See above.) The length from the original vertex to where this perpendicular intersects this leg is then $\cos\theta$, so that $\cos\theta$ is constructible if θ is.

Conversely, given a length $\cos\theta$, extend it to have length longer than 1, and draw a circle of radius 1 centered at the other endpoint:



Construct the perpendicular to this line that passes through the non-center endpoint of the original line segment, and mark the intersect of this perpendicular with the circle. (See above.) Connect this point to the center of the circle and we have constructed θ.

Trisections and polygons. We can now argue that angle trisection by straightedge and compass is not always possible. This is not to say that it is *never* possible, since for example an angle of measure 180° *can* be trisected because we know that an angle of measure 60° (if we accept the fact that equilateral triangles are constructible) is constructible, but just that there are constructible angles that cannot be trisected. In particular, 60° is constructible, but we claim that its trisection 20° is not, or equivalently that $\cos 20°$ is not constructible. Indeed, using the trig identity:

$$\cos(3\alpha) = 4\cos^3\alpha - 3\cos\alpha$$

with $\alpha = 20°$ (this identity can be derived by taking the real parts in the complex-exponential equality $(e^{i\alpha})^3 = e^{3i\alpha}$), we can see that $\cos 20°$ is a root of $4x^3 - 3x - \frac{1}{2}$. Then $2\cos 20°$ is a root of $x^3 - 3x - 1$, which is irreducible over $\mathbb{Q}$. (No rational roots.) Thus $2\cos 20°$ generates a degree 3 extension of $\mathbb{Q}$, so it is not constructible, and thus neither is $\cos 20°$.

As for constructing a regular n-gon, the key observation is that a regular n-gon is constructible if and only if the "interior" angle $2\pi/n$ (in radians) is constructible. Indeed, if the polygon is constructible, drawing line segments from the vertices to the center will construct the angle $2\pi/n$, and conversely if this angle is constructible, constructing it n times in a row in a "circular" manner where each new angle is adjacent to the previous one and then drawing a circle will give the vertices of the n-gon, which we can then connect by line segments. So, n-gon construction is equivalent to construction of $\cos(2\pi/n)$. But this is the real part of the primitive root of unity $\zeta_n = e^{2\pi i/n}$, so the point is that the constructibility of the n-gon comes down to properties of the n-th cyclotomic extension $\mathbb{Q}(\zeta_n)$ over $\mathbb{Q}$. Galois theory will provide the final ingredient needed to understand this extension in full.

From field extensions to groups. We have seen how questions about "constructing", or more generally "expressing", numbers in a certain way can be turned into questions about field extensions; for example, expressing α in terms of rationals and $+, -, \cdot, \div, \sqrt{\ }$ turns into a question about obtaining $\mathbb{Q}(\alpha)$ from $\mathbb{Q}$ via intermediate quadratic extensions. Similarly, it should seem plausible now that the problem of expressing roots of polynomials via "nice" formulas can similarly be rephrased in terms of field extensions, but we will clarify this in detail later. Our goal now is to understand the final piece of this puzzle, which is the problem of turning questions about field extensions into questions about groups:

$$\text{constructions} \rightsquigarrow \text{field extensions} \rightsquigarrow \text{groups}.$$

The upshot is that we already know quite a lot about groups, and this will help to understand fields and their extensions more thoroughly.

Here is the key definition. Given a field extension E/F, we define the *automorphism group* of the extension to be the group $\operatorname{Aut}(E/F)$ of those (field) automorphisms of E which *fix* the base field F:

$$\operatorname{Aut}(E/F) := \{\sigma \in \operatorname{Aut}(E) \mid \sigma(a) = a \text{ for all } a \in F\}.$$

We also say that $\sigma \in \operatorname{Aut}(E/F)$ is an automorphism of E *over* F. This is a subgroup of the full automorphism group $\operatorname{Aut}(E)$ of E (the group operation is composition), but is one that attempts to encode information about the base field somehow: not all automorphisms of E will behave in any particularly special way with regard to F, so we take F into consideration by requiring that automorphisms fix the elements of F. In some sources this group is already called the *Galois group* of E/F, but the phrase "Galois group" is more commonly only used in a special case to be clarified in a bit, and we will do so as well.

Examples and root preservation. An automorphism σ of $\mathbb{Q}(\sqrt{D})$ over $\mathbb{Q}$ (where D is a non-square in $\mathbb{Q}$) is determined by its effect on $\sqrt{D}$ since it is required to fix the elements of $\mathbb{Q}$: for $a, b \in \mathbb{Q}$, we have

$$\sigma(a + b\sqrt{D}) = \sigma(a) + \sigma(b)\sigma(\sqrt{D}) = a + b\sigma(\sqrt{D}).$$

At first glance $\sigma(\sqrt{D})$ is an element of $\mathbb{Q}(\sqrt{D})$, and so is of the form $\sigma(\sqrt{D}) = c + d\sqrt{D}$, but there are other restrictions on what this element can actually be. In particular, since $\sqrt{D}$ satisfies $(\sqrt{D})^2 - D = 0$, we can see that $\sigma(\sqrt{D})$ must also satisfy the same equation:

$$(\sqrt{D})^2 - D = 0 \implies \sigma((\sqrt{D})^2 - D) = \sigma(0) \implies (\sigma(\sqrt{D}))^2 - D = 0.$$

That is, $\sqrt{D}$ is a root of $x^2 - D \in \mathbb{Q}[x]$ and $\sigma(\sqrt{D})$ must then be a root of the same polynomial. This gives only two possibilities: $\sigma(\sqrt{D}) = \sqrt{D}$ (which gives the identity automorphism overall) and $\sigma(\sqrt{D}) = -\sqrt{D}$, so $\mathrm{Aut}(\mathbb{Q}(\sqrt{D})/\mathbb{Q})$ is a group of order 2 and is thus isomorphic to $\mathbb{Z}/2\mathbb{Z}$.

In general, if $p(x) \in F[x]$, then any $\sigma \in \mathrm{Aut}(E/F)$ must sends roots of $p(x)$ to roots of $p(x)$, and moreover gives a bijection (i.e. a permutation) of the roots contained in E: if $\alpha \in E$ satisfies

$$c_0 + c_1\alpha + \cdots + c_n\alpha^n = 0 \text{ for } c_i \in F,$$

then applying $\sigma \in \mathrm{Aut}(E/F)$ gives

$$c_0 + c_1\sigma(\alpha) + \cdots + c_n\sigma(\alpha)^n = 0,$$

where we use the fact that the elements c_i of F are fixed under σ. The fact that *all* the roots of $p(x)$ are permuted in this way is a consequence of the fact that there are finitely many roots: σ gives an injective map from the set of roots to itself, so since this set is finite this map must actually be bijective. These observations will make computing elements of $\mathrm{Aut}(E/F)$ simple—although still possibly nontrivial—in many examples. (Note for later that if $p(x)$ has m roots in E, then this observation produces a homomorphism $\mathrm{Aut}(E/F) \to S_m$, where S_m is a symmetric group, by sending each automorphism to the bijection it induces on the roots.)

As a second example, we compute $\mathrm{Aut}(\mathbb{Q}(\sqrt[3]{2})/\mathbb{Q})$. Again any $\sigma \in \mathbb{Q}(\sqrt[3]{2})$ is determined by its value on a generator, which here we can take to be $\sqrt[3]{2}$. Since $\sqrt[3]{2}$ is a root of $x^3 - 2 \in \mathbb{Q}[x]$, $\sigma(\sqrt[3]{2})$ must be a root of $x^3 - 2$ as well. But the only root of $x^3 - 2$ contained in $\mathbb{Q}(\sqrt[3]{2})$ is $\sqrt[3]{2}$ itself, since the other two roots are non-real complex. (They are $\zeta_3\sqrt[3]{2}$ and $\zeta_3^2\sqrt[3]{2}$ where ζ_3 is a primitive complex third root of unity.) Thus the only possibility is $\sqrt{(\sqrt[3]{2})} = \sqrt[3]{2}$, which forces σ to be the identity map. Hence $\mathrm{Aut}(\mathbb{Q}(\sqrt[3]{2})/\mathbb{Q}) = \{id\}$ is the trivial group.

Fixed fields. This final example $\mathrm{Aut}(\mathbb{Q}(\sqrt[3]{2})/\mathbb{Q}) = \{id\}$ shows that some caution should be taken if we are trying to use these groups to study field extensions, since in this case the automorphism group of the extension is not actually able to detect the extension, or more precisely the *base field*, at all: considering $\mathbb{Q}(\sqrt[3]{2})$ as an extension of itself also produces a trivial automorphism group, and so this group cannot tell exactly what we are considering $\mathbb{Q}(\sqrt[3]{2})$ to be an extension of. We would like for $\mathrm{Aut}(E/F)$ to truly detect F so that we know definitively we are considering E as an extension of F and not of something else, and the upshot is that not all field extensions will be suitable in this regard. What we really want is to consider those extensions where the *only* elements of E that are fixed by all elements of $\mathrm{Aut}(E/F)$ are those of F.

Given a subgroup H of the full field automorphism group $\mathrm{Aut}(E)$, we defined the *fixed field* E^H of H to be the set of elements of E fixed by all elements of H:

$$E^H := \{a \in E \mid \sigma(a) = a \text{ for all } \sigma \in H\}.$$

This is in fact a subfield of E since the fact that automorphisms preserve addition and multiplication implies that E^H is closed under the field operations of E, so we can thus consider E to be an extension of E^H. Now, here are two observations which hint at the deeper relation between field extensions and groups we are working towards. First, if $H_1 \leq H_2 \leq \mathrm{Aut}(E)$ are two subgroups of $\mathrm{Aut}(E)$, one contained in the other, then anything fixed by all elements of H_2 is necessarily fixed by all elements of H_1 as well, so $E^{H_2} \subseteq E^{H_1}$. Thus, subgroups of automorphism groups correspond to subextensions of fields, only with the order reversed. Second, if $F_1 \subseteq F_2 \subseteq E$ is a tower of extensions, then anything that fixes all of F_2 will also fix all of F_1, so $\mathrm{Aut}(E/F_2) \leq \mathrm{Aut}(E/F_1)$. Thus, subextensions of fields correspond to subgroups of the automorphism group, again with the order reversed.

This is good stuff (!), and as stated before hints at a deep relation between these two types of objects. For example, we can ask things like: to what types of field extensions do *normal* subgroups correspond? What do quotient groups then correspond to? Can we *classify* all possible subextensions of a given field extension by classifying the corresponding groups instead? Galois theory will give us all the answers we need.

Galois groups and extensions. If we truly want $\mathrm{Aut}(E/F)$ to detect the actual fact that E is an extension of F in particular, then we should require that the fixed field of E corresponding to this full group is F itself. We say that E is a *Galois extension* of F when this is the case, so E/F is Galois if $E^{\mathrm{Aut}(E/F)} = F$. Thus, for example, $\mathbb{Q}(\sqrt{D})$ is a Galois extension of $\mathbb{Q}$, but $\mathbb{Q}(\sqrt[3]{2})$ is not a Galois extension of $\mathbb{Q}$. When E is a Galois extension of F, then we call $\mathrm{Aut}(E/F)$ the *Galois group* of the extension, and denote it by $\mathrm{Gal}(E/F)$ instead. (As said earlier, some sources use "Galois group" and the notation $\mathrm{Gal}(E/F)$ for arbitrary extensions, but we will reserve these only for the case of an extension which is actually Galois.)

We should note that this definition of "Galois extension" is one of many equivalent ones that can be given, and indeed the book gives a different definition first in terms of the relation between $\mathrm{Aut}(E/F)$ and the degree $[E : F]$. The Galois property is also equivalent to saying that E is the splitting field of a separable polynomial over F, which using other terminology is the same as saying that the extension is *normal* and *separable.* We have chosen to give the fixed field definition first since it highlights the idea that we want elements of the group to detect the actual *extension* including the base field, but we will show the equivalence of this with other definitions soon enough.

Lecture 11: More on Galois Groups

Warm-Up 1. We compute $\mathrm{Aut}(\mathbb{Q}(\sqrt{2},\sqrt{3})/\mathbb{Q})$ and determine if $\mathbb{Q}(\sqrt{2},\sqrt{3})$ is a Galois extension of $\mathbb{Q}$. Since $\sqrt{2}$ and $\sqrt{3}$ generate this extension, we know that any $\sigma \in \mathrm{Aut}(\mathbb{Q}(\sqrt{2},\sqrt{3})/\mathbb{Q})$ is determined by its values on $\sqrt{2}$ and $\sqrt{3}$. Furthermore, since these two are roots of $(x^2-2)(x^2-3)$ over $\mathbb{Q}$, we know that any σ must permute them among the other roots, which means that the only possibilities for $\sigma(\sqrt{2})$ and $\sigma(\sqrt{3})$ are $\pm\sqrt{2}, \pm\sqrt{3}$. At first glance this appears to give 12 possible choices for σ: pick one of four values for $\sigma(\sqrt{2})$, and then pick one of the three remaining values for $\sigma(\sqrt{3})$.

But there are some more restrictions we can derive, since $\sqrt{2}$ is *also* a root of x^2-2 alone, and $\sqrt{3}$ of x^3-2. The point is that *all* polynomial equations must be preserved, not only the one whose roots are the given generators. This means that σ must permute $\sqrt{2}$ among the roots of x^2-2, and similarly for $\sqrt{3}$ and x^3-2, so that the only possibilities are actually:

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

This gives four such automorphisms, so $\mathrm{Aut}(\mathbb{Q}(\sqrt{2},\sqrt{3})/\mathbb{Q})$ is a group of order 4. If we denote by σ_1 the element that exchanges $\sqrt{2}$ and $-\sqrt{2}$ but fixes $\sqrt{3}$, and by σ_2 the element fixing $\sqrt{2}$ and exchanging $\sqrt{3}$ with $-\sqrt{3}$, then the four elements are $1, \sigma_1, \sigma_2$, and $\sigma_1\sigma_2$. In fact, we have

$$\mathrm{Aut}(\mathbb{Q}(\sqrt{2},\sqrt{3})/\mathbb{Q}) = \langle\sigma_1\rangle\langle\sigma_2\rangle \cong \langle\sigma_1\rangle \times \langle\sigma_2\rangle \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z},$$

where each $\mathbb{Z}/2\mathbb{Z}$ factor keeps track of what is happening to each generator: leave it alone, or send to its negative. (We're using group-theoretic notation here, so that $\langle\sigma_1\rangle\langle\sigma_2\rangle$ denotes the product of two cyclic subgroups in the sense of the fall quarter.)

To determine if this extension is Galois we need to compute the fixed field of this entire automorphism group. Elements of this extension look like

$$a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6} \text{ with } a, b, c, d \in \mathbb{Q}.$$

To determine which such elements are fixed by all $\sigma \in \mathrm{Aut}(\mathbb{Q}(\sqrt{2}, \sqrt{3})/\mathbb{Q})$, it is enough to check which are fixed by the generators σ_1 and σ_2. Applying σ_1 gives

$$a + b\sigma_1(\sqrt{2}) + c\sigma_1(\sqrt{3}) + d\sigma_1(\sqrt{6}) = a - b\sqrt{2} + c\sqrt{3} - d\sqrt{6},$$

where we use the fact that $\sqrt{6} = \sqrt{2}\sqrt{3}$ in order to compute $\sigma_1(\sqrt{6})$. This forces $b = 0$ and $d = 0$ for an element fixed by σ_1, and then to also be fixed by σ_2 forces $c = 0$:

$$a + c\sqrt{3} = \sigma_2(a + c\sqrt{3}) = a - c\sqrt{3} \iff c = 0.$$

Thus the only elements of the extension fixed by all elements of the automorphism group are those in $\mathbb{Q}$. The extension $\mathbb{Q}(\sqrt{2}, \sqrt{3})/\mathbb{Q}$ is hence Galois, and its Galois group is $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$.

Note that proper subgroups of the full Galois group can have larger fixed fields. For example, the composite map $\sigma_1\sigma_2$ sends both of $\sqrt{2}$ and $\sqrt{3}$ to their negatives, and this ends up fixing $\sqrt{6}$:

$$\sigma_1\sigma_2(\sqrt{6}) = \sigma_1\sigma_2(\sqrt{2}\sqrt{3}) = [\sigma_1\sigma_2(\sqrt{2})][\sigma_1\sigma_2(\sqrt{3})] = (-\sqrt{2})(-\sqrt{3}) = \sqrt{6}.$$

The fixed field of the subgroup $\langle \sigma_1\sigma_2 \rangle$ is thus $\mathbb{Q}(\sqrt{6})$. (It is only after imposing the requirement that σ_1 and σ_2 individually fix field elements that we reduce to a fixed field of $\mathbb{Q}$.) This hints at the way in which subextensions of $\mathbb{Q}(\sqrt{2}, \sqrt{3})$ over $\mathbb{Q}$ can be extracted from subgroups of the full Galois group, a fact which will be clarified in the *Fundamental Theorem of Galois Theory* later.

Warm-Up 2. We do the same thing as above for the extension $\mathbb{Q}(\sqrt[4]{2})$ of $\mathbb{Q}$. An element σ of the automorphism group $\mathrm{Aut}(\mathbb{Q}(\sqrt[4]{2})/\mathbb{Q})$ must send $\sqrt[4]{2}$ to one of the following roots of $x^4 - 2$ in $\mathbb{C}$:

$$\sqrt[4]{2},\ i\sqrt[4]{2},\ -\sqrt[4]{2},\ -i\sqrt[4]{2}.$$

(Note i is a primitive fourth root of unity.) But the only such roots that lie in $\mathbb{Q}(\sqrt[4]{2})$ are $\pm\sqrt[4]{2}$ since the other two are not real. Thus there are only two possibilities for σ: the identity, and $\sqrt[4]{2} \mapsto -\sqrt[4]{2}$. Hence $\mathrm{Aut}(\mathbb{Q}(\sqrt[4]{2})/\mathbb{Q}) \cong \mathbb{Z}/2\mathbb{Z}$.

If $\sigma(\sqrt[4]{2}) = -\sqrt[4]{2}$, then we have:

$$\sigma(\sqrt[4]{4}) = \sigma(\sqrt[4]{2}\,\sqrt[4]{2}) = \sigma(\sqrt[4]{2})\sigma(\sqrt[4]{2}) = (-\sqrt[4]{2})(-\sqrt[4]{2}) = \sqrt[4]{4}$$

and

$$\sigma(\sqrt[4]{8}) = \sigma(\sqrt[4]{2}\,\sqrt[4]{4}) = \sigma(\sqrt[4]{2})\sigma(\sqrt[4]{4}) = (-\sqrt[4]{2})(\sqrt[4]{4}) = -\sqrt[4]{8}.$$

An element in this extension is a linear combination of 1, $\sqrt[4]{2}$, $\sqrt[4]{4} = \sqrt{2}$, and $\sqrt[4]{8}$, so we see that the elements fixed by the entire automorphism group are those of the form $a + b\sqrt[4]{4}$ with $a, b \in \mathbb{Q}$, so the fixed field is $\mathbb{Q}(\sqrt[4]{4}) = \mathbb{Q}(\sqrt{2})$. Thus $\mathbb{Q}(\sqrt[4]{2})$ is not Galois over $\mathbb{Q}$. (Note, however, that $\mathbb{Q}(\sqrt[4]{2})$ *is* Galois over $\mathbb{Q}(\sqrt[4]{4}) = \mathbb{Q}(\sqrt{2})$ and $\mathrm{Gal}(\mathbb{Q}(\sqrt[4]{2})/\mathbb{Q}(\sqrt{2})) \cong \mathbb{Z}/2\mathbb{Z}$.)

Orders and degrees. As a first step towards attaining a better understanding of the structure of automorphism/Galois groups, and in particular of restrictions we can use to make the determination of these groups less labor intensive, we have the following relation between the order of an automorphism group and the degree of the given field extension:

> If E is the splitting field of a polynomial $p(x) \in F[x]$, then $|\mathrm{Aut}(E/F)| \leq [E : F]$. (So, the degree bounds the order.) If moreover $p(x)$ is separable over F, then we have equality: $|\mathrm{Aut}(E/F)| = [E : F]$.

The condition that $|\operatorname{Aut}(E/F)| = [E : F]$ is, as we'll see, one of the equivalent ways of saying what it means for E/F to be Galois, and is how the book first defines the notion of a Galois extension. To say that E is the splitting field of a polynomial over F is the book's definition of what it means for E to be normal over F, which we alternatively defined as an extension in which any irreducible polynomial that has a root splits completely. (A problem on the Discussion 2 Problems sheet proves that these two notions of "normal" are equivalent.) Thus, said another way, the claim is that if E is a (finite) normal extension of F, then $|\operatorname{Aut}(E/F)| \leq [E : F]$, and if E is also separable over F, then $|\operatorname{Aut}(E/F)| = [E : F]$. The first inequality is in fact true without the normality/splitting field assumption, but we'll save this general case for later.

As a quick sanity check, let us see what this looks like in the examples we've done so far: $\operatorname{Gal}(\mathbb{Q}(\sqrt{D})/\mathbb{Q}) \cong \mathbb{Z}/2\mathbb{Z}$ (D not a square) has order 2, which agrees with the degree $[\mathbb{Q}(\sqrt{D}) : \mathbb{Q}]$; $\operatorname{Gal}(\mathbb{Q}(\sqrt{2}, \sqrt{3})\mathbb{Q}) \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ has order 4, which also agrees with the degree $[\mathbb{Q}(\sqrt{2}, \sqrt{3}) : \mathbb{Q}]$; and $\operatorname{Aut}(\mathbb{Q}(\sqrt[4]{2})/\mathbb{Q}) \cong \mathbb{Z}/2\mathbb{Z}$ has order 2, which is strictly less than the degree $[\mathbb{Q}(\sqrt[4]{2}) : \mathbb{Q}] = 4$. (Of course, this final strict inequality reflects the fact that $\mathbb{Q}(\sqrt[4]{2})$ is not Galois over $\mathbb{Q}$.) Also, the example $|\operatorname{Aut}(\mathbb{Q}(\sqrt[3]{2})/\mathbb{Q})| = 1$ from last time also works here, since $\mathbb{Q}(\sqrt[3]{2})$ has degree 3 over $\mathbb{Q}$.

Proof of order/degree relation. To prove the claim above, we argue by induction on the degree $[E : F]$. In the base case $[E : F] = 1$, we have $E = F$ and so $\operatorname{Aut}(E/E)$ is trivial, and thus has order at most (in fact equal to) $[E : E] = 1$. Now, if $[E : F] > 1$, pick an irreducible factor $q(x)$ of $p(x)$ and a root $\alpha \in E$ of $q(x)$. (Note $\alpha \notin F$.) For any $\sigma \in \operatorname{Aut}(E/F)$, $\sigma(\alpha)$ is also a root of $q(x)$, and σ restricts to an automorphism

$$\tau : F(\alpha) \to F(\text{root})$$

that sends α to the root. The point is that to count the possible $\sigma \in \operatorname{Aut}(E/F)$, we can proceed via a two-step process: first count the possible such τ, which are isomorphisms between $F(\alpha)$ and other fields obtained by adjoining a root of $q(x)$ to F, and then count the number of ways of extending such a τ up to a full automorphism $\sigma : E \to E$, whose restriction to $F(\alpha)$ is the given τ:

$$|\operatorname{Aut}(E/F)| = (\text{number of } E \to E \text{ extending a given } \tau)(\text{number of possible } \tau\text{'s}).$$

Visually, the idea is that the possible σ's fit into the following diagram:

$$\begin{tikzcd}
E \arrow[rr, "\sigma"] \arrow[d, dash] & & E \arrow[d, dash] \\
F(\alpha) \arrow[rr, "\tau"] \arrow[dr, dash] & & F(\text{root}) \arrow[dl, dash] \\
& F &
\end{tikzcd}$$

and we can "construct" σ by first lifting up from the base field F up to $\tau = \sigma|_{F(\alpha)} : F(\alpha) \to F(\text{root})$, and then lifting further up to $\sigma : E \to E$.

Now, since τ is completely determined by sending α to a root of $q(x)$, the number of such τ is the number of distinct roots of $q(x)$, which is bounded by the degree of $q(x)$ itself:

$$\text{number of possible } \tau = \text{number of distinct roots of } q(x) \leq \deg q(x).$$

If $q(x)$, and hence $p(x)$ of which $q(x)$ is a factor, is separable and thus has $\deg q(x)$ distinct roots in E, then we have equality here. Since $q(x)$ is irreducible over F, the degree of $F(\alpha)$ over F is precisely $\deg q(x)$, so we get

$$\text{number of } \tau \leq \deg q(x) = [F(\alpha) : F]$$

with equality if $q(x)$ (or $p(x)$) is separable.

Next we consider the problem of lifting a given τ. We can use such a $\tau : F(\alpha) \to F(\text{root})$ to identity $F(\alpha)$ with $F(\text{root})$, so that the lifting problem becomes that of extending the *identity* map $id : F(\alpha) \to F(\alpha)$ to $\sigma : E \to E$. (To be clearer, the point is that the number of σ which extend the given τ is the same as the number of σ extending the identity on $F(\alpha)$. This is because given two σ and σ' extending the same τ, we have that $\sigma^{-1}\sigma' : E \to E$ extends $\tau^{-1}\tau = id : F(\alpha) \to F(\alpha)$, and given σ extending the identity $F(\alpha) \to F(\alpha)$ and σ' extending τ, $\sigma'\sigma : E \to E$ also extends $\tau \circ id = \tau$. This sets up a one-to-one correspondence between σ extending τ and σ extending the identity on $F(\alpha)$.) But the $\sigma : E \to E$ which extend the identity on $F(\alpha)$ are precisely the elements of $\operatorname{Aut}(E/F(\alpha))$, so the number of $E \to E$ extending a given τ is $|\operatorname{Aut}(E/F(\alpha))|$. We thus have

$$|\operatorname{Aut}(E/F)| = |\operatorname{Aut}(E/F(\alpha))|(\deg q(x)) \leq |\operatorname{Aut}(E/F(\alpha))|[F(\alpha) : F],$$

with equality if $p(x)$ is separable.

Since $[E : F(\alpha)] < [E : F]$ by the tower law and because we can still view E as the splitting field of $p(x)$ only now over the base field $F(\alpha)$, we may assume by induction (on the extension degree) that the conclusion of our claim holds for the extension $E/F(\alpha)$. Thus $|\operatorname{Aut}(E/F(\alpha))| \leq [E : F(\alpha)]$ with equality if $p(x)$ is separable. Putting this all together gives

$$|\operatorname{Aut}(E/F)| \leq |\operatorname{Aut}(E/F(\alpha))|[F(\alpha) : F] \leq [E : F(\alpha)][F(\alpha) : F]$$

with equality if $p(x)$ is separable. Since $[E : F(\alpha)][F(\alpha) : F] = [E : F]$, we have our desired result.

Why separability? As a quick example, we illustrate why separability is an important condition to require in the theory we're building up. Certainly the equality $|\operatorname{Aut}(E/F)| = [E : F]$ in the result above might not hold without it (we technically only proved that separability implies this equality, but not that this equality implies separability), but we can also see what can go wrong in the fixed field definition of Galois extension.

For p prime, we can consider the extension $\mathbb{F}_p(x)(\sqrt[p]{x})$ of $\mathbb{F}_p(x)$, which as we've seen is the splitting field of the irreducible polynomial $X^p - x$. (So, $\mathbb{F}_p(x)(\sqrt[p]{x})$ is literally defined to be $\mathbb{F}_p[x]/(X^p - x)$.) This extension is normal since it is a splitting field extension, but it is not separable since $X^p - x$ does not have distinct roots: since

$$X^p - x = (X - \sqrt[p]{x})^p$$

by the freshman's dream, $\sqrt[p]{x}$ is the only root and it thus has multiplicity p. The point is that an element of $\operatorname{Aut}(\mathbb{F}_p(x)(\sqrt[p]{x}), \mathbb{F}_p(x))$ must send roots to roots, so the only option is to send $\sqrt[p]{x}$ to itself, which means that any such automorphism fixes everything. Thus $\operatorname{Aut}(\mathbb{F}_p(x)(\sqrt[p]{x}), \mathbb{F}_p(x))$ is trivial with fixed field all of $\mathbb{F}_p(x)(\sqrt[p]{x})$, so the extension is not Galois. In this case, the size of the automorphism group is strictly less than the degree of the extension: $1 < p$. The upshot is that without separability there aren't enough roots to permute in order to get a good Galois group!

Galois groups of finite fields. We finish by giving a first application of the result proved above. We claim that any finite extension $\mathbb{F}_{p^n}$ of $\mathbb{F}_p$ is Galois, and that its Galois group over $\mathbb{F}_p$ is straightforward to describe explicitly. Since $\mathbb{F}_{p^n}$ is normal and separable over $\mathbb{F}_p$—it is the splitting field of the separable polynomial $x^{p^n} - x \in \mathbb{F}_p[x]$, as we've seen before—the result above gives

$$|\operatorname{Aut}(\mathbb{F}_{p^n}/\mathbb{F}_p)| = [\mathbb{F}_{p^n} : \mathbb{F}_p] = n.$$

We claim that this group is actually cyclic, which we can verify by finding an element of order n.

But we have seen this element before: it is the Frobenius map $\sigma : a \mapsto a^p$. This is an automorphism of $\mathbb{F}_{p^n}$, and we showed as a Warm-Up previously that its fixed field is precisely $\mathbb{F}_p \subseteq \mathbb{F}_{p^n}$, so it is indeed an element of $\operatorname{Aut}(\mathbb{F}_{p^n}/\mathbb{F}_p)$. If this element has order $k < n$, then $\sigma^k = id$, which means that

$$\sigma^k(a) = a \text{ for all } a \in \mathbb{F}_{p^n}.$$

But $\sigma^k(a) = a^{p^k}$ since each application of σ takes another p^{th} power, so saying that $\sigma^k(a) = a$ is the same as saying that a is a root of $x^{p^k} - x$, which has p^k roots. Since $\mathbb{F}_{p^n}$ has p^n elements, it is not possible for $\sigma^k(a) = a$ for all $a \in \mathbb{F}_{p^n}$ if $k < n$, so we conclude that the smallest power of σ which is the identity map is the n^{th} power. Thus σ has order n and thus generates all of $\operatorname{Aut}(\mathbb{F}_{p^n}/\mathbb{F}_p)$ as claimed. Moreover, since the fixed field of σ alone is $\mathbb{F}_p$, the fixed field of all of $\operatorname{Aut}(\mathbb{F}_{p^n}/\mathbb{F}_p)$ is also $\mathbb{F}_p$ since any automorphism here is a power of σ. Thus $\mathbb{F}_{p^n}$ is a Galois extension of $\mathbb{F}_p$, with cyclic Galois group generated by Frobenius. The moral of the story is that finite fields are simple Galois-theoretic objects to study!

Lecture 12: Fundamental Theorem of Galois Theory

Warm-Up. We define the *Galois group* of a separable polynomial over a field F to be the Galois group of its splitting field. (For now we're taking it for granted that the various ways of defining what it means for an extension to be "Galois" are equivalent, and being the splitting field of a separable polynomial is one of them.) Let us determine the Galois group of $x^3 - 2$ over $\mathbb{Q}$. As we have seen before, the splitting field is $\mathbb{Q}(\sqrt[3]{2}, \zeta_3)$ and is of degree 6 over $\mathbb{Q}$. (The splitting field is the composite of $\mathbb{Q}(\sqrt[3]{2})$ and $\mathbb{Q}(\zeta_3)$, so it has degree at most $3 \cdot 2 = 6$—recall that the minimal polynomial of ζ_3 is $\phi_3(x) = x^2 + x + 1$—but at the same time the degree is divisible by the degrees 3 and 2 of the subextensions $\mathbb{Q}(\sqrt[3]{2})$ and $\mathbb{Q}(\zeta_3)$, so the degree is exactly 6.)

Thus by the result from last time, since $x^3 - 2$ is separable over $\mathbb{Q}$ we have

$$\operatorname{Gal}(\mathbb{Q}(\sqrt[3]{2}, \zeta_3)/\mathbb{Q}) = [\mathbb{Q}(\sqrt[3]{2}, \zeta_3) : \mathbb{Q}] = 6.$$

Now, we get a map $\operatorname{Gal}(\mathbb{Q}(\sqrt[3]{2}, \zeta_3)/\mathbb{Q}) \to S_3$ by having each automorphism act on the roots of $x^3 - 2$ by permutation. Moreover, this map is injective since if σ acts as the identity permutation on the roots it must be the identity automorphism on $\operatorname{Gal}(\mathbb{Q}(\sqrt[3]{2}, \zeta_3)/\mathbb{Q})$: if σ fixes all the roots, it fixes all elements generated by the roots, but this is the entirety of $\mathbb{Q}(\sqrt[3]{2}, \zeta_3)$ since the roots generate the splitting field. Thus $\operatorname{Gal}(\mathbb{Q}(\sqrt[3]{2}, \zeta_3)/\mathbb{Q})$ is isomorphic to a subgroup of S_3 of order 6, so the only possibility is $\operatorname{Gal}(\mathbb{Q}(\sqrt[3]{2}, \zeta_3)/\mathbb{Q}) \cong S_3$, which is thus the Galois group of $x^3 - 2$ over $\mathbb{Q}$.

But, we can see more explicitly that the Galois group should be S_3 by determining what its elements are concretely in terms of cycle notation. Let us denote by "1", "2", and "3" the roots $\sqrt[3]{2}$, $\zeta_3 \sqrt[3]{2}$, and $\zeta_3^2 \sqrt[3]{2}$ (in that order) of $x^3 - 2$. An element of the Galois group will permute these amongst themselves, *and* will permute ζ_3 amongst the roots of $\phi_3(x) = x^2 + x + 1$ (i.e. the primitive third roots of unity), of which the only additional root is ζ_3^2. Thus, to start, any element of the Galois group will send ζ_3 either to itself or to ζ_3^2.

Let us determine first the possibilities where $\zeta_3 \mapsto \zeta_3$. The root 1 (not the number 1, but the root we've labeled "1") can be sent to either root 2 or root 3. If $1 \mapsto 1$, then we can work out that

2 and 3 are fixed as well:

$$\zeta_3\sqrt[3]{2} \mapsto (\text{image of } \zeta_3)(\text{image of } \sqrt[3]{2}) = \zeta_3\sqrt[3]{2} \quad \text{and} \quad \zeta_3^2\sqrt[3]{2} \mapsto (\zeta_3^2)(\sqrt[3]{2}) = \zeta_3^2\sqrt[3]{2}.$$

Thus in this case we get the identity automorphism $(1) \in S_3$. If $1 \mapsto 2$, then

$$\zeta_3\sqrt[3]{2} \mapsto (\zeta_3)(\zeta_3\sqrt[3]{2}) = \zeta_3^2\sqrt[3]{2} \quad \text{and} \quad \zeta_3^2\sqrt[3]{2} \mapsto (\zeta_3^2)(\zeta_3\sqrt[3]{2}) = \zeta_3^3\sqrt[3]{2} = \sqrt[3]{2},$$

so 2 is sent to 3 and 3 to 1. Hence this element of the Galois group gives $(123) \in S_3$. Finally, if $1 \mapsto 3$, then we can work out that $3 \mapsto 2$ and $2 \mapsto 1$, so this element is $(132) \in S_3$.

Now we consider the possibilities where $\zeta_3 \mapsto \zeta_3^2$. If $1 \mapsto 1$, we have:

$$\zeta_3\sqrt[3]{2} \mapsto (\zeta_3^2)(\sqrt[3]{2}) = \zeta_3^2\sqrt[3]{2} \quad \text{and} \quad \zeta_3^2\sqrt[3]{2} \mapsto (\zeta_3^2)^2(\sqrt[3]{2}) = \zeta_3\sqrt[3]{2},$$

so this is (23). If $1 \mapsto 2$, we get:

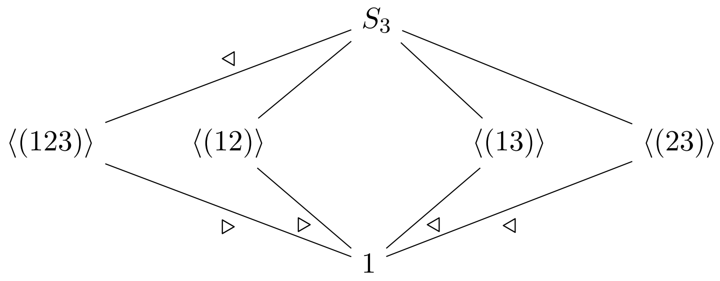
$$\zeta_3\sqrt[3]{2} \mapsto (\zeta_3^2)(\zeta_3\sqrt[3]{2}) = \sqrt[3]{2} \quad \text{and} \quad \zeta_3^2\sqrt[3]{2} \mapsto (\zeta_3^2)^2(\zeta_3\sqrt[3]{2}) = \zeta_3^5\sqrt[3]{2} = \zeta_3^2\sqrt[3]{2},$$

so we have (12). Finally, for $1 \mapsto 3$:

$$\zeta_3\sqrt[3]{2} \mapsto (\zeta_3)^2(\zeta_3^2\sqrt[3]{2}) = \zeta_3\sqrt[3]{2} \quad \text{and} \quad \zeta_3^2\sqrt[3]{2} \mapsto (\zeta_3^2)^2(\zeta_3^2\sqrt[3]{2}) = \sqrt[3]{2},$$

which is $(13) \in S_3$. Thus $\mathrm{Gal}(\mathbb{Q}(\sqrt[3]{2}, \zeta_3)/\mathbb{Q}) = \{(1), (123), (132), (23), (12), (13)\} = S_3$ as expected.

Lattices and towers. The group S_3 has six subgroups (including itself and the trivial group), which we can arrange in the following *subgroup lattice*, where each node is a subgroup of the ones connecting to it above:



(The triangles indicate which group is normal in which.) Now, to each of these groups H we can associate their fixed field $\mathbb{Q}(\sqrt[3]{2}, \zeta_3)^H$ as a subfield of $\mathbb{Q}(\sqrt[3]{2}, \zeta_3)$. For example, the fixed field of the entire Galois group S_3 is the base field $\mathbb{Q}$ since the extension $\mathbb{Q}(\sqrt[3]{2}, \zeta_3)/\mathbb{Q}$ is Galois, and the fixed field of the trivial subgroup 1 is all of $\mathbb{Q}(\sqrt[3]{2}, \zeta_3)$ since the identity automorphism fixes everything. For the remaining fixed fields, we have that:

- the only generator fixed by (123) is ζ_3 (recall that (123) came from the case $\zeta_3 \to \zeta_3$ in the Warm-Up), so the fixed field of $\langle(123)\rangle$ is $\mathbb{Q}(\zeta_3)$,
- the only generator fixed by (12) is $\zeta_3^2\sqrt[3]{2}$ (this came from the $\zeta_3 \mapsto \zeta_3^2$ case, so ζ_3 is not fixed), so the fixed field of $\langle(12)\rangle$ is $\mathbb{Q}(\zeta_3^2\sqrt[3]{2})$,
- (13) fixes $\zeta_3\sqrt[3]{2}$ and no other generator, so this fixed field is $\mathbb{Q}(\zeta_3\sqrt[3]{2})$, and
- (23) fixes $\sqrt[3]{2}$ and no other generator, so this fixed field is $\mathbb{Q}(\sqrt[3]{2})$.

We can arrange all the resulting fixed fields in a *tower diagram*, where each node is now an extension of the ones connecting to it above:

$$\begin{tikzcd}
 & & \mathbb{Q} \arrow[dll, dash, "\triangleleft"'] \arrow[dl, dash] \arrow[dr, dash] \arrow[drr, dash] & & \\
\mathbb{Q}(\zeta) \arrow[drr, dash, "\triangleright"'] & \mathbb{Q}(\zeta_3^2\sqrt[3]{2}) \arrow[dr, dash, "\triangleright"'] & & \mathbb{Q}(\zeta_3\sqrt[3]{2}) \arrow[dl, dash, "\triangleleft"] & \mathbb{Q}(\sqrt[3]{2}) \arrow[dll, dash, "\triangleleft"] \\
 & & \mathbb{Q}(\sqrt[3]{2}, \zeta_3) & &
\end{tikzcd}$$

(The triangles indicate which extensions are Galois, but note that this is not at all standard notation.) The clear similarity between these two diagrams is precisely what the statement of the *Fundamental Theorem of Galois Theory* gives us, which at its core tells us how to study properties of fields via the corresponding Galois groups. A first basic observation is that the second diagram in fact contains *all* possible subextensions of $\mathbb{Q}(\sqrt[3]{2}, \zeta_3)$ over $\mathbb{Q}$: any intermediate field $\mathbb{Q} \subseteq E \subseteq \mathbb{Q}(\sqrt[3]{2}, \zeta_3)$ *must* be the fixed field of some subgroup of $\text{Gal}(\mathbb{Q}(\sqrt[3]{2}, \zeta_3)/\mathbb{Q}) \cong S_3$. (One reason why this theorem is so powerful is that, in general, subfields of a given field are actually quite difficult to classify completely. In this case, showing that $\mathbb{Q}(\sqrt[3]{2}, \zeta_3)$ only contains six subfields extending $\mathbb{Q}$—including itself and $\mathbb{Q}$—directly would be a bit labor intensive. The problem is that a subfield could by generated by *anything*, and there is no obvious reason at first why a random generator would give the same field as one of the ones used above. For example, it is not immediately obvious that $\mathbb{Q}(\zeta_3\sqrt[3]{4})$ is indeed included in the tower diagram above, but it is because it equals $\mathbb{Q}(\zeta_3^2\sqrt[3]{2})$ since $(\zeta_3^2\sqrt[3]{2})^2 = \zeta_3\sqrt[3]{4}$ and $(\zeta_3\sqrt[3]{4})^2 = 2\zeta_3^2\sqrt[3]{2}$. The idea is that classifying subgroups of a given group is a much more tractable problem, and if we have this then we get a classification on the field side as well.)

We can recover the group diagram from the field diagram by taking Galois groups: each subgroup in the subgroup lattice is $\text{Gal}(\mathbb{Q}(\sqrt[3]{2}, \zeta_3)/E)$ where E is the field in the corresponding spot in the tower diagram. Moreover, containments between fields is reflected by the containments between groups, only with the direction reversed: for example, $\mathbb{Q} \subseteq \mathbb{Q}(\sqrt[3]{2}) \subseteq \mathbb{Q}(\sqrt[3]{2}, \zeta_3)$ corresponds to $S_3 \geq \langle (23) \rangle \geq 1$. (We stated this observation when we first introduced automorphism groups of extensions: $E_1 \subseteq E_2 \subseteq K$ implies $\text{Aut}(K/E_2) \leq \text{Aut}(K/E_1)$, and $H_1 \leq H_2 \leq \text{Aut}(K)$ implies $K^{H_2} \subseteq K^{H_1} \subseteq K$.) Another fact: the degrees $[\mathbb{Q}(\sqrt[3]{2}, \zeta_3) : E]$ in the lowest rungs of the tower diagram are exactly the orders of the groups $Gal(\mathbb{Q}(\sqrt[3]{2}, \zeta_3)/E)$. Indeed, $\mathbb{Q}(\sqrt[3]{2}, \zeta_3)$ has degree 3 over $\mathbb{Q}(\zeta)$, which is the order of $\langle (123) \rangle$, and $\mathbb{Q}(\sqrt[3]{2}, \zeta_3)$ has degree 2 over each of $\mathbb{Q}(\zeta_3^2\sqrt[3]{2})$, $\mathbb{Q}(\zeta_3\sqrt[3]{2})$, and $\mathbb{Q}(\sqrt[3]{2})$, which is the order of each of $\langle (12) \rangle$, $\langle (13) \rangle$, and $\langle (23) \rangle$. And finally, normality: the fact that $\langle (123) \rangle$ is normal in S_3 reflects the fact that $\mathbb{Q}(\zeta_3)$ is Galois over $\mathbb{Q}$ (it is the splitting field of $\phi_3(x) = x^2 + x + 1$), and the fact that, for instance, $\langle (23) \rangle$ is not normal in S_3 reflects the fact that $\mathbb{Q}(\sqrt[3]{2})$ is not Galois over $\mathbb{Q}$. To top it all off: the *quotient* of S_3 by the normal subgroup $\langle (123) \rangle$, which is $\mathbb{Z}/2\mathbb{Z}$, is precisely the Galois group of the corresponding Galois extension $\mathbb{Q}(\zeta_3)/\mathbb{Q}$)—generated by the map $\zeta_3 \mapsto \zeta_3^2$—and of course the degree of this extension is the size of the Galois group. (In the non-normal cases, it is the *indices* of the subgroups that correspond to the degrees of the corresponding extensions. For example, $\langle (23) \rangle$ has index 3 in S_3, and this is the degree of $\mathbb{Q}(\sqrt[3]{2})$ over $\mathbb{Q}$.)

So, to summarize: all field-theoretic data in the tower diagram is reflected by the group-theoretic data in the subgroup lattice, and vice-versa!

Fundamental Theorem of Galois Theory. The observations made above are not unique to the example at hand, and are reflective of a more general phenomenon. Here, then, is the statement of the *Fundamental Theorem of Galois Theory* in all its glory. Suppose K is a Galois extension of F. Then:

(0) There is a bijective correspondence between subextensions of K/F and subgroups of $\operatorname{Gal}(K/F)$ given by:

$$\begin{aligned} \{F \subseteq E \subseteq K\} &\to \{H \leq \operatorname{Gal}(K/F)\} \\ E &\mapsto \operatorname{Aut}(K/E) \\ K^H &\hookleftarrow H \end{aligned}$$

(1) These mappings are inclusion reversing, as we've seen.
(2) The degree $[K : E]$ is the order $\operatorname{Aut}(K/E)$ (note here the *base* field E is what varies), and the degree $[E : F]$ is the index $[\operatorname{Gal}(K/F) : \operatorname{Aut}(K/E)]$ (note here the *extension* E varies).
(3) K is always Galois over E, so that all automorphism groups $\operatorname{Aut}(K/E)$ above are actually Galois groups $\operatorname{Gal}(K/E)$.
(4) E is Galois over F if and only if $\operatorname{Gal}(K/E)$ is a normal subgroup of $\operatorname{Gal}(K/F)$. In this case, the Galois group $\operatorname{Gal}(E/F)$ is isomorphic to the quotient $\operatorname{Gal}(K/F)/\operatorname{Gal}(K/E)$.
(5) Intersections of fields correspond to joins of subgroups (i.e. the group the two subgroups generate), and intersections of groups correspond to composites of fields:

$$\begin{aligned} E_1 \cap E_2 &\mapsto \langle(\operatorname{Gal}(K/E_1), \operatorname{Gal}(K/E_2))\rangle \\ K^{H_1}K^{H_2} &\hookleftarrow H_1 \cap H_2. \end{aligned}$$

Note that the book does not give this first part of the statement a number, so I'm calling it part "zero". I think this is important enough to emphasize in its own right since the claim that these mappings are *inverse* to one another is highly non-trivial.

And there you have it! We will work towards the proof of this over the next few days. The full statement might seem a bit daunting at first—particularly part (4)—but the more we use it the more natural it will become. Speaking of part (4), note that since K/F is separable, K/E is always separable as well since whether or not a polynomial has a repeated root does not depend on the extension we are in. So, the only thing missing in order for E/F to be "Galois" is the condition that E be normal over F, and the claim is that this is equivalent to normality on the group side; indeed, this is where the name "normal" for a normal extension comes from! (I don't know where the name "normal" in the group case comes from.)

Biquadratic extensions. We finish with a quick and easy example. Recall that an extension $\mathbb{Q}(\sqrt{D_1}, \sqrt{D_2})$ of $\mathbb{Q}$ is *biquadratic* if none of D_1, D_2, D_1D_2 are squares in $\mathbb{Q}$. (This notion was introduced on the first homework.) Just as in the case of $\mathbb{Q}(\sqrt{2}, \sqrt{3})$ we saw previously, the Galois group of a biquadratic extension is $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$, where each $\mathbb{Z}/2\mathbb{Z}$ factors tells us whether we send

$\sqrt{D_i}$ to itself or its negative. The subgroup lattice is:

$$\begin{tikzcd}
& \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} \arrow[dl, dash] \arrow[d, dash] \arrow[dr, dash] & \\
\mathbb{Z}/2\mathbb{Z} \times \{0\} \arrow[dr, dash] & \langle (1,1) \rangle \arrow[d, dash] & \{0\} \times \mathbb{Z}/2\mathbb{Z} \arrow[dl, dash] \\
& \{0\} &
\end{tikzcd}$$

(In this case all subgroups are normal since $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ is abelian, so we omit $\lhd$ from the notation. By the way, a Galois extension with an abelian Galois group is called an *abelian extension*, so $\mathbb{Q}(\sqrt{D_1}, \sqrt{D_2})$ is an abelian extension of $\mathbb{Q}$ for example, whereas $\mathbb{Q}(\sqrt[3]{2}, \zeta_3)$ is not an abelian extension of $\mathbb{Q}$.) The corresponding tower diagram is:

$$\begin{tikzcd}
& \mathbb{Q} \arrow[dl, dash] \arrow[d, dash] \arrow[dr, dash] & \\
\mathbb{Q}(\sqrt{D_2}) \arrow[dr, dash] & \mathbb{Q}(\sqrt{D_1 D_2}) \arrow[d, dash] & \mathbb{Q}(\sqrt{D_1}) \arrow[dl, dash] \\
& \mathbb{Q}(\sqrt{D_1}, \sqrt{D_2}) &
\end{tikzcd}$$

Indeed, $\mathbb{Z}/2\mathbb{Z} \times \{0\}$ does not fix $\sqrt{D_1}$ because of first $\mathbb{Z}/2\mathbb{Z}$ factor but does fix $\sqrt{D_2}$ because of the second zero factor, and vice-versa for $\{0\} \times \mathbb{Z}/2\mathbb{Z}$, while $\langle (1,1) \rangle$ changes the sign of both $\sqrt{D_1}$ and $\sqrt{D_2}$, so that $\sqrt{D_1}\sqrt{D_2} \mapsto (-\sqrt{D_1})(-\sqrt{D_2}) = \sqrt{D_1}\sqrt{D_2}$ is fixed. (Again, try proving that these are in fact *all* the subfields of $\mathbb{Q}(\sqrt{D_1}, \sqrt{D_2})$ extending $\mathbb{Q}$ without using Galois theory—it is possible but tedious!) All intermediate fields have degree 2 "above" and "below", and all of the "middle" groups in the subgroup lattice have order 2 "below" and index 2 "above".

More generally, the same is true for biquadratic extensions of any field F of characteristic not equal to 2. (We need $\operatorname{char} F \neq 2$ to guarantee that $-\sqrt{D}$ is not the same as $\sqrt{D}$.) In fact, this gives a Galois-theoretic definition of "biquadratic extension": a biquadratic extension of F is a Galois extension with Galois group $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$. Using the subgroup lattice we can produce the tower diagram, and the fact that each resulting intermediate field in the middle has degree 2 over F guarantees that it is of the form $F(\sqrt{D})$ for some D, and the fact that there are only three such intermediate fields guarantees that one of them is indeed the form $F(\sqrt{D_1 D_2})$ where $\sqrt{D_1}$ and $\sqrt{D_2}$ generate the other two.

Lecture 13: More on Galois Extensions

Warm-Up. We determine the Galois group of $x^4 - 2$ over $\mathbb{Q}$, and over $\mathbb{F}_5$. (We didn't do the $\mathbb{F}_5$ example in class.) First, the splitting field of $x^4 - 2$ over $\mathbb{Q}$ is $\mathbb{Q}(\sqrt[4]{2}, i)$ (note i is a primitive fourth root of unity), so the group we want is $\operatorname{Gal}(\mathbb{Q}(\sqrt[4]{2}, i)/\mathbb{Q})$. By thinking of this field as the composite of $\mathbb{Q}(\sqrt[4]{2})$ and $\mathbb{Q}(i)$, we see that it has degree at most $4 \cdot 2 = 8$ over $\mathbb{Q}$. But this degree is also divisible by 4 via the subextension

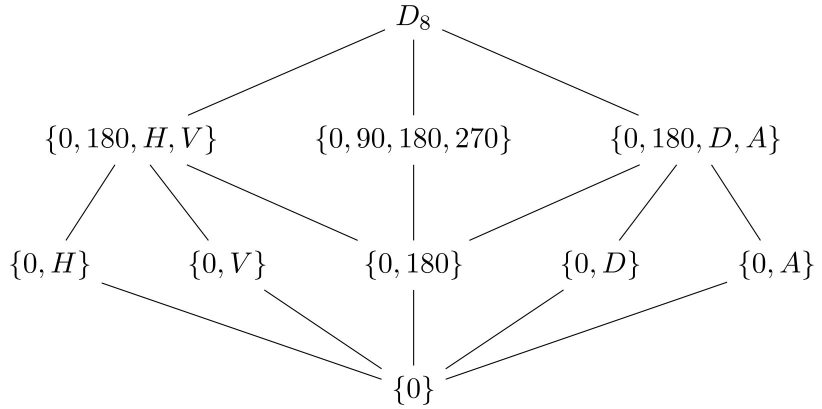
$$\mathbb{Q} \subseteq \mathbb{Q}(\sqrt[4]{2}) \subseteq \mathbb{Q}(\sqrt[4]{2}, i),$$

so it is either 4 or 8. If it was 4 then we would necessarily have $\mathbb{Q}(\sqrt[4]{2}, i) = \mathbb{Q}(\sqrt[4]{2})$, which is not true since the latter does not contain i, so we have $[\mathbb{Q}(\sqrt[4]{2}, i) : \mathbb{Q}] = 8$. This can also be seen from the tower law applied to the subextension above using the fact that $\mathbb{Q}(\sqrt[4]{2}, i)$ has degree 2 over $\mathbb{Q}(\sqrt[4]{2})$ since i has minimal polynomial $x^2 + 1$ over $\mathbb{Q}(\sqrt[4]{2})$.

Thus $\mathrm{Gal}(\mathbb{Q}(\sqrt[4]{2}, i)/\mathbb{Q})$ is a group of order 8. To see which group of order 8 it is, note that it can be viewed as a subgroup of S_4 via permuting the four roots of $x^4 - 2$. A subgroup of order 8 of S_4 is a Sylow 2-subgroup, and hence must be isomorphic to any other subgroup of order 8 since all Sylow 2-subgroups are conjugate to one another. Since D_8 (recall dihedral groups!) is a subgroup of S_4 of order 8 (view elements of D_8 as rotations and reflections of a square and permute the vertices), we have that any subgroup of S_4 of order 8 is isomorphic to D_8, so we conclude that $\mathrm{Gal}(\mathbb{Q}(\sqrt[4]{2}, i)/\mathbb{Q}) \cong D_8$ is the desired Galois group.

When viewed as a polynomial over $\mathbb{F}_5$, $x^4 - 2$ is irreducible. Linear factors are ruled out by not having a root in $\mathbb{F}_5$, and quadratic factors can be ruled out by brute force or by using the fact derived on the most recent homework that $x^4 - 2$ is irreducible over $\mathbb{F}_5$ if it is relatively prime to $x^5 - x$ and $x^{5^2} - x$, which can be verified using the Euclidean algorithm. Thus $x^4 - 2$ has a root in $\mathbb{F}_{5^4} = \mathbb{F}_5[x]/(x^4 - 2)$, and in fact splits completely in this extension since it is normal. (We will show definitively in a bit that Galois implies normal.) Thus $\mathbb{F}_{5^4}$ is the splitting field of $x^4 - 2$ over $\mathbb{F}_5$, so its Galois group is $\mathrm{Gal}(\mathbb{F}_{5^4}/\mathbb{F}_5) \cong \mathbb{Z}/4\mathbb{Z}$ generated by Frobenius. (We determined the Galois group $\mathrm{Gal}(\mathbb{F}_{p^n}/\mathbb{F}_p) \cong \mathbb{Z}/n\mathbb{Z}$ in general a few days ago.)

Another lattice/tower diagram. Let us use the example of the splitting field of $x^4 - 2$ over $\mathbb{Q}$ to illustrate again the relations encoded by the Fundamental Theorem of Galois Theory. Using the rotation/reflection notation for $D_8 = \{0, 90, 180, 270, H, V, D, A\}$, the subgroup lattice is:



To determine the fixed field of each subgroup, let us describe the elements of the Galois group as explicit cycles. Each element of the Galois group permutes $\sqrt[4]{2}$ among the roots of $x^4 - 2$:

$$\sqrt[4]{2} \mapsto \sqrt[4]{2},\ i\sqrt[4]{2},\ -\sqrt[4]{2},\ -i\sqrt[4]{2} \quad \text{(call these } 1, 2, 3, 4 \text{ in that order)}$$

and permutes i among the roots of $x^2 + 1$: $i \mapsto i,\ -i$. The possibilities when $i \mapsto i$ are:

$$(1),\ (1234),\ (13)(24),\ (1432).$$

For example, if $1 \mapsto 4$, so that $\sqrt[4]{2} \mapsto -i\sqrt[4]{2}$, then $-i\sqrt[4]{2} \mapsto -(i)(-i\sqrt[4]{2}) = -\sqrt[4]{2}$, so $4 \mapsto 3$, and so on, recalling that i is fixed in this case. In terms of the rotation/reflection notation, these four elements are $0, 90, 180, 270$ respectively. The possibilities for $i \mapsto -i$ are:

$$(12)(34),\ (13),\ (14)(23),\ (24),$$

which are V, A, H, D respectively. For example, if $1 \mapsto 2$, then $i\sqrt[4]{2} \mapsto (-i)(i\sqrt[4]{2}) = \sqrt[4]{2}$, so $2 \mapsto 1$, and $-\sqrt[4]{2} \mapsto -(i\sqrt[4]{2}) = -i\sqrt[4]{2}$, so $3 \mapsto 4$ and you can check that $4 \mapsto 3$. As another example, if $1 \mapsto 3$, so that $\sqrt[4]{2} \mapsto -i\sqrt[4]{2}$, then $i\sqrt[4]{2} \mapsto (-i)(-\sqrt[4]{2}) = i\sqrt[4]{2}$, so 3 is fixed, as is 4 in this case.

With this notation, we can compute some fixed fields. Since $180 = (13)(24)$, none of the roots $\sqrt[4]{2},\ i\sqrt[4]{2},\ -\sqrt[4]{2},\ -i\sqrt[4]{2}$ are fixed by 180, but $\sqrt[4]{4} = \sqrt{2}$ is fixed:

$$\sqrt{2} = (\sqrt[4]{2})(\sqrt[4]{2}) \mapsto (-\sqrt[4]{2})(-\sqrt[4]{2}) = \sqrt[4]{4} = \sqrt{2}.$$

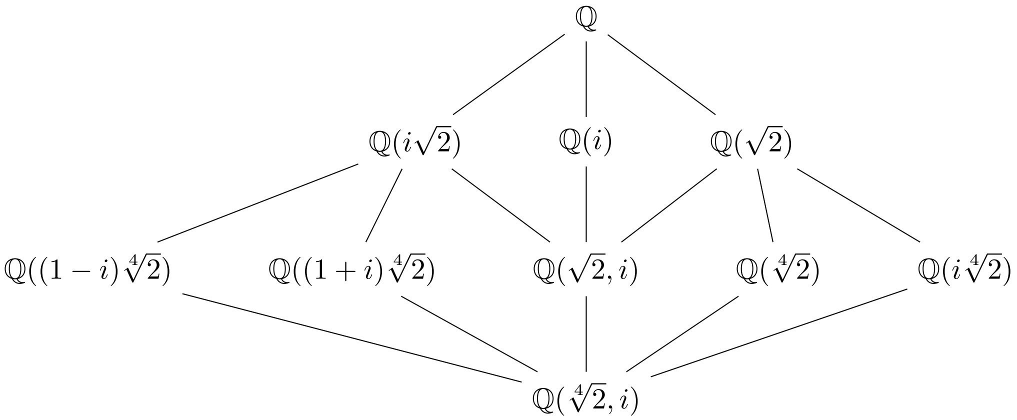
This element also fixes i since it came from the $i \mapsto i$ case, so the fixed field of $\{0, 180\}$ is $\mathbb{Q}(\sqrt{2}, i)$. (We can use the degree restrictions in the Fumdamental Theorem to argue that the fixed field is not larger: since $\{0, 180\}$ has index 4 in D_8, the degree of the fixed field over $\mathbb{Q}$ should be 4, and $\mathbb{Q}(\sqrt{2}, i)$ already has degree 4 over $\mathbb{Q}$.) Since $D = (24)$ fixes $\sqrt[4]{2}$ and not i (this came from the $i \mapsto -i$ case), the fixed field of $\{0, D\}$ is $\mathbb{Q}(\sqrt[4]{2})$. If we want elements fixed by both $D = (24)$ and $180 = (13)(24)$, note that $\sqrt[4]{2}$ no longer works since this is not fixed by 180, but $\sqrt{2} = (\sqrt[4]{2})^2$ does work: this is fixed by 180 as shown above, and also by D since D fixes $\sqrt[4]{2}$. (Note that now i is not fixed because D does not fix i.) Thus the fixed field of $\{0, 180, D, A\}$ (A introduces nothing new since $A = 180 \circ D$) is $\mathbb{Q}(\sqrt{2})$, and so the chain of subgroups

$$\{0\} \subseteq \{0, 180\} \subseteq \{0, 180, D, A\} \subseteq D_8$$

corresponds to the chain of towers

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

Computing all fixed fields gives the following tower diagram:


$\mathbb{Q}$
$\mathbb{Q}(i\sqrt{2})$
$\mathbb{Q}(i)$
$\mathbb{Q}(\sqrt{2})$
$\mathbb{Q}((1-i)\sqrt[4]{2})$
$\mathbb{Q}((1+i)\sqrt[4]{2})$
$\mathbb{Q}(\sqrt{2}, i)$
$\mathbb{Q}(\sqrt[4]{2})$
$\mathbb{Q}(i\sqrt[4]{2})$
$\mathbb{Q}(\sqrt[4]{2}, i)$

Note that some of these, in particular $\mathbb{Q}((1-i)\sqrt[4]{2})$ and $\mathbb{Q}((1+i)\sqrt[4]{2})$, take some effort to find explicitly. For example, the fixed field of $\{0, H\}$, or equivalently just H, has degree 4 over $\mathbb{Q}$ since this is the index of $\{0, H\}$ in D_8. So, to describe this explicitly we need an element of $\mathbb{Q}(\sqrt[4]{2}, i)$ fixed by H which has a minimal polynomial of degree 4 over $\mathbb{Q}$. Finding such an element requires possibly more guess and check than earlier examples, but we should be looking for someting expressible in terms of the roots of $x^4 - 2$, in this case as a *sum* of such roots. We can verify that $(1-i)\sqrt[4]{2} = \sqrt[4]{2} - i\sqrt[4]{2}$ is indeed fixed by H: $H = (14)(23)$ sends $\sqrt[4]{2} \mapsto -i\sqrt[4]{2}$ and $i\sqrt[4]{2} \mapsto -\sqrt[4]{2}$, so

$$\sqrt[4]{2} - i\sqrt[4]{2} \mapsto (-i\sqrt[4]{2}) - (-\sqrt[4]{2}) = \sqrt[4]{2} - i\sqrt[4]{2}.$$

If $x = (1-i)\sqrt[4]{2}$, then $x^4 = 2(1-i)^4 = 2(-2i)^2 = -8$, so $x^4 + 8$ is the minimal polynomial of $(1-i)\sqrt[4]{2}$. Thus $\mathbb{Q}((1-i)\sqrt[4]{2})$ is a degree 4 extension of $\mathbb{Q}$ fixed by H, so it must be the desired fixed field since we already know this fixed field should have degree 4 over $\mathbb{Q}$.

Finally, let us point out a few spots in this diagram that illustrate how normality comes in. The subgroup $\{0, 90, 180, 270\}$ is normal in D_8 (it has index 2), so the corresponding fixed field $\mathbb{Q}(i)$ should be a Galois extension of $\mathbb{Q}$, which it is since it is the splitting field of $x^2 + 1$. The subgroup $\{0, 180\}$ is also normal in D_8 (it is the center of D_8), so the fixed field $\mathbb{Q}(\sqrt{2}, i)$ should be Galois over $\mathbb{Q}$, and it is since it is the splitting field of $(x^2 - 2)(x^2 + 1)$. In this case, the Galois group of $\mathbb{Q}(\sqrt{2}, i)$ over $\mathbb{Q}$, which is $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ since this extension is biquadratic, should be the quotient of $D_8 \cong \mathrm{Gal}(\sqrt[4]{2}, i)$ by $\mathbb{Z}/2\mathbb{Z} \cong \{0, 180\}$, which it is. Finally, the subgroup $\{0, D\}$ is not normal in D_8, so the corresponding fixed field $\mathbb{Q}(\sqrt[4]{2})$ is not a Galois extension of $\mathbb{Q}$, as we have seen before: the fixed field of $\mathbb{Q}(\sqrt[4]{2})/\mathbb{Q}$ is actually $\mathbb{Q}(\sqrt{2})$, and the fact that $\mathbb{Q}(\sqrt[4]{2})$ *is* Galois over $\mathbb{Q}(\sqrt{2})$ reflects the fact that $\{0, D\}$ *is* a normal subgroup of $\{0, 180, D, A\}$, just not of all of D_8.

Degrees and fixed fields. Now we start working towards a proof of the Fundamental Theorem of Galois Theory. As a first step, we state the following result:

> If H is a finite subgroup of $\mathrm{Aut}(K)$, then the degree of K over the fixed field K^H is the order of H: $|H| = [K : K^H]$.

We will take this for granted for the time being, and will say something about the proof—which is quite involved—later. Certainly, this claim is a consequence of the Fundamental Theorem of Galois Theory, but of course since we are wanting to use it to prove the Fundamental Theorem, we must give an independent proof. This result really is the key to making it all work.

But assuming this for now, we can now justify the fact that the size of any automorphism group $\mathrm{Aut}(K/F)$ is bounded by the degree $[K : F]$. We proved this previously in the case where K is the splitting field of a polynomial over F, but now we make no such assumption. Moreover, we also claim that equality $|\mathrm{Aut}(K/F)| = [K : F]$ holds if and only if F is the fixed field of $\mathrm{Aut}(K/F)$. (In the splitting field case, equality was true under the assumption that the polynoimal in question was separable.) So, assume K is an extension of F, and let E be the fixed field of $\mathrm{Aut}(K/F)$, which contains F as a subfield since, by definition, anything in $\mathrm{Aut}(K/F)$ fixes F. By the result above, we have $|\mathrm{Aut}(K/F)| = [K : E]$. (Take H to be $\mathrm{Aut}(K/F)$ in teh result above, so that $E = K^H$.) Since $[K : F] = [K : E][E : F]$, we have

$$|\mathrm{Aut}(K/F)| = [K : E] \leq [K : F],$$

which is the desired bound. Moreover, equalilty $|\mathrm{Aut}(K/F)| = [K : F]$ holds if and only if $[K : F] = [K : E]$, which given that $F \subseteq E$ holds if and only if $E = F$, or equivalently F is the fixed field of $\mathrm{Aut}(K/F)$.

Revisting Galois extensions. The condition $|\mathrm{Aut}(K/F)| = [K : F]$ is the book's definition of what it means for K/F to be a Galois extension, and so now we see that this definition is equivalent to the one we gave: K/F is Galois if F is the fixed field of $\mathrm{Aut}(K/F)$. Given the bound $|\mathrm{Aut}(K/F)| \leq [K : F]$, we thus see that Galois extensions are precisely those which have the *maximum* number of automorphisms possible, which in the end is what makes it possible to use automorphisms to distinguish between the extension and the base field—the original motivation we gave for introducing "Galois" extensions.

Now, as mentioned above, when K is the splitting field of a separable polynomial over F, we have already proved that we have equality $|\mathrm{Aut}(K/F)| = [K : F]$. Thus such extensions are always Galois. In fact, we can now prove that this condition is equivalent to being Galois, and to being normal and separable. We claim that the following conditions on a finite extension K/F are equivalent:

(i) K/F is Galois (take either the fixed field definition or the degree definition)

(ii) K/F is normal and separable

(iii) K is the splitting field of a separable polynomial over F.

We have already proved $(iii) \Rightarrow (i)$, and $(ii) \Rightarrow (iii)$ was essentially on the problem set for Discussion 2 (showing that my definition of "normal" and the book's definition are equivalent), but we will prove it here again anyway. (The term "normal" in (ii) is my definition: whenever an irreducible polynomial over F has a root in K, it splits completely in K.)

To prove $(i) \Rightarrow (ii)$, suppose K/F is Galois and suppose $p(x) \in F[x]$ is irreducible with a root $\alpha \in K$. Denote the elements of $\mathrm{Gal}(K/F)$ by $id, \sigma_1, \ldots, \sigma_m$. Then

$$\alpha,\ \sigma_1(\alpha),\ \ldots,\ \sigma_m(\alpha) \in K$$

are all roots of $p(x)$ since Galois group elements send roots to roots. (Note we do not yet know that these give *all* the roots of $p(x)$, only that each one is in fact a root.) Let

$$\alpha,\ \beta_1,\ \ldots,\ \beta_k \in K$$

denote the *distinct* elements from the list above, and set $f(x)$ to be the polynomial

$$f(x) = (x-\alpha)(x-\beta_1)\cdots(x-\beta_k) \in K[x].$$

We claim that $f(x) = p(x)$, which, if true, gives us what we want: $p(x)$ splits completely in K since, by construction, $f(x)$ splits in K, and the roots $\alpha, \beta_1, \ldots, \beta_k$ of $f(x) = p(x)$ are all distinct, so that $p(x)$ is separable; since $p(x)$ was an arbtirary irreducible polynomial over F with a root in K, this shows that the entire extension K/F is both normal and separable.

By construction, any element of $\mathrm{Gal}(K/F)$ permutes the elements $\alpha, \beta_1, \ldots, \beta_k \in K$, so it permutes the roots of $f(x)$. This means that any Galois automorphism simply permutes the factors in the definition of $f(x)$, so that the polynomial obtained after applying Galois automorphisms to the roots is $f(x)$ itself. But if we expand f, this means precisely that each coefficient of $f(x)$ (made up of the α, β_i) must be fixed by the action of the entire Galois group. Thus each coefficient is in the fixed field of the Galois group, which is F because K/F is Galois, so that $f(x)$ is actually an element of $F[x]$. Any polynomial in $F[x]$ having α as a root is divisible by $p(x)$ since $p(x)$, being irreducible, generates the ideal of all polynomials with α as a root; hence $p(x) \mid f(x)$ since $f(x)$ has α as a root. Moreover, since each α, β_i is a root of $p(x)$, we have that $f(x) \mid p(x)$: the factorization of $p(x)$ in its splitting field (whatever it is) must include at least each of the linear factors making up $f(x)$, so that the product $f(x)$ of these linear factors must divide the factorizatoin of $p(x)$. Thus we have $p(x) \mid f(x)$ and $f(x) \mid p(x)$, so since both of these are *monic* polynomials, we get $p(x) = f(x)$ as claimed. As explained above, this shows K/F is normal and separable, so $(i) \Rightarrow (ii)$.

Finally we show $(ii) \Rightarrow (iii)$, which completes the proof of equivalence. Suppose K/F is normal and separable, and let $\alpha_1, \ldots, \alpha_n \in K$ be generators for K over F. (Recall K/F is finite.) Let $p_i(x) \in F[x]$ be the minimal polynomial of α_i over F. Note that each $p_i(x)$ has distinct roots since K/F is separable. Since each $p_i(x)$ has a root in K (namely α_i) and K/F is normal, each $p_i(x)$ splits completely over K, so K contains all the roots of the $p_i(x)$. Let $q(x)$ be the product $p_1(x)\cdots p_m(x)$ with any repeated factors excluded (in case some $p_i(x)$ are the same), so that $q(x)$ has distinct roots and is thus separable. Since K contains all roots of $q(x)$, K is then the splitting field of the separable polynomial $q(x)$ over F, as desired. (Note the splitting field cannot be a proper subfield of K since it must contain all α_i because these are all roots of $q(x)$, and K is the smallest field extending F that contains all α_i.)

Lecture 14: More on the Fundamental Theorem

Warm-Up. We prove (what I called) part (0) of the Fundamental Theorem of Galois Theory, which is the claim that the correspondence between subextensions and subgroups is bijective: if K/F is a Galois extension, then the maps

$$\begin{aligned} \{F \subseteq E \subseteq K\} &\to \{H \leq \operatorname{Gal}(K/F)\} \\ E &\mapsto \operatorname{Aut}(K/E) \\ K^H &\leftarrow\!\shortmid H \end{aligned}$$

are inverses of one another. First, take a subextension $F \subseteq E \subseteq K$. From this we get the subgroup $\operatorname{Aut}(K/E)$ of $\operatorname{Gal}(K/F)$, and then from this we get the fixed field $K^{\operatorname{Aut}(K/E)}$. The claim is that this fixed field is precisely E itself, so that composing the maps above forwards and then backwards produces the identiy map on the set of subextensions. To say that the fixed field of $\operatorname{Aut}(K/E)$ is E is just what it means for K/E to be a Galois extension, which is what we prove. We use the fact that Galois is equivalent to being the splitting field of a separable polynomial over the base. Since K/F is Galois, K is the splitting field of a separable polynomial $f(x) \in F[x]$. But then also $f(x) \in E[x]$ since E contains F, so K is also the splitting field of a separable polynomial over E, meaning that K/E is Galois as desired.

Now, take a subgroup $H \leq \operatorname{Gal}(K/F)$. Then we get the fixed field K^H, and then the group $\operatorname{Aut}(K/K^H)$. The claim is that this automorphism group is just H again, so that composing backwards and then forwards above gives the identity map on the set of subgroups. Note that H is a subgroup of $\operatorname{Aut}(K/K^H)$ since, by definition, any element of H fixes any element of K^H. The question is whether there can be *more* automorphisms that fix K^H other than those in H, and the answer is no: we have $|H| = [K : K^H]$ by the still unproven key fact from last time (that we took for granted), so $|H| = [K : K^H] = \operatorname{Aut}(K/K^H)$, where the second equality follows from the fact that K/K^H is Galois by the reasoning we gave in the "forwards then backwards" argument above. Thus H is a subgroup of $\operatorname{Aut}(K/K^H)$ with order equal to that of the entire group, so $H = \operatorname{Aut}(K/K^H)$ as claimed. Hence the correspondence between subextensions and subgroups in the Fundamental Theorem of Galois Theory is indeed bijective.

Proof of the Fundamental Theorem. We are now ready to prove the Fundamental Theorem of Galois Theory. Actually, we have proven many of the parts already, so here we go. Fix a subextension $F \subseteq E \subseteq K$. Then we check each part of the theorem:

(0) This is the claim that the correspondence between subextensions and subgroups is bijecive, which we just proved.

(1) This is the claim that this correspondence is inclusion-reversing, which we pointed out when we first introduced Galois groups.

(2) The first claim here is that the degree of K/E is the order of $\operatorname{Gal}(K/E)$, which is just the claim that K/E is Galois and was proved in the Warm-Up. (This is why we now use the notation of $\operatorname{Gal}(K/E)$ instead of $\operatorname{Aut}(K/E)$.) The second claim is that the degree of the smaller extension E/F is equal to the index of $\operatorname{Gal}(K/E)$ in $\operatorname{Gal}(K/F)$, and comes from the tower law together with Lagrange's Theorem for groups from the fall:

$$[E : F] = \frac{[K : F]}{[K : E]} = \frac{|\operatorname{Gal}(K/F)|}{|\operatorname{Gal}(K/E)|} = [\operatorname{Gal}(K/F) : \operatorname{Gal}(K/E)].$$

(3) This is the claim that K/E, which is just a restatement of the first part of (2), nothing new.

(4) This is the claim that the smaller extension E/F is Galois if and only if $\mathrm{Gal}(K/E)$ is a *normal* subgroup of $\mathrm{Gal}(K/F)$, and that in this case the Galois group of E/F is (isomorphic to) the quotient of $\mathrm{Gal}(K/F)$ by $\mathrm{Gal}(K/E)$. This will take some work to prove, and we will come back to this after the final part.

(5) Finally, this is the claim that intersections of fields correspond to joins of subgroups, and that intersections of subgroups correspond to composites of fields. Suppose we have two subextensions E_1, E_2 of K/F. Then $E_1 \cap E_2$ is also a subextension (simple to verify) of K/F, and is fixed by both groups $\mathrm{Gal}(K/E_1)$ and $\mathrm{Gal}(K/E_2)$, which are each subgroups of $\mathrm{Gal}(K/F)$. Thus $E_1 \cap E_2$ is fixed by the subgroup these two generate, which is the *join* $\langle \mathrm{Gal}(K/E_1), \mathrm{Gal}(K/E_2)\rangle$. If $a \in K$ is an element not in $E_1 \cap E_2$, then it is excluded form E_1 or E_2, so that there is an element in either $\mathrm{Gal}(K/E_1)$ or $\mathrm{Gal}(K/E_2)$ not fixing a. But this means that the join $\langle \mathrm{Gal}(K/E_1), \mathrm{Gal}(K/E_2)\rangle$ then does not fix a, so we conclude that the only elements of K fixed by the join are those in $E_1 \cap E_2$. Hence we get our first claim:

$$\mathrm{Gal}(K/E_1 \cap E_2) = \langle \mathrm{Gal}(K/E_1), \mathrm{Gal}(K/E_2)\rangle.$$

Going the other way, take two subgroups H_1, H_2 of $\mathrm{Gal}(K/F)$. Then elements of $H_1 \cap H_2$ fix both individual fixed fields K^{H_1} and K^{H_2} (since the intersection is contained in both H_1 and H_2), so $H_1 \cap H_2$ fixes all elements of the composite $K^{H_1}K^{H_2}$. (Note that elements of this composite are all expressible in terms of elements of K^{H_1} and K^{H_2} alone, in particular they are explicitly quotients of sums of products of an element of K^{H_1} with an element of K^{H_2}.) If $\sigma \in \mathrm{Gal}(K/F)$ is not in $H_1 \cap H_2$, then it is not in at least one of H_1 or H_2, so there is an element of K^{H_1} or of K^{H_2} it does not fix. But then σ does not fix every element of $K^{H_1}K^{H_2}$, so we conclude that the only elements of $\mathrm{Gal}(K/F)$ fixing the composite $K^{H_1}K^{H_2}$ are those in $H_1 \cap H_2$. This gives the second claim we need:

$$K^{H_1}K^{H_2} = K^{H_1 \cap H_2}.$$

And there's our proof, except of course for part (4), and the still unproven claim that $[K : K^H] = |H|$ for a finite subgroup H of $\mathrm{Aut}(K)$ upon which everything has depended since it was crucial in deriving the various characterizations of "Galois extension" we gave last time. We will come back to this claim next time.

Conjugates and embeddings. Part (4) of the Fundamental Theorem is best approached using some new terminology. Given $\alpha \in K$ and $\sigma \in \mathrm{Gal}(K/F)$, we call $\sigma(\alpha)$ a *(Galois) conjugate* of α. (Note that, with this terminology, the roots of an irreducible polynomial over F are all conjugates of one another.) Given $F \subseteq E \subseteq K$ and $\sigma \in \mathrm{Gal}(K/F)$, we call the image $\sigma(E)$ of E under σ a *(Galois) conjugate* of E; this is always a subfield of K and contains the conjugates of all the elements of E. (The reason for using the term "conjugate" will soon become clear.)

Now, each $\sigma \in \mathrm{Gal}(K/F)$ restricts to an isomorphism $\sigma|_E : E \to \sigma(E)$ from E to the corresponding conjugate field. This map can be viewed as an injective homomorphism $E \to K$ with image $\sigma(E)$, and so gives an *embedding* of E into K over F, where "over" F means that it restricts to the identity on F. (An embedding is simply an injective field homomorphism. We pointed out in the first week or so that any homomorphism between fields is either the zero map or an embedding.) We denote the set of such embeddings by $\mathrm{Emb}(E/F)$, and our goal for now is to determine the number of such embeddings. Not only is it true that restricting an element of $\mathrm{Gal}(K/F)$ to E gives an embedding of E into K, but we claim that *all* such embeddings arise in this way: any embedding $E \to K$ over F can be "extended" to an automorphism $K \to K$ over F, or in other words an element of $\mathrm{Gal}(K/F)$.

To see this, note that since K/E is actually Galois, K is the splitting field of a separable polynomial over E, so we can express K as being generated by the roots of this polynomial over E: $K = E(\alpha_1, \dots, \alpha_m)$. Now, given an embedding $\tau : E \to K$ over F, we argue that we can extend it first to $E(\alpha_1)$, and then to $E(\alpha_1, \alpha_2)$, and so on until we have extended it to $K = E(\alpha_1, \dots, \alpha_m)$ as desired. We have actually seen this type of argument a few times before—when showing that splitting fields are unique and when bounding the order of $\mathrm{Aut}(E/F)$ by $[E : F]$ when E is the splitting field of a polynomial over F—but let us be clear. Using the minimal polynomial of α_1 over E, we can identity $E(\alpha_1)$ with a quotient $E[x]/(m(x))$, which we can in turn identity with $\tau(E)[x]/(\text{image of } m(x)) \cong \tau(E)(\text{root})$ where "root" is any root of the image of $m(x)$. By sending α_1 to this "root", we thus get a map

$$E(\alpha_1) \to \tau(E)(\text{root})$$

that extends $E \to \tau(E) \subseteq K$. Now we do the same thing with α_2: use the minimal polynomial of α_2 over $E(\alpha_1)$ and the corresponding quotient $E(\alpha_1)[x]/(\text{polynomial})$ to extend $E(\alpha_1) \to \tau(E)(\text{root})$ to a map

$$E(\alpha_1, \alpha_2) \to \tau(E)(\text{root}, \text{another root}) \subseteq K.$$

Continuing in this way then gives a map $K = E(\alpha_1, \dots, \alpha_m) \to K$ extending the original $\tau : E \to K$. (There is one subtle point here, in that we need to know all the "roots" we adjoin on the right side at each step are actually in K, so that the resulting map indeed has image in K as opposed to simply an algebraic closure of K. But this follows from the fact that K is a splitting field: the minimal polynomial of α_1 is sent to the minimal polynomial of the corresponding root under the first extension to $E(\alpha_1)$, so that K, which contains all the root of the first minimal polynomial, also contains the roots of the second. This is true at each step, so K contains all the required roots.) This resulting map is an automorphism since, being nonzero, it is injective, and the fact that $[K : E] = [\text{image of } K : E]$ implies that it is surjective as well. Thus, as claimed, any embedding $E \to K$ is the restriction of an element of $\mathrm{Gal}(K/F)$.

Now, suppose $\sigma_1, \sigma_2 \in \mathrm{Gal}(K/F)$. We want to determine when it is that these two elements give the *same* restriction to E. But $\sigma_1|_E = \sigma_2|_E$ if and only if $\sigma_1^{-1}\sigma_2|_E$ is the identity map on E, which just means that $\sigma^{-1}\sigma_2$ fixes E and is thus an element of $\mathrm{Gal}(K/E)$. Recalling some things about cosets from the fall, we see that this condition is the same as saying that σ_1 and σ_2 determine the same coset of $\mathrm{Gal}(K/E)$ in $\mathrm{Gal}(K/F)$. Thus, σ_1 and σ_2 give the same restriction to E, or in other words the same embedding in $\mathrm{Emb}(E/F)$, if and only if they become equal in the set of cosets. Hence, the number of such embeddings is precisely the number of cosets:

$$|\,\mathrm{Emb}(E/F)| = [\mathrm{Gal}(K/F) : \mathrm{Gal}(K/E)].$$

Since the index $[\mathrm{Gal}(K/F) : \mathrm{Gal}(K/E)]$ equals the degree $[E : F]$ by part (3) of the Fundamental Theorem of Galois Theory, we thus get that $|\,\mathrm{Emb}(E/F)| = [E : F]$.

Normality. To say that E/F is Galois is the same as saying that $|\,\mathrm{Aut}(E/F)| = [E : F]$, which, based on the equality derived above, is the same as saying that the number of automorphisms of E over F equals the number of embeddings of E over F: $|\,\mathrm{Aut}(E/F)| = |\,\mathrm{Emb}(E/F)|$. But note that any automorphism ϕ is in particular an embedding as well, in this case with conjugate field $\phi(E)$ equal to E itself. Thus $\mathrm{Aut}(E/F)$ is always a subset of $\mathrm{Emb}(E/F)$, so the condition that E/F is Galois is equivalent to saying that *every* embedding of E over F actually has image equal to E, or in other words that every conjugate field of E is E itself:

$$E/F \text{ is Galois } \iff \sigma(E) = E \text{ for all } \sigma \in \mathrm{Gal}(K/F).$$

We thus seek to understand when it is that this is true.

The key observation is that $\sigma(E)$ is the fixed field of the conjugate subgroup $\sigma\operatorname{Gal}(K/E)\sigma^{-1}$ of $\operatorname{Gal}(K/F)$. (Hence why we use the term "conjugate" in the field setting.) Indeed, if $a \in E$ and $\tau \in \operatorname{Gal}(K/E)$, then

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

where we use the fact that $\tau(a) = a$ since $\tau \in \operatorname{Gal}(K/E)$ fixes elements of E. This says that elements of $\sigma\operatorname{Gal}(K/E)\sigma^{-1}$ fix elements of $\sigma(E)$, so

$$\sigma(E) \subseteq K^{\sigma\operatorname{Gal}(K/E)\sigma^{-1}}.$$

But the degree of K over $\sigma(E)$ is the same as the degree of K over E since $\sigma(E)$ is isomorphic to E (over F), and the order of $\sigma\operatorname{Gal}(K/E)\sigma^{-1}$ is equal to the order of $\operatorname{Gal}(K/E)$. Thus

$$[K:\sigma(E)] = [K:E] = |\operatorname{Gal}(K/E)| = |\sigma\operatorname{Gal}(K/E)\sigma^{-1}| = [K:K^{\sigma\operatorname{Gal}(K/E)\sigma^{-1}}|,$$

so we must have $\sigma(E) = K^{\sigma\operatorname{Gal}(K/E)\sigma^{-1}}$ as claimed.

Thus E/F is Galois if and only if $\sigma(E) = E$ for all $\sigma \in \operatorname{Gal}(K/F)$ if and only if $E = K^{\sigma\operatorname{Gal}(K/E)\sigma^{-1}}$ for all $\sigma \in \operatorname{Gal}(K/F)$. But E is the fixed field of $\operatorname{Gal}(K/E)$, so by bijectivity of the Galois correspondence:

$$K^{\operatorname{Gal}(K/E)} = K^{\sigma\operatorname{Gal}(K/E)\sigma^{-1}} \iff \operatorname{Gal}(K/E) = \sigma\operatorname{Gal}(K/E)\sigma^{-1}$$

for all $\sigma \in \operatorname{Gal}(K/F)$, which is precisely the condition that $\operatorname{Gal}(K/E)$ be a *normal* subgroup of $\operatorname{Gal}(K/F)$. Hence "Galois" on the field side is equivalent to "normal" on the group side, which is the first part of (4) in the Fundamental Theorem.

Finally, when E/F is Galois, we have a homomorphism $\operatorname{Gal}(K/F) \to \operatorname{Gal}(E/F)$ of groups given by restriction: $\sigma \mapsto \sigma|_E$. This is surjective since every embedding (in particular automorphism) of E over F is the restriction of an element of $\operatorname{Gal}(K/F)$, and its kernel consists of those element of $\operatorname{Gal}(K/F)$ that become the identity map on E, which are precisely the elements of $\operatorname{Gal}(K/E)$. Thus the first isomorphism theorem for groups gives

$$\operatorname{Gal}(K/F)/\operatorname{Gal}(K/E) \cong \operatorname{Gal}(E/F),$$

which is the rest of part (4) of the Fundamental Theorem of Galois Theory. This completes our (elaborate!) proof of this theorem.

Lecture 15: More on Galois Theory

Warm-Up. Suppose K/F is Galois of prime-power degree p^n for some prime p. We show that there exists a chain of extensions

$$K \subseteq K_1 \subseteq K_2 \subseteq \ldots \subseteq K_{n-1} \subseteq K$$

such that each field is of degree p and Galois over the previous one. The point is that trying to justify this in a strictly field-theoretic manner (no groups) is likely to be quite challenging if not impossible, since it would require constructing the desired extensions seemingly out of nowhere. But, this is the type of thing which Galois theory makes feasible, essentially because we have already solved the corresponding problem for groups. We'll note that going forward we will be using much of the group theory we did in the fall, some of which involved results for very specific types of groups, and that it is not expected that you will immediately be able to recall all of this. But, the goal is to revisit these topics in the course of looking at various applications, so that hopefully your memory will be refreshed. You can check the lecture notes for the fall quarter to recall why these results are true if you'd like, but we will treat them here as a black box.

Since $[K : F] = p^n$, the Galois group $\text{Gal}(K/F)$ has order p^n, so it is what we called in the fall a p-group. A key property of p-groups is that they have subgroups of any possible allowed order, meaning that for each $1 \leq k \leq n$ there exists a subgroup of order p^k. (This was a consequence of the *class equation* and properties of the *center* of a group. Again, check the notes from the fall if you want to see more.) Thus in particular $\text{Gal}(K/F)$ has a subgrouop H of order p^{n-1}. H then has index $\frac{p^n}{p^{n-1}} = p$ in $\text{Gal}(K/F)$, so since this index is the smallest prime divising $|\text{Gal}(K/F)| = p^n$, H must actually be a normal subgroup. (That H being normal is implied by $[G : H]$ being the smallest prime dividing $|G|$ is another group-theoretic result to recall.) By the Fundamental Theorem of Galois Theory, the fixed field K^H of H is then a Galois extension of F of degree p, so this is our sought-after "K_1''. Set $K_1 := K^H$.

Now consider the Galois extension K/K_1. This has degree p^{n-1} by the tower law, so $\text{Gal}(K/K_1)$ has order p^{n-1}. This is still a p-group, so there exists a subgroup A of order p^{n-2}. This subgroup then has index $\frac{p^{n-1}}{p^{n-2}} = p$, so it is normal in $\text{Gal}(K/K_1)$. Thus if K_2 denotes the fixed field K^A of $A \leq \text{Gal}(K/K_1)$, K_2 is a Galois extension of K_1 of degree p. And so on, continuing in this manner produces the desired K_3 as the fixed field of a subgroup of order p^{n-3} of $\text{Gal}(K/K_2)$, then K_4, and onward as required.

Revisiting constructibility. In particular, we see that if $[K : F]$ is a power of 2, then K can be constructed from F by a sequence of successive quadratic extensions. This fact gives the converse to the claim that if α is a constructible real number (in the straightedge and compass sence), then $[\mathbb{Q}(\alpha) : \mathbb{Q}]$ is a power of 2. Recall that α is constructible if and only if it can be expressed in terms of rational numbers using only the operators of addition, subtraction, multiplication, diviison, and (repeated) square root extractions, and we argued previously that if this is true, then $\mathbb{Q}(\alpha)$ has degree over $\mathbb{Q}$ which is a power of 2, since each new square root we introduce in expressing α requires moving to a new quadratic extension. We now claim that if $[\mathbb{Q}(\alpha) : \mathbb{Q}]$ is a power of 2, then α is constructible.

Indeed, if $[\mathbb{Q}(\alpha) : \mathbb{Q}]$ is a power of 2, then the Warm-Up gives

$$\mathbb{Q} \subseteq K_1 \subseteq \ldots \subseteq K_{n-2} \subseteq K_{n-1} \subseteq \mathbb{Q}(\alpha)$$

where each field is of degree 2 over the previous one. Any extension of degree 2 over $\mathbb{Q}$ is obtained by adjoining a square root (we did this as a Warm-Up early in the quarter), so $K_i = K_{i-1}(\sqrt{D_i})$ for some non-square $D_i \in K_{i-1}$. (Set $K_0 = \mathbb{Q}$ and $K_n = \mathbb{Q}(\alpha)$ so that this notation works for all fields above.) Thus, we have $\mathbb{Q}(\alpha) = K_{n-1}(\sqrt{D_n})$, so

$$\alpha = a + b\sqrt{D_n} \text{ for some } a, b \in K_{n-1}.$$

But now a, b, D_n are all in $K_{n-1} = K_{n-2}(\sqrt{D_{n-1}})$, so each is expressible as

$$a = x + y\sqrt{D_{n-1}},\ b = c + d\sqrt{D_{n-1}},\ D_n = s + t\sqrt{D_{n-1}}$$

for some $x, y, c, d, s, t \in K_{n-2}$. Plugging these into $\alpha = a + b\sqrt{D_n}$ gives an expression for α in terms of $x, y, c, d, s, t, D_{n-1}$ and square root extractions. Now do the same for $x, y, c, d, s, t, D_{n-1} \in K_{n-2} = K_{n-3}(\sqrt{D_{n-2}})$, expressing each in terms of elements of K_{n-3}, $D_{n-2} \in K_{n-2}$, and square root extractions. Continuing down all the way to $K_0 = \mathbb{Q}$ will give an expression for α in terms of rationals and the basic operations (including square root extractions), showing that α is constructible. Thus, we conclude that α is constructible if and only if $[\mathbb{Q}(\alpha) : \mathbb{Q}]$ is a power of 2. (Proving this converse direction without Galois theory is bound to be quite tedious, since knowing only that $\mathbb{Q}(\alpha)$ has degree 2^n over $\mathbb{Q}$ does not at all immediately suggest how to actually express α in a certain way, let alone via repeated square root extractions. I think that it is probably possible to do this directly—indeed, something along these lines was likely the proof that Gauss originally gave of this fact in the 19th century—but it for sure won't be easy.)

Bound on degree. Now we come back to the one unproven claim we have left: if H is a finite subgroup of $\text{Aut}(K)$, then $[K : K^H] = |H|$. Recall that this formed the basis of our starting point in the proof of the Fundamental Theorem of Galois Theory, but we postponed the proof until now. Part of the reason why we did so is because the first step, namely that $[K : K^H] \leq |H|$, is a bit technical and I felt it would detract us from the key ideas in Galois theory. The proof does illustrate some nice ideas, in particular how the fact that K^H is a fixed field comes in, but is not all that illustrative of crucial techniques going forward. Nonetheless, here we go. We give essentially the same proof as the book, only cleaned up a bit.

Denote the elements of H by $H = \{id, \sigma_1, \ldots, \sigma_k\}$ and suppose $\alpha_1, \ldots, \alpha_m \in K$ are nonzero with $m > |H|$. We claim that these elements are linearly *dependent* over K^H (i.e. considering K as a vector space over K^H), which if true gives the desired inequality $[K : K^H] \leq |H|$: no linearly independent set of elements can have size larger than $|H|$, so the size $[K : K^H]$ of a basis for K over K^H must in particular be at most $|H|$. Apply each element of H to the α_j and form the following system of linear equations with coefficients the resulting elements of K:

$$\begin{aligned}
\alpha_1 x_1 + \cdots + \alpha_m x_m &= 0 \\
\sigma_1(\alpha_1) x_1 + \cdots + \sigma_1(\alpha_m) x_m &= 0 \\
\sigma_2(\alpha_1) x_1 + \cdots + \sigma_2(\alpha_m) x_m &= 0 \\
\vdots \qquad\qquad\qquad \vdots \quad &\ \ \vdots \\
\sigma_k(\alpha_1) x_1 + \cdots + \sigma_k(\alpha_m) x_m &= 0.
\end{aligned}$$

So, the first equation is the one with coefficients $id(\alpha_j)$, and the i-th equation ($i > 1$) has coefficients $\sigma_i(\alpha_j)$. Since $m > |H|$, this system has more variables than equations, and hence there exists a nontrivial solution $(x_1, \ldots, x_m)$ where each $x_i \in K$. We may assume that this solution has the *minimal* number of nonzero values (there is at least one nonzero value since the solution is nontrivial), and by rearranging the x_j's and α_j's if necessary we can assume the nonzero values occur at the beginning, so that in particular $x_1 \neq 0$.

Now, the nonzero product $x_1\sigma_1(x_1)\sigma_2(x_1)\cdots\sigma_m(x_1) \in K$ is fixed by all elements of H, since applying any element of H will simply permuate the factors. (Note that since H is a group, left multiplication by any element in it is a permutation of the elements.) Thus $x_1\sigma_1(x_1)\sigma_2(x_1)\cdots\sigma_m(x_1)$ belongs to the fixed field K^H. After multiplying the first equation in our system through by $\sigma_1(x_1)\sigma_2(x_1)\cdots\sigma_m(x_1)$ we thus obtain a solution to our system of linear equations whose first entry is an element of K^H, so we may as well assume that x_1 was already in K^H. The claim is then that *all* x_j are actually in K^H. If so, then the first equation of our system

$$x_1\alpha_1 + \cdots + x_m\alpha_m = 0$$

expresses 0 as a nontrivial linear combination of the α_j over K^H, which shows they are linearly dependent over K^H as required.

If there is some x_t not in K^H, then there is some σ_ℓ that does not fix x_t: $\sigma_\ell(x_t) \neq x_t$. Applying σ_ℓ to all equations in our system produces a new system with the *same* coefficients, only with the x_j replaced by $\sigma_\ell(x_j)$; in other words, we get the same system of equations but with solution $(\sigma_\ell(x_1), \dots, \sigma_\ell(x_m))$ instead. Indeed, the point is that the set of products $\sigma_\ell \sigma_i \in H$ is just H itself with the elements permuted, so the effect of applying σ_ℓ is to simply permute the equations of our system. For example, the new first equation is

$$\sigma_\ell(\alpha_1)(\sigma_\ell(x_1)) + \cdots + \sigma_\ell(\alpha_m)(\sigma_\ell(x_m)) = 0,$$

which is the original $(\ell+1)$-st equation evaluated at $(\sigma_\ell(x_1), \dots, \sigma_\ell(x_m))$.

So we now have two solutions of our system of equations:

$$(x_1, \dots, x_m) \quad \text{and} \quad (\sigma_\ell(x_1), \dots, \sigma_\ell(x_m)).$$

Since our system of equations is homogeneous (meaning we have all zeroes to the right side of the equal signs), the difference of these two solutions is still a solution, so

$$(x_1 - \sigma_\ell(x_1), \dots, x_m - \sigma_\ell(x_m))$$

is a solution of our system of equations. But since $x_1 \in K^H$, $x_1 = \sigma_\ell(x_1) = 0$, so the first value in this new solution is zero. This solution is nontrivial since the entry $x_t - \sigma_\ell(x_t)$ is nonzero by the choice of σ_ℓ, so this is thus a solution with a fewer number of nonzero values than the original $(x_1, \dots, x_m)$. (Note that any x_j which were originally zero still gives $x_j - \sigma_i(x_j) = 0$.) This contradicts the choice of $(x_1, \dots, x_m)$ as a solution with a minimal number of nonzero components, so we conclude (finally!) that all x_j are in fact in K^H. As explained above, this then shows that $\alpha_1, \dots, \alpha_m \in K$ are linearly dependent over K^H, as desired. (Maybe now you can see why I saved this proof until the end!)

Bound on order. The remaining inequality, $|H| \leq [K : K^H]$, is simpler to justify. We will give a different proof than the book does, which uses ideas we've already seen quite a bit. (The book's proof is phrased using the language of *characters*, which are homomorphisms from a group into the multiplicative group of a field. This is certainly an important topic in further study of algebra, but is not something we will need in our course, hence why I don't want to take the time to introduce them.) By what we just proved above, we know that K is a finite extension of K^H. Let $\alpha_1, \dots, \alpha_n$ be a basis, so that $K = K^H(\alpha_1, \dots, \alpha_n)$.

We count (or rather bound) the number of embeddings of K into, say, the algebraic closure of K^H over K^H. (We use the algebraic closure because it is guaranteed to be a field that contains the roots of all polynomials over K^H.) Any embedding of $K \to \overline{K^H}$ restricts to an embedding of $K^H(\alpha_1)$, and to an embedding of $K^H(\alpha_1, \alpha_2)$, and to an embedding of $K^H(\alpha_1, \alpha_2, \alpha_3)$, and so on. Thus we can bound the number of embeddings by bounding the number of such restrictions at each step. At the first step, we have argued (in the proof that splitting fields have automorphism group order bounded by the degree) that the number of embeddings of $K^H(\alpha_1)$ is bounded by the number of roots of the minimal polynomial of α_1 over K^H (since α_1 in particular has to be sent to such a root), which in turn is bounded by the degree of this polynomial, which is equal to $[K^H(\alpha_1) : K^H]$.

In the same way, at the next step the number of embeddings of $K^H(\alpha_1, \alpha_2)$ (given that we have already specified what happens to α_1) is bounded by the number of roots of the minimal polynomial of α_2 over $K^H(\alpha_1)$, which is bounded by $[K^H(\alpha_1, \alpha_2) : K^H(\alpha_1)]$. At the next stage, the number of embeddings of $K^H(\alpha_1, \alpha_2, \alpha_3)$ is bounded by $[K^H(\alpha_1, \alpha_2, \alpha_3) : K^H(\alpha_1, \alpha_2)]$, and so on as we go "up" the tower. In the end, the number of embeddings of $K = K^H(\alpha_1, \dots, \alpha_n)$ is bounded by the product of the number of embeddings at each step, which is:

$$[K : K^H(\alpha_1, \dots, \alpha_{n-1})] \cdots [K^H(\alpha_1, \alpha_2) : K^H(\alpha_1)][K^H(\alpha_1) : K^H] = [K : K^H].$$

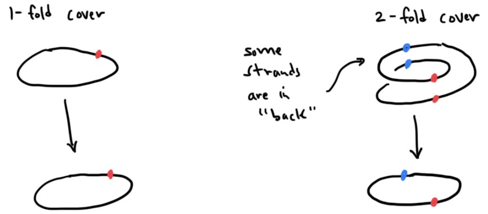
Thus $|\operatorname{Emb}(K/K^H)| \leq [K : K^H]$, and since automorphisms are special types of embeddings, we have:

$$|H| \leq |\operatorname{Aut}(K/K^H)| \leq |\operatorname{Emb}(K/K^H)| \leq [K : K^H],$$

as desired. We thus conclude that $[K : K^H] = |H|$, our last unproven claim. (This same reasoning shows that for any finite extension K/F, splitting field or not, we have $|\operatorname{Aut}(K/F)| \leq [K : F]$. We instead derived this fact a few days ago as a consequence of the $[K : K^H] = |H|$ claim.)

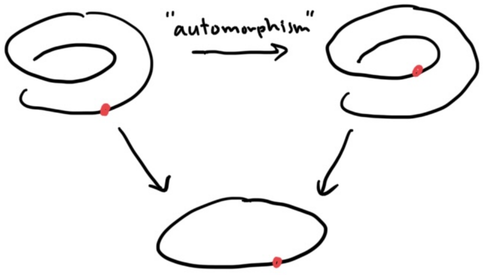
Galois correspondence in topology. Next time we will begin to look at more concrete applications of the Fundamental Theorem of Galois Theory, but we finish for now with a brief digression meant to illustrate how the ideas of Galois theory, in terms of creating a "dictionary" between certain mathematical objects (fields in the case at hand) and groups, can show up in other areas.

Recall that on the last day of the fall quarter we spent a brief amount of time talking about the concept of the *fundamental group* of a space. (The fundamental group of a space is a group constructed out of the *loops* in that space, but you can check the notes from the fall for clarification.) The intent then was simply to illustrate an interesting way in which groups show up in topology, and to hint at how *free groups* in particular can be studied via topological means. Now, an important notion in topology is that of a *covering space*, which is space that in a sense "covers" another space. For example, we can view the circle as a "1-fold" cover of itself, which means that we can project a circle onto a circle so that any point below corresponds to one point above:



By doubling up the circle above into a "double helix" we get a 2-fold cover of the circle, as in the second picture above: each point in the circle below now corresponds to two points in the (twisted) circle above. And so on, we can form a 3-fold cover of the circle, a 4-fold cover, etc. Other spaces besides circles can be "covered" by other "covering spaces" as well.

One can then study the group of "automorphisms" of a cover, which are continuous mappings of the cover to itself that "preserve" the base space being covered:



It turns out that this automorphism group can be viewed as a subgroup of the fundamental group of the base space, and lo-and-behold there ends up being a one-to-one correspondence between covers of X and subgroups of the fundamental group of X. This correspondence is known as the *Galois correspondence* in topology, so named because of the obvious analogy with the relation between field extensions and groups given by Galois theory. (In a sense, a covering space is somehow analogous to an extension field.) Numerical data on the topology side, such as the "degree" of the cover, corresponds to numerical data on the group side—orders, indices—and certain types of covers correspond to normal subgroups. So, precisely the types of things one would expect a "Galois correspondence" to preserve! Take a topology course like MATH 344-2 to learn more.

That's all we'll say about this topic, except for this final crazy thing. The idea of viewing covers and extensions fields as analogous to one another is purely meant in a figurative and not literal sense... or so one would think! Crazy as it may sound, given how disparate fields and spaces seem to be, there *is* actually a way to think about field extensions as if they were literal covering spaces, in the setting of algebraic geometry! (Recall from a brief discussion last quarter that algebraic geometry provides a way to study algebra—in that case rings—by treating them as if they were geometric objects.) The claim is that a Galois extension K/F can be viewed as a literal type of "covering space", only that you have to greatly generalize what you mean by "cover". (This requires a whole bunch of the subject called *category theory* to make precise, but it can be done!) One you do this, it turns out that the Galois group of K/F becomes the literal (well, almost literal) "fundamental group" of the "covering space" K/F, thereby merging the two notions of "Galois correspondence" together. Pretty awesome (albeit highly technical) stuff! (If you have ever seen the movie *A Beautiful Mind* about John Nash, there is a scene towards the end were Nash—played by Russell Crowe—is talking to a graduate student, who says something like "I believe I can prove that Galois extensions *are* covering spaces." Now, of course, the person who wrote the movie just through this in in order to include some cool-sounding mathematical buzzwords, but there's no doubt that real mathematicians were consulted when drafting this since the idea that "Galois extensions are covering spaces" is an actual thing that people study! Fun fact: this is now the second time I make reference to this movie in my teaching—I also mention it when teaching MATH 291-3!)

Lecture 16: Simple Extensions

Warm-Up. Suppose E_1 and E_2 are both Galois extensions of F. We show that the intersection $E_1 \cap E_2$ and the composite E_1E_2 are Galois over F as well. For the intersection, we use the characterization of Galois as being normal and separable. If E_1 and E_2 are both separable over F, then $E_1 \cap E_2$ is too since repeated roots in $E_1 \cap E_2$ of an irreducible polynomial over F whould also be repeated in both E_1 and E_2. If $p(x) \in F[x]$ is irreducible and has a root $\alpha \in E_1 \cap E_2$, then it also has a root in each E_i and thus splits in E_i since E_i is normal over F. Thus $p(x)$ splits in $E_1 \cap E_2$, so $E_1 \cap E_2$ is Galois over F.

For the composite, for $i = 1, 2$ take a separable polynomial $f_i(x) \in F[x]$ whose splitting field is E_i. (This uses the splitting field characterization of Galois.) Then the splitting field of the product $f_1(x)f_2(x)$ is E_1E_2: all roots of $f_i(x)$ lie in E_i, so all roots of the product are in E_1E_2, and $f_1(x)f_2(x)$ does not split in any proper subfield of E_1E_2 since the splitting field of the product must contain both E_1 and E_2 (since $f_i(x)$ splits in E_i), and hence must contain the smallest field extending both E_1 and E_2, which is precisely E_1E_2. Now, $f_1(x)f_2(x)$ might not be separable, but by factor each $f_i(x)$ into linear terms and removing any duplicate factors we obtain a separable polynomial over F whose splitting field is still E_1E_2, so E_1E_2 is Galois over F.

Remark. After discussing the Fundamental Theorem of Galois Theory, the book spends a section talking about finite fields, and then has more to say about composites in the following section. All of the material here on finite fields is something we've already discussed in class or on the homework, except for one minor point we will clarify in a bit. So, we will skip this discussion. Similarly, we are only highlighting the aspects of the Galois theory of composite fields that are important for our purposes as we need them, and will omit some of the less crucial results.

Galois closures. The results of the Warm-Up allow us to construct, for any finite separable extension of a field F, a Galois extension which contains it. This is good since Galois extensions are the ones to which the Fundamental Theorem of Galois Theory applies, and we are saying that we can study *arbitrary* finite and separable extensions via Galois ones. If K/F is finite and separable, say generated by $\alpha_1, \ldots, \alpha_n \in K$, let E_i denote the splitting field of the minimal polynomial of α_i. Then the composite $E_1 \ldots E_n$ is a Galois extension of F (by the Warm-Up and induction) that contains K.

If there is *a* Galois extension of F containing K, then we can ask for the *smallest* such extension, and this is what we call the *Galois closure* of K/F. This can be constructed as the intesection of all Galois extensions of K containing F, where we again use the Warm-Up (or rather, its generilzation to more than 2 Galois extensions) to argue that this intersection is Galois. The upshot is that we can always a nice Galois extension containing a given (finite and separable) K/F to work with.

In practice, the Galois closure is not to difficult to determine when given explicit fields. For example, we claim that the Galois closure of $\mathbb{Q}(\sqrt[3]{2})$ over $\mathbb{Q}$ is $\mathbb{Q}(\sqrt[3]{2}, \zeta_3)$, or in other words the splitting field of $x^3 - 2 \in \mathbb{Q}[x]$. Indeed, this splitting field is certainly a Galois extension of $\mathbb{Q}$ that contains $\mathbb{Q}(\sqrt[3]{2})$. To see that it is the smallest such extension, note that any Galois extension containing $\mathbb{Q}(\sqrt[3]{2})$ contains the root of $\sqrt[3]{2}$ of $x^3 - 2$, and thus, being normal, must contain all the roots. Hence any Galois extension containing $\mathbb{Q}(\sqrt[3]{2})$ contains all of $\mathbb{Q}(\sqrt[3]{2}, \zeta_3)$, so this is indeed the Galois closure. More generally, the Galois closure of $\mathbb{Q}(\alpha)$ over $\mathbb{Q}$ will be the splitting field of the minimal polynomial of α over $\mathbb{Q}$.

Simple extensions. Now we come to justify a fact we've mentioned before, namely that any finite extension of $\mathbb{Q}$ is simple. (Recall that a *simple* extension of F is one $F(\alpha)$ which is generated by a single element. We call this element a *primiitive element* for the extension.) More generally, the claim is that any finite separable extension of any field is simple.

But before proving this, we first state the following characterization of simple extensions, at least in the finite case: if K/F is finite, then K is simple if and only if there are only finitely many intermediate fields between F and K. (Note that there is no separaibility assumption here.) For the forwards direction (we give this direction for completeness, but it is not essential), suppose that α generates K over F. If E is an intermediate field $F \subseteq E \subseteq K$, then the minimal polynomial of α over E has degree $[K : E]$. If E' is the field generated by the coefficients of this minimal polynomial, then the minimal polynomial of α over E' is the same as that over E, so $[K : E']$ is also $[K : E]$. But $E' \subseteq E$ since E already contains the coefficients of this minimal polynomial, so $E' = E$ and we conclude that all the only intermediate fields $F \subseteq E \subseteq K$ are those generated by minimal polynomials of α over subextensions of K. Any such minimal polynomial divides the minimal polynomial of α over F, so since this latter polynomial has only finitely many factors, there can be only finitely many such intermediate fields as desired.

For the backwards direction, which is the one we actually care more about, we must distinguish between the case of finite fields vs infinite fields. Since any finite field is a finite extension of some $\mathbb{F}_p$, it is enough to show that any finite field (of characteristic p) is a simple extension of $\mathbb{F}_p$. (Then when $\mathbb{F}_{p^m} \subseteq \mathbb{F}_{p^n}$, $\mathbb{F}_{p^n}$ being simple over $\mathbb{F}_p$ implies that it is simple over $\mathbb{F}_{p^m}$ as well with the same generator.) To see that $K = \mathbb{F}_{p^n}$ is simple over $\mathbb{F}_p$, let $\alpha \in \mathbb{F}_{p^n}$ be a generator of the cyclic multiplicative group $\mathbb{F}_{p^n}^\times$. Then we claim that α generates all of $\mathbb{F}_{p^n}$ over $\mathbb{F}_p$. Indeed, if $\mathbb{F}_p(\alpha)$ was contained in a proper subfield $\mathbb{F}_{p^m}$ of $\mathbb{F}_{p^n}$, then α would belong to the multiplicative group $\mathbb{F}_{p^m}^\times$ and would have order dividing $p^m - 1$, so that it could not generate $\mathbb{F}_{p^n}^\times$. Hence $\mathbb{F}_{p^n} = \mathbb{F}_p(\alpha)$ for this α as claimed.

Now suppose F is infinite and let $\alpha, \beta \in K$. The fields $F(\alpha + c\beta)$ for $c \in F$ are all intermediate between F and K, so since by assumption there can only be finitely many of these, we must have

$$F(\alpha + c_1\beta) = F(\alpha + c_2\beta)$$

for some $c_1 \neq c_2 \in F$. But then the difference $(c_1 - c_2)\beta$ of the generators is in $F(\alpha + c_1\beta)$, and thus so is β after dividing by $c_1 - c_2 \in F$, and hence so is $\alpha = (\alpha + c_1\beta) - c_1\beta$. We conclude that $F(\alpha, \beta) \subseteq F(\alpha + c_1\beta)$, so we have equality since the other containment is clear. By induction, we thus have that any $F(\alpha_1, \ldots, \alpha_n)$ is simple over F, and K is of this form.

Primitive Element Theorem. The *primitive element theorem* states any finite separable extension of a field is simple. (In particular then, any finite extension of a field of characteristic zero, such as $\mathbb{Q}$, is simple.) The proof is now a quick application of Galois theory and the characterization of simple extensions given above. If K/F is finite and separable, then consider the Galois closure L of K/F. Then $\mathrm{Gal}(L/F)$ is a finite group (Galois extensions are always of finite degree), so it has only finitely many subgroups. But any field intermediate between F and K is the fixed field of one of these subgroups, so there can only be finitely many such intermediate extensions. Thus K is simple over F by the characterization of simple extensions above.

We should note that there are ways of proving the primitive element theorem without using Galois theory, but they are fairly tedious and require some heavy computations with the Euclidean algorithm. (One approach to Galois theory is to first prove the primitive element theorem without it, and then derive the Fundamental Theorem of Galois Theory from it.) The approach we have followed is, I think at least, a better one conceptually.

Fundamental Theorem of Algebra. We finish by giving a proof of the Fundamental Theorem of Algebra using Galois theory and the primitive element theorem. This will be, for the most part, a purely algebraic proof, except for one property of polynomials over $\mathbb{R}$ that requires some analysis (or calculus, really) to understand. We will point this out when we get there. We should point out that the book gives two proofs of the Fundamental Theorem of Algebra later on when discussing Galois groups of polynomials over $\mathbb{Q}$, but the proof we give here is not quite the same. This proof is a "tour de force" in using the machinery of group theory to avoid doing any hard computations with fields.

Recall the claim is that $\mathbb{C}$ is algebraically closed, which means that any polynomial over $\mathbb{C}$ has a root in $\mathbb{C}$. As we mentioned a while back, this is equivalent to the claim that there are no proper finite extensions of $\mathbb{C}$, which is what we actually prove. (A proper finite extension would be generated by an element with minimal polynoial of degree at least 2, and so would be a polynoimal with no root in $\mathbb{C}$.) So, suppose K is a finite extension of $\mathbb{C}$. Our goal is to show that $K = \mathbb{C}$.

Since K is finite over $\mathbb{C}$ and $\mathbb{C}$ is finite over $\mathbb{R}$, K is finite over $\mathbb{R}$ and

$$[K : \mathbb{R}] = [K : \mathbb{C}][\mathbb{C} : \mathbb{R}] = 2[K : \mathbb{C}].$$

The extension $K/\mathbb{R}$ is finite and separable, so by replacing K by its Galois closure over $\mathbb{R}$ we can assume that K is Galois over $\mathbb{R}$. Then $\mathrm{Gal}(K/\mathbb{R})$ has order $[K : \mathbb{R}]$, which is even. Write this order has $|\,\mathrm{Gal}(K/\mathbb{R})| = 2^n m$ where m is odd, and let H be a Sylow 2-subgroup of $\mathrm{Gal}(K/\mathbb{R})$, so that

$|H| = 2^n$. The index of H in $\text{Gal}(K/\mathbb{R})$ is then m, so the Fundamental Theorem of Galois Theory says that the fixed field K^H has degree m over $\mathbb{R}$.

But m is odd, and we claim that the only extension of $\mathbb{R}$ of odd degree is $\mathbb{R}$ itself. Indeed, $K^H = \mathbb{R}(\alpha)$ for some α by the primitive element theorem, and the minimal polynomial of α over $\mathbb{R}$ then has degree m. Here is the fact from analysis/calculus we need: any polynomial $p(x)$ of odd degree over $\mathbb{R}$ in fact has a *real* root! This is a consequence of the Intermediate Value Theorem: since the degree is odd, one of $\lim_{x\to\infty} p(x)$ and $\lim_{x\to-\infty} p(x)$ is $+\infty$ and the other $-\infty$, so since $p(x)$ is continuous it must attain the value 0 at some $c \in \mathbb{R}$, which is then a real root. (We won't give a proof of this here—it is a standard part of any analysis course.) In our case, this means that the minimal polynomial of α has degree 1 since if it had larger degree it could not be irreducible. This implies $\alpha \in \mathbb{R}$, so $K^H = \mathbb{R}(\alpha) = \mathbb{R}$, and hence $m = 1$.

Thus $|\text{Gal}(K/\mathbb{R})| = 2^n$, and in turn $|\text{Gal}(K/\mathbb{C})| = 2^{n-1}$. If this order is larger than 1, then $\text{Gal}(K/\mathbb{C})$ is a 2-group, so there exists an extension of $\mathbb{C}$ of degree 2. (See the Warm-Up and discussion about constructibility from last time.) But this is not possible since the square root of any complex number is complex by the quadratic formula, so we conclude that $|\text{Gal}(K/\mathbb{C})| = 1$. Thus K has degree 1 over $\mathbb{C}$, so $K = \mathbb{C}$ as claimed.

Lecture 17: More on Cyclotomic Extensions

Warm-Up. Suppose p is prime. We show that $\mathbb{F}_p(x, y)$ is a finite but non-simple extension of $\mathbb{F}_p(x^p, y^p)$. Here, $\mathbb{F}_p(x, y)$ is the field of two-variable rational functions (quotients of two-variable polynomials) over $\mathbb{F}_p$, and $\mathbb{F}_p(x^p, y^p)$ is the same only with x and y appearing with exponents that are multiples of p. Note that the Primitive Element Theorem implies that non-simple finite extensions do not exist over fields of characteristic 0 nor over finite fields since such extensions are always separable, so if we want a non-simple finite extension we must work with infinite fields of prime characteristic. ($\mathbb{F}_p(x, y)$ is inseparable over $\mathbb{F}_p(x^p, x^p)$ since $X^p - x^p = (X - x)^p$ is irreducible over $\mathbb{F}_p(x^p, y^p)$ and has repeated root $x \in \mathbb{F}_p(x, y)$.)

First, the fact that $\mathbb{F}_p(x, y)$ is finite over $\mathbb{F}_p(x^p, y^p)$ comes (among other ways) from viewing it as the composite of $\mathbb{F}_p(x, y^p)$ and $\mathbb{F}_p(x^p, y)$: this composite should be the smallest extension of $\mathbb{F}_p(x^p, y^p)$ containing both x and y, and this is $\mathbb{F}_p(x, y)$. The fields $\mathbb{F}_p(x, y^p) = \mathbb{F}_p(x^p, y^p)(x)$ and $\mathbb{F}_p(x^p, y) = \mathbb{F}_p(x^p, y^p)(y)$ each have degree p over $\mathbb{F}_p(x^p, y^p)$, since the minimal polynomial the generator (x or y) over $\mathbb{F}_p(x^p, y^p)$ is $X^p - x^p$ in the first case and $X^p - y^p$ in the second. The composite $\mathbb{F}_p(x, y)$ thus as degree at most $[\mathbb{F}_p(x, y^p) : \mathbb{F}_p(x^p, y^p)][\mathbb{F}_p(x^p, y) : \mathbb{F}_p(x^p, y^p)] = p^2$, so it is finite over $\mathbb{F}_p(x^p, y^p)$. Actually, the degree is exactly p^2. We can see this by working out the degree of $\mathbb{F}_p(x, y)$ over $\mathbb{F}_p(x, y^p)$, for instance, using minimal polynomials: the minimal polynomial of $y \in \mathbb{F}_p(x, y)$ over $\mathbb{F}_p(x, y^p)$ divides $X^p - y^p$, so this extension has degree dividing p, and it is not 1 since $\mathbb{F}_p(x, y) \neq \mathbb{F}_p(x, y^p)$.

Now, to see that $\mathbb{F}_p(x, y)$ is not simple over $\mathbb{F}_p(x^p, y^p)$, let $\alpha \in \mathbb{F}_p(x, y)$. Then

$$\alpha = \frac{f(x, y)}{g(x, y)}$$

for some $f(x, y), g(x, y) \in \mathbb{F}_p[x, y]$. By the freshman's dream and the fact that Frobenius fixes $\mathbb{F}_p$, we have $f(x, y)^p = f(x^p, y^p)$ and $g(x, y)^p = g(x^p, y^p)$, so that

$$\alpha^p = \frac{f(x, y)^p}{g(x, y)^p} = \frac{f(x^p, y^p)}{g(x^p, y^p)} \in \mathbb{F}_p(x^p, y^p).$$

Hence α is a root of $X^p - \alpha^p$ over $\mathbb{F}_p(x^p, y^p)$, so its minimal polynomial has degree at most p and hence $\mathbb{F}_p(x, y)(\alpha)$ has degree at most p over $\mathbb{F}_p(x^p, y^p)$. Thus no elmeent of $\mathbb{F}_p(x, y)$ can generate an extension of degree p^2 over $\mathbb{F}_p(x^p, y^p)$, so $\mathbb{F}_p(x, y)$ is not simple over $\mathbb{F}_p(x^p, y^p)$.

Degrees of composites. Before moving on, we highlight one fact about composites of Galois extensions. If E_1, E_2 are Galois over F, we already know that $[E_1E_2 : F]$ is at most $[E_1 : F][E_2 : F]$, but now we claim that we can give the following exact value for $[E_1E_2 : F]$:

$$[E_1E_2 : F] = \frac{[E_1 : F][E_2 : F]}{[E_1 \cap E_2 : F]}.$$

The book derives this in the more general setting where only one of E_1 or E_2 is assumed to be Galois over F, but we give a simpler proof here assuming both are Galois. This is in some sense the field-theoretic analog of the equality

$$|AB| = \frac{|A||B|}{|A \cap B|}$$

we saw for subgroups A, B of a group G in the fall. Note we showed last time that if E_1 and E_2 are Galois over F, then so are E_1E_2 and $E_1 \cap E_2$.

First, the join $\langle \operatorname{Gal}(E_1E_2/E_1), \operatorname{Gal}(E_1E_2, E_2)\rangle$ (as a subgroup of $\operatorname{Gal}(E_1E_2/F)$) is simply the product $\operatorname{Gal}(E_1E_2/E_1)\operatorname{Gal}(E_1E_2/E_2)$: this product *is* a subgroup given that each factor is normal in $\operatorname{Gal}(E_1E_2/F)$, so it is indeed the smallest subgroup containing both factors, as the join should be. The join is the Galois group of E_1E_2 over $E_1 \cap E_2$ by the Fundamental Theorem of Galois Theory, so since $\operatorname{Gal}(E_1 \cap E_2/F) \cong \operatorname{Gal}(E_1E_2/F)/\operatorname{Gal}(E_1E_2/E_1 \cap E_2)$, we have:

$$|\operatorname{Gal}(E_1 \cap E_2/F)| = \frac{|\operatorname{Gal}(E_1E_2/F)|}{|\operatorname{Gal}(E_1E_2/E_1 \cap E_2)|} = \frac{|\operatorname{Gal}(E_1E_2/F)|}{|\operatorname{Gal}(E_1E_2/E_1)||\operatorname{Gal}(E_1E_2/E_2)|}.$$

Now, $\operatorname{Gal}(E_i/F) \cong \operatorname{Gal}(E_1E_2/F)/\operatorname{Gal}(E_1E_2/E_i)$, so

$$|\operatorname{Gal}(E_i/F)| = \frac{|\operatorname{Gal}(E_1E_2/F)|}{|\operatorname{Gal}(E_1E_2/E_i)|}.$$

Putting it all together gives

$$|\operatorname{Gal}(E_1 \cap E_2/F)| = \frac{|\operatorname{Gal}(E_1E_2/F)|}{|\operatorname{Gal}(E_1E_2/E_1)||\operatorname{Gal}(E_1E_2/E_2)|} = \frac{|\operatorname{Gal}(E_1/F)||\operatorname{Gal}(E_2/F)|}{|\operatorname{Gal}(E_1E_2/F)|}.$$

Rearranging and replacing these orders by degrees of field extensions gives the desired equality.

Note in the Warm-Up that this equality does happen to give the right answer p^2 for the degree of $\mathbb{F}_p(x, y)$ over $\mathbb{F}_p(x^p, y^p)$ even though the extensions involved are not Galois, where we use the fact that $\mathbb{F}_p(x, y^p) \cap \mathbb{F}_p(x^p, y) = \mathbb{F}_p(x^p, y^p)$. Thus this equality can hold in settings other than the Galois case, even though the Galois case will be our primary focus going forward.

Revisiting cyclotomic extensions. Before moving on to focusing more heavily on Galois groups of polynomials next time, we revisit the topic of cyclotomic fields and definitively settle the problem of constructing regular n-gons with straightedge and compass. Recall that the *n-th cyclotomic extension* of $\mathbb{Q}$ is $\mathbb{Q}(\zeta_n)$ where ζ_n is a primitive (complex) n-th root of unity. The minimal polynomial of ζ_n over $\mathbb{Q}$ is the n-th cyclotomic polynomial $\phi_n(x)$, whose roots are precisely the primitive n-th roots, which can all be written as ζ_n^a for some a relatively prime to n:

$$\phi_n(x) = \prod_{\substack{(a,n)=1 \\ 1 \le a \le n}} (x - \zeta_n^a).$$

The degree of $\phi_n(x)$, and hence the degree of the extension $\mathbb{Q}(\zeta_n)/\mathbb{Q}$ is $\varphi(n)$, the number of positive integers less than n that are relatively prime to n.

Let us now determine the Galois group $\text{Gal}(\mathbb{Q}(\zeta_n)/\mathbb{Q})$. Any element of this group is determined by its value on the generator ζ_n, which has to be sent to another primitive n-th root of unity since the Galois group permutes the roots of $\phi_n(x)$. Thus an element of the Galois group is determined by some $1 \leq a \leq n, (a, n) = 1$ via $\zeta_n \mapsto \zeta_n^a$, and for any such a we get such a map. This thus gives a bijection between $\text{Gal}(\mathbb{Q}(\zeta_n)/\mathbb{Q})$ and $(\mathbb{Z}/n\mathbb{Z})^\times$, the group of elements of $\mathbb{Z}/n\mathbb{Z}$ that are relatively prime to n:

$$\text{Gal}(\mathbb{Q}(\zeta_n)/\mathbb{Q}) \to (\mathbb{Z}/n\mathbb{Z})^\times \text{ defined by } \sigma \mapsto \text{the exponent } a \text{ in } \sigma(\zeta_n) = \zeta_n^a.$$

We claim this bijection is actually a group isomorphism, meaning that composition in the Galois group corresponds to multiplication in the multiplicative group. Indeed, if σ_a and σ_b are the maps corresponding to $a, b \in (\mathbb{Z}/n\mathbb{Z})^\times$ respectively, then

$$(\sigma_a\sigma_b)(\zeta_n) = \sigma_a(\zeta_n^b) = (\zeta_n^b)^a = \zeta_n^{ab} = \sigma_{ab}(\zeta_n),$$

so that the composition $\sigma_a\sigma_b$ corresonds to ab. Thus $\text{Gal}(\mathbb{Q}(\zeta_n)/\mathbb{Q}) \cong (\mathbb{Z}/n\mathbb{Z})^\times$ as groups.

A final observation comes from the Chinese Remainder Theorem: if $n = p_1^{k_1} \cdots p_m^{k_m}$ is the prime factorization of n with the p_i distinct, we saw last quarter that

$$(\mathbb{Z}/n\mathbb{Z})^\times \cong (\mathbb{Z}/p_1^{k_1}\mathbb{Z})^\times \times \cdots \times (\mathbb{Z}/p_m^{k_m}\mathbb{Z})^\times$$

as a consequence of the Chinese Remainder Theorem. By the work above, $(\mathbb{Z}/p_i^{k_i}\mathbb{Z})^\times$ is the Galois group of $\mathbb{Q}(\zeta_{p_i^{k_i}})/\mathbb{Q}$, so this isomorphism turns into

$$\text{Gal}(\mathbb{Q}(\zeta_n)/\mathbb{Q}) \cong \text{Gal}(\mathbb{Q}(\zeta_{p_1^{k_1}})/\mathbb{Q}) \times \cdots \times \text{Gal}(\mathbb{Q}(\zeta_{p_m^{k_m}})/\mathbb{Q}).$$

This reflects the fact that $\mathbb{Q}(\zeta_n)$ is the composite of the fields $\mathbb{Q}(\zeta_{p_i^{k_i}})$, which can be proved using the formula for the degree of composities we derived earlier. Check the book for details if interested.

Constructing polygons. We can now address the constructibility of regular polygons. Recall the problem is to characterize the values of n for which the regular n-gon is constructible using straightedge and compass alone. We previously argued that this is equivalent to constructing the center angle $2\pi/n$ of such a polygon, and that this in turn is equivalent to constructing $\cos(2\pi/n)$. This is the real part of ζ_n, so this is equivalent to constructing the complex number ζ_n, which is finally equivalent to the degree $[\mathbb{Q}(\zeta_n) : \mathbb{Q}]$ being a power of 2. (The direction saying that constructible implies power of 2 uses the tower law, and the direction saying that power of 2 implies constructible was an application of Galois theory we saw a few days ago.)

Now, the degree of $\mathbb{Q}(\zeta_n)$ over $\mathbb{Q}$ is $\varphi(n)$, so the problem comes down to determing the n for which $\varphi(n)$ is a power of 2, which is now a number-theoretic computation. As a first step, note that when $n = p_1^{k_1} \cdots p_m^{k^m}$ with p_i distinct primes, the Chinese Remainder isomorphism above implies that

$$\varphi(n) = \varphi(p_1^{k_1}) \cdots \varphi(p_m^{k_m})$$

since $\varphi(a)$ is the order of $(\mathbb{Z}/a\mathbb{Z})^\times$. Thus for $\varphi(n)$ to be a power of 2, it is equivalent that each $\varphi(p_i^{k_i})$ be a power of 2

The value $\varphi(p^k)$ can be determined by a simple counting argument. List the numbers 1 through p^k as follows:

$$1, 2, \ldots, p, p+1, p+2, \ldots, 2p, \ldots, 3p, \ldots, p^{k-1}p = p^k.$$

The numbers here that are relatively prime to p^k are those not divisible by p. There are $p-1$ such numbers between 1 and p, another $p-1$ between $p+1$ and $2p$, and so on: between np and $(n+1)p$ there are $p-1$ numbers not divisible by p. There are p^{k-1} such "intervals" overall in the list above, so we get $p^{k-1}(p-1)$ numbers in this list not divisible by p. Thus

$$\phi(p^k) = p^{k-1}(p-1) \text{ for } p \text{ prime.}$$

So, $\phi(p^k)$ is a power of 2 precisely when $p = 2$ (no restriction on k here), or when $k = 1$ and p is an odd prime such that $p-1$ is a power of 2. (If $k > 1$ and p is odd, then the p^{k-1} term prevents $\phi(p^k)$ from being a power of 2.) Now that to say $k = 1$ means that we have distinct odd primes in the prime factorization of n.

If p is an odd prime, then $\phi(p) = p-1$ is a power of 2, say 2^ℓ, if and only if $p = 2^\ell + 1$. But in fact, it turns out that being prime places restrictions on ℓ, in that if $2^\ell + 1$ is prime then ℓ must itself actually be a power of 2! Indeed, if ℓ is divisible by some odd number $m > 1$, say $\ell = mb$ for some $b > 1$, then $2^b + 1$ divides $(2^b)^m + 1 = 2^\ell + 1$, so that $2^\ell + 1$ would not be prime. (The divisibility comes here from the fact that the polynomial $x+1$ divides x^m+1 when m is odd, since -1 is a root of $x^m + 1$ in this case.) Hence if $2^\ell + 1$ is prime, ℓ cannot be divisible by an odd number larger than 1, so it must be a power of 2. The conclusion is that $\phi(p)$ is a power of 2 if and only if p is a prime of the form $2^{2^t} + 1$–such primes are called *Fermat primes*.

Thus, in summary, $\phi(n)$ is a power of 2 if and only if n is the product of a power of 2 and distinct Fermat primes, so that the regular n-gon is constructible by straighedge and compass if and only if n is a power of 2 ($2^0 = 1$ allowed) times a product of distinct Fermat primes. Since $2^{2^0} + 1 = 3$, $2^{2^1} + 1 = 5$, and $2^{2^2} + 1 = 17$ are Fermat primes, we thus get for example that the regular 5-gon is construtible, and so is the regular 17-gon, so is the regular 15-gon, and tons of other examples. (Fun fact: the only known Fermat primes are those above, $2^{2^3} + 1 = 257$, and $2^{2^4} + 1 = 65537$. It turns out that $2^{2^t} + 1$ is not prime for $5 \leq t \leq 30$, and the primality of larger such numbers is still an open question.) The regular 7-gon is the first example of a nonconstructible polygon, since 7 is not a Fermat prime.

To actually construct the resulting polygons is a different matter, but can be handled by Galois theory as well. What is required is to produce a sequence of extensions:

$$\mathbb{Q} \subseteq E_1 \subseteq E_2 \subseteq \ldots \subseteq \mathbb{Q}(\zeta_n),$$

each quadratic over the previous one, which one can do by working out the subgroups of the Galois group $(\mathbb{Z}/n\mathbb{Z})^\times$. We will look at an example of this next time.

Abelian extensions. The Galois group of $\mathbb{Q}(\zeta_n)$ over $\mathbb{Q}$ is abelian, and for this reason we call $\mathbb{Q}(\zeta_n)$ an *abelian* extension of $\mathbb{Q}$. (In general, an abelian extension of a field is a Galois extension with abelian Galois group. We will soon consider *cyclic* extensions of fields—the case of a cyclc Galois group—in relation to the solvability of polynomials.) We will not say much about abelian extensions apart from how they show up in questions about polynomials, but let us mention here some interesting facts without proof anyway.

Not only are cyclotomic extensions of $\mathbb{Q}$ abelian, but one of the most basic (basic in the sense of a being cornerstone of the theory, not in terms of having a "basic" proof) results is that in some sense these are the building blocks of *all* abelian extensions of $\mathbb{Q}$. The key fact here is the *Kronecker-Weber Theorem* (highly, highly nontrivial to prove), which states that any abelian extension of $\mathbb{Q}$ is contained in a cyclotomic extension: if $E/\mathbb{Q}$ is abelian, then $E \subseteq \mathbb{Q}(\zeta_n)$ for some n. There is a basic procedure for constructing subfields of cyclotomic fields (again, we will see an example next time), so in some ways this gives a description of all abelian extensions of $\mathbb{Q}$.

Now, fix the prime p and consider the cyclotomic extensions $\mathbb{Q}(\zeta_{p^k})$ for varying k. The composites of these fields is the subfield of $\mathbb{C}$ generated by all p^{th}-power roots of unity. This is in fact an infinite extension of $\mathbb{Q}$, so Galois theory in the sense we have developed does not apply to it since our definition of "Galois extension" is restricted to finite extensions only, but there is a version of "infinite Galois theory" which works in much the same way in this case. The key observation is that the "Galois group" of this extension over $\mathbb{Q}$ arises as a type of "limit" of the Galois groups of the finite extensions $\mathbb{Q}(\zeta_{p^k})/\mathbb{Q}$; these Galois groups are $(\mathbb{Z}/p^k\mathbb{Z})^\times$, and the resulting "limit" of these turns out to be the multiplicative group of the ring $\mathbb{Z}_p$ of p-adic integers we briefly introduced last quarter! (In fact, there was a homework problem last quarter alluding to the idea of viewing $\mathbb{Z}_p$ as a type of "limit" of the $\mathbb{Z}/p^k\mathbb{Z}$.) Thus, the p-adics appear in field theory in describing the Galois groups of certain infinite abelian extensios of $\mathbb{Q}$.

The "maximal" abelian extension of $\mathbb{Q}$ (no finiteness condition) is the composite of all the p^{th}-power roots of unity fields above, and is generated by *all* roots of unity, regardless of which n they come from. The Galois group of this maximal abelian extension turns out to be the product of the $(\mathbb{Z}/p\mathbb{Z})^\times$ as p ranges among *all* primes, and this group turns out to be a crucial object in number theory. Indeed, much of modern number theory is devoted to studying this and related groups that arise as Galois groups of infinite extensions of $\mathbb{Q}$. For reasons beyond the scope of this course, the study of such Galois groups turns out to be intimately connected to the problem from last quarter of determining which rings $\mathbb{Z}$[something] were actually UFDs, where in particular the rings $\mathbb{Z}[\zeta_n]$ obtained by adjoining a root of unity to $\mathbb{Z}$ (the "ring of integers" of the field $\mathbb{Q}(\zeta_n)$) are the important ones needed to understand Fermat's Last Theorem.

Inverse Galois theory. Let us mention one more topic as a tangent, which we will very very briefly mention in the next few days abut which we won't study in any more depth. The subject of *inverse Galois theory* is concerned with determining which groups can actually arise as examples of Galois groups. (So, given the group construct the extension, hence the name "inverse".) It is a fact that any finite *abelian* group is indeed the Galois group (over $\mathbb{Q}$) of some subfield of a cyclotmic field, which gives a sort of converse to the Kronecker-Weber theorem mentioned above. This converse is much easier to prove, and indeed the book gives a proof, which you can check if interested. (The idea is to write your finite abelian group as a product of cyclic groups, then to realize each of these cyclic factors as a quotient of some $(\mathbb{Z}/p\mathbb{Z})^\times$, and then to take a fixed field.)

We will be able to show shortly that, if we are only asking for *some* field extension with Galois group equal to a given group, the answer is that this is always true: for any finite group G, there exists a Galois extension K/F of some field F with $\text{Gal}(K/F) \cong G$. The more difficult problem, and is what inverse Galois theory is primarily concerned with, is whether there exists a Galois extension of $\mathbb{Q}$ specifically whose Galois group is G. The answer is "yes" for finite abelian groups as alluded to above, but it is in general an open problem for arbitrary finite groups. The answer is not likely to be all that important for its own sake, but the point is that developing an answer will likely require new fundamental insights into fields and Galois theory, and that these insights are bound to give important techniques applicable to other problems going forward. Tough stuff!

Lecture 18: General Polynomials

Warm-Up. We determine an explicit sequence of extensions

$$\mathbb{Q} \subseteq E_1 \subseteq E_2 \subseteq \mathbb{Q}(\zeta_{15})$$

where each field is quadratic over the previous one. This is possible since $15 = 3 \cdot 5$ is a product of distinct Fermat primes, so that ζ_{15} is a constructible complex number. (Hence the regular 15-gon is constructible with straightedge and compass.) Since $\varphi(15) = 8$, we can only fit two intermediate fields E_1 and E_2 into our desired chain, so that $[E_1 : \mathbb{Q}] = 2$, $[E_2 : \mathbb{Q}] = 4$, and $[\mathbb{Q}(\zeta_{15}) : \mathbb{Q}] = 8$.

Explicitly, we have $(\mathbb{Z}/15\mathbb{Z})^\times = \{1, 2, 4, 7, 8, 11, 13, 14\}$ under multiplication mod 15. The field E_1 should be the fixed field of a subgroup of order 4 since its index in $(\mathbb{Z}/15\mathbb{Z})^\times$ will be $[E_1 : \mathbb{Q}] = 2$, and then E_2 should be the fixed field of a subgroup of order 2 (in order to have index $[E_2 : \mathbb{Q}] = 4$) contained in this first subgroup (in order to guarantee $E_1 \subseteq E_2$). Note that both of these fields should be simple extensions of $\mathbb{Q}$ by the Primitive Element Theorem. One set of possible choices for these subgroups are

$$\{1, 2, 4, 8\} \supseteq \{1, 4\}.$$

To determine the fixed field of $\{1, 2, 4, 8\}$, all we need is an element fixed under the action of this subgroup but not anything larger. This element should be expressible in terms of ζ_{15} alone since this generates all of $\mathbb{Q}(\zeta_{15})$. We can see that

$$\zeta_{15} + \zeta_{15}^2 + \zeta_{15}^4 + \zeta_{15}^8$$

works. Indeed, acting by $2 \in (\mathbb{Z}/15\mathbb{Z})^\times$, which generates this entire subgroup, gives:

$$\zeta_{15} + \zeta_{15}^2 + \zeta_{15}^4 + \zeta_{15}^8 \mapsto (\zeta_{15}^2) + (\zeta_{15}^2)^2 + (\zeta_{15}^2)^4 + (\zeta_{15}^2)^8 = \zeta_{15}^2 + \zeta_{15}^4 + \zeta_{15}^8 + \zeta_{15},$$

so $\zeta_{15} + \zeta_{15}^2 + \zeta_{15}^4 + \zeta_{15}^8$ is fixed. (Recall that $a \in (\mathbb{Z}/15\mathbb{Z})^\times$ denotes the element of the Galois group defined by $\zeta_{15} \mapsto \zeta_{15}^a$.) In the same way, the generator for the fixed field of $\{1, 4\}$ can be taken to be $\zeta_{15} + \zeta_{15}^4$. Thus $E_1 = \mathbb{Q}(\zeta_{15} + \zeta_{15}^2 + \zeta_{15}^4 + \zeta_{15}^8)$ and $E_2 = \mathbb{Q}(\zeta_{15} + \zeta_{15}^4)$, so

$$\mathbb{Q} \subseteq \mathbb{Q}(\zeta_{15} + \zeta_{15}^2 + \zeta_{15}^4 + \zeta_{15}^8) \subseteq \mathbb{Q}(\zeta_{15} + \zeta_{15}^4) \subseteq \mathbb{Q}(\zeta_{15})$$

is the tower of extensions we want.

Now, to go a bit further, to actually *construct* ζ_{15}, or equivalently the regular 15-gon, requires knowing how to explicitly write ζ_{15} in terms of rationals and square root extractions. For this we need to know the square roots that generate each of the extensions above, starting with $\mathbb{Q}(\zeta_{15} + \zeta_{15}^2 + \zeta_{15}^4 + \zeta_{15}^8)$ as a quadratic extension of $\mathbb{Q}$. This is not so straightforward, and essentially requires determining the minimal polynomial of $\zeta_{15} + \zeta_{15}^2 + \zeta_{15}^4 + \zeta_{15}^8$ over $\mathbb{Q}$, which we know should be a quadratic polynomial. We won't go through the details here, but it turns out that

$$\mathbb{Q}(\zeta_{15} + \zeta_{15}^2 + \zeta_{15}^4 + \zeta_{15}^8) = \mathbb{Q}(\sqrt{5}).$$

Next, $\mathbb{Q}(\zeta_{15} + \zeta_{15}^4)$ should be quadratic over $\mathbb{Q}(\sqrt{5})$, and some more tough computations will show that

$$\mathbb{Q}(\zeta_{15} + \zeta_{15}^4) = \mathbb{Q}(\sqrt{5}, \sqrt{30 - 6\sqrt{5}}).$$

Finally, we can work out that in fact $\mathbb{Q}(\zeta_{15}) = \mathbb{Q}(\sqrt{5}, \sqrt{30 - 6\sqrt{5}}, \sqrt{7 + \sqrt{5} - \sqrt{30 + 6\sqrt{5}}})$, which is quadratic over $\mathbb{Q}(\sqrt{5}, \sqrt{30 - 6\sqrt{5}})$. From this one can work out how to explicitly express $\zeta_{15} = \cos(2\pi/15) + i\sin(2\pi/15)$ as a constructible number, where the answer ends up being:

$$\cos(\tfrac{2\pi}{15}) = \tfrac{1}{8}\left(\sqrt{30 - 6\sqrt{5}}\right) + \sqrt{5} + 1 \text{ and } \sin(\tfrac{2\pi}{15}) = \tfrac{1}{4}\sqrt{7 + \sqrt{5} - \sqrt{30 + 6\sqrt{5}}}.$$

Note that constructing $\cos(\frac{2\pi}{15})$ only requires going up to the third field in our tower, whereas constructing $\sin(\frac{2\pi}{15})$, and hence ζ_{15}, requires going all the way to $\mathbb{Q}(\zeta_{15})$. To actually construct the regular 15-gon, you would first construct $\sqrt{5}$, then $30 - 6\sqrt{5}$, then $\sqrt{30 - 6\sqrt{5}}$, and so on.

Hints of solvability. The goal for the remainder of the course is to generalize the discussion of constructiblity in terms of square root extractions, to the setting where we will now allow *arbitrary* root extractions. We will give precise definitions later, but let us see now the basic idea. It is these arbitrary root extractions which are the ones to consider when looking for an analog of the quadratic formula for the roots of polynomials of higher degree. (For example, the "cubic equation" requires both square root and cube root extractions).

Analogously to what whappened with constructible numbers, to express a number α using rationals, $+, -, \cdot, \div$, and arbitrary root extractions requires coming up with a tower of extensions of the form

$$\mathbb{Q} \subseteq K_1 \subseteq K_2 \subseteq \ldots \subseteq \mathbb{Q}(\alpha)$$

where now K_i is generated over K_{i-1} by some n_i-th root: $K_i = K_{i-1}(\sqrt[n_i]{D_i})$. (No restriction on the the types of roots $\sqrt[n_i]{}$ being used here.) We want each extension to be Galois over the previous one, and then this sequence corresponds to a chain of subgroups:

$$G \trianglerighteq G_1 \trianglerighteq G_2 \trianglerighteq \ldots \trianglerighteq 1,$$

each normal in the previous. The key property which will make this all work out is that each extension K_i/K_{i-1} be a *cyclic* extension, meaning that the Galois group is cyclic. These Galois groups are quotients of the G_i in the chain above, so we want each quotient G_i/G_{i-1} to be cyclic. Cyclic groups are abelian, and having a chain like this where each quotient is abelian is precisely by the definition we gave in the fall of a *solvable* group. Thus, we see that solvability of groups is directly related to "solvability" of polynomials. (In the finite case, we will see that having a chain with abelian quotients is equivalent to having a chain with cyclic quotients, due to the fact that finite abelian groups are products of cyclic groups.)

Galois groups of polynomials. So, that is where we are headed. But before we can get there, we need a better understanding of Galois groups of polynomials. Recall that if $f(x)$ is a separable polynomial over a field F, its *Galois group* is by definition the Galois group of its splitting field over F. This group permutes the roots of $f(x)$, so if $f(x)$ has degree n then the Galois group can be realized as a subgroup of S_n. We worked out previously that for $x^3 - 2$ over $\mathbb{Q}$, the Galois group is all of S_3—the largest it can be for a polynomial of degree 3—while for $x^4 - 2$ over $\mathbb{Q}$, the Galois group is D_8, which is strictly smaller than S_4. Why do we get the full symmetric group in one case, but not the other?

Our first goal is to understand the type of scenario where the Galois group is indeed the largest it can be, i.e. S_n. Then we can think about when the Galois group will be the next largest thing it can be, which is the alernating group A_n since A_n is the largest proper subgroup of S_n. (This was a result from the fall: A_n is the only subgroup of index 2 in S_n.) Ideally, we want a way to determine these things without having to determine the Galois group in full, since this can be challenging given that the roots of an arbitrary polynomial are not straightforward to write down. We want theorems that dictate, without too much work, what Galois group we will get, or at least which narrow down the choices.

General polynomials. Let F be a field, and let $x_1, \ldots, x_n$ be a set of indeterminates. (So, these are independent variables, but we will think of them as elements in a field rather than variables. They are literally just n symbols.) We then define the *general polynomial* of degree n to be the one that has $x_1, \ldots, x_n$ as roots:

$$f(x) = (x - x_1)(x - x_2) \cdots (x - x_n).$$

(Here x *is* a variable.) This can be viewed as a polynomial over $F(x_1, \ldots, x_n)$, the field of rational functions over F in $x_1, \ldots, x_n$. The point is that this represents the most "generic" polynomial one can write down, where there are absolutely no predetermined relations among the roots.

The coefficients (without signs) of $f(x)$ are called the *elementary symmetric polynomials* in the indeterminants $x_1, \ldots, x_n$. These are "symmetric" in the sense that they are invariant under any permutation of the x_i. Concretely, we have:

$$f(x) = x^n - s_1 x^{n-1} + s_2 x^{n-2} - \cdots + (-1)^n s_n$$

where the elementary symmetric polynomials $s_1, \ldots, s_n$ are

$$\begin{aligned} s_1 &= x_1 + \cdots + x_n = \sum_i x_i \\ s_2 &= x_1x_2 + x_1x_3 + \cdots + x_1x_n + x_2x_3 + \cdots + x_{n-1}x_n = \sum_{i<j} x_ix_j \\ &\vdots \\ s_n &= x_1 \cdots x_n. \end{aligned}$$

The upshot is that $f(x)$ is actually a polynomial in $F(s_1, \ldots, s_n)[x]$, where $F(s_1, \ldots, s_n)$ is the field of rational functions in the $s_1, \ldots, s_n$, and $F(x_1, \ldots, x_n)$ is then the splitting field of $f(x)$ over $F(s_1, \ldots, s_n)$. (Note $f(x)$ is separable simply because the roots $x_1, \ldots, x_n$ are distinct by definition. Also, $F(x_1, \ldots, x_n)$ is finite over $F(s_1, \ldots, s_n)$ since, being the splitting field of a polynomial of degree n, its degree is bounded by $n!$.) Thus $F(x_1, \ldots, x_n)$ is a Galois extension of $F(s_1, \ldots, s_n)$, and we would like to understand its Galois group, which we claim is precisely S_n.

Invariants of the permutation action. By permuting $x_1, \ldots, x_n$, we can view S_n as a subgroup of the automorphism group of $F(x_1, \ldots, x_n)$. The fixed field of this subgroup is a subfield of $F(x_1, \ldots, x_n)$ over which this latter field has degree $n!$. (This is the claim that if H is a finite subgroup of $\mathrm{Aut}(K)$, then $[K : K^H] = |H|$.) Now, since the elementary symmetric polynomials $s_1, \ldots, s_n$ are invariant under this action of S_n, we see that the entire field $F(s_1, \ldots, s_n)$ is contained in the fixed field of S^n:

$$F(s_1, \ldots, s_n) \subseteq \text{fixed field} \subseteq F(x_1, \ldots, x_n).$$

By the tower law, we thus have that $[F(x_1, \ldots, x_n) : F(s_1, \ldots, s_n)]$ is divisible by $[F(x_1, \ldots, x_n) :$ fixed field$] = n!$. But, we also have from before that $[F(x_1, \ldots, x_n) : F(s_1, \ldots, s_n)] \leq n!$ since $F(x_1, \ldots, x_n)$ is the splitting field of the degree n polynomial $f(x)$, so we must thus have that $[F(x_1, \ldots, x_n) : F(s_1, \ldots, s_n)]$ is exactly $n!$. Hence $F(s_1, \ldots, s_n)$ is indeed the entire fixed field of $S_n \leq \mathrm{Aut}(F(x_1, \ldots, x_n))$:

$$F(x_1, \ldots, x_n)^{S_n} = F(s_1, \ldots, s_n).$$

Therefore, we have $\mathrm{Gal}(F(x_1, \ldots, x_n)/F(s_1, \ldots, s_n)) \cong S_n$, meaning that S_n is also the Galois group of the general polynomial of degree n, as claimed.

Let us highlight the result of the work above: any rational function in the $x_1, \ldots, x_n$ that is invariant under all permutations must be expressible in terms of the $s_1, \ldots, s_n$ alone. In particular, any polynomial that is invariant under all permutations—called a *symmetric* polynomial—is a linear combination of $s_1, \ldots, s_n$, and indeed these form a *basis* for the vector space of symmetric polynomials of degree n. (This is why these are the "elementary" symmetric polynomials. This fact is usually known as the *fundamental theorem of symmetric polynomials.*) For example, the polynomial $(x_1-x_2)^2$ is invariant under the action of S_2, so it should be possible to write it solely in terms of $s_1 = x_1+x_2$ and $s_2 = x_1x_2$, which we can work out explicitly:

$$(x_1-x_2)^2 = x_1^2 - 2x_1x_2 + x_2^2 = (x_1+x_2)^2 - 4x_1x_2 = s_1^2 - 4s_2.$$

(There was actualy a homework problem we had in the fall hinting at all this, phrased in terms of "stabilizers", which dealt directly with finding fixed elements under this permutation action.)

What makes a polynomial "generic"? The moral is that a "generic" separable polynomial of degree n over a field F should have Galois group S_n, where "generic" is not a term we will define precisely but which we should think of as saying that there are no nontrivial "algebraic relations" among the roots (or coefficients!), as is the case with $f(x) = (x-x_1)\dots(x-x_n)$. It is this type of polynomial which will give the largest possible Galois group, and for other "non-generic" polynomials we should get something smaller.

For example, consider again x^3-2 and x^4-2 over $\mathbb{Q}$. The roots of these, respectively, are

$$\sqrt[3]{2},\ \zeta_3\sqrt[3]{2},\ \zeta_3^2\sqrt[3]{2} \quad \text{and} \quad \sqrt[4]{2},\ i\sqrt[4]{2},\ -\sqrt[4]{2},\ -i\sqrt[4]{2}.$$

In what sense is the first set of roots "generic", but not the second? Again, we will not give a precise meaning to this, but here is an observation. For the roots of x^3-2, note that the quotient of any two is a primitive third root of unity:

$$\frac{\zeta_3\sqrt[3]{2}}{\sqrt[3]{2}} = \zeta_3 = \frac{\zeta_3^2\sqrt[3]{2}}{\zeta_3\sqrt[3]{2}} \text{ and } \frac{\zeta_3^2\sqrt[3]{2}}{\sqrt[3]{2}} = \zeta_3^2.$$

The idea is that all of these satisfy the same "algebraic relation", namely $x^3-1=0$, and thus, there is no way to "algebraically distinguish" between any pairs of roots. That is, all roots "behave" in the same way. (Again, this is the case with the general polynomial $f(x)$ as well.)

But for x^4-2, we have for instance

$$\frac{i\sqrt[4]{2}}{\sqrt[4]{2}} = i \quad \text{and} \quad \frac{-\sqrt[4]{2}}{\sqrt[4]{2}} = -1.$$

The first is a primitive fourth root of unity, but the second is a primitive *second* root of unity; the first satisfies $x^4-1=0$ and the second $x^2-1=0$. Thus, pairs of roots can indeed be "algebraically distinguished" from one another, via these specific "algebraic relations". It is essentially this fact that places restrictions on the types of permutations allowed in the Galois group, resulting in a smaller group than all of S_4. We thus seek to find ways of detecting these types of behaviors more efficiently, since, again, the explicit roots are not always easy to work with or even find.

Lecture 19: Discriminants and Cubics

Warm-Up 1. We write the polynomial $x_1^2+x_2^2+x_3^2$ in terms of elementary symmetric polynomials. This is possible since $x_1^2+x_2^2+x_3^2$ is invariant under the action of S_n which permutes the x_i, so it belongs to the fixed field of this action. By the work from last time, this fixed field is generated by the elementary symmetric polynomials s_1, s_2, s_3, which in this case are given by:

$$s_1 = x_1+x_2+x_3 \quad s_2 = x_1x_2+x_1x_3+x_2x_3 \quad s_3 = x_1x_2x_3.$$

We have:

$$s_1^2 = (x_1+x_2+x_3)^2 = x_1^2+x_2^2+x_3^2+2x_1x_2+2x_1x_3+2x_3x_3,$$

so

$$x_1^2 + x_2^2 + x_3^2 = s_1^2 - 2(x_1x_2 + x_1x_3 + x_2x_3) = s_1^2 - 2s_2$$

is our desired expression.

Warm-Up 2. We show that any finite group arises as a Galois group. Suppose G is finite of order n. By Cayley's Theorem in group theory, G is (isomorphic to) a subgroup of S_n. (Recall that this comes from the action of G on itself by left multiplication, which gives a permutation of its n elements.) Since S_n is the Galois group of the Galois extension

$$F(s_1, \ldots, s_n) \subseteq F(x_1, \ldots, x_n),$$

the Fundamental Theorem of Galois Theory says that the fixed field of $G \leq S_n$ is an intermediate field E with $\operatorname{Gal}(F(x_1, \ldots, x_n)/E) \cong G$, as desired. (The tougher, and open, question is whether G can be obtained as the Galois group of an extension of $\mathbb{Q}$.)

Discriminants. The general polynomial $f(x) = (x - x_1) \cdots (x - x_n)$ gives the "generic" setting in which the Galois group is the largest it can be, namely S_n, so we now seek to understand when the Galois group sits inside of the *next* largest it can be, which is A_n. (Recall that A_n is the only subgroup of S_n of index 2, so it is indeed the proper subgroup of S_n of largest size. Note that this does not mean all other subgroups are contained in A_n—$D_8 \leq S_4$ is not contained in A_4 for example—only that all other subgroups have a strictly smaller size.) We will see that determining whether or not the Galois group sits inside of A_n is quite straightforward.

Define the *discriminant* D of the general polynomial $f(x)$ to be the product of the squares of differences of its roots:

$$D = \prod_{i<j} (x_i - x_j)^2 \in F(x_1, \ldots, x_n).$$

Note that reordering the roots does not alter the discriminant, since this will only change the sign of some $x_i - x_j$, which does not matter after squaring. In fact, this also means that D is fixed under all permutations of the roots x_i, so that D lies in the fixed field for the full Galois group S_n of $F(x_1, \ldots, x_n)$ over $F(s_1, \ldots, s_n)$, which is $F(s_1, \ldots, s_n)$. In particular, this shows that the discriminant D, even though defined using the roots of $f(x)$, is expressible solely in terms of the coefficients $\pm s_i$ of $f(x)$.

Now, the square root of the discriminant is

$$\sqrt{D} = \prod_{i<j} (x_i - x_j).$$

The action of S_n now *can* affect the sign since we are no longer squaring; for example, the transposition (12) which exchanges x_1 and x_2 will turn $x_1 - x_2$ into $x_2 - x_1 = -(x_1 - x_2)$, and this leads to $\sqrt{D}$ being replaced by $-\sqrt{D}$. (Here we need to assume $\operatorname{char} F \neq 2$ since otherwise $\sqrt{D}$ and $-\sqrt{D}$ mean the same thing.) By keeping track of the number of such transpositions, we thus see that an element of S_n will fix $\sqrt{D}$ if and only if that element actually belongs to A_n. (In fact, this was how the book originally defined "even permutation" back in the group theory portion, only without using the term "discriminant".)

We can now readily apply all this to a concrete polynomial over any field. Suppose $f(x) \in F[x]$ (this is no longer necessarily a 'general polynomial") has degree n and let K be its splitting field. Denote the roots of $f(x)$ by $\alpha_1, \ldots, \alpha_n$. Then replacing x_i by α_i gives the *discriminant* of $f(x)$:

$$D = \prod_{i<j} (\alpha_i - \alpha_j)^2.$$

This discriminant is zero as an element of K if and only if at least one root is repeated, so in the separable case, which we will assume is the case going forward, $D \neq 0$. Then D is invariant under $\mathrm{Gal}(K/F) \leq S_n$, and so belongs to the fixed field, which is F; in other words, D belongs to the field $F(s_1, \ldots, s_n)$ generated by the coefficients of $f(x)$, which is just F itself since all coefficients of $f(x)$ are in F. The discriminant is thus expressible solely in terms of the coefficients of $f(x)$, so we have some hope of being able to compute it even without knowing the roots α_i explicitly.

The square root $\sqrt{D}$ of the discriminant is then an element of K, and by applying the conclusion we drew in the general polynomial case, we have that $\sqrt{D} \in F$ if and only if $\sqrt{D}$ is fixed by $\mathrm{Gal}(K/F)$ if and only if $\mathrm{Gal}(K/F) \leq A_n$. Thus, we can tell whether or not the Galois group of $f(x)$ is a subgroup of A_n simply by seeing if its discriminant—which we hope to be able to compute using only the coefficients—is a square in F.

Quadratics. Let us see how the above plays out in the simple case of quadratic polynomials. Here we already know the answer based on the quadratic formula, so the point is really just to phrase what we know in terms of the discriminant. One first point to make is that for a polynomial over a *finite* field, the answer is easy: we saw before that the Galois group of any finite field over another is always cyclic and generated by Frobenius, so this characterizes all possible Galois groups of such polynomials. Thus, we really focus only on the characteristic zero case going forward.

The possible Galois groups of $f(x) = x^2 + ax + b \in F[x]$ are $S_2 \cong \mathbb{Z}/2\mathbb{Z}$ and A_2, which is the trivial group. The roots of $f(x)$ are

$$\frac{-a \pm \sqrt{a^2 - 4b}}{2}.$$

(Note that writing these down requires $\operatorname{char} F \neq 2$.) The discriminant of $f(x)$ is thus:

$$D = \left(\frac{-a + \sqrt{a^2 - 4b}}{2} - \frac{-a - \sqrt{a^2 - 4b}}{2} \right)^2 = a^2 - 4b.$$

If $a^2 - 4b$ is a square in F, then $\sqrt{a^2 - 4b} \in F$ and the two roots of $f(x)$ are in F, whereas if $a^2 - 4b$ is not a square in F, then the two roots are in the quadratic extension $F(\sqrt{D})$. In the former case, F is already the splitting field of $f(x)$ and so the Galois group is trivial, and in the latter case the Galois group is $S_2 \cong \mathbb{Z}/2\mathbb{Z}$. Thus we do see that the Galois group is a subgroup of A_2 if and only if D is a square in F. (Note also that $D = 0$ if and only if there is only one repeated root, so this is the inseparable case.)

Cubics. Next we consider cubic polynomials. Our aim is to show how an explicit formula for the discriminant in terms of the coefficients can be found, but the actual derivation should not be the main takeaway. We will get an explicit formula, and there is also an explicit formula in the degree 4 we will consider next time, but there is really no need to memorize these formulas since it is easy to look them up when needed. Ultimately, we care more about using the discriminant and related objects to classify the Galois group, but nonetheless it is nice to see in the cubic how formulas for discriminants actually come about.

Consider the cubic $x^3 + ax^2 + bx + c \in F[x]$. A first observation is that by making a change of variables $x = y - \frac{a}{3}$ (only possible in characteristic $\neq 3$), this cubic can be turned into one which has no quadratic term:

$$f(x) = x^3 + ax^2 + bx + c \rightsquigarrow g(y) = y^3 + py + q$$

where $p = \frac{1}{3}(3b - a^2)$ and $q = \frac{1}{27}(2a^3 - 9ab + 27c)$. (We will leave it to you to verify that this does work. A lot of what is needed in these types of computations is coming up with clever "tricks" that simplify the work, but how to actually come up with these tricks is not so enlightening.) A key thing to note is that $f(x)$ and $g(y)$ have the *same* the discriminant: the roots of $g(y)$ are the roots of $f(x)$ translated by $\frac{a}{3}$, so the *difference* between the roots of $g(y)$ is the same as the difference between the corresponding roots of $f(x)$ because we add and then subtract $\frac{a}{3}$. Thus we can focus only on $g(y)$ when computing the discriminant of $f(x)$.

Let $\alpha_1, \alpha_2, \alpha_3$ denote the roots of $g(y)$, so that

$$g(y) = (y - \alpha_1)(y - \alpha_2)(y - \alpha_3).$$

A simple computation using the product rules gives the following values of the derivative $g'(y)$:

$$\begin{aligned} g'(\alpha_1) &= (\alpha_1 - \alpha_2)(\alpha_1 - \alpha_3) \\ g'(\alpha_2) &= (\alpha_2 - \alpha_1)(\alpha_2 - \alpha_3) \\ g'(\alpha_3) &= (\alpha_3 - \alpha_1)(\alpha_3 - \alpha_2), \end{aligned}$$

and from this we can see that the discriminant of $g(y)$ (and hence of $f(x)$) is

$$D = (\alpha_1 - \alpha_2)^2(\alpha_1 - \alpha_3)^2(\alpha_2 - \alpha_3)^2 = -g'(\alpha_1)g'(\alpha_2)g'(\alpha_3).$$

(Check the signs!) But the derivative of $g(y) = y^3 + py + q$ is $g'(y) = 3y^2 + p$ (having such a simple form for the derivative is the reason why the change of variables is useful), so we get:

$$\begin{aligned} D &= -(3\alpha_1^2 + p)(3\alpha_2^2 + p)(3\alpha_3^2 + p) \\ &= -27\alpha_1^2\alpha_2^2\alpha_3^2 - 9p(\alpha_1^2\alpha_2^2 + \alpha_1^2\alpha_3^2 + \alpha_2^2\alpha_3^2) - 3p^2(\alpha_1^2 + \alpha_2^2 + \alpha_3^2) - p^3. \end{aligned}$$

The point is that each of the expressions here in terms of the roots are all invariant under the action of S_3, and hence should be expressible in terms of the "elementary symmetric polynomials" s_1, s_2, s_3 in the coefficients of $g(y)$! For instance, $\alpha_1^2\alpha_2^2\alpha_3^2$ is precisely s_3^2, and $\alpha_1^2 + \alpha_2^2 + \alpha_3^2$ is $s_1^2 - 2s_2$ according to the first Warm-Up. Similarly, the expression for $\alpha_1^2\alpha_2^2 + \alpha_1^2\alpha_3^2 + \alpha_2^2\alpha_3^2)$ in terms of s_1, s_2, s_3 can be found, and then using the fact that the coefficients in our case

$$g(y) = y^3 + py + q = y^3 - s_1y^2 + s_2y - s_3$$

are $s_1 = 0$, $s_2 = p$, and $s_3 = -q$, we get that the discriminant is explicitly

$$D = -27(-q)^2 - 9p(p^2) - 3p^2(-2p) - p^3 = -4p^3 - 27q^2.$$

After using the expressions for p and q in terms of a, b, c a formula for the discriminant in terms of the original coefficients of $f(x)$ can be found, but it is fairly common to leave the discriminant in terms of the shifted coefficients p and q since this expression is much simpler.

Galois groups of cubics. The upshot of the work above is that the discriminant of a cubic can be computed solely using the coefficients. This is good since, as we will see, the concrete roots of a cubic in general are quite messy to write down. But, with the discriminant at hand, we can now classify all possible Galois groups in the cubic case. Any Galois group of a cubic in $F[x]$ (F of characteristic zero) is a subgroup of S_3, so the possibilities are: the trivial group, $\mathbb{Z}/2\mathbb{Z}$ generated by a 2-cycle, $A_3 \cong \mathbb{Z}/3\mathbb{Z}$ generated by a 3-cycle, and all of S_3. Note that if a cubic factors, it either factors into three linear terms or a linear term and an irreducible quadratic term. We have:

- if the cubic has three roots in F, then the splitting field is just F and hence the Galois group is trivial;

- if the cubic has only one root in F, then it factors into a linear term and irreducible quadratic term, and the Galois group is the Galois group of the quadratic factor, which is $\mathbb{Z}/2\mathbb{Z}$ by what we said about quadratics before;
- if the cubic has no roots in F and the discriminant is a square in F, then the Galois group is $A_3 \cong \mathbb{Z}/3\mathbb{Z}$; and
- if the cubic has no roots in F and the discriminant is a non-square in F, then the Galois group is S_3.

(The first two cover the reducible case, and the latter two the irreducible case.) Thus, Galois groups of cubics are easily determined.

Here are two examples. We have already worked out before that the Galois group of $x^3 - 2$ over $\mathbb{Q}$ is S_3. Indeed, now we can compute that the discriminant of $x^3 - 2$ is $-27(-2)^2 = -3^3 2^2$, which is not a square in $\mathbb{Q}$. No roots of $x^3 - 2$ are in $\mathbb{Q}$, so the Galois group matches with what we derived before. For $x^3 - 3x + 1$ over $\mathbb{Q}$, however, the discrminant works out to be 81, which is a square in $\mathbb{Q}$. Thus the Galois group of $x^3 - 2$ is $A_3 \cong \mathbb{Z}/3\mathbb{Z}$. Indeed, it turns out that adjoining a single root θ to $\mathbb{Q}$ already gives the splitting field and that the other two roots are expressible in terms of θ. (We will talk about the "cubic formula" for expressing such roots in a few days.)

Lecture 20: Quartics and Resolvents

Warm-Up. We describe all splitting fields of cubic polynomials over a field F (characteristic zero), following the classification of Galois groups we derived last time. First, if the cubic has three roots in F, then the splitting field is F, as we observed last time. (This is the trivial Galois group case.) Second, if the cubic has only one root in F, then it factors into a linear term and an irreducible quadratic, so the splitting field is the splitting field of this quadratic part. Certainly adjoining a root of this quadratic gives the correct field, but we note more importantly that this field is also obtained by adjoining $\sqrt{D}$ where D is the discriminant of the cubic. Indeed, if K denotes the splitting field of the cubic in this case, then $\sqrt{D} \in K$ since $\sqrt{D}$ is expressible in terms of the roots generating K, so that we have

$$F \subseteq F(\sqrt{D}) \subseteq K.$$

But here both $F(\sqrt{D})$ and K have degree 2 over F (note that $\sqrt{D} \notin F$ since the Galois group $\mathbb{Z}/2\mathbb{Z}$ is not a subgroup of $A_3 \cong \mathbb{Z}/3\mathbb{Z}$), so $K = F(\sqrt{D})$.

Now, if the cubic is irreducible and $\sqrt{D} \in F$, the case of Galois group A_3, then adjoining one root of the cubic to F produces the Galois group. Indeed, if θ is a root, then $F(\theta)$ has degree 3 over F, which matches the degree of the splitting field K since $\text{Gal}(K/F) \cong A_3$ has order 3. Here then the other two roots can be expressed in terms of the chosen root θ alone. Finally, in the case of an irreducible cubic with $\sqrt{D} \notin F$ and Galois group S_3, adjoining one root θ gives a degree 3 extension sitting inside the splitting field:

$$F \subseteq F(\theta) \subseteq K.$$

The splitting field has degree 6 over F since $|S_3| = 6$, so K has degree 2 over $F(\theta)$. Since $\sqrt{D} \notin F(\theta)$ (otherwise $\sqrt{D}$ generates a degree 2 extension sitting inside the degree 3 extension $F(\theta)$, which is not possible), we have that $F(\theta, \sqrt{D})$ has degree 6 over F, so we must have $F(\theta, \sqrt{D}) = K$. Another way of saying this is that both $F(\theta)$ and $F(\sqrt{D})$ are subfields of K, so their composite is as well, and this composite has degree $2 \cdot 3 = 6$, so it must be the splitting field.

What about the roots? Describing the splitting fields in the irreducible cubic case above required adjoining a root to the base field. So, in order for this to be an *explicit* description, one might argue that we need an explicit root. We will see next time that the explicit roots of a cubic are given by *Cardano's formulas.* Just to give a sense for what they look like now, consider the example of $x^3 - 3x + 1$ over $\mathbb{Q}$ from last time, with Galois group A_3. Set A and B to be the following:

$$A = \sqrt[3]{-\frac{27}{2} + \frac{3}{2}\sqrt{-243}} \quad \text{and} \quad B = \sqrt[3]{-\frac{27}{2} - \frac{3}{2}\sqrt{-243}}.$$

Now, there is some ambiguity here in that cube roots are not unique, so there are three possible choices for each of A and B. It turns out that there is a choice which also satisfies $AB = 9$, and these are the ones we pick. Then the fact is that the roots of $x^3 - 3x + 1$ are:

$$\frac{A+B}{3}, \ \frac{\zeta_3 A + \zeta_3^2 B}{3}, \ \text{and} \ \frac{\zeta_3^2 A + \zeta_3 B}{3}$$

where ζ_3 is a primitive third root of unity. We will discuss the derivation of these a bit next time, highlighting some ideas that will be important going foward.

Quartic polynomials. Now we study the Galois groups of quartics $x^4 + ax^3 + bx^2 + cx + d$. As a first step, as in the cubic case we can make a change of variables to simplify the polynomial a bit: setting $x = y - \frac{a}{4}$ gives

$$x^4 + ax^3 + bx^2 + cx + d \rightsquigarrow g(y) := y^4 + py^2 + qy + r,$$

with some explicit expression for p, q, r in terms of a, b, c, d that we don't really care about here. As in the cubic case, the discriminants of these two quartics are the same. If $g(y)$ is reducible, then we essentially reduce to previous cases: if $g(y)$ factors into a linear and an irreducible cubic, the Galois group of $g(y)$ is the Galois group of the irreducible cubic, so either A_3 or S_3 depending on whether the discriminant of the cubic is a square; if $g(y)$ factors into two irreducible quadratics, then the splitting field is $F(\sqrt{D_1}, \sqrt{D_2})$ where D_1, D_2 are the discriminant of these quadratics. If this extension is biquadratic, so $D_1 D_2$ is not a square in F, then the Galois group is $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$, whereas if $D_1 D_2$ is a square, then $F(\sqrt{D_1}, \sqrt{D_2}) = F(\sqrt{D_1})$ and the Galois group is $\mathbb{Z}/2\mathbb{Z}$.

Transitivity. We are thus left with the case where the quartic $g(y)$ is irreducible. A key observation in this case is that the Galois group must act *transitively* on the roots: given any two roots, there is a Galois group element that sends one to the other. Indeed, if α_1 and α_2 are two roots, then via

$$F(\alpha_1) \cong F[y]/(g(y)) \cong F(\alpha_2)$$

we have a map that sends α_1 to α_2, which can then be extended to an element of the Galois group.

This transitivity places restrictions on which subgroup of S_4 the Galois group can be. For example, the Galois group cannot be $\langle (12) \rangle$ since there is not element here that will send the third root to the fourth root (once we label the roots $1, 2, 3, 4$), and the Galois group cannot be of order 3 (so generated by a 3-cycle) for the same reason. Moreover this also rules out an isomorphic copy of S_3 in S_4 as the Galois group: if we consider the version of S_3 that fixes a root a and permutes the others b, c, d, then nothing sends a to b, so this is not transitive. We can work out that the only transitive subgroups of S_4 are the following:

$$S_4, \ A_4, \ \text{a copy of } D_8, \ \text{a copy of } \mathbb{Z}/4\mathbb{Z}, \ \text{and the normal subgroup } \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} \text{ of } A_4.$$

(The copy of D_8 we get depends on how we label the vertices of a square: for the usual $1, 2, 3, 4$ counterclockwise ordering, we get the usual version of D_8 where the smallest rotation is (1234), but if we use the labeling $1, 3, 4, 2$ for example, then we get a copy where the smallest rotation is (1342). The copy of $\mathbb{Z}/4\mathbb{Z}$ we get depends on the 4-cycle we use to generate it, and by $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ we mean $\{(1), (12)(34), (13)(24), (14)(23)\} \trianglelefteq A_4$.) These are thus the only possible Galois groups we can get in the irreducible quartic case.

Resolvents. We are left having to distinguish between the possibilities above. The basic idea we will use, and which will also play a role in the derivation of Cardano's formulas and in our eventual discussion regarding solvability, is to consider actions of the sought-after Galois group on *other* sets apart from simply the roots of the quadratic at hand. After all, we saw examples in the fall where multiple actions of a groups were used to derive information about its structure, so it only makes sense that the same should be true for Galois groups. The strategy is as follows: first, use field theory to produce sets on which the Galois group should act, and then use group theory to determine the Galois group.

To this end, denote the roots of our irreducible quartic $g(y)$ by $\alpha_1, \alpha_2, \alpha_3, \alpha_4$. Define the elements $\theta_1, \theta_2, \theta_3$ (in the splitting field K of the quartic) by

$$\begin{aligned}\theta_1 &= (\alpha_1 + \alpha_2)(\alpha_3 + \alpha_4)\\ \theta_2 &= (\alpha_1 + \alpha_3)(\alpha_2 + \alpha_4)\\ \theta_3 &= (\alpha_1 + \alpha_4)(\alpha_2 + \alpha_3).\end{aligned}$$

Since the Galois group permutes the α_i, we see that it *also* permutes the θ_i! Thus, we can view the Galois group as now acting on these three elements instead. We define the *resolvent cubic* of $g(y)$ to be the cubic which has $\theta_1, \theta_2, \theta_3$ as the roots:

$$h(y) = (y - \theta_1)(y - \theta_2)(y - \theta_3).$$

The upshot is that, by using the Galois group of this cubic, which we can fully classify, we can extract information about the Galois group of our quartic. One thing to note is that the cubic $h(y)$ and quartic $g(y)$ have the same discriminant: the differences $\theta_i - \theta_j$ in the discriminant of $h(y)$ give the differences (up to sign) in the discriminant of $g(y)$. For example, we have

$$\theta_1 - \theta_2 = (\alpha_1 + \alpha_2)(\alpha_3 + \alpha_4) - (\alpha_1 + \alpha_3)(\alpha_2 + \alpha_4) = (\alpha_1 - \alpha_4)(\alpha_3 - \alpha_2).$$

After squaring, all the sign differences go away and we get the same discriminants.

A second fact is that we can determine the coefficients of the resolvent cubic explicitly in terms of the coefficients of the quartic: the coefficients of $h(y)$ are the elementary symmetric polynomials in the θ_i, which are in the base field F since they are fixed under the action of the Galois group $\mathrm{Gal}(K/F)$, and $F = F(s_1, s_2, s_3, s_4)$ where the s_i elementary symmetric polynomials in the α_i, which are the coefficients of $g(y)$. Going through the necessary computations in terms of symmetric polynomials results in

$$h(y) = y^3 + 2py^2 + (p^2 - 4r)y - q^2.$$

We omit the details since they are not so enlightening, and in the end the explicit resolvent cubic is something we can simply look up when needed.

The irreducible resolvent case. We finish by determining the Galois group $\mathrm{Gal}(K/F)$ in the cases where the resolvent cubic is irreducible over F. (The reducible cases will be left for next time.) If the resolvent cubic is irreducible and $\sqrt{D} \notin F$, where D here denotes either the discriminant of our quartic or of the resolvent cubic (they are the same!), then the Galois group of the resolvent cubic is S_3. The splitting field of this cubic is $F(\theta_1, \theta_2, \theta_3)$, so we have the tower

$$F \subseteq F(\theta_1, \theta_2, \theta_3) \subseteq K.$$

Since $F(\theta_1, \theta_2, \theta_3)$ has degree $|S_3| = 6$ over F, 6 must divide $\text{Gal}(K/F)$. Looking at the candidates for $\text{Gal}(K/F)$ (i.e. the transitive subgroups of S_4), the only ones which have orders divisible by 6 are A_4 and S_4. But since $\sqrt{D} \notin F$, the Galois group is not a subgroup of A_4, so we must have $\text{Gal}(K/F) \cong S_4$ in this case.

If the resolvent is irreducible and $\sqrt{D}$ is an element of F, the Galois group of the resolvent cubic is A_3. Thus in this case $F(\theta_1, \theta_2, \theta_3)$ has degree $|A_3| = 3$, so 3 divides the order of $\text{Gal}(K/F)$. But now $\text{Gal}(K/F)$ is a subgroup of A_4 since $\sqrt{D} \in F$, and the only transitive subgroup of A_4 with order divisible by 3 is A_4 itself. Hence $\text{Gal}(K/F) \cong A_4$ in this case. Thus, in the case of an irreducible resolvent, we can fully determine the Galois group using the discriminant alone.

Examples. Here are two quick examples. The resolvent cubic of $x^4 - x - 1$ over $\mathbb{Q}$ is

$$x^3 + 4x - 1.$$

This cubic is irreducible over $\mathbb{Q}$ (rational root test), and its discriminant (also the discriminant of $x^4 - x - 1$) is -283. Since this is not a square in $\mathbb{Q}$, the Galois group of $x^4 - x - 1$ is S_4. The resolvent cubic of $x^4 + 8x + 12$ over $\mathbb{Q}$ is

$$x^3 - 48x - 64.$$

The rational root also shows that this is irreducible over $\mathbb{Q}$, and its discriminant 576^2 is a square in $\mathbb{Q}$, so $x^4 + 8x + 12$ has Galois group A_4.

Lecture 21: Cardano's Formulas

Warm-Up. We determine the Galois groups of $x^4 + 2x + 2$ and $x^4 + 24x + 36$ over $\mathbb{Q}$. These are both irreducible over $\mathbb{Q}$, the former by Eisenstein's Criterion with the prime 2, and the second by ruling out rational roots and then quadratic factors through a brute-force check. (The details of checking are straightforward, but somewhat tedious.) Thus we are in the scenario where the Galois group is a transitive subgroup of S_4.

The resolvent cubic of $x^4 + 2x + 2$ is $x^3 - 8x - 4$, which is irreducible over $\mathbb{Q}$ since it has no rational roots. The discriminant of this cubic (which is also the discriminant of the quartic) is 1616, which is not a square in $\mathbb{Q}$. Thus $x^4 + 2x + 2$ has Galois group S_4 over $\mathbb{Q}$. The resolvent cubic of $x^4 + 24x + 36$ is $x^3 - 144x - 576$. This can be shown to be irreducible by the rational root test, or perhaps more quickly by noting that its reduction mod 5 is $x^3 + x + 4$, which is irreducible over $\mathbb{F}_5$ since it has no root there. The discriminant in this case is $2985984 = 1728^2$, which is a square, so the Galois group of $x^4 + 24x + 36$ is A_4.

The reducible resolvent case. We finish our determination of Galois groups of quartics by considering irreducible quartics with reducible resolvent cubics. First, suppose the resolvent factors into three linear terms over F, so that the roots $\theta_1, \theta_2, \theta_3$ of the resolvent are all in F. Since F is the fixed field of the Galois group, this means that the Galois group fixes each θ_i, so that the Galois group is contained in the *stabilizer* of each θ_i in S_4, and hence is contained in the intersection of these stabilizers. Now, a homework problem from the fall (!) actually determined what these stabilizers are: the stabilizer of θ_i is a copy of D_8 in S_4. (This was exercise 2.2.12 in our book, which was on Homework 3 in the fall. Actually, that problem only explicitly dealt with $\theta_1 = (\alpha_1+\alpha_2)(\alpha_3+\alpha_4)$, but the answer is the same for θ_2 and θ_3 simply by relabeling the roots.) To be sure, the "copy" of D_8 we get depends on how we label the vertices of the square we permute to get D_8: for θ_1, we get the copy where the minimal rotation is (1234); for θ_2 the copy where the minimal rotation is (1324); and for θ_3 the version where the rotation is (1423). The intersection of all these copies of D_8 can be worked out to be

$$\{(1), (12)(34), (13)(24), (14)(23)\} \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z},$$

which thus contains our Galois group. Since no proper subgroup of this group is among our transitive candidates, we get that the Galois group is precisely $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ in this case.

Finally we have the case where the resolvent cubic factors into a linear term and an irreducible quadratic, so that there is only one root in F. The Galois group is then in the stabilizer D_8 of this root, but not in $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ since it does not stabilize all roots, so the candidates are D_8 and $\mathbb{Z}/4\mathbb{Z}$ (i.e. the subgroup of rotations in D_8) since these are the only transitive subgroups contained in D_8 not equal to $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$. To distinguish between these two possibilities, consider the tower

$$F \subseteq F(\sqrt{D}) \subseteq K$$

where D is the discriminant. (Note that $\sqrt{D} \notin F$ in this case since the Galois group, either D_8 or $\mathbb{Z}/4\mathbb{Z}$, is not contained in A_4.) Elements in the Galois group of K over $F(\sqrt{D})$ are elements in $\operatorname{Gal}(K/F)$ that *do* fix $\sqrt{D}$. But to fix $\sqrt{D}$ requires belong to A_4, so we see that the Galois group of the smaller extension $K/F(\sqrt{D})$ is $\operatorname{Gal}(K/F) \cap A_4$. In the case where $\operatorname{Gal}(K/F) = D_8$, we can work out that this intersection is $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} = \{(1), (12)(34), (13)(24), (14)(24)\}$, whereas when $\operatorname{Gal}(K/F) = \mathbb{Z}/4\mathbb{Z}$, this intersection ends up having order 2. In particular, $\operatorname{Gal}(K/F) = D_8$ if and only if $\operatorname{Gal}(K/F(\sqrt{D})$ is *transitive*. But now we are done: if our quartic is irreducible over $F(\sqrt{D})$, then $\operatorname{Gal}(K/F(\sqrt{D}))$ must act transitively on the roots, so $\operatorname{Gal}(K/F) = D_8$ in this case; while if our quartic is reducible, $\operatorname{Gal}(K/F(\sqrt{D}))$ does not act transitively on the roots since it must permute the roots of each irreducible factor amongst themselves, so we must have $\operatorname{Gal}(K/F) = \mathbb{Z}/4\mathbb{Z}$. (This latter observation is analogous to how the four roots of $(x^2-2)(x^2-3)$ are actually permuted in pairs, since the roots of the factor x^2-2 are permuted among themselves as the roots of the x^2-3 factor, so that no Galois group element can send a root of one factor to a root of another.) Thus, we distinguish between D_8 and $\mathbb{Z}/4\mathbb{Z}$ as the Galois group in this final case by seeing whether or not our quartic is irreducible over $F(\sqrt{D})$.

Examples. The classification of Galois groups of quartic polynomials is not something that is of crucial importance in general, but it was worth going through since it highlights some ideas—such as that of having our Galois group act on *other* sets apart from the just roots themselves—that are important, and also because it forces us to recall some group theory we will need going forward. Ultimately, the classification is something we can always look up again when needed, so recalling each case off the top-of-our-head is not very important either.

But, let us look at a few more examples to see how this final bit of the classification works. First, the quartic $x^4+36x+63$ is irreducible over $\mathbb{Q}$, and its resolvent factors as $(x-18)(x+6)(x+12)$. Since this thus has 3 roots in $\mathbb{Q}$, $x^4+36x+63$ has Galois group $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$. Next, x^4-2 is an example we previously worked out directly, where we found the Galois group to be D_8 by explicitly computing all elements in terms of cycles. Using our new classification we get D_8 as follows. The resolvent of x^4-2 is x^3+8x, whose only root in $\mathbb{Q}$ is 0. So we are in the D_8 or $\mathbb{Z}/4\mathbb{Z}$ case. Now, the discriminant here is -2^11, so $\mathbb{Q}(\sqrt{D}) = \mathbb{Q}(i\sqrt{2})$. Since x^4-2 factors into quadratics as

$$x^4-2 = (x^2-\sqrt{2})(x^2+\sqrt{2}),$$

we see that x^4-2 remains irreducible over $\mathbb{Q}(i\sqrt{2})$ since these quadratic factors are not in $\mathbb{Q}(i\sqrt{2})[x]$. Thus since x^4-2 is irreducible over $\mathbb{Q}(\sqrt{D})$, the Galois group is indeed D_8.

Finally, consider x^4+5x+5, which has resolvent cubic $(x-5)(x^2+5x+5)$. This cubic has one root in $\mathbb{Q}$, so again we are in the D_8 or $\mathbb{Z}/4\mathbb{Z}$ case. The discriminant is $5^3 11^2$, so $\mathbb{Q}(\sqrt{D}) = \mathbb{Q}(\sqrt{5})$. But now, we have:

$$x^4+5x+5 = [x^2+\sqrt{5}x+\tfrac{1}{2}(5-\sqrt{5})][x^2-\sqrt{5}x+\tfrac{1}{2}(5+\sqrt{5})],$$

which *is* a valid factorization over $\mathbb{Q}(\sqrt{5})$. Thus x^4+5x+5 is reducible over $\mathbb{Q}(\sqrt{D})$, so its Galois group is $\mathbb{Z}/4\mathbb{Z} \leq S_4$.

Cardano's formulas. We now consider the problem of explicitly finding the roots of a cubic polynomial, which are given by what are called *Cardano's formulas*. Ultimately, these formulas are not very practically useful since it is rare that we need to know such explicit roots (indeed, modern algorithms for finding or approximating roots rely on calculus instead, such as in the technique called *Newton's method*), but theoretically the ideas that go into the derivation of Cardano's formulas *are* useful. In particular, we will see a first example of a *Lagrange resolvent*, which will be crucial to understanding the solvability of polynomials in general. With this in mind, we will omit most of the computational details involved and will focus on highlighting the key aspects. (I am doing all this earlier than the book does, since I want to use this to introduce the keys ideas first.)

Let x^3+px+q be our cubic over F. (Assume we've already made the change of variables needed to get rid of the quadratic term.) The Galois group is then a subgroup of S_3. Let $\theta_1, \theta_2, \theta_3$ denote the roots we want to find. Take ζ to be a primitive third root of unity, and set A in the splitting field K to be

$$A = \theta_1 + \zeta\theta_2 + \zeta^2\theta_3.$$

(We will see later that we lose nothing by assuming K does contain ζ.) This A is called a *Lagrange resolvent*, and it provides something on which the Galois group can act, which we will use to extract information about the structure of K. Note that A is not invariant under all permutations of the roots, but for the element $(123) \in S_3$ we have:

$$(123) \cdot A = \theta_2 + \zeta\theta_3 + \zeta^2\theta_1 = \zeta^2 A$$

and similarly for (132) we have

$$(132) \cdot A = \theta_3 + \zeta\theta_1 + \zeta^2\theta_2 = \zeta A.$$

The upshot is that, even though A is not fixed by these elements, A^3 *is* fixed, or in other words A^3 is fixed by the alternating group $A_3 = \langle (123) \rangle$. (Cubing will introduce ζ^6 in the first expression and ζ^3 in the second, both of which are 1 since ζ is a cube root of unity.)

In the same way, set $B \in K$ to be

$$B = \theta_1 + \zeta^2\theta_2 + \zeta\theta_3.$$

(This is also a "Lagrange resolvent".) Then we get that B^3 is also fixed by A_3. The idea is to then find explicit expressions for A^3 and B^3—exploiting the fact that they are fixed by A_3—and to then take cube roots to find A and B. Once we have A and B, we throw in the equation

$$0 = \theta_1 + \theta_2 + \theta_3,$$

which comes from the fact that the first elementary symmetric polynomial s_1 is zero (the coefficient of x^2) in our case, to get the following system of *linear* equations:

$$\begin{aligned} \theta_1 + \zeta\theta_2 + \zeta^2\theta_3 &= A \\ \theta_1 + \zeta^2\theta_2 + \zeta\theta_3 &= B \\ \theta_1 + \theta_2 + \theta_3 &= 0. \end{aligned}$$

We can solve this system using the inverse of the matrix

$$\begin{bmatrix} 1 & \zeta & \zeta^2 \\ 1 & \zeta^2 & \zeta \\ 1 & 1 & 1 \end{bmatrix}$$

to thus get what we want: explicit formulas for θ_1, θ_2, and θ_3 in terms of A and B, which, as we'll see, can be given in terms of the coefficients p, q of our cubic.

The structure of the splitting field. Consider the tower

$$F \subseteq F(\sqrt{D}) \subseteq K$$

where D is the discriminant of our cubic. The fact that A^3 lies in the fixed field of A_3 but not all of S_3 suggests that, if the Galois group were the full S_3, A^3 should lie in an extension between F and K. In the case where the full Galois group is S_3, then in fact the tower above corresponds to the subgroup chain given by:

$$S_3 \geq A_3 \geq 1,$$

where $F(\sqrt{D})$ is precisely the fixed field of A_3. (In the case where the Galois group is A_3, we have $F = F(\sqrt{D})$ and what follows still applies, and in the case where the Galois group is trivial or $\mathbb{Z}/2\mathbb{Z}$, we have $F = K$ or $F(\sqrt{D}) = K$ respectively, and we don't need Cardano's formulas to find the roots since at worst only the quadratic formula is needed.)

The conclusion is that A^3, being fixed by A_3, must be an element of $F(\sqrt{D})$, so that A^3 should be expressible in terms of $\sqrt{D}$, and the same is true of B^3 also. The overarching idea is that K, by adjoining a cube root, can be obtained as a *cubic* extension of $F(\sqrt{D})$, which itself is at worst a quadratic extension of F, so that we should be able to express A, B and hence our roots θ_i using only cube roots, square roots, and the coefficients of our polynomial. Essentially, it is the subgroup chain $S_3 \geq A_3 \geq 1$ that allows for this to happen, and specificaly the fact that the quotients S_3/A_3 and $A_3/1$ in this chain are cyclic of orders 2 and 3 respectively. (In other words, it's the fact that S_3 and A_3 are *solvable* that is key, as we will see formally next week.)

Deriving the roots. Deriving the explicit roots now comes down to an exercise in working with symmetric polynomials. (We will omit most of the details of the computations needed.) First, we can compute A^3 directly by cubing $A = \theta_1 + \zeta\theta_2 + \zeta^2\theta_3$. The result will involve expressions in the θ_i, some of which are symmetric in these roots and can hence be expressed in terms of elementary symmetric polynomials s_1, s_2, s_3, or in other words the coefficients of the cubic in question. Concretely, we get:

$$A^3 = s_1^3 - 3s_1s_2 + 9s_3 + 3\zeta\underbrace{(\theta_1^2\theta_2 + \theta_2^2\theta_3 + \theta_3^2\theta_1)}_{R_1} - 3(\zeta+1)\underbrace{(\theta_1\theta_2^2 + \theta_2\theta_3^2 + \theta_3\theta_1^2)}_{R_2}.$$

The terms labeled R_1 and R_2 above are not symmetric since they are not invariant under all permutations in S_3, but notice that $R_1 + R_2$ and R_1R_2 *are* symmetric! Hence these expressions can be written in terms of the coefficients of our cubic. But these are the coefficients of the quadratic polynomial

$$x^2 - (R_1 + R_2)x + R_1R_2,$$

whose roots are precisely R_1 and R_2. Thus if we know how to express this quadratic in terms of the coefficients of our cubic, then the values of R_1 and R_2 can be found using the quadratic formula, and we will get a expressions for R_1, R_2 involving the coefficients of our cubic and square roots. We can then repeat this process for B^3.

Thus putting it all together, we find the roots of our cubic explicitly as follows:

- write down the quadratic $x^2 - (R_1 + R_2)x + R_1R_2$ whose coefficients are expressible in terms of the coefficients of our cubic;
- use the quadratic formula to then find R_1 and R_2 in terms of our cubic coefficients and a square root;
- use these values to write down the expression for A^3 in terms of our cubic coefficients and a square root;
- do the same for B^3;
- find A and B by taking a cube root, resulting in an expressions involving our cubic coefficients, a square root, and a cube root;
- use A and B to write down the 3×3 system of linear equations we had at the start for the roots of our cubic, and
- solve this linear system using an inverse matrix to find our roots in terms of the cubic coefficients, a square root, and a cube root.

(Again, for us actually carrying this out is not as important as recognizing that what allows is to happen is really the solvable structure of S_3, and in particular the ability to write A^3 and B^3 in terms a square root of the discriminant. A similar idea, using generalizations of the Lagrange resolvents A and B, will work in the general setting.) If you carry this out, you end up with the following expressions:

$$A = \sqrt[3]{-\frac{27}{2}q + \frac{3}{2}\sqrt{-3D}} \quad \text{and} \quad B = \sqrt[3]{-\frac{27}{2}q - \frac{3}{2}\sqrt{-3D}},$$

where the cube roots chosen are the ones satisfying $AB = -3p$, and the roots of $x^3 + px + q$ are

$$\theta_1 = \frac{A+B}{3}, \;\; \theta_2 = \frac{\zeta A + \zeta^2 B}{3}, \;\; \text{and } \theta_3 = \frac{\zeta^2 A + \zeta B}{3}.$$

These are Cardano's formulas.

Roots of quartics. So, we have the quadratic formula for roots of a quadratic, Cardano's formulas (i.e. the "cubic formula") for the roots of a cubic, and all that remains is a "quartic formula" for the roots of a quartic. (There is no "quintic formula" or anything along these lines of higher degree.) In fact, the quartic case (assuming we've written it to not have a cubic term) reduces to the cubic case using the resolvent cubic, and thus also uses Cardano's formulas. Recall that the roots of the resolvent cubic where

$$\theta_1 = (\alpha_1 + \alpha_2)(\alpha_3 + \alpha_4) \quad \theta_2 = (\alpha_1 + \alpha_3)(\alpha_2 + \alpha_4) \quad \theta_3 = (\alpha_1 + \alpha_4)(\alpha_2 + \alpha_3)$$

where $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ are the roots of the quartic. If you play around with these equations you can work out that it is possible to solve for $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ solely in terms of $\theta_1, \theta_2, \theta_3$. For example, one value is

$$\alpha_1 = \frac{1}{2}(\sqrt{-\theta_1} + \sqrt{-\theta_2} + \sqrt{-\theta_3}),$$

and the others are similar. Thus, we can write down the resolvent cubic, find its roots using Cardano's formulas, and then find the roots of our quartic. We end up with an expression involving the coefficients of our quartic, a square root, a cube root, and more square roots.

Lecture 22: Solvability by Radicals

Warm-Up. A cubic over $\mathbb{Q}$ has either one real root and two non-real complex conjugate roots, or three real roots since the non-real complex conjugate roots have to come in pairs. We show that the three real root case occurs if and only if the discriminant is nonnegative. This is a straightforward argument using only the square root of the discriminant and is not dependent on Cardano's formulas nor any Galois theory, but we will use it to highlight the types of numbers which Cardano's formula can produce.

If the roots $\theta_1, \theta_2, \theta_3$ of our cubic are all real, then

$$\sqrt{D} = (\alpha_1 - \alpha_2)(\alpha_1 - \alpha_3)(\alpha_2 - \alpha_3)$$

is an element of $\mathbb{R}$, which requires that $D \geq 0$ since the square roots of a negative real number are imaginary. (The $D = 0$ case only happens when there are repeated roots, so for a separable polynomial with three real roots, we must have $D > 0$.) Conversely (or rather, the contrapositive of the converse), if there is one real root θ and a pair of non-real complex conjugate roots $c \pm id$, then:

$$\begin{aligned}\sqrt{D} &= [\theta - (c+id)][\theta - (c-id)][(c+id) - (c-id)] \\ &= [(\theta - c)^2 + d^2]2id.\end{aligned}$$

Since $\theta, c, d \in \mathbb{R}$, the expression for $\sqrt{D}$ is purely imaginary, which thus requires that $D < 0$. Hence there are three real roots if and only if $D \geq 0$ as claimed.

The point is that, even though Cardano's formulas explicitly use non-real complex numbers, as in the primitive cube root of unity ζ used or in $\sqrt{-3D}$ if $D \geq 0$, at least one and potentially all three of the resulting roots will be real. If $D < 0$, then $\sqrt{-3D}$ is real and A and B in Cardano's formulas can be chosen to be real, so that $\frac{A+B}{3}$ will be the real root in this case while $\frac{\zeta A + \zeta^2 B}{3}$ and $\frac{\zeta^2 A + \zeta B}{3}$ are the complex conjugate roots, whereas if $D > 0$, $\sqrt{-3D}$ is imaginary and A and B are non-real, but the resulting roots are in fact real. The moral is: it is a more subtle issue to determine if a given expression is real than solely by seeing whether or not any non-real numbers are used.

Solvable by radicals. Recall that the underlying field theory behind Cardano's formulas comes from using the tower

$$F \subseteq F(\sqrt{D}) \subseteq K$$

to construct the roots of a cubic. We wrote down some element $A = \theta_1 + \zeta\theta_2 + \zeta^2\theta_3$ on which our Galois group can act, argued that A^3 would have to belong to $F(\sqrt{D})$ so that it would be expressible in terms of $\sqrt{D}$, and then took a cube root to get A and eventually an expression for the roots. Group theoretically, this all corresponds to having either the chain of subgroups

$$S_3 \trianglerighteq A_3 \trianglerighteq 1$$

or just $A_3 \trianglerighteq 1$ (either way $F(\sqrt{D})$ is the fixed field of A_3) depending on whether the Galois group is S_3 or A_3 (we will comment on the other possibilities later), where we note that in both cases each successive quotient (such as S_3/A_3 or $A_3/1$) is cyclic.

We now seek to generalize this picture to the quartic and higher-order cases, to the extent possible. But for this we need some proper definitions. We say that a polynomial $p(x) \in F[x]$ is *solvable by radicals* over F if each of its roots can be written in terms of elements of F using $+, -, \cdot, \div$ and root extractions. More formally, this means that each root lies in a field K obtained via a sequence of extensions

$$F \subseteq K_1 \subseteq K_2 \subseteq \cdots \subseteq K_{t-1} \subseteq K$$

where K_i is of the form $K_i = K_{i-1}(\sqrt[n_i]{D_i})$ for some $D_i \in K_{i-1}$. (Set $K_0 = F$ and $K_t = K$.) This is analogous to what we saw earlier for constructible numbers, only that we now allow arbitrary roots and not simply 2^k-th roots. So, we have for instance $K_1 = F(\sqrt[n_1]{D_1})$ for some $D_1 \in F$, and then

$$K_2 = K_1(\sqrt[n_2]{D_2}) = F(\sqrt[n_1]{D_1}, \sqrt[n_2]{D_2})$$

for some $D_2 \in K_1$, and so on until $K = F(\sqrt[n_1]{D_1}, \sqrt[n_2]{D_2}, \ldots, \sqrt[n_t]{D_t})$. We will call extensions of the form $K_{i-1}(\sqrt[n_i]{D_i})$ *simple radical extensions*, so $p(x)$ is solvable by radicals if all its roots lie in fields obtained via sequences of simple radical extensions. Our goal is to understand what this condition corresponds to in term of the Galois group. (Spoiler alert: we already saw the relevant concept in the fall, and hopefully the use of the term "solvable" gives it away.)

Quadratics and cubics. First up, quadratics and cubics are always solvable by radicals. In the quadratic case, this comes exactly from the quadratic formula: the roots of $x^2 + bx + c$ are

$$\frac{-b \pm \sqrt{b^2 - 4c}}{2},$$

and so lie in the radical extension $F(\sqrt{D})$ of F. (Here, $D = b^2 - 4c$ is the discriminant. Also, let us suppose F has characteristic zero throughout, so that division by any nonzero element is possible.) The Galois group in this case is either trivial (when $\sqrt{D} \in F$ so that $F = F(\sqrt{D})$) or $\mathbb{Z}/2\mathbb{Z}$.

The fact that cubics are solvable by radicals comes from Cardano's formulas, which, as we've said, do express the roots in terms of radicals and elements of the base field. In the case where the cubic is reducible, Cardano's formulas do give the correct roots, but are not needed since in this case we can make use of the quadratic formula (in the linear times irreducible quadratic case) or nothing at all (three linear factors) to get the roots. So, we really only need Cardano's formulas in the irreducible cubic case, in which case we use one of the two towers

$$F \subseteq F(\sqrt{D}) \subseteq K \quad \text{or} \quad F = F(\sqrt{D}) \subseteq K$$

with Galois groups S_3 or A_3 respectively, as discussed previously. In particular, if $A = \theta_1 + \zeta\theta_2 + \zeta^2\theta_3$ is our "Lagrange resolvent", with $A^3 \in F(\sqrt{D})$, then if we set $\beta = A^3 \in F(\sqrt{D})$, so that $A = \sqrt[3]{\beta}$, then our simple radical extensions look like

$$F \subseteq F(\sqrt{D}) \subseteq F(\sqrt{D}, \sqrt[3]{\beta}) = K.$$

The corresponding sequence of Galois groups is, as stated before, either

$$S_3 \trianglerighteq A_3 \trianglerighteq 1 \quad \text{or} \quad A_3 \trianglerighteq 1$$

depending on whether $F \neq F(\sqrt{D})$ or $F = F(\sqrt{D})$. The fact that only a square root and cube root is needed to get all the roots is encoded in the fact that the successive quotients in these two chains are $\mathbb{Z}/2\mathbb{Z}$ and $\mathbb{Z}/3\mathbb{Z}$ in the S_3 case, and just $\mathbb{Z}/3\mathbb{Z}$ in the A_3 case. (So, in the A_3 case, the square root used in Cardano's formulas takes the square root a square, and so the square root "goes away" in the end.) This relation between the type of root needed and the type of (cyclic) quotient appearing will be crucial to understanding the general situation.

We will note here that both the quadratic formula and Cardano's formulas actually work for the "general polynomials" of degree 2 and 3. In the degree 2 case, the general polynomial is

$$(x-x_1)(x-x_2)=x^2-s_1x+s_2\in F(s_1,s_2)$$

where s_1, s_2 are the elementary symmetric polynomials. The quadratic formula gives the roots as

$$\frac{s_1\pm\sqrt{s_1^2-4s_2}}{2},$$

which, if you plug in $s_1=x_1+x_2$ and $s_2=x_1x_2$ and simplify, become precisely x_1 and x_2. Thus, the roots of the general polynomial of degree 2 can indeed be given in terms of the coefficients and radicals, as expected. If you do the same for the general polynomial of degree 3:

$$(x-x_1)(x-x_2)(x-x_3)=x^3-s_1x^2+s_2x-s_3\in F(s_1,s_2,s_3),$$

Cardano's formulas will boil down precisely to x_1, x_2, x_3, again as expected. So, even these general polynomials with indeterminate roots are indeed "solvable by radicals".

The quartic case. In the quartic case, as we briefly discussed last time, the expression for the roots $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ can be derived from Cardano's formula applied to the resolvent cubic, where the roots $\theta_1, \theta_2, \theta_3$ of the resolvent cubic give, for example,

$$\alpha_1=\frac{1}{2}(\sqrt{-\theta_1}+\sqrt{-\theta_2}+\sqrt{-\theta_3})$$

and similar expressions for $\alpha_2, \alpha_3, \alpha_4$. The point is that this too shows that quartics are solvable by radicals. In particular, to construct α_1 as above, we need a square root and a cube root to get the "A" and "B" needed in Cardano's formula, then another square root to get $\sqrt{-\theta_1}$, and finally another square root to get $\sqrt{-\theta_2}$ in the expression for α_1. (It turns out the remaining $\sqrt{-\theta_3}$ can be derived from the first two $\sqrt{-\theta_i}$.) This then means that the roots of a quartic come from extensions of the form

$$F\subseteq F(\sqrt{D})\subseteq F(\sqrt{D},\sqrt[3]{\cdot})\subseteq F(\sqrt{D},\sqrt[3]{\cdot},\sqrt{\cdot})\subseteq F(\sqrt{D},\sqrt[3]{\cdot},\sqrt{\cdot},\sqrt{\cdot})$$

with the dots $\cdot$ indicating some unspecified elements of which to take roots. This is indeed a sequence of simple radical extensions as desired.

But now let us think about the corresponding groups. Again let us focus on the case of irreducible quartics, since the reducible quartics are handled by the quadratic and cubic cases after factoring. The possible Galois groups of irreducible quartics are: $S_4, A_4, D_8, \mathbb{Z}/2\mathbb{Z}\times\mathbb{Z}/2\mathbb{Z}, \mathbb{Z}/4\mathbb{Z}$. In the S_4 case, we have the following chain of subgroups, each normal in the previous one:

$$S_4\trianglerighteq A_4\trianglerighteq\mathbb{Z}/2\mathbb{Z}\times\mathbb{Z}/2\mathbb{Z}\trianglerighteq\mathbb{Z}/2\mathbb{Z}\trianglerighteq 1.$$

The fields in the tower of radical extensions above are precisely the fixed fields of these subgroups, and the type of root we have to adjoin at each step comes from the type of *cyclic* group we get as quotients in this subgroup chain: $S_4/A_4\cong\mathbb{Z}/2\mathbb{Z}$ is cyclic of order 2, hence why we first adjoint a square root; $A_4/(\mathbb{Z}/2\mathbb{Z}\times\mathbb{Z}/2\mathbb{Z})\cong\mathbb{Z}/3\mathbb{Z}$ has order 3, hence why we adjoin a cube root next; $(\mathbb{Z}/2\mathbb{Z}\times\mathbb{Z}/2\mathbb{Z})/(\mathbb{Z}/2\mathbb{Z})\cong\mathbb{Z}/2\mathbb{Z}$ has order 2, so we adjoin a square root next; and finally one more square root from $(\mathbb{Z}/2\mathbb{Z})/1 \cong \mathbb{Z}/2\mathbb{Z}$. If our Galois group was A_4, we essentially start with the second field in our tower since $F = F(\sqrt{D})$. For Galois group D_8 with subgroup chain

$$D_8 \trianglerighteq \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} \trianglerighteq \mathbb{Z}/2\mathbb{Z} \trianglerighteq 1$$

and quotients $\mathbb{Z}/2\mathbb{Z}$, $\mathbb{Z}/2\mathbb{Z}$, and $\mathbb{Z}/2\mathbb{Z}$, we can in effect get way with only adjoining three successive square roots, since essentially the cube root used in the "quartic formula" takes the cube root of a cube and so isn't really there. A similar process works for the other possible Galois groups.

Revisiting solvable groups. The discussions above are meant to suggest that, indeed, there is a strong relation between polynomials which are solvable by radicals and subgroup chains of their Galois groups with cyclic quotients. We saw such chains in the fall, under the more general setting of successive *abelian* quotients, when discussing *solvable* groups. To recall the definition, a group G is solvable if there is a chain of subgroups

$$1 \trianglelefteq G_1 \trianglelefteq G_2 \trianglelefteq \ldots \trianglelefteq G_{n-1} \trianglelefteq G,$$

each normal in the next, such that each quotient G_i/G_{i-1} is abelian. (Set $G_0 = 1$ and $G_n = G$.) The types of Galois groups we've seen in the quadratic, cubic, and quartic cases are all solvable since cyclic quotients are also abelian quotients. The big theorem we will prove this week is that this relation holds in general:

A (separable) polynomial is solvable by radicals if and only if its Galois group is solvable.

(Indeed, this is where the term "solvable" for this group property comes from.) The reason why all polynomials of degree less than 5 are solvable by radicals—and hence why there exist quadratic, cubic (Cardano), and quartic (modified Cardano) formulas—is because S_4 and all of its subgroups are solvable groups. The smallest non-solvable group is A_5 in the order 60 case, and we will discuss what happens with quintics and solvability soon enough.

Now, there is one point of clarification here. The definition of "solvable" uses abelian quotients, but the examples we've seen have quotients that are cyclic, and moreover we have alluded to the idea that *cyclic* quotients are what are truly needed, since it is the cyclic structure which (as we'll see) gives rise to a simple radical extension. So, there might seem to be a mismatch at first in using abelian vs cyclic quotients in the definition of solvable, but in fact there is no difference in the case of finite groups: a finite group G is solvable if and only if there is a chain as in the definition but with cyclic groups as the successive quotients. (This is not true for infinite groups in general. The reason why the definition of "solvable" is phrased in terms of abelian and not cyclic quotients in general has to do with other areas where solvable groups come up—in particular in the study of what are called *Lie groups*—but this will be of no concern for us.)

To see why having abelian quotients is equivalent to having cyclic quotients in the finite group case, suppose G is a finite solvable group with chain

$$1 \trianglelefteq G_1 \trianglelefteq G_2 \trianglelefteq \ldots \trianglelefteq G_{n-1} \trianglelefteq G$$

where each quotient G_i/G_{i-1} is abelian. (The cyclic quotient implies abelian quotient direction is easy since cyclic groups are abelian.) Then G_i/G_{i-1} is a finite abelian group, so it is a product of finitely many finite cyclic groups. Take one such cyclic factor, which is a cyclic (normal) subgroup of the form

$$H_i/G_{i-1} \trianglelefteq G_i/G_{i-1} \text{ where } G_{i-1} \trianglelefteq H_i \trianglelefteq G_i.$$

(This is something we saw in the fall: every subgroup of a quotient A/B is the quotient of a subgroup C sitting between B and A.) Then in the chain

$$G_{i-1} \trianglelefteq H_i \trianglelefteq G_i,$$

the first quotient H_i/G_{i-1} is cyclic and the second G_i/H_i is still finite and abelian. We then repeat this same process for G_i/H_i, to "break off" another cyclic piece, and so on until we've filled in various subgroups between G_{i-1} and G_i with now successive cyclic quotients. More formally, we can use induction: the quotient G_i/H_i is isomorphic to

$$G_i/H_i \cong (G_i/G_{i-1})/(H_i/G_{i-1})$$

by the third isomorphism theorem in group theory, and so its decomposition as a product of cyclic groups uses one fewer factor than does the decomposition for G_i/G_{i-1}, since we are modding out G_i/G_{i-1} exactly by one such factor H_i/G_{i-1}. (This is like saying that $\mathbb{Z}/k_1\mathbb{Z} \times \mathbb{Z}/k_2\mathbb{Z} \times \cdots \times \mathbb{Z}/k_t\mathbb{Z}$ mod the first factor $\mathbb{Z}/k_1\mathbb{Z}$ leaves $\mathbb{Z}/k_2\mathbb{Z} \times \cdots \times \mathbb{Z}/k_t\mathbb{Z}$.)

Thus, by induction of the number of cyclic factors in such a decomposition, we may assume that the chain $H_i \trianglelefteq G_i$ has already been filled in with more subgroups

$$H_i \trianglelefteq H_i' \trianglelefteq \ldots \trianglelefteq G_i$$

with cyclic successive quotients. This then gives a refinement of $G_i \trianglelefteq H_i \trianglelefteq G_i$ into

$$G_i \trianglelefteq H_i \trianglelefteq H_i' \trianglelefteq \ldots \trianglelefteq G_i$$

with cyclic quotients all along the way. Doing this for each $G_{i-1} \trianglelefteq G_i$ in our original abelian quotient chain then gives a chain with cyclic quotients as desired.

No quintic formula. We will start working towards the proof that "solvable by radicals" is the same as "solvable Galois group" next time, but for now we can give a definite example of a polynomial that is not solvable by radicals, which then means that there is no "quintic formula" for the roots of a quintic polynomial in general. Indeed, take the general polynomial of degree 5:

$$(x - x_1)(x - x_2)(x - x_3)(x - x_4)(x - x_5) \in F(s_1, s_2, s_3, s_4, s_5).$$

This has Galois group S_5, which we saw in the fall is not solvable. (The point is that the only possible "normal subgroup chain" one can have in this case is $1 \trianglelefteq A_5 \trianglelefteq S_5$ since A_5 is *simple*—meaning it has no nontrivial proper normal subgroups—and this does not work as a "solvable chain" since the first quotient is A_5, which is not abelian/cyclic.) Thus, since the Galois group of this general polynomial is not solvable, it is not solvable by radicals. (The same is true of the general polynomial in higher degrees, since S_n is not solvable for $n \geq 5$. So, there is no formula for the roots of a general polynomial in any degree larger than 5 either.)

To be clear, we are not saying that the roots are not possible to write down explicitly—indeed, they are simply x_1, x_2, x_3, x_4, x_5—but rather that the roots cannot be expressed in terms of radicals/root extractions and elements of the base field $F(s_1, s_2, s_3, s_4, s_5)$, which is what a "quintic formula" would require. If we want a more concrete example of a polynomial involving actual *numbers* (and not indeterminates) that is not solvable by radicals, we will show next time that $x^5 - 4x^4 + 2x + 2$ is one such example over $\mathbb{Q}$.

Lecture 23: Cyclic Extensions

Warm-Up. We show that the polynomial $x^5 - 4x^4 + 2x + 2$ is not solvable by radicals over $\mathbb{Q}$. This will come from the fact that its Galois group is S_5, which is not solvable. The key properties here are that $x^5 - 4x^4 + 2x + 2$ is irreducible over $\mathbb{Q}$ (Eisenstein with the prime $p = 2$) and that it has exactly three real roots. The argument we will give in fact applies to *any* irreducible quintic with three real roots, giving more examples of polynomials with Galois group S_5 and hence not solvable by radicals.

To see that $f(x) = x^5 - 4x^4 + 2x + 2$ has exactly three real roots we use some calculus. (Drawing the graph should convince you that this is true, but we can be more precise.) First, we have

$$f(-1) = -5 \quad f(0) = 2 \quad f(2) = -26 \quad f(4) = 10.$$

Since $f(x)$ defines a continuous function and $f(-1) < 0 < f(0)$, the intermediate value theorem implies there exists r_1 between -1 and 0 such that $f(r_1) = 0$, so this is one root. In the same way, there is a root r_2 between 0 and 2, and one r_3 between 2 and 4. Hence $f(x)$ has at least three real roots. Since non-real complex roots have to come in complex conjugate pairs, if there were more than three real roots there would have to be five. But now the Mean Value Theorem would imply that the derivative $f'(x)$ had at least four real roots, and then that the second derivative $f''(x)$ had at least three. But $f''(x) = 20x^3 - 48x^2 = x^2(20 - 48x)$ only has two roots, so $f(x)$ cannot have five real roots. Thus $f(x)$ has exactly three real roots as claimed.

Now, since $f(x)$ is irreducible, $\mathbb{Q}(r_1)$ (where r_1 is a real root) has degree 5 over $\mathbb{Q}$. But $[\mathbb{Q}(r_1) : \mathbb{Q}] = 5$ divides the degree of the splitting field over $\mathbb{Q}$ by the tower law, and hence 5 divides the order of the Galois group of $f(x)$. By Cauchy's theorem in group theory (a finite group has an element of order p for any prime p dividing its order), this Galois group (a subgroup of S_5) has an element of order 5, which is thus a 5-cycle $(abcde)$. But complex conjugation, i.e. the sending one non-real complex root β to the other $\overline{\beta}$, is also an element of this Galois group, so the Galois group contains a 2-cycle. Label the non-real complex roots by 1 and 2, so that the 2-cycle is (12). By relabeling the other elements if needed, we may assume our 5-cycle looks like $(1bcde)$. But all powers of this 5-cycle are also 5-cycles, and one of these powers will send 1 to 2, so we may assume our 5-cycle looks like $(12cde)$ after relabeling. Then label the real roots 3, 4, and 5, so that the 5-cycle is (12345). The upshot is that the Galois group of $x^5 - 4x^4 + 2x + 2$ is a subgroup of S_5 containing (12345) and (12), but we showed in the fall that these two elements generate all of S_5, so we conclude that the Galois group must be all of S_5 as claimed.

Thus $x^5 - 4x^4 + 2x + 2$ is not solvable by radicals over $\mathbb{Q}$. Take note of what this means: this does not mean that this polynomial has no roots, nor that the roots aren't expressible in *some* way, but rather it means the roots are not expressible in terms of (possibly iterated) radical expressions involving rational numbers and basic algebraic operations. (So, again, there can be no "quintic formula" analogous to the quadratic formula.)

Cyclic extensions. We now start working towards the proof that solvable by radicals is equivalent to solvable Galois group, focusing here on the forward direction. As alluded to last time, the key idea is that simple radical extensions of fields should in some sense correspond to cyclic quotients in a group solvability chain. Actually, this is not literally true as stated, but it will be "morally" true and in fact literally true under a basic assumption we will soon clarify.

Here is a key definition: we say that K/F is a *cyclic extension* if it is Galois and $\text{Gal}(K/F)$ is a cyclic group. These are the types of extension we will need to make our correspondence work. But, it is not true that simple radical extensions $F(\sqrt[n]{a})$ are always cyclic in this sense. For example, we have seen before that $\mathbb{Q}(\sqrt[3]{2})$ (which is a simple radical extension) is not Galois over $\mathbb{Q}$, so it is certainly not cyclic. Here the issue comes down to the fact that $\sqrt[3]{2}$ is a root of $x^3 - 2$ but $\mathbb{Q}(\sqrt[3]{2})$ is not the full splitting field of this polynomial. To get the full splitting field we have to adjoin the cube roots unity, so that we can get all the roots. It turns out that this the *only* thing preventing a simple radical extension from being cyclic in general, and that as soon as we adjoin appropriate roots of unity the trouble goes away. This is not an issue we see with quadratic extensions $\mathbb{Q}(\sqrt{D})$ of $\mathbb{Q}$, which *are* cyclic, precisely because $\mathbb{Q}$ *does* contain the square roots of unity ± 1.

The precise claim we want is that if F contains the n-th roots of unity (suppose the characteristic is zero for our purposes, but more generally the claim holds as long as the characteristic does not divide n), then any simple radical extension $F(\sqrt[n]{a})$ is cyclic over F. (We will see next time that the converse is also true: if F contains all n-th roots of unity and K is cyclic over F, then K is of the form $F(\sqrt[n]{a})$.) First, if F contains all n-th roots of unity, then it contains all roots of $x^n - a$ since these roots are of the form $\zeta \sqrt[n]{a}$ for ζ an n-th root of unity. Thus $F(\sqrt[n]{a})$ is the splitting field of $x^n - a$ over F, so it is a Galois extension. Second, an element σ of the Galois group $\operatorname{Gal}(F(\sqrt[n]{a})/F)$ is fully characterized by its value on $\sqrt[n]{a}$, and this value must be some other root of $x^n - a$:

$$\sigma(\sqrt[n]{a}) = \zeta_\sigma \sqrt[n]{a}$$

for some n-th root of unity ζ_σ. We thus have a map

$$\operatorname{Gal}(F(\sqrt[n]{a})/F) \to \mu_n \text{ defined by } \zeta \mapsto \zeta_\sigma$$

where μ_n denotes the cyclic group of n-th roots of unity. To see that this is a homomorphism, take another Galois group element τ and compute:

$$(\sigma\tau)(\sqrt[n]{a}) = \sigma(\zeta_\tau \sqrt[n]{a}) = \sigma(\zeta_\tau)\sigma(\sqrt[n]{a}) = \zeta_\tau\zeta_\sigma \sqrt[n]{a},$$

where $\sigma(\zeta_\tau) = \zeta_\tau$ because ζ_τ is in the fixed field F of the full Galois group. Thus composition in the Galois group corresponds to multiplication in μ_n. Finally, the map $\operatorname{Gal}(F(\sqrt[n]{a})/F) \to \mu_n$ is injective since an element σ in the kernel satisfies $\zeta_\sigma = 1$, so that $\sigma(\sqrt[n]{a}) = 1\sqrt[n]{a} = \sqrt[n]{a}$, which forces σ to be the identity. Hence we can identity $\operatorname{Gal}(F(\sqrt[n]{a})/F)$ with a subgroup of μ_n, so $\operatorname{Gal}(F(\sqrt[n]{a})/F)$ is cyclic because subgroups of cyclic groups are cyclic.

Manipulating root extensions. In order to be able to apply the result above, we are forced to work with fields containing roots of unity. But if we take the type of extension used in the definition of "solvable by radicals", namely one obtained via a sequence of simple radical extensions:

$$F \subseteq K_1 \subseteq K_2 \subseteq \ldots \subseteq K_{t-1} \subseteq K,$$

there is no assumption that any of these fields contain roots of unity, apart from ± 1. (Let us use the term *root extension* for an extension K/F of this form.) In particular then, there is no guarantee that each K_i/K_{i-1} will be cyclic. The way around this is to essentially force these fields to contain the required roots of unity by simply adjoining them. We will come back to the details of doing so next time, where the upshot is that adjoining these roots does not in fact alter "sovability", either of the polynomial or of the Galois group. It is in this way that we will still be able relate arbitrary simple radical extensions on the field side to cyclic quotients on the group side. The main technical fact we'll need, which we state without proof for the time being, is that any root extension K/F as above can be turned into one where K is Galois over F and where each extension K_i/K_{i-1} is in fact cyclic. So, we may as well assume that all root extensions we work with are of this form already, at least when it comes to questions dealing with solvability.

A key observation we'll need in order to be able to prove this next time is that composites of root extensions are also root extensions. To see this, suppose

$$F \subseteq F(\sqrt[n_1]{a_1}) \subseteq F(\sqrt[n_1]{a_1}, \sqrt[n_2]{a_2}) \subseteq \ldots \subseteq F(\sqrt[n_1]{a_1}, \sqrt[n_2]{a_2}, \ldots, \sqrt[n_t]{a_t})$$

and

$$F \subseteq F(\sqrt[m_1]{b_1}) \subseteq F(\sqrt[m_1]{b_1}, \sqrt[m_2]{b_2}) \subseteq \ldots \subseteq F(\sqrt[m_1]{b_1}, \sqrt[m_2]{b_2} \ldots, \sqrt[m_s]{b_s})$$

are two root extensions. Then we can adjoin $\sqrt[m_1]{b_1}$ to the end of the first chain, then $\sqrt[m_2]{b_2}$ to that, and so until we adjoin $\sqrt[m_s]{b_s}$:

$$F \subseteq \ldots \subseteq F(\sqrt[n_1]{a_1}, \ldots, \sqrt[n_t]{a_t}) \subseteq F(\sqrt[n_1]{a_1}, \ldots, \sqrt[n_t]{a_t}, \sqrt[m_1]{b_1}) \subseteq F(\sqrt[n_1]{a_1}, \ldots, \sqrt[n_t]{a_t}, \sqrt[m_1]{b_1}, \ldots, \sqrt[m_s]{b_s}).$$

The final field is the composite of the original root extensions, obtained itself via a sequence of simple radical extensions. Inductively we get that the same is true for composites of any (finite) number of root extensions, and after we prove that arbitrary root extensions can be turned into Galois ones with cyclic intermediate extensions, we will have that composites of *these* types of root extensions are also of the same type—namely Galois with cyclic intermediate extensions.

Solvable by radicals implies solvable. With the technical fact outlined above, we can now prove that if a polynomial is solvable by radicals, then its Galois group is solvable. Suppose $p(x)$ is solvable by radicals over F. Then each root α_i of $p(x)$ lies in a root extension L_i of F, which we can assume is Galois with cyclic intermediate extensions. Set L be the composite of all these L_i, which then is also a Galois root extension of F with cyclic intermediate extensions. Note that L contains all α_i since $\alpha_i \in K_i$, so L contains the splitting field of $p(x)$. (In general L will be larger than the splitting field, but we'll see that this does not matter!)

Suppose concretely that

$$F \subseteq K_1 \subseteq K_2 \subseteq \ldots \subseteq K_{t-1} \subseteq L$$

is the sequence of simple radical (and cyclic!) extensions which get us from F to L. Taking automorphism groups over F gives the following sequence of groups:

$$\mathrm{Gal}(L/F) \trianglerighteq G_{t-1} \trianglerighteq G_{t-2} \trianglerighteq \ldots \trianglerighteq G_1 \trianglerighteq 1.$$

(So, K_i is the fixed field of G_{t-i}. The difference in how we write the indices comes from the fact that taking automorphisms groups reverses containments. Each subgroup here is normal in the previous one because each intermediate extension in our field tower is Galois.) The Galois group of K_i over K_{i-1} is cyclic since K_i/K_{i-1} is a cyclic extension. But this Galois group is also the quotient of $\mathrm{Gal}(L/K_{i-1})$ by $\mathrm{Gal}(L/K_i)$, which are what we are calling G_{t-i-1} and G_{t-i} above:

$$\mathrm{Gal}(K_i/K_{i-1}) \cong \mathrm{Gal}(L/K_{i-1})/\,\mathrm{Gal}(L/K_i) = G_{t-(i-1)}/G_{t-i}.$$

Thus each successive quotient (working right to left) in

$$\mathrm{Gal}(L/F) \trianglerighteq G_{t-1} \trianglerighteq G_{t-2} \trianglerighteq \ldots \trianglerighteq G_1 \trianglerighteq 1$$

is cyclic, so $\mathrm{Gal}(L/F)$ is solvable.

As said above, L is in general larger than the splitting field K of $p(x)$, but we do have the tower

$$F \subseteq K \subseteq L.$$

Since K/F is Galois (it is a splitting field extension after all), the Galois group $\mathrm{Gal}(K/F)$ is a quotient of $\mathrm{Gal}(L/F)$. But quotients of solvable groups are solvable, so the Galois group $\mathrm{Gal}(K/F)$

of $p(x)$ is solvable as claimed. (The fact that quotients of solvable groups are solvable is from the fall, but here's the argument, modulo some group-theoretic claims you can check on your own. Take a "solvability chain" for G:

$$1 \trianglelefteq G_1 \trianglelefteq G_2 \trianglelefteq \ldots \trianglelefteq G$$

with cyclic/abelian quotients. If H is normal in G, then H is normal in G_iH and $G_{i-1}H$ is normal in G_iH, so we have the normal subgroup chain

$$1 = H/H \trianglelefteq G_1H/H \trianglelefteq G_2H/H \trianglelefteq \ldots \trianglelefteq GH/H = G/H.$$

The successive quotients are

$$(G_iH/H)/(G_{i-1}H/H) \cong G_iH/G_{i-1}H$$

by the third isomorphism theorem for groups, and $G_iH/G_{i-1}H$ is isomorphic to $G_i/(G_i \cap G_{i-1}H)$ by the second isomorphism theorem, which is a quotient of G_i/G_{i-1} and is hence cyclic/abelian.)

Lecture 24: More on Cyclic Extensions

Warm-Up 1. We show that if a simple radical extension $F(\sqrt[n]{a})/F$ is Galois, then $F(\sqrt[n]{a})$ must contain the n-th roots of unity. Since $F(\sqrt[n]{a})$ is Galois over F, it is normal. Thus since $x^n - a$ has a root in $F(\sqrt[n]{a})$, it must split completely in this extension, so $F(\sqrt[n]{a})$ contains all roots of $x^n - a$. These roots are

$$\sqrt[n]{a},\ \zeta\sqrt[n]{a},\ \zeta^2\sqrt[n]{a}, \ldots,\ \zeta^{n-1}\sqrt[n]{a}$$

where ζ is a primitive n-th root of unity, so in particular $F(\sqrt[n]{a})$ contains

$$\zeta = \frac{\zeta\sqrt[n]{a}}{\sqrt[n]{a}}.$$

Thus $F(\sqrt[n]{a})$ contains all n-th roots of unity as claimed.

Warm-Up 2. If θ is a root of $x^3 - 3x - 1$ over $\mathbb{Q}$, we show that $\mathbb{Q}(\theta)$ is an example of a cyclic extension of $\mathbb{Q}$ that is not a simple radical extension. We will show in a bit that cyclic always implies simple radical if our base field contains the appropriate roots of unity, so this highlights what goes wrong if this is not the case. The Galois group of $x^3 - 3x - 1$ is $A_3 \cong \mathbb{Z}/3\mathbb{Z}$, so this is a cyclic extension, and moreover all roots of $x^3 - 3x - 1$ are real since the existence of non-real complex conjugate roots would imply that complex conjugation would be an element of order 2 in the Galois group, which $\mathbb{Z}/3\mathbb{Z}$ does not have.

If $\mathbb{Q}(\theta)$ is a simple radical extension $\mathbb{Q}(\theta) = \mathbb{Q}(\sqrt[n]{a})$ of $\mathbb{Q}$, then the first Warm-Up implies that $\mathbb{Q}(\theta)$ contains the n-th roots of unity. Here $n > 2$ (we don't know yet that it would necessarily have to be 3), since $n = 2$ only gives a degree 2 extension and $\mathbb{Q}(\theta)$ is degree 3 over $\mathbb{Q}$. But for $n > 2$, the primitive n-th root of unity in particular is not real, so $\mathbb{Q}(\theta) = \mathbb{Q}(\sqrt[n]{a})$ would contain a non-real complex number. This is not possible since θ is real (as are all roots of $x^3 - 3x - 1$), so $\mathbb{Q}(\theta)$ could not have been simple radical over $\mathbb{Q}$.

Note that if we adjoin ζ_3 (primitive third root of unity) to our fields $\mathbb{Q} \subseteq \mathbb{Q}(\theta)$ all of this trouble goes away: $\mathbb{Q}(\theta, \zeta_3)$ is in fact still cyclic over $\mathbb{Q}(\zeta_3)$ (we will see next time that adjoining the same element to both fields does not affect whether the Galois group is cyclic), and, since $\mathbb{Q}(\zeta_3)$ does contain the third roots of unity, what we will prove next shows that $\mathbb{Q}(\theta, \zeta_3)$ *is* simple radical over $\mathbb{Q}(\zeta_3)$: $\mathbb{Q}(\theta, \zeta_3)$ can be obtained from $\mathbb{Q}(\zeta_3)$ by adjoining a single cube root, as indicated by

Cardano's formulas. (The discriminant of $x^3 - 3x - 1$ is 81, which is a square in $\mathbb{Q}$, so no square root term is needed in the expression for the roots.)

Cyclic implies radical. We now show that if F contains the n-th roots of unity (either F has $\operatorname{char} F$ zero or not dividing n), then any cyclic extension K/F is of the form $K = F(\sqrt[n]{a})$. This is the converse of something we showed last time, and finishes the "cyclic corresponds to radical" idea we've hoped for, at least when our fields contain appropriate roots of unity.

To get a feel for where this proof comes from, let us recall the derivation of Cardano's formulas. When looking at the splitting field K/F of, say, an irreducible cubic with Galois group S_3, the point was to consider the intermediate field $F(\sqrt{D})$ (D the discriminant), which is the fixed field of $A_3 \leq S_3$:

$$F \subseteq F(\sqrt{D}) \subseteq K.$$

The first extension has Galois group $\mathbb{Z}/2\mathbb{Z}$ (or trivial in the case where the Galois group is A_3, so that $F = F(\sqrt{D})$), and the second has Galois group $A_3 \cong \mathbb{Z}/3\mathbb{Z}$. The fact that K is cyclic over $F(\sqrt{D})$ (with Galois group of order 3) suggests that K should be obtainable from $F(\sqrt{D})$ via adjoining a single cube root, so that it should be simple radical. (Again, not literally true until we introduce cube roots of unity, which we'll discuss afterwards.) To obtain a generator, we took the element

$$A = \theta_1 + \zeta\theta_2 + \zeta^2\theta_3$$

where θ_i were the roots of the cubic and ζ a primitive third root of unity, and argued that A^3 had to be fixed by A_3, so that $A^3 \in F(\sqrt{D})$. Then for $\beta = A^3$, we have $A = \sqrt[3]{\beta}$ and $K = F(\sqrt{D})(\sqrt[3]{\beta})$, giving K as a radical extension of $F(\sqrt{D})$ as desired.

We wish to mimic the same idea in the case K/F at hand, by finding an element A of K whose n-th power is in F, where n is order of the cyclic Galois group $\operatorname{Gal}(K/F) = \langle\sigma\rangle$. To this end, take the roots θ_i of a polynomial of degree n over F for which K is the splitting field (such a thing exists since K is Galois over F), and set

$$A = \theta_1 + \zeta\theta_2 + \zeta^2\theta_3 + \cdots + \zeta^{n-1}\theta_n$$

where ζ is a primitive n-th root of unity. (Such an element is called a *Lagrange resolvent* for the extension. Another way of expressing this is as

$$A = \theta_1 + \zeta\sigma(\theta_1) + \cdots + \zeta^{n-1}\sigma^{n-1}(\theta_1)$$

where σ generates the cyclic Galois group, since applying σ repeatedly cycles through the roots $\theta_i = \sigma^{i-1}(\theta_1)$.) Then we can compute that

$$\sigma(A) = \sigma(\theta_1) + \zeta\sigma(\theta_2) + \cdots + \zeta^{n-1}\sigma(\theta_n) = \theta_2 + \zeta\theta_3 + \cdots + \zeta^{n-1}\theta_1 = \zeta^{n-1}A,$$

which uses the fact that $\sigma(\zeta) = \zeta$ because $\zeta \in F$. This gives $\sigma(A^n) = (\zeta^{n-1})^n A = A$ since ζ^{n-1} is an n-th root of unity, so A^n is fixed by σ and hence by the full Galois group $\operatorname{Gal}(K/F) = \langle\sigma\rangle$. We then have $A = \sqrt[n]{A^n}$, or $A = \sqrt[n]{\beta}$ where $\beta = A^n \in F$, and we claim that $K = F(A) = F(\sqrt[n]{\beta})$ is the simple radical extension we want. To see that $K = F(A)$, we can argue that if $F(A)$ were a proper subfield of K, it would be fixed by some nontrivial subgroup of $\operatorname{Gal}(K/F) = \langle\sigma\rangle$, but

$$\sigma^k(A) = (\zeta^{n-1})^k A$$

shows that no power of σ smaller than n will fix A, since the smallest power such that $(\zeta^{n-1})^k = 1$ is $k = n$ because $\zeta^{n-1} = \zeta^{-1}$ is a primitive n-th root of unity. Thus $F(A)$ is only fixed by the trivial subgroup of $\operatorname{Gal}(K/F)$, so $F(\sqrt[n]{\beta}) = F(A) = K$ as claimed.

Thus K/F being cyclic implies K/F being a simple radical extension when F contains the appropriate roots of unity, except for one subtle point we ignored above! The issue is that if the Lagrange resolvent A used is actually zero, then A *is* already in F, so that $F(A) = F$ is in fact a proper subfield of K and does not equal K. The argument above breaks down here because any σ^k does fix $A = 0$, regardless of what $(\zeta^{n-1})^k$ is. So, we need to guarantee that we use a nonzero Lagrange resolvent to make this work. But this should be possible: the Lagrange resolvent we wrote down is not the only one we could have written down, since for example we can relabel the θ_i in a different way and maybe use

$$\theta_4 + \zeta\theta_1 + \zeta^2\theta_n + \cdots + \zeta^{n-1}\theta_{n-2},$$

for example, instead, or maybe use the roots θ_i of some *different* polynomial altogether. All we need is *some* Lagrange resolvent constructed in this way that is nonzero. The book justifies that this is possible using *linear independence* of the σ^i, which is not a concept we looked at previously. (This is what the book used to prove the equality $[K : K^H] = |H|$ in the setup of the Fundamental Theorem of Galois Theory, but we took a different approach.)

Back to Cardano. Let us revisit Cardano's formulas once again. As stated before, we would like to apply the "cyclic implies radical" idea to a tower like

$$F \subseteq F(\sqrt{D}) \subseteq K,$$

but this doesn't quite work as is since F might not contain a primitive third root of unity. The Lagrange resolvent $A = \theta_1 + \zeta\theta_2 + \zeta^2\theta_3$ isn't even guaranteed to belong to K then! But the fix is easy: we simply adjoin the required root of unity ζ to each, so that the tower are *really* using is

$$F(\zeta) \subseteq F(\sqrt{D}, \zeta) \subseteq K(\zeta).$$

We will show soon enough that doing so does not alter the cyclic nature of Galois groups, so that $F(\sqrt{D}, \zeta)$ over $F(\zeta)$ still has a cyclic Galois group since $\text{Gal}(F(\sqrt{D})/F)$ is cyclic (which is either $\mathbb{Z}/2\mathbb{Z}$ or trivial depending on whether or not $F = F(\sqrt{D})$), and $K(\zeta)$ over $F(\sqrt{D}, \zeta)$ still has a cyclic Galois group since $\text{Gal}(K/F(\sqrt{D})) = A_3$ is cyclic. Each intermediate extension is thus cyclic, and thus a radical extension since the base field $F(\zeta)$ does contain the third roots of unity.

In the definition of a polynomial being solvable by radicals we really should be using extensions of F, and not of $F(\zeta)$, but this we fix by tacking F on at the start:

$$F \subseteq F(\zeta) \subseteq F(\sqrt{D}, \zeta) \subseteq K(\zeta).$$

This is then the full sequence of extensions used in the derivation of Cardano's formulas: the first extension introduces $\zeta = \frac{1}{2} + i\frac{\sqrt{3}}{2}$, or equivalently introduces $\sqrt{3}$; the second introduces $\sqrt{D}$, so that now we have $\sqrt{-3D}$; and the third introduces the cube root needed to get the A and B (Lagrange resolvents) used in Cardano's formulas.

Galois root extensions. We now justify the lingering fact from last time, used in the proof that solvable by radicals implies solvable Galois group, that any root extension K/F (i.e. one obtained via a sequence of simple radical extensions) can be turned into one that is Galois with each intermediate extension cyclic. Thus, defining "solvable by radicals" using only such Galois extensions loses nothing.

Suppose that we have a root extension

$$F \subseteq K_1 \subseteq \ldots \subseteq K_{t-1} \subseteq K.$$

The first step is to produce a Galois root extension. Pick a Galois closure L of K/F, and take $\sigma_\ell \in \mathrm{Gal}(L/F)$. Applying σ_ℓ to each term in our tower, and noting that σ_ℓ fixes F, gives

$$F \subseteq \sigma_\ell(K_1) \subseteq \ldots \subseteq \sigma_\ell(K_{t-1}) \subseteq \sigma_\ell(K).$$

Each extension here is still simple radical: if K_i is obtained from K_{i-1} by adjoining $\sqrt[n_i]{a_i}$, then $\sigma_\ell(K_i)$ is obtained from $\sigma_\ell(K_{i-1})$ by adjoining $\sigma(\sqrt[n_i]{a_i}) = \sqrt[n_i]{\sigma_\ell(a_i)}$, where $\sqrt[n_i]{\sigma_\ell(a_i)}$ is a root of the polynomial $x^n - \sigma_\ell(a_i)$ obtained by applying σ_ℓ to $x^n - a_i$. Now, take the composite of all such extensions $\sigma_\ell(K)$:

$$\sigma_1(K) \cdots \sigma_n(K).$$

This composite is still a root extension, as we showed last time, and now it is also Galois over F since it is invariant under the action of $\mathrm{Gal}(L/F)$: multiplication by any σ permutes the σ_ℓ, so

$$\sigma(\sigma_1(K) \cdots \sigma_n(K)) = \sigma_1(K) \cdots \sigma_n(K).$$

(Recall when proving the "Galois $\iff$ normal subgroup" part of the Fundamental Theorem of Galois Theory that we showed Galois is equivalent to being invariant under embeddings, which here we can take to be embeddings into the Galois closure L, which all extend to elements of $\mathrm{Gal}(L/F)$.) Thus this composite is a Galois root extension of F as desired.

Let us thus just assume that our original root extension K/F was Galois. The second step is to obtain cyclic intermediate extensions. Since each intermediate extension is simple radical, we get cyclic extensions as soon as we have the required roots of unity in our base field. Suppose that $n_1, \ldots, n_t$ are the orders of the roots needed in our root extension tower (i.e. the n_i in $K_{i-1}(\sqrt[n_i]{a_i})$), and take ζ to be a primitive $n_1 \cdots n_t$-th root of unity. (Then all n_i-th roots of unity are also n-th roots of unity.) We then adjoin ζ to the fields in our tower to get:

$$F(\zeta) \subseteq K_1(\zeta) \subseteq \ldots \subseteq K_{t-1}(\zeta) \subseteq K(\zeta).$$

Each extension here is still simple radical, obtained by adjoining the same $\sqrt[n_i]{a_i}$ as before, and since the base field $F(\zeta)$ now contains all appropriate roots of unity, each extension here is cyclic.

So, we have a Galois root extension $K(\zeta)/F(\zeta)$ with intermediate cyclic extensions. But, we want a Galois root extension of F, so we simply tack on F at the start:

$$F \subseteq F(\zeta) \subseteq K_1(\zeta) \subseteq \ldots \subseteq K_{t-1}(\zeta) \subseteq K(\zeta).$$

We already know that each intermediate extension here is cyclic, *except* for the initial one $F \subseteq F(\zeta)$. This does not have to be cyclic, but the point is that it can be broken down further:

$$F \subseteq F_1 \subseteq F_2 \subseteq \ldots \subseteq F(\zeta)$$

into ones which *are* cyclic. This will come from the fact that the Galois group of $F(\zeta)$ over F is abelian, which we will prove next time. Once we have this, we use the fact that abelian groups are solvable to come up with a chain

$$\mathrm{Gal}(F(\zeta)/F) \trianglerighteq \ldots \trianglerighteq G_1 \trianglerighteq 1$$

with cyclic quotients, and then take fixed fields to get the intermediate fields F_i we want. Putting this entire extended tower between F and $F(\zeta)$ in

$$F \subseteq F(\zeta) \subseteq K_1(\zeta) \subseteq \ldots \subseteq K_{t-1}(\zeta) \subseteq K(\zeta)$$

then gives a Galois root extension of F with intermediate cyclic extensions, as desired.

Solvable implies solvable by radicals. We finally prove, modulo one detail we will get to next time, that if the Galois group of a polynomial is solvable, then the polynomial is solvable by radicals. Suppose $p(x)$ is our polynomial over F and K is its splitting field. Take a solvability chain for $\mathrm{Gal}(K/F)$:

$$\mathrm{Gal}(K/F) \trianglerighteq \ldots \trianglerighteq G_1 \trianglerighteq 1$$

with each quotient cyclic. Then take fixed fields to get a tower

$$F \subseteq K_1 \subseteq \ldots \subseteq K$$

with each intermediate extension cyclc. We want to get a tower with *radical* intermediate extensions, so we adjoin an appropriate primitive root of unity ζ (using the same $n_1 \cdots n_t$ root of unity as above, where the n_i now are the orders of the intermediate cyclic Galois groups), to get

$$F(\zeta) \subseteq K_1(\zeta) \subseteq \ldots \subseteq K(\zeta).$$

The fact we will prove next time is that each intermediate extension $K_i(\zeta)/K_{i-1}(\zeta)$ is still cyclic. (We also mentioned this fact earlier when discussing the "real" way to derive Cardano's formula.) Taking this for granted for now, we now have that each cyclic extension $K_i(\zeta)/K_{i-1}(\zeta)$ is a radical extension since the base field $F(\zeta)$ contains the required roots of unity, so the tower above is a root extension of $F(\zeta)$.

Finally, as with Cardano's formulas, we tack on F at the start:

$$F \subseteq F(\zeta) \subseteq K_1(\zeta) \subseteq \ldots \subseteq K(\zeta),$$

and note that the first extension $F \subseteq F(\zeta)$ is *already* a simple radical extension since it is obtained by adjoining a root ζ of 1. Thus, $K(\zeta)$ is a root extension of F that contains all roots of $p(x)$, so $p(x)$ is solvable by radicals.

Lecture 25: More on Solvability

Warm-Up 1. Suppose F is a field of characteristic zero and that ζ is a primitive n-th root of unity. We prove that $F(\zeta)$ is an abelian extension of F, which verifies a claim we made last time. This showed up in the proof that arbitrary root extensions can always be replaced by ones which are Galois and have cyclic intermediate extensions, namely in justifying the fact that the first step $F \subseteq F(\zeta)$ of the resulting tower could be enlarged to one with cyclic intermediate extensions. Once we know that $\mathrm{Gal}(F(\zeta)/F)$ is abelian, hence solvable, we can take a solvability chain

$$\mathrm{Gal}(F(\zeta)/F) \trianglerighteq \ldots \trianglerighteq 1$$

with cyclic quotients, and then take fixed fields to get the enlargement

$$F \subseteq F_1 \subseteq \ldots \subseteq F(\zeta)$$

we want with cyclic intermediate extensions. Note that $F(\zeta)$ is Galois over F since we can view it as the splitting field of $x^n - 1 \in F[x]$.

To see that $\mathrm{Gal}(F(\zeta)/F)$ is abelian, we note (as we did in the case $F = \mathbb{Q}$) that any element σ of the Galois group is determined by its action on ζ, and that $\sigma(\zeta)$ has to be some other primitive n-th root of unity, so that $\sigma(\zeta) = \zeta^a$ for some $(n, a) = 1$. This then gives a map

$$\mathrm{Gal}(F(\zeta)/F) \to (\mathbb{Z}/n\mathbb{Z})^\times \text{ defined by } \sigma \mapsto \text{the } a \text{ such that } \sigma(\zeta) = \zeta^a.$$

This map is a homomorphism by the same argument we gave in the case where $F = \mathbb{Q}$. Moreover, it is injective since if $\sigma \mapsto 1$, then $\sigma(\zeta) = \zeta$, which forces σ to be the identity on all of $F(\zeta)$, so that the map above has trivial kernel. Thus $\mathrm{Gal}(F(\zeta)/F)$ is isomorphic to a subgroup of $(\mathbb{Z}/n\mathbb{Z})^\times$, and so is in particular abelian.

This is all we need to get a tower from F to $F(\zeta)$ with cyclic intermediate extensions, but let us comment on how this general case is different from $F = \mathbb{Q}$. In the case of $\mathrm{Gal}(\mathbb{Q}(\zeta)/\mathbb{Q})$, we actually got the full multiplicative group $(\mathbb{Z}/n\mathbb{Z})^\times$ as the Galois group, but this is not necessarily the case for general F. The issue is that, although with $\mathbb{Q}$ the map $\mathrm{Gal}(\mathbb{Q}(\zeta)/\mathbb{Q}) \to (\mathbb{Z}/n\mathbb{Z})^\times$ is surjective since any $a \in (\mathbb{Z}/n\mathbb{Z})^\times$ gives rise to a Galois group element via $\zeta \mapsto \zeta^a$, the requirement that elements of $\mathrm{Gal}(F(\zeta)/F)$ *fix* elements of F in general places restrictions on the values of a we can actually get. For example, consider $\mathbb{R}(\zeta_7)$ over $\mathbb{R}$. Then $\mathrm{Gal}(\mathbb{R}(\zeta_7)/\mathbb{R})$ is a subgroup of $(\mathbb{Z}/7\mathbb{Z})^\times$. But for $2 \in (\mathbb{Z}/7\mathbb{Z})^\times$, we claim that $\zeta \mapsto \zeta^2$ does not give a valid element of the Galois group. Indeed, we have

$$\zeta = \cos(\tfrac{2\pi}{7}) + i\sin(\tfrac{2\pi}{7}) \quad \text{and} \quad \zeta^2 = \cos(\tfrac{4\pi}{7}) + i\sin(\tfrac{4\pi}{7}).$$

But $\cos(2\pi/7), \sin(2\pi/7) \in \mathbb{R}$ are in the base field, so any Galois group element sends these two to themselves. Also, $i\sin(2\pi/7) = \zeta - \cos(2\pi/7)$ is in $\mathbb{R}(\zeta_7)$, and thus so is $i = i\sin(2\pi/7)/\sin(2\pi/7)$, and i must be sent to a root of x^2+1 under the Galois group, so $i \mapsto \pm i$. Altogether this gives

$$\zeta = \cos(2\pi/7) + i\sin(2\pi/7) \mapsto \cos(2\pi/7) \pm i\sin(2\pi/7),$$

which cannot equal ζ^2. (A more succinct way of saying this is that $\mathbb{R}(\zeta_7)$ is actually $\mathbb{C}$ because $\mathbb{C}$ is the only non-trivial algebraic extension of $\mathbb{R}$ since $\mathbb{C}$ is algebraically closed, and the only nontrivial element of $\mathrm{Gal}(\mathbb{C}/\mathbb{R})$ is complex conjugation, and so cannot be all of $(\mathbb{Z}/n\mathbb{Z})^\times$ when $n \geq 5$.) In the case where $F = \mathbb{Q}$ this does not apply because Galois group elements do not have to fix $\cos(2\pi/7)$ and $\sin(2\pi/7)$.

Warm-Up 2. We show that if K/F is Galois, then for any α we have that $K(\alpha)$ is Galois over $F(\alpha)$ and that

$$\mathrm{Gal}(K(\alpha)/F(\alpha)) \cong \mathrm{Gal}(K/K \cap F(\alpha)).$$

In particular, and what is relevant to what we did last time, if $\mathrm{Gal}(K/F)$ is cyclic, this implies that $\mathrm{Gal}(K(\alpha)/F(\alpha))$ is cyclic since $\mathrm{Gal}(K/K\cap F(\alpha))$ is a subgroup of $\mathrm{Gal}(K/F)$ via the tower

$$F \subseteq K \cap F(\alpha) \subseteq K.$$

This was used in the proof that solvable Galois group implies solvable by radicals, where after adjoining an appropriate root of unity to a given tower we needed to know that doing so did not affect the fact that we had intermediate extensions which were cyclic.

First, if K is the splitting field of $p(x)$ over F, then $K(\alpha)$ is the splitting field of $p(x)(x-\alpha)$ (we can remove any repeated factors of $x-\alpha$ if need be in order to get something separable) over $F(\alpha)$, so that $K(\alpha)$ is Galois over $F(\alpha)$. We have a homomorphism $\mathrm{Gal}(K(\alpha)/F(\alpha)) \to \mathrm{Gal}(K/K\cap F(\alpha))$ given by restriction: $\sigma \mapsto \sigma|_K$. If $\sigma|_K$ is the identity on K, then since σ is also meant to fix $F(\alpha)$ (the base field of $K(\alpha)/F(\alpha)$), we have that σ fixes anything of the form

$$\frac{b_0 + b_1\alpha + \cdots + b_n\alpha^n}{c_0 + c_1\alpha + \cdots + c_m\alpha^m},$$

which make up all the elements of $K(\alpha)$. This shows that the kernel of the restriction map above is trivial, so it is injective. Moreover, given $\tau \in \mathrm{Gal}(K/K\cap F(\alpha))$, we can construct a corresponding

$\sigma \in \text{Gal}(K(\alpha)/F(\alpha))$ which restricts to τ simply by applying τ to elements of K:

$$\frac{b_0 + b_1\alpha + \cdots + b_n\alpha^n}{c_0 + c_1\alpha + \cdots + c_m\alpha^m} \mapsto \frac{\tau(b_0) + \tau(b_1)\alpha + \cdots + \tau(b_n)\alpha^n}{\tau(c_0) + \tau(c_1)\alpha + \cdots + \tau(c_m)\alpha^m}.$$

This says that $\text{Gal}(K(\alpha)/F(\alpha)) \to \text{Gal}(K/K \cap F(\alpha))$ is surjective, so it is an isomorphism. As explained above, this completes the proof that a polynomial is solvable by radicals if and only if its Galois group is solvable.

Back to the quartic formula. When first introducing solvability, we briefly discussed the process of deriving the quartic formula (for the roots of a quartic) from Cardano's formulas from the "solvable by radicals" perspective. What we said back then wasn't *quite* completely true, because at no point did we adjoin roots of unity, which we've now seen is actually necessary. So, let us briefly revisit this from the correct point of view.

Suppose we have a quartic with Galois group S_4. Since S_4 is solvable, this quartic should be solvable by radicals. The chain that gives us solvability of S_4 is:

$$S_4 \trianglerighteq A_4 \trianglerighteq \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} \trianglerighteq \mathbb{Z}/2\mathbb{Z} \trianglerighteq 1,$$

where $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ is $\{(1), (12)(34), (13)(24), (14)(23)\}$. Taking fixed fields gives

$$F \subseteq F(\sqrt{D}) \subseteq K_1 \subseteq K_2 \subseteq K$$

where K is the full splitting field. The new step is that, in order to get simple radical extensions, we must adjoint an appropriate root of unity. In this case, the quotients in the chain for S_4 above are $\mathbb{Z}/2\mathbb{Z}$, $\mathbb{Z}/3\mathbb{Z}$, $\mathbb{Z}/2\mathbb{Z}$, and $\mathbb{Z}/2\mathbb{Z}$, so via the proof that solvable groups implies solvable by radicals, we should adjoin an n-th root of unity where $n = 2 \cdot 3 \cdot 2 \cdot 2$, given by the orders of the cyclic quotients. But actually, we can do a bit better here, since in fact the second roots of unity ± 1 are already present and don't need to be adjoined. The only new root of unity needed is ζ_3, so it is enough to adjoin this to get:

$$F \subseteq F(\zeta_3) \subseteq F(\sqrt{D}, \zeta_3) \subseteq K_1(\zeta_3) \subseteq K_2(\zeta_3) \subseteq K(\zeta_3)$$

with tacked on at the start. Each extension here is cyclic, since by the second Warm-Up the Galois group of each extension is a subgroup of the corresponding cyclic one in the original tower. (Some of these extensions might now be trivial: for example, originally $\text{Gal}(K_1/F(\sqrt{D})$ had order 3, so the new $\text{Gal}(K_1(\zeta_3)/F(\sqrt{D}, \zeta_3))$ either has order 3 or 1, and in the latter case $K_1(\zeta_3) = F(\sqrt{D}, \zeta_3)$. This is fine, since all that matters is that we still have cyclic extensions.) The existence of this tower then says that all roots of the quartic can be expressed in terms of elements of F, ζ_3, $\sqrt{D}$, some cube root to get to $K_1(\zeta_3)$, then two more square roots to get to $K_2(\zeta_3)$ and finally $K(\zeta_3)$. These radicals are precisely the ones showing up in the roots of the quartic based off of Cardano's formulas we outlined earlier.

Some more examples. Let us finish by looking at a few more solvable examples, and see what form we expect the roots to take. A problem on the fourth homework asked for the Galois group of $x^6 - 2$ over $\mathbb{Q}$, where the answer turns out to be D_{12}. This is solvable (recall that the smallest non-solvable group is A_5 of order 60, so all smaller groups are solvable), and in particular we have

$$D_{12} \trianglerighteq \langle r \rangle \trianglerighteq \langle r^3 \rangle \trianglelefteq 1$$

where r denotes the basic rotation of order 6. The quotients here are $\mathbb{Z}/2\mathbb{Z}$, $\mathbb{Z}/3\mathbb{Z}$, and $\mathbb{Z}/2\mathbb{Z}$. Let us thus adjoin primitive sixth $(6 = 2 \cdot 3)$ root of unity to the fixed fields, so that

$$\mathbb{Q} \subseteq K_1 \subseteq K_2 \subseteq K$$

becomes

$$\mathbb{Q} \subseteq \mathbb{Q}(\zeta_6) \subseteq K_1(\zeta_6) \subseteq K_2(\zeta_6) \subseteq K(\zeta_6).$$

The second extension is cyclic of degree dividing 2, the third cyclic of degree dividing 3, and the final one cyclic of degree dividing 2, so at worst (when the degrees are as large as possible) we can express our roots using $\mathbb{Q}$, ζ_6, a square root, a cube root, and another square root. But for this polynomial we already know what the roots are:

$$\sqrt[6]{2},\ \zeta_6\sqrt[6]{2},\ \zeta_6^2\sqrt[6]{2},\ \zeta_6^3\sqrt[6]{2},\ \zeta_6^4\sqrt[6]{2},\ \zeta_6^5\sqrt[6]{2}.$$

So it's not so much that we need solvability to actually find the roots, but more to notice that the roots we have match what we expect: we need ζ_6, and the sixth root $\sqrt[6]{2} = \sqrt[3]{\sqrt{2}}$ we need can be interpreted as a cube root of a square root. (It seems that the remaining square root isn't actually needed! This solvability idea is not meant to give the most "efficient" way of expressing the roots, just a way that will work.)

Next, the polynomial $x^5 - 2$ over $\mathbb{Q}$ is also one that was covered on the homework, as the $p = 5$ case of $x^p - 2$ for p prime in general. The Galois group turned out to be of order $5 \cdot 4 = 20$, realizable as either the group of 2×2 matrices

$$\begin{bmatrix} a & b \\ 0 & 1 \end{bmatrix} \text{ with } a \in \mathbb{F}_5^\times \text{ and } b \in \mathbb{F}_5,$$

or as the semi-direct product $\mathbb{Z}/5\mathbb{Z} \rtimes \mathbb{Z}/5\mathbb{Z}^\times$ given the action of $\mathbb{Z}/5\mathbb{Z}^\times$ on $\mathbb{Z}/5\mathbb{Z}$ by multiplication. (This group is called the *Frobenius group* of order 20.) This group is solvable, since

$$\mathbb{Z}/5\mathbb{Z} \rtimes \mathbb{Z}/5\mathbb{Z}^\times \trianglerighteq \mathbb{Z}/5\mathbb{Z} \rtimes 1$$

has quotients $\mathbb{Z}/5\mathbb{Z}^\times \cong \mathbb{Z}/4\mathbb{Z}$ and $\mathbb{Z}/5\mathbb{Z}$. Thus we can adjoin a primitive 20-th root of unity to the fixed fields to get

$$\mathbb{Q} \subseteq \mathbb{Q}(\zeta_{20}) \subseteq K_1(\zeta_{20}) \subseteq K(\zeta_{20}).$$

The second extension is cyclic with order dividing 4 (it could be 2!) and the final cyclic of order dividing 5. Thus at worst we can express the roots of $x^5 - 2$ using $\mathbb{Q}$, ζ_{20}, a fourth root, and a fifth root. (Note that if the order of the second extension was in fact 2, we can deal with this with what we already have since squaring a fourth root gives a square root: $(\sqrt[4]{a})^2 = \sqrt{a}$.) Now again, we do in fact know what the roots are in this case: $\zeta_5^i\sqrt[5]{2}$. So, we don't actually need a fourth (nor square) root after all, and the ζ_5 comes from ζ_{20}^4.

Finally, consider $x^5 - 5x + 12$ over $\mathbb{Q}$. This has Galois group D_{10}. (We won't justify this here, but will look at some similar examples on the homework and at the start of next time.) We have

$$D_{10} \trianglerighteq \mathbb{Z}/5\mathbb{Z} \trianglerighteq 1,$$

where $\mathbb{Z}/5\mathbb{Z}$ is the subgroup of rotations, with successive quotients $\mathbb{Z}/2\mathbb{Z}$ and $\mathbb{Z}/5\mathbb{Z}$. Thus we need only adjoin a fifth root of unity to the fixed fields:

$$\mathbb{Q} \subseteq \mathbb{Q}(\zeta_5) \subseteq K_1(\zeta_5) \subseteq K(\zeta_5).$$

The roots should thus be expressible in terms of $\mathbb{Q}$, ζ_5, a square root, and a fifth root. Now, I don't know what the roots actually are here, and how to actually express them in this way, but we know for sure that it is possible to do so!

Lecture 26: Infinite Galois Theory

Warm-Up. We show that the Galois group of $x^5+20x+16$ over $\mathbb{Q}$ is A_5. Doing so will require the following new result: if the prime p does not divide the discriminant of $f(x)$ (monic and separable with integer coefficients), then the Galois group of $f(x)$ mod p over $\mathbb{F}_p$ is isomorphic to a subgroup of the Galois group of $f(x)$ over $\mathbb{Q}$, and moreover if $f(x)$ mod p factors into irreducible polynomials of degrees $n_1,\ldots,n_k$, then the Galois group of $f(x)$ over $\mathbb{Q}$ contains an element of cycle type $(n_1,\ldots,n_k)$. This is known as *Dedekind's theorem*, and we will say a bit about it and its proof afterwards. The point is that we can deduce information about the Galois group of $f(x)$ over $\mathbb{Q}$, and the types of elements it can contain, by reducing $f(x)$ mod various primes.

We will rely on WolframAlpha to do our computations mod p for us. First, the discriminant of $f(x)=x^5+20x+16$ (via WolframAlpha) is $2^{16}5^6 = 1.024$ billion, which is a square in $\mathbb{Q}$ and so we know that the Galois group of $f(x)$ is at least a subgroup of A_5. Now, reducing mod 3 and factoring into irreducibles (WolframAlpha) gives

$$x^5+20x+16 \equiv x^5+2x+1 \text{ mod } 3,$$

so that $f(x)$ remains irreducible over $\mathbb{F}_3$. The splitting field of this reduction is thus $\mathbb{F}_{3^5}$, so that the Galois group of the reduction is $\mathrm{Gal}(\mathbb{F}_{3^5}/\mathbb{F}_3)\cong\mathbb{Z}/5\mathbb{Z}$. (Recall that all Galois groups of finite fields are cyclic and generated by Frobenius.) Thus the Galois group of $f(x)$ over $\mathbb{Q}$ contains a copy of $\mathbb{Z}/5\mathbb{Z}$, and so contains a 5-cycle. The existence of a 5-cycle also comes from the second part of Dedekind's theorem, where based on the degree 5 of the factorization of $f(x)$ into irreducibles mod 3, the Galois group contains an element of cycle type (5), meaning a single 5-cycle.

Reducing instead mod 7 give

$$x^5+20x+16 \equiv (x+2)(x+3)(x^3+2x^2+5x+5) \text{ mod } 7,$$

so the Galois group of $f(x)$ over $\mathbb{Q}$ contains an element of cycle type $(1,1,3)$, which means a 3-cycle. (Alternatively, the splitting field of $f(x)$ mod 7 is $\mathbb{F}_{3^3}$, so the reduced Galois group is $\mathrm{Gal}(\mathbb{F}_{3^3}/\mathbb{F}_3)\cong\mathbb{Z}/3\mathbb{Z}$. This is then a subgroup of our desired Galois group, so the Galois group contains an element of order 3, which is necessarily a 3-cycle.) Actually, knowing now that the Galois group over $\mathbb{Q}$ contains a 5-cycle and a 3-cycle is enough to deduce that it is A_5, but let us see what happens for a few more primes anyway. Reducing mod 11 also gives a factorization into two irreducible linear terms and one cubic, which also gives an element of cycle type $(1,1,3)$, so nothing new. Reducing mod 13 gives an irreducible quintic, which gives a 5-cycle, so nothing new again. If we keep trying more primes we eventually find that mod 23 we get

$$x^5+20x+16 \equiv (x+17)(x^2+12x+14)(x^2+17x+2) \text{ mod } 23,$$

so our Galois group contains an element of cycle type $(1,2,2)$, which is the product of two 2-cycles, consistent with the expectation that our group should be A_5. We could keep going, trying more primes and deducing the cycle types of more elements, but this is not necessary in this example.

Since our Galois group G contains a 3-cycle and a 5-cycle, of orders 3 and 5 respectively, its order must be divisible by both 3 and 5 and hence by 15. Thus $|G|\geq 15$, so the index of G inside A_5 is at most 4:

$$[A_5:G]=\frac{|A_5|}{|G|}\leq\frac{60}{15}=4.$$

The action of A_5 on the cosets of G then gives a homomorphism from $A_5\to S_n$, where n is $1,2,3,$ or 4. Since A_5 has more elements than S_n in these cases, this map is not injective, so it has a non-trivial kernel. But A_5 is simple, so this kernel must be all of A_5, and hence the action of A_5 on the cosets of G is trivial. This means that $\sigma G = G$ for all $\sigma \in A_5$, which requires that $\sigma \in G$, and hence we conclude that $G = A_5$ as claimed.

Dedekind's theorem. The new result used above is an important tool in the determination of Galois groups over $\mathbb{Q}$. Such new tools are needed because, as it is hopefully becoming clear, computing Galois groups in the quintic and higher degree cases is not as straightforward as in the cubic and quartic cases. There is no nice algorithm that works for all quintic polynomials, and methods become more ad-hoc. Even Dedekind's theorem can't give us all the answers, since it is often the case that knowing cycle types of elements is still not enough to determine the group precisely. So, we will not dwell on computing Galois groups of polynomials beyond what we've already done, and will take the "solvable by radicals $\iff$ solvable Galois group" fact as our final main result. This is fitting, since this addresses the main motivation for group theory we proposed back on the first day of the fall quarter. In our remaining time we will instead push Galois theory in a different direction, and highlight some key uses elsewhere.

But, we should say a bit about the proof of Dedekind's theorem anyway for the sake of completeness. The idea of reducing mod different primes, one-at-a-time, is a cornerstone of modern number theory, and indeed it is number theory that provides the proper context behind Dedekind's theorem, where the key notion is that of a *decomposition group*. We can nonetheless give a proof of Dedekind's theorem that avoids the full machinery of number theory, and this is what the final set of discussion problems will be concerned with.

Here is the basic idea. Take $K := \mathbb{Q}(\alpha_1, \dots, \alpha_n)$ to be the splitting field of our polynomial, with $\alpha_1, \dots, \alpha_n$ the roots. Then we can find a prime ideal P of the ring $\mathbb{Z}[\alpha_1, \dots, \alpha_n]$ which "lifts" p in the sense that $P \cap \mathbb{Z} = p\mathbb{Z}$. It is this prime ideal that allows us to turn results over $\mathbb{F}_p$ into results over $\mathbb{Q}$. Define the group D_P to be the set of elements of the Galois group $\mathrm{Gal}(K/\mathbb{Q})$ which fix P:

$$D_P := \{\sigma \in \mathrm{Gal}(K/\mathbb{Q}) \mid \sigma(P) = P\}.$$

This is a subgroup of $\mathrm{Gal}(K/\mathbb{Q})$, called the *decomposition subgroup* at P. It is this subgroup which is isomorphic to the Galois group of the reduction of our polynomial mod p, and which will contain the element (something analogous to a "Frobenius" element) of the desired cycle type. (See the final discussion problems for the details behind all this.) In the end, Dedekind's theorem is really a result about lifting primes, which is a big deal in number theory.

Infinite Galois extensions. For the end of our course, we will delve into the topic of *infinite Galois theory*, which is concerned with infinite field extensions. Most of the theory mimics the finite degree case we have been considering thus far pretty well, but with one important difference we will get to fairly soon. The ultimate goal for us to understand the Galois group of the algebraic closure of $\mathbb{Q}$, to the extent possible.

To start, we can define the notion of an infinite Galois extension in the same way as in the finite case, with a slight difference when it comes to the splitting field definition. We say that an infinite algebraic extension K is *Galois* over F if it is normal and separable over F, using the same definitions of normality and separability we had before. This is equivalent to saying that the fixed field of the automorphism group $\mathrm{Aut}(K/F)$ is F, and is also equivalent to saying that K is the splitting field of an *infinite* collection of separable polynomials over F. (We cannot require that it be the splitting field of a single polynomial or even a finite set only, since such splitting fields are necessarily of finite degree.) When K is Galois over F, the Galois group $\mathrm{Gal}(K/F)$ is just the group $\mathrm{Aut}(K/F)$ of automorphisms of K fixing F, as it was before. If F is any field, its algebraic closure $\overline{F}$ is an example of an infinite Galois extension, and the Galois group $\mathrm{Gal}(\overline{F}/F)$ is called the *absolute Galois group* of F.

The main question to ask is whether, as in the finite case, we have a bijective correspondence between intermediate subfields of K/F and subgroups of $\mathrm{Gal}(K/F)$. The answer in the infinite case is actually "no", but we can get around this and make the answer "yes" by slightly restricting the types of subgroups we consider, as we will discuss shortly.

Example. To see why the answer to the question above is "no", consider the following example. Take K to be the composite of all quadratic extensions of $\mathbb{Q}$ inside the algebraic closure $\overline{\mathbb{Q}}$. In fact, it is enough to consider only extensions of the form $\mathbb{Q}(\sqrt{\pm p})$ for p prime, since if $D = \pm p_1^{k_1} \cdots p_n^{k_n}$ then $\mathbb{Q}(\sqrt{D})$ is the composite of $\mathbb{Q}(\sqrt{\pm p_i})$ for those p_i with k_i odd. There are thus countably many such quadratic extensions, all contained in K.

For any $\sigma \in \mathrm{Gal}(K/\mathbb{Q})$, σ must permute the roots $x^2 \pm p$ for each prime p, so σ either fixes $\sqrt{\pm p}$ or sends it to its negative. This implies that σ^2 will fix each $\sqrt{\pm p}$, so that σ^2 is the identity. Hence every element of $\mathrm{Gal}(K/\mathbb{Q})$ has order dividing 2. From this one can show that $\mathrm{Gal}(K/\mathbb{Q})$ will have uncountably many subgroups of index 2. But the fixed field of such a subgroup is then quadratic over $\mathbb{Q}$, and we said above there are only countably many such quadratic extensions. Hence there can be no bijective correspondence between the countably many intermediate quadratic extensions $\mathbb{Q} \subseteq \mathbb{Q}(\sqrt{D}) \subseteq K$ and the uncountably many subgroups of $\mathrm{Gal}(K/\mathbb{Q})$ of index 2. There will in fact be many subgroups with the *same* fixed field, so that the subgroup is not recoverable from the field. This is indicative of what happens in general: the mapping from subgroups to fixed fields will not be injective, and the mapping from fields to subgroups will not be surjective.

Topology to the rescue. So, the statement of the Fundamental Theorem of Galois Theory does not hold as cleanly in the infinite case as it did in the finite case. To get an actual bijective correspondence that relates all the things we expect (normality, degrees, etc), we have to restrict the types of subgroups we consider. It turns out that there is a *topology* one can define on $\mathrm{Gal}(K/F)$ which fixes everything. In particular, there is a sense in which we can talk about open and closed subsets and subgroups of $\mathrm{Gal}(K/F)$. (No worries if you haven't seen "topology", "open", or "closed" elsewhere. The precise definitions will not be important for our purposes.)

The main result is that the Fundamental Theorem holds as we would expect as long as we only consider *closed* (in the topological sense) subgroups of the Galois group. It turns out that the fixed fields of these are precisely the intermediate extensions of K/F that are *finite* over F, and here is where we get the bijective correspondence we want:

$$\{F \subseteq E \subseteq K \mid E/F \text{ finite}\} \longleftrightarrow \{\text{closed proper subgroups of } \mathrm{Gal}(K/F)\}.$$

The (finite) degree of E/F corresponds to the (finite) index of $\mathrm{Aut}(K/E)$ in $\mathrm{Gal}(K/F)$, and E/F is Galois if and only if $\mathrm{Aut}(K/E)$ is normal in $\mathrm{Gal}(K/F)$, which case $\mathrm{Gal}(E/F)$ is the corresponding quotient of $\mathrm{Gal}(K/F)$ by $\mathrm{Aut}(K/E)$.

Limits. Ultimately, our goal is to understand not only finite intermediate extensions E/F, but the full infinite algebraic extension K/F. Note that any element $\alpha \in K$ does actually lie in a finite intermediate extension, since α being algebraic over F implies that $F(\alpha)$ is finite over F. Thus we can think of K as the union of all of its intermediate finite extensions.

On the group side, this suggests that it should be possible to characterize $\mathrm{Gal}(K/F)$ using only those finite quotients $\mathrm{Gal}(E/F)$ which correspond to finite intermediate extensions $F \subseteq E \subseteq K$. Indeed, we can certainly take any $\sigma \in \mathrm{Gal}(K/F)$ and restrict it to to get an element $\sigma|_E \in \mathrm{Gal}(E/F)$, and moreover these restrictions altogether should completely characterize σ itself, precisely because any element α on which σ can act is contained in a finite extension. Thus, we can describe σ by specifying an infinite collection of elements from the various finite Galois groups $\mathrm{Gal}(E/F)$ with E ranging over finite intermediate extensions:

$$\sigma \longleftrightarrow (\sigma|_E)_E.$$

(The object on the right is a "tuple" of elements indexed by the E's.) But these elements satisfy some compatibilities, since if one finite extension E is contained in another E', restricting the element $\sigma|_{E'}$ occurring at index E' to $E \subseteq E'$ will produce the element $\sigma|_E$ occurring at index E.

The resulting set of tuples is called the *inverse limit* of the finite groups $\mathrm{Gal}(E/F)$, and so the conclusion is that we can recover our infinite Galois group $\mathrm{Gal}(K/F)$ as a type of "limit" of its finite quotients:

$$\mathrm{Gal}(K/F) = \varprojlim \mathrm{Gal}(E/F).$$

We will not go into the definition of "inverse limit" (denoted by lim with an "inverse arrow" as above) in too much detail in general, and will instead focus on some concrete examples. The inverse limit is a subgroup of the direct product of all the $\mathrm{Gal}(E/F)$, where the elements in the limit are those which satisfy appropriate compatibilities. Thus we get, for example, that the absolute Galois group $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ of $\mathbb{Q}$ is the limit of the finite Galois groups $\mathrm{Gal}(E/\mathbb{Q})$ as E ranges over all finite extensions of $\mathbb{Q}$. This is actually an incredibly complicated group, but we will say something interesting about it next time.

The finite field case. To see a first example of the discussion above, consider the union K_q (or composite) of all extensions of $\mathbb{F}_p$ of the form $\mathbb{F}_{p^{q^n}}$ with q prime. So, we are looking at

$$\mathbb{F}_p \subseteq \mathbb{F}_{p^2} \subseteq \mathbb{F}_{p^4} \subseteq \mathbb{F}_{p^8} \subseteq \ldots \text{ when } q = 2,$$

or

$$\mathbb{F}_p \subseteq \mathbb{F}_{p^3} \subseteq \mathbb{F}_{p^9} \subseteq \mathbb{F}_{p^{27}} \subseteq \ldots \text{ when } q = 3,$$

for example. Each finite Galois group $\mathrm{Gal}(\mathbb{F}_{p^{q^n}}/\mathbb{F}_p) \cong \mathbb{Z}/q^n\mathbb{Z}$ is cyclic and generated by Frobenius, so we get that

$$\mathrm{Gal}(K_q/\mathbb{F}_p) = \varprojlim \mathbb{Z}/q^n\mathbb{Z}.$$

This inverse limit is actually one we briefly considered last quarter: it is the additive group of q-adic integers $\mathbb{Z}_q$! Indeed, we define the q-adics first using a power series approach, but a homework problem from back then described the q-adics as this inverse limit, only without using the phrase "inverse limit". This inverse limit consists of infinite tuples

$$(a_1, a_2, a_3, \ldots) \in \mathbb{Z}/q\mathbb{Z} \times \mathbb{Z}/q^2\mathbb{Z} \times \mathbb{Z}/q^3\mathbb{Z} \times \cdots,$$

where the desired "compatibility" is that reducing the j-th coordinate mod q^i produces the i-th coordinate:

$$a_j \equiv a_i \bmod q^i \text{ when } j > i.$$

(Check that old homework problem to see how this matches up with the "power series" approach to defining $\mathbb{Z}_q$.)

Going further, we argued earlier in this course that the algebraic closure of $\mathbb{F}_p$ could be characterized as the union of all finite extensions of $\mathbb{F}_p$: $\overline{\mathbb{F}} = \bigcup_n \mathbb{F}_{p^n}$. This is then then union/composite of all K_q above, and in fact it turns out that

$$\mathrm{Gal}(\overline{\mathbb{F}_p}/\mathbb{F}_p) \cong \prod_q \mathrm{Gal}(K_q/\mathbb{F}_p) \cong \prod_q \mathbb{Z}_q.$$

This (additive) group is called the *profinite completion* of $\mathbb{Z}$ and is usually denoted by $\widehat{\mathbb{Z}}$. Its elements can concretely be described as infinite tuples

$$(x_2, x_3, x_4, \ldots) \in \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z} \times \cdots = \prod_n \mathbb{Z}/n\mathbb{Z}$$

satisfying the compatibility that whenever d divides n, x_n should be congruent to x_d mod d. This is precisely the inverse limit of the finite groups $\mathbb{Z}/n\mathbb{Z}$:

$$\operatorname{Gal}(\overline{\mathbb{F}_p}, \mathbb{F}_p) \cong \widehat{\mathbb{Z}} = \varprojlim \mathbb{Z}/n\mathbb{Z}.$$

(In general, the *profinite completion* $\widehat{G}$ of a group G is the inverse limit of all its finite quotients G/N, as N ranges among the normal subgroups of finite index.)

Lecture 27: Absolute Galois Groups

Roots of unity. Let us consider another example of an infinite Galois extension. For p prime, let μ_{p^∞} denote the set of all p-power roots of unity, so

$$\mu_{p^\infty} = \mu_p \cup \mu_{p^2} \cup \mu_{p^3} \cup \ldots$$

where μ_n is the group of n-th roots of unity. The field $\mathbb{Q}(\mu_{p^\infty})$ generated by all p-power roots of unity is an infinite Galois extension of $\mathbb{Q}$ since it is the splitting field of the infinite collection of polynomials $\{x^{p^n} - 1\}_{n \in \mathbb{N}}$. The p-power cyclotomic fields $\mathbb{Q}(\zeta_{p^n})$ are finite intermediate extensions, and we get that the infinite Galois group of $\mathbb{Q}(\mu_{p^\infty})$ over $\mathbb{Q}$ is

$$\operatorname{Gal}(\mathbb{Q}(\mu_{p^\infty})/\mathbb{Q}) = \varprojlim \operatorname{Gal}(\mathbb{Q}(\zeta_{p^n})/\mathbb{Q}) = \varprojlim (\mathbb{Z}/p^n\mathbb{Z})^\times = \mathbb{Z}_p^\times,$$

where $\mathbb{Z}_p^\times$ is the multiplicative group of p-adic integers—i.e. the group of units of the ring $\mathbb{Z}_p$.

Now take the extension $\mathbb{Q}^{ab}$ of $\mathbb{Q}$ generated by *all* roots of unity, regardless of their degree. (The reason for the $\mathbb{Q}^{ab}$ notation will be explained below.) This extension contains all $\mathbb{Q}(\mu_{p^\infty})$ from above, and is in fact the composite of these since the n-th roots of unity for $n = p_1^{k_1} \cdots p_m^{k_m}$ are contained in $\mathbb{Q}(\zeta_{p_1^{k_1}}) \cdots \mathbb{Q}(\zeta_{p_m^{k_m}})$. This implies that

$$\operatorname{Gal}(\mathbb{Q}^{ab}/\mathbb{Q}) \cong \prod_p \operatorname{Gal}(\mathbb{Q}(\mu_{p^\infty})/\mathbb{Q}) = \prod_p \mathbb{Z}_p^\times,$$

which is then the group of units $\widehat{\mathbb{Z}}^\times$ of the ring of profinite integers $\widehat{\mathbb{Z}}$ we introduced last time.

(Side remark. We saw in the fall on a homework problem that the group of all roots of unity in $\mathbb{C}$ was isomorphic to the quotient $\mathbb{Q}/\mathbb{Z}$. It is true that the group of automorhpisms of $\mathbb{Q}/\mathbb{Z}$ is actually the same $\widehat{\mathbb{Z}}^\times$ as above, which is essentially another way of phrasing the result above about the Galois group of $\mathbb{Q}^{ab}$. (The ring of endomorphisms of $\mathbb{Q}/\mathbb{Z}$ is $\widehat{\mathbb{Z}}$.)

Abelianization. The field $\mathbb{Q}^{ab}$ generated by all roots of unity over $\mathbb{Q}$ is actually the maximal abelian extension of $\mathbb{Q}$, hence the notation $\mathbb{Q}^{ab}$ for this field. This is essentially the statement of the Kronecker-Weber Theorem we mentioned back when discussing cyclotomic extensions, which says that any abelian extension $\mathbb{Q}$ is contained in a cyclotomic extension. This one field then encodes all finite abelian extensions of $\mathbb{Q}$.

By the Fundamental Theorem of (Infinite) Galois Theory, we can rephrase saying that $\mathbb{Q}^{ab}$ is the maximal abelian extension of $\mathbb{Q}$ as saying that the Galois group $\operatorname{Gal}(\mathbb{Q}^{ab}/\mathbb{Q})$ is the maximal abelian quotient of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$. The maximal abelian quotient of a group G is known as its *abelianization*, and is the quotient of G by its *commutator subgroup* $[G, G]$, which we briefly introduced in the fall on a discussion problem: $[G, G]$ is the subgroup of G generated by all comutators $xyx^{-1}y^{-1}$, and the quotient $G/[G, G]$ measures the extent to which G fails to be abelian, where G is abelian if and only if $[G, G]$ is trivial if and only if $G/[G, G] \cong G$. The upshot is that, although the absolute Galois group $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ is a fairly complicated group, is maximal abelian quotient is possible to describe explicitly as $\mathrm{Gal}(\mathbb{Q}^{ab}/\mathbb{Q}) = \widehat{\mathbb{Z}}^{\times}$. In general, if F is a finite extension of $\mathbb{Q}$, the abelianization of the absolute Galois group $\mathrm{Gal}(\overline{F}/F)$ can also be described explicitly in number-theoretic terms.

Absolute Galois group of $\mathbb{Q}$. The algebraic closure $\overline{\mathbb{Q}}$ of $\mathbb{Q}$ contains all finite extensions of $\mathbb{Q}$. Hence the absolute Galois group of $\mathbb{Q}$ also encodes all finite extensions of $\mathbb{Q}$; in particular, if K is finite over $\mathbb{Q}$, the Galois group of the Galois closure of $K/\mathbb{Q}$ occurs as a quotient of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$. Finite extensions of $\mathbb{Q}$, called *algebraic number fields*, are the central object of study in much of modern number theory and related areas of algebraic geometry. Thus, $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ is an important object of study in number theory and geometry as well.

(Speaking of geometry, let us just mention the following. Back in Lecture 15, we briefly spoke about a "Galois corrspondence" in topology using the notion of a fundamental group, that was in some sense analogous to the correspondence between fields and groups in Galois theory. We vaguely alluded to the idea that these two "Galois correspondences" can actually be viewed as literally the same from the correct point of view, and that there is a way to study Galois field extensions in a "topological" way via algebraic geometry. If we take $\mathbb{Q}$ and turn it into a "space" by using is spectrum—i.e. its set of prime ideals—as we briefly outlined last quarter, it turns out that the "fundamental group" of the spectrum of $\mathbb{Q}$ is actually $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$. This generalizes to other fields in place of $\mathbb{Q}$. We will say no more about his here, but it is truly a fascinating story.)

Fermat's Last Theorem. We will finish our course by giving two ways in which this absolute Galois group is useful for understanding number theoretic questions, both related to Fermat's Last Theorem. We will say will be quite abrupt and vague since we do not have the tools needed to make any of this really precise, but that's ok since our goal is merely to illustrate the types of problems where this Galois group pops up. First let us give some context. If F is a finite extension of $\mathbb{Q}$, which is meant to be viewed as a "souped-up" version of $\mathbb{Q}$, then, as we briefly mentioned earlier this quarter, F contains a "ring of integers", which is a "souped-up" version of $\mathbb{Z}$ inside F. For example, the ring of integers of the number field $\mathbb{Q}(\sqrt{2})$ is $\mathbb{Z}[\sqrt{2}]$; the ring of integers of $\mathbb{Q}(\sqrt[3]{2})$ is $\mathbb{Z}[\sqrt[3]{2}]$; and the ring of integers of the cyclotomic field $\mathbb{Q}(\zeta_n)$ is $\mathbb{Z}[\zeta_n]$.

We spoke last quarter when motivating the idea of "unique factorization" about how one can approach Fermat's Last Theorem from this perspective. To recall, the equation $x^n + y^n = z^n$ of Fermat's Last Theorem can be written as $x^n = z^n - y^n$, and using ζ_n we can factor the right side as

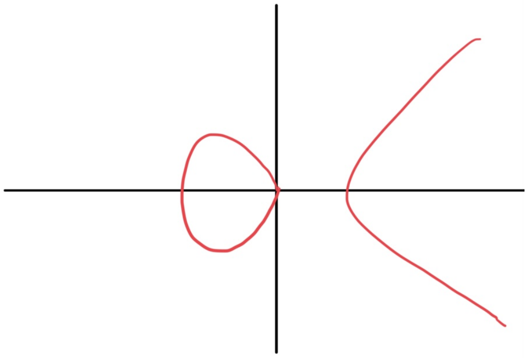
$$x^n = z^n - y^n = (z - y)(z - \zeta_n y)(z - \zeta_n^2 y) \cdots (z - \zeta_n^{n-1} y),$$

which is a factorization valid in the cyclotomic ring $\mathbb{Z}[\zeta_n]$. The idea is that since the left side is an n-th power, any prime occurring in the factorization of a potential value of x has to occur n times overall in the prime factorization of the entire right side, and by studying the possible prime factors of each $z - \zeta_n^i y$ one can try to show that this is not possible. But, as we said last quarter, this argument only works if we can compare prime factorizations on both sides, which requires something like unique factorization in $\mathbb{Z}[\zeta_n]$. Not all such rings are UFDs, so this approach to Fermat's Last Theorem does not work for all values of n.

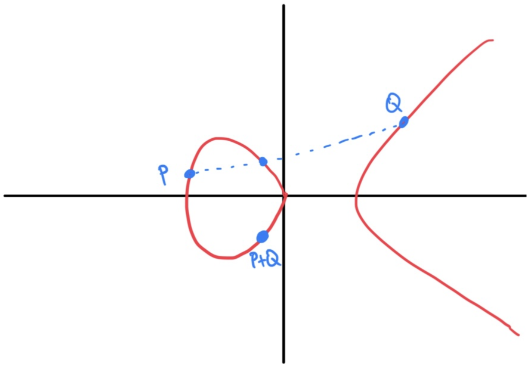
The problem of determing which $\mathbb{Z}[\zeta_n]$ are actually UFDs can be approached using the absolute Galois group $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$. One can construct a certain group made up out of the ideals of $\mathbb{Z}[\zeta_n]$, called the *ideal class group* of $\mathbb{Q}(\zeta_n)$, which essentially controls how close to being a UFD the ring $\mathbb{Z}[\zeta_n]$ is, since this class group is trivial if and only if $\mathbb{Z}[\zeta_n]$ is a UFD. (This group takes the product IJ of ideals I, J we defined last quarter and turns into an honest group operation by introducing a certain equivalence relation and taking equivalence classes.) It turns out that there is a natural action of the absolute Galois group $\text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ on this ideal class group, and that this action can be used to determine when the class group is in fact trivial. (The details are beyond our reach, but the idea comes from "decomposing" the ideal class group into something analogous to "eigenspaces" for this Galois action, and studying the resulting pieces.) Thus, the structure of $\text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ gives one way of understanding this approach to Fermat's Last Theorem.

Galois representations. In the cases of Fermat's Last Theroem where the approach above does not work, the absolute Galois group of $\mathbb{Q}$ is still essential to making progress. If you've ever read anything about the history of Fermat's Last Theorem (which you should as it is quite fascinating), you no doubt would have seen that one of the key ingredients which finally led to a proof in the 1990's was the notion of a *Galois representation*. This is nothing but a certain type of action of the absolute Galois group of $\mathbb{Q}$.

To give a sense for what a Galois representation is, we must say something about the subject of *elliptic curves*, which we will do via a single example. The curve defined by $y^2 = x(x-1)(x+1)$, an example of an elliptic curve, looks like:



Consider the following operation on the points of the curve. Take two points P and Q such as those drawn here:



The line passing through P and Q will intersect the curve at a third point, and we then denote the *reflection* of this point across the x-axis by $P+Q$. The reason why we use this additive notation for this resulting point is because this actually defines a *group* operation on the set of points of this curve! (This is one of properties that makes elliptic curves of great importance in number theory.

The identity of this group operation is actually not drawn in the picture above, as it is a "point at infinity" and should be visualized as being in some sense infinitely far away from all other points on the curve. The correct setting in which to view the curve then is in the context of *projective* geometry.)

With this group operation, we can talk about points of finite order on the curve, i.e. its *torsion* points. By looking at the p-tortion points (p prime), then the p^2-torsion points, p^3-torsion points, and so on we get a whole sequence of points on the curve. It turns out that there is a natural way in which $\text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ can act on this sequence of torsion points, and all of this data can then be arranged into a map

$$\text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to GL_2(\mathbb{Z}_p)$$

where $\mathbb{Z}_p$ is the ring of p-adic integers. (The GL_2 comes from viewing the action as acting on a 2-dimensional vector whose coordinates are elements of $\mathbb{Z}_p$, which itself arises by taking some kind of inverse limit of p^i-torsion points.) This map/action is what constitutes a *Galois represenation.*

The modern approach to Fermat's Last Theorem using Galois representations thus proceeds as follows: assume $x^n + y^n = z^n$ had a nontrivial solution; use it to construct a certain elliptic curve; use the elliptic curve to construct Galois representations of $\text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$; and finally use the structure of $\text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ to show that such Galois representations cannot actually exist, and thus the proposed solution of $x^+y^n = z^n$ does not exist either. This is by now far beyond the scope of our course and leads to incredibly deep and complicated mathematics, but hopefully gives a sense for how the infinite Galois group $\text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ shows up in important ways in mathematics. Thanks for reading!