FALL 2021: MATH 830 CLASS NOTES

D. KATZ

Lecture 1: Monday, August 23. We begin with a brief introduction to solving polynomial equations by radicals and then look at some examples that illustrate the interaction between groups and fields, which is the main theme of Galois theory. In the examples below, we will assume that all of our fields are contained in $\mathbb{C}$, the field of complex numbers.

Let us start with a degree two polynomial, $p(x) = ax^2 + bx + c$, with coefficients in a field, say $\mathbb{Q}$, the field of rational numbers. The quadratic formula says that the solutions to the equation $p(x) = 0$ are given by the expression

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

In other words, the roots of $p(x)$ can be expressed algebraically in terms of the coefficients of $p(x)$, using the operations: $+$, -, $\cdot$, $\div$ and $\sqrt{\ }$. More generally, if $p(x)$ is an arbitrary polynomial, our intuition tells us that to say $p(x)$ is solvable by radicals means that we can express the roots of $p(x)$ algebraically in terms of the coefficients of $p(x)$ using the operations $+$, -, $\cdot$, $\div$ together with the extraction of roots of various orders. It turns out that this is also possible for polynomials of degrees three and four, and in fact, in each of these cases, there exist universal formulas that work simultaneously for all polynomials of degree three or degree four. However, in the early 19th century, Abel showed that there exist degree five polynomials that *cannot* be solved by radicals. Thus, in particular, there cannot exist a universal formula for degree five polynomials expressing roots in terms of radicals. Later, by associating a group to a polynomial, Galois characterized when a polynomial was solvable by radicals in terms of a particular group property (which is also called *solvable.*). Using Galois' characterization, in principle, one can determine group theoretically whether or not a polynomial $p(x)$ is solvable by radicals.

The examples below give a small taste of the subject's flavor. In particular, we can see how to associate a group with a given extension obtained by adjoining roots of a polynomial.

Example 1.1. Consider the field extension $\mathbb{R} \subseteq \mathbb{C} = \{a + bi \mid a, b \in \mathbb{R}\}$. Note that $\mathbb{C}$ is obtained from $\mathbb{R}$ by adjoining a root, namely i, of the polynomial $f(x) = x^2 + 1$ to $\mathbb{R}$. We also note that, in the process of adjoining one root of $f(x)$, we have also adjoined the second root, namely $-i$.[1]

Now consider the map $\sigma : \mathbb{C} \to \mathbb{C}$ given by complex conjugation, i.e., $\sigma(\alpha) = \overline{\alpha}$, for all $\alpha \in \mathbb{C}$. Then σ satisfies the following familiar properties:

(i) $\sigma(\alpha) = \alpha$, for all $\alpha \in \mathbb{R}$.
(ii) $\sigma(\alpha + \beta) = \sigma(\alpha) + \sigma(\beta)$, for all $\alpha, \beta \in \mathbb{C}$.
(iii) $\sigma(\alpha\beta) = \sigma(\alpha)\sigma(\beta)$, for all $\alpha, \beta \in \mathbb{C}$.

Note that since $\overline{\overline{\alpha}} = \alpha$ for all $\alpha \in \mathbb{C}$, σ is also one-to-one and onto. We call σ an *automorphism of* $\mathbb{C}$ *fixing* $\mathbb{R}$. Now suppose τ were another automorphism of $\mathbb{C}$ fixing $\mathbb{R}$. Then:

$$0 = \tau(0) = \tau(i^2 + 1) = \tau(i^2) + \tau(1) = (\tau(i))^2 + 1.$$

In other words, $\tau(i)$ is a root of $x^2 + 1$, thus $\tau(i) = i$ or $\tau(i) = -i$. It follows that $\{id, \sigma\}$ are the only automorphisms of $\mathbb{C}$ fixing $\mathbb{R}$.[2] Moreover, since $\sigma^2 = id$, these automorphisms form a group.

[1] Of course, something even more remarkable is true. By adjoining i to $\mathbb{R}$, we actually obtain *all* of the roots to *all* of the polynomials with coefficients in $\mathbb{R}$. This is the Fundamental Theorem of Algebra, first rigorously proven by Gauss. We will see a proof of this theorem later in the semester.

[2] However, there are *uncountably* many automorphisms of $\mathbb{C}$ fixing $\mathbb{Q}$.

Putting this all together we have that the number of automorphisms of $\mathbb{C}$ fixing $\mathbb{R}$ equals the number of roots of x^2+1 in $\mathbb{C}$, which equals the *degree* of the field extension $\mathbb{R} \subseteq \mathbb{C}$, i.e., the dimension of $\mathbb{C}$ as a vector space over $\mathbb{R}$.

Example 1.2. This example is very similar to the previous one. Consider $x^2 - 2 \in \mathbb{Q}[x]$, which has the positive real root $\sqrt{2}$. Consider the set of real numbers $\mathbb{Q}(\sqrt{2}) := \{a + b\sqrt{2} \mid a, b \in \mathbb{Q}\}$. This set is clearly closed under addition and multiplication of real numbers. Moreover, if $0 \neq a + b\sqrt{2}$, then $\frac{1}{a^2+2b^2} \cdot (a - b\sqrt{2})$ is the multiplicative inverse of $a + b\sqrt{2}$. Thus, $\mathbb{Q}(\sqrt{2})$ is a field extension of $\mathbb{Q}$ having degree two. Moreover, $-\sqrt{2} \in \mathbb{Q}(\sqrt{2})$, so that both roots of $x^2 - 2$ belong to $\mathbb{Q}(\sqrt{2})$.

Let us check that $\sigma(a + b\sqrt{2}) = a - b\sqrt{2}$ is an automorphism of $\mathbb{Q}(\sqrt{2})$ fixing $\mathbb{Q}$. The map σ is clearly additive and fixes $\mathbb{Q}$. More over,

$$\sigma((a+b\sqrt{2})\cdot(c+d\sqrt{2})) = \sigma((ac+2bd)+(ad+bc)\sqrt{2})) = (ac+2db)-(ad+bc)\sqrt{2} = \sigma(a+b\sqrt{2})\cdot\sigma(c+d\sqrt{2}).$$

Finally, if we think of σ as a linear transformation from the vector space $\mathbb{Q}(\sqrt{2})$ over $\mathbb{Q}$ to itself, the matrix of σ with respect to the basis $\{1, \sqrt{2}\}$ is just $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, so σ is 1-1 and onto. Thus, σ is an automorphism of $\mathbb{Q}(\sqrt{2})$ fixing $\mathbb{Q}$. As in the previous example, if τ is an automorphism of $Q(\sqrt{2})$ fixing $\mathbb{Q}$, then $\tau(\sqrt{2})$ is also a root of $x^2 - 2$. Thus, $\tau(\sqrt{2}) = \pm\sqrt{2}$. In the first case, $\tau(a + b\sqrt{2}) = \tau(a) + \tau(b)\tau(\sqrt{2}) = a + b\sqrt{2}$, while in the second case, $\tau(a + b\sqrt{2}) = \tau(a) + \tau(b)\tau(\sqrt{2}) = a - b\sqrt{2}$. Thus, either $\tau = id$, the identity map, or $\tau = \sigma$. Since $\sigma^2 = id$, it follows that the group of automorphisms of $\mathbb{Q}(\sqrt{2})$ fixing $\mathbb{Q}$ has two elements (and thus is isomorphic to $\mathbb{Z}_2$), and so the number of elements in the Galois group of $x^2 - 2$ over $\mathbb{Q}$ equals the number of roots of $x^2 - 2$ in $\mathbb{Q}(\sqrt{2})$, which equals the degree of the extension $\mathbb{Q} \subseteq \mathbb{Q}(\sqrt{2})$.

Example 1.3. Let us now consider $f(x) = x^3 - 2 \in \mathbb{Q}[x]$. The graph of $y = x^3 - 2$ in the plane $\mathbb{R}^2$ shows that $f(x)$ has one real root, namely $\sqrt[3]{2}$. Set $\mathbb{Q}(\sqrt[3]{2}) := \{a + b\sqrt[3]{2} + c(\sqrt[3]{2})^2 \mid a, b, c \in \mathbb{Q}\}$. Then $\mathbb{Q}(\sqrt[3]{2})$ is a field containing $\sqrt[3]{2}$. It is not as easy to show that $\mathbb{Q}(\sqrt[3]{2})$ is a field as in the previous examples, but here is one way to see this. It is relatively easy to see that $\mathbb{Q}(\sqrt[3]{2})$ is closed under addition and multiplication. The crucial point is that non-zero elements have multiplicative inverses. For this, I leave it to you to check that $x^3 - 2$ is irreducible over $\mathbb{Q}$. Thus, $(x^3 - 2)\mathbb{Q}[x]$, the ideal in $\mathbb{Q}[x]$ generated by $x^3 - 2$, is a maximal ideal in the principle ideal domain $\mathbb{Q}[x]$. There is a a natural surjective ring homomorphism $\phi : \mathbb{Q}[x] \to \mathbb{Q}(\sqrt[3]{2})$, defined by $\phi(g(x)) = g(\sqrt[3]{2})$. Since $x^3 - 2$ belongs to the kernel of ϕ and generates a maximal ideal, this ideal must be the kernel of ϕ. Therefore, $\mathbb{Q}(\sqrt[3]{2}) \cong \mathbb{Q}[x]/(x^3 - 2)$ is a field. Another way to see that $\mathbb{Q}(\sqrt[3]{2})$ is a field would be to use the division algorithm (or more properly, the Euclidean algorithm). If $\alpha \in \mathbb{Q}(\sqrt[3]{2})$, $\alpha = g(\sqrt[3]{2})$, for some $g(x) \in \mathbb{Q}[x]$ having degree two or less. Thus, $g(x)$ and $x^3 - 2$ are relatively prime polynomials. By the Euclidean algorithm, there exist $a(x), b(x) \in \mathbb{Q}[x]$ such that $1 = a(x)g(x) + b(x)(x^3 - 2)$. Substituting $x = \sqrt[3]{2}$, we have $1 = a(\sqrt[3]{2})g(\sqrt[3]{2})$, which shows that $\alpha = g(\sqrt[3]{2})$ has a multiplicative inverse, and so $\mathbb{Q}(\sqrt[3]{2})$ is a field.

Now suppose that τ is an automorphism of $\mathbb{Q}(\sqrt[3]{2})$ fixing $\mathbb{Q}$. As before, we have

$$0 = \tau(0) = \tau((\sqrt[3]{2})^3 - 2) = (\tau(\sqrt[3]{2}))^3 - \tau(2) = (\tau(\sqrt[3]{2}))^3 - 2,$$

which shows that $\tau(\sqrt[3]{2}) \in \mathbb{Q}(\sqrt[3]{2})$ is a root of $x^3 - 2$. However, $x^3 - 2$ has just one real root, and therefore only one root in $\mathbb{Q}(\sqrt[3]{2})$. To see this, if we set $\epsilon := e^{\frac{2\pi i}{3}}$, then $\epsilon^3 = 1$. Moreover, $(\epsilon^2)^3 = 1$. It follows that the three roots of $x^3 - 2$ are: $\sqrt[3]{2}$, $\sqrt[3]{2}\epsilon$, $\sqrt[3]{2}\epsilon^2$. Since ϵ, ϵ^2 are not real numbers, the roots $\sqrt[3]{2}\epsilon$ and $\sqrt[3]{2}\epsilon^2$ are not real. It follows that the only automorphism of $\mathbb{Q}(\sqrt[3]{2})$ fixing $\mathbb{Q}$ is the identity automorphism. Thus, the Galois group of $x^3 - 2$ over $\mathbb{Q}$ is the identity, and its order equals the number of roots of $x^3 - 2$ in $\mathbb{Q}(\sqrt[3]{2})$. On the other hand, in this case, the order of the Galois group is *not* equal to the degree of the extension $\mathbb{Q} \subseteq \mathbb{Q}(\sqrt[3]{2})$.

To summarize: In each of the examples above, we started with an irreducible polynomial $p(x)$ with coefficients in a field F and we created a field extension $F \subseteq F(\alpha)$, where α is a root of $p(x)$. We then associated to this extension a finite group (the Galois group) which has the property that the order of this group equals the number of roots of $p(x)$ in the field $F(\alpha)$. Thus this group reflects an algebraic property of the extension. This fact holds in general, and is a key point in our development of Galois theory. Finally, extensions $F \subseteq F(\alpha)$, like Examples 1.1 and 1.2, with the property that the degree of the extension equals the order of the Galois group (i.e, *Galois extensions*) exhibit many more connections between the extension and the corresponding Galois group. These connections are elucidated by the Galois Correspondence Theorem, which is the main goal of the first part of the course.

Lecture 2: Wednesday, August 25. In this lecture we prove the following proposition which establishes, in general, the points mentioned in the summary of the last lecture.

Proposition 2.1. *Let F be a field and $f(x) \in F[x]$ an irreducible, monic polynomial of degree d. Let α be a root of $f(x)$, with $\alpha \in K$, a field containing F. Set $F(\alpha) := \{a_0 + a_1\alpha + \cdots + a_{d-1}\alpha^{d-1} \mid a_0, \ldots, a_{d-1} \in F\}$ and let $[F(\alpha) : F]$ denote the dimension of $F(\alpha)$ as a vector space over F. Then:*

(i) $F(\alpha)$ is a field. In fact, it is the smallest subfield of K containing F and α.
(ii) $[F(\alpha) : F] = d$.
(iii) Let $F \subseteq L$ be an extension of fields and $\sigma : F(\alpha) \to L$ a field homomorphism fixing F. Then $\sigma(\alpha)$ is a root of $f(x)$.
(iv) The number of field homomorphisms from $F(\alpha) \to L$ fixing F equals the number of roots of $f(x)$ in L. In particular, the number of automorphisms of $F(\alpha)$ fixing F equals the number of roots of $f(x)$ in $F(\alpha)$.
(v) The number of automorphisms of $F(\alpha)$ fixing F is less than or equal to $[F(\alpha) : F]$.

Before giving a proof of Proposition 2.1, we need a few preliminary remarks.

Remarks 2.2. (a) Let K_1 and K_2 be fields. The map $\sigma : K_1 \to K_2$ is a *field homomorphism* if for all $\alpha, \beta \in K_1$, $\sigma(\alpha + \beta) = \sigma(\alpha) + \sigma(\beta)$ and $\sigma(\alpha\beta) = \sigma(\alpha)\sigma(\beta)$. The following properties of σ are easily checked:

(i) $\sigma(0) = 0$ and $\sigma(1) = 1$.
(ii) $\sigma(-\alpha) = -\sigma(\alpha)$ and $\sigma(\alpha^{-1}) = \sigma(\alpha)^{-1}$, for all $\alpha \neq 0$ in K_1.

These properties have the following important consequence: Any field homomorphism σ is one-to-one. To see this, suppose $\sigma(\alpha) = 0$, with $\alpha \in K_1$. If $\alpha \neq 0$, then

$$1 = \sigma(1) = \sigma(\alpha\alpha^{-1}) = \sigma(\alpha)\sigma(\alpha^{-1}) = 0 \cdot \sigma(\alpha^{-1}) = 0,$$

a contradiction, so $\alpha = 0$. Since σ is a homomorphism, this shows σ is one-to-one.

(b) For a field F, we will need some standard properties of the polynomial ring $F[x]$ with coefficients in F. Almost all of the standard properties follow from the division algorithm: Give $f(x), g(x) \in F[x]$, there exist unique polynomials $h(x), r(x) \in F[x]$ such that $g(x) = f(x)h(x) + r(x)$, where $r(x)$ is either 0 (in which case $f(x)$ divides $g(x)$), or the degree of $r(x)$ is strictly less than the degree of $f(x)$. We will sketch the proof of this algorithm below, but here are some consequences that follow from it:

(i) The Euclidean algorithm to find greatest common divisors (GCDs). Given $f(x)$ and $g(x)$, a greatest common divisor is a polynomial $d(x) \in R[x]$ of largest degree such that $d(x)$ divides both $f(x)$ and $g(x)$. To find a GCD, we may assume the degree of $g(x)$ greater than or equal to the degree of $f(x)$ and write $g(x) = f(x)h(x) + r(x)$ according to the division algorithm. If $r(x) = 0$, then $f(x)$ divides $g(x)$ and $f(x)$ is a GCD. If not, apply the division algorithm again writing $f(x) = r_1(x)h_1(x)+r_2(x)$, where $r_1(x) := r(x)$ and the degree of $r_2(x)$ is strictly less that the degree of $r_1(x)$. If $r_2(x) = 0$, then $r_1(x)$ is a GCD of $f(x)$ and $r_1(x)$, and hence also a GCD of $f(x)$ and $g(x)$. Continuing this process, the last non-zero remainder is a GCD of $f(x)$ and $g(x)$. The proof of this is by induction on the degree of $g(x)$ using the fact that given an expression of the form $a = bc + d$, the pairs a, b and b, d have the same set of common divisors. Finally, the GCD of $f(x)$ and $g(x)$ is unique, if we take the greatest common divisor that is a monic polynomial.
(ii) If one uses backwards substitution in the equations generated by the Euclidean algorithm, one obtains *Bezout's Principle*: If $d(x)$ is the GCD of $f(x)$ and $g(x)$, then there exist $a(x), b(x)$ such that $d(x) = a(x)f(x) + b(x)g(x)$. In particular, if the GCD of $f(x)$ and $g(x)$ is 1, then there exist $a(x), b(x)$ such that $1 = a(x)f(x) + b(x)g(x)$
(iii) Every non-zero, non constant polynomial in $F[x]$ can be written uniquely (up to unit multiples) as a product of irreducible polynomials in $F[x]$.
(iv) $F[x]$ is a Principal Ideal Domain (PID).

.

Sketch of proof of the division algorithm. We may assume $f(x)$ does not divide $g(x)$, otherwise there is nothing to prove. If the degree of $g(x)$ is less than the degree of $f(x)$, then we can write $g(x) = f(x) \cdot 0 + g(x)$, which has the required form. Now suppose $g(x)$ has degree greater than or equal to the degree of $f(x)$. Write $g(x) = ax^m + g_0(x)$ and $f(x) = bx^n + f_0(x)$, where m is the degree of $g(x)$, n is the degree of $f(x)$ and $g_0(x)$ the lower degree terms of $g(x)$ and $f_0(x)$ the lower degree terms of $f(x)$. Then

$$g(x) = \frac{b}{a}x^{m-n}f(x) + (g_0(x) - \frac{b}{a}x^{m-n}f_0(x)),$$

where the degree of $g_0(x) - \frac{b}{a}x^{m-n}f_0(x)$ is strictly less than m, the degree of $g(x)$. If $m = n$, this gives the required expression. If $m > n$, then by induction, we can write $g_0(x) - \frac{b}{a}x^{m-n}f_0(x) = f(x)h(x) + r(x)$, with the degree of $r(x)$ less than n. Thus, $g(x) = (\frac{b}{a}x^{m-n} + h(x)) \cdot f(x) + r(x)$, as required.

Now, if $g(x) = f(x)h_1(x) + r_1(x) = f(x)h_2(x) + r_2(x)$, with both $r_i(x)$ having degree less than n, then

$$r_2(x) - r_1(x) = f(x)(h_1(x) - h_2(x)).$$

Thus $f(x)$ divides $r_2(x) - r_1(x)$. Since this latter polynomial has degree less than n, this can only happen if $r_2(x) - r_1(x) = 0$, i.e., $r_1(x) = r_2(x)$. But this implies $f(x)h_1(x) = f(x)h_2(x)$, so $h_1(x) = h_2(x)$. □

Proof of Proposition 2.1. For part (i), since $F(\alpha)$ is contained in the field K, it suffices to show that $F(\alpha)$ is closed under addition, multiplication and the existence of multiplicative inverses. $F(\alpha)$ is clearly closed under addition. Suppose $A, B \in F(\alpha)$. Then, by definition, there exist $A(x), B(x) \in F[x]$ each having degree less than d such that $A = A(\alpha)$ and $B = B(\alpha)$. Write $A(x)B(x) = f(x)h(x) + r(x)$, with the degree of $r(x)$ less than d. Then

$$A \cdot B = A(\alpha) \cdot B(\alpha) = f(\alpha)h(\alpha) + r(\alpha) = r(\alpha),$$

which shows that $A \cdot B \in F(\alpha)$. If $A \neq 0$, then $A^{-1} \in K$. We must see that A^{-1} is a polynomial expression in α having degree less than d, i.e., $A^{-1} \in F(\alpha)$. For this, we note that since $A(x)$ has degree less than d and $f(x)$ is irreducible, then the GCD of $A(x)$ and $f(x)$ must equal 1. Thus we can write $1 = b(x)A(x) + c(x)f(x)$, for $b(x), c(x) \in F[x]$. If $b(x)$ has degree less than d, we substitute $x = \alpha$ to obtain $1 = b(\alpha)A(\alpha)$, which shows that $b(\alpha)$, which belongs to $F(\alpha)$, is the multiplicative inverse of A. If $b(x)$ has degree greater than d, we can divide $b(x)$ by $f(x)$ to get a remainder $r(x)$, and then obtain an equation $1 = r(x)A(x) + d(x)f(x)$ for some $d(x) \in F[x]$. (Do you see how?) Substituting $x = \alpha$ shows that $r(\alpha) \in F(\alpha)$ is the multiplicative inverse of A. Thus, $F(\alpha)$ is a field. To see that it is the smallest subfield of K containing F and α, note that if L is a subfield of K containing $F$and α, then since L is closed under addition and multiplication, L contains all expressions of the form $a_0 + a_1\alpha + \cdots + a_{d-1}\alpha^{d-1}$, with each $a_j \in F$. Thus, L contains $F(\alpha)$, which gives what we want.

Lecture 3: Friday, August 27. We continue with the proof of Proposition 2.1.

For part (ii), clearly $1, \alpha, \ldots, \alpha^{d-1}$ span $F(\alpha)$ as a vector space over F. If these elements were not linearly independent over F, then some non-trivial linear combination of them would equal zero. Thus, there exists $t(x)$ in $F[x]$ such that the degree of $t(x)$ is less than d and $t(\alpha) = 0$. As in the previous paragraph, we may write $1 = a(x)t(x) + b(x)f(x)$, for some $a(x), b(x) \in F[x]$. Substituting $x = \alpha$ yields $1 = 0$, a contradiction. Thus, $1, \alpha, \ldots, , \alpha^{d-1}$ are linearly independent over F and therefore form a basis for $F(\alpha)$ over F. Hence, $[F(\alpha) : F] = d$.

For part (iii), suppose $\sigma : F(\alpha) \to L$ is a field homomorphism fixing F. If we write $f(x) = \sum_{i=0}^{d} a_{d-i}x^{d-i}$, with $a_d = 1$, then

$$0 = \sigma(0) = \sigma(\sum_{i=0}^{d} a_{d-i}\alpha^{d-i}) = \sum_{i=0}^{d} \sigma(a_{d-i})(\sigma(\alpha))^{d-i} = \sum_{i=0}^{d} a_{d-i}(\sigma(\alpha))^{d-i} = f(\sigma(\alpha)),$$

which shows that $\sigma(\alpha)$ is a root of $f(x)$ in L.

In light of part (iii) of the proposition, to prove part (iv), it suffices to prove the following statement. If $\beta \in L$ is a root of $f(x)$, then there exists a field homomorphism $\sigma : F(\alpha) \to L$ fixing F such that $\sigma(\alpha) = \beta$. It turns out that the obvious map works. If $A = a_0 + a_1\alpha + \cdots + a_{d-1}\alpha^{d-1}$, we define $\sigma : F(\alpha) \to L$ by

$\sigma(A) = a_0 + a_1\beta + \cdots + a_{d-1}\beta^{d-1}$. Clearly σ is additive. To see that σ respects multiplication, we use the notation from part (i), so that given $A, B \in F(\alpha)$, we have $A(x), B(x) \in F[x]$ such that $A = A(\alpha)$, $B = B(\alpha)$. We also have $A(x)B(x) = f(x)h(x) + r(x)$, with the degree of $r(x)$ less than d. Thus $AB = r(\alpha)$ as an element of $F(\alpha)$.

Now, on the one hand, $\sigma(AB) = \sigma(r(\alpha)) = r(\beta)$. On the other hand, $\sigma(A)\sigma(B) = A(\beta)B(\beta) = r(\beta)$, so $\sigma(AB) = \sigma(A)\sigma(B)$. Since σ clearly fixes F, σ is a field homomorphism fixing F that takes α to β. The second statement in part (iv) follows immediately from the first statement.

Finally, to see (v), since the number of automorphisms of $F(\alpha)$ fixing F equals the number of roots of $f(x)$ in $F(\alpha)$, the number of such automorphisms is less than or equal to d, the degree of $f(x)$. Since $d = [F(\alpha) : F]$, this finishes the proof. □

We now would like to see how the various ideas presented so far can be applied to more general field extensions $F \subseteq K$, rather than just simple extensions of the form $F \subseteq F(\alpha)$. We begin with some preliminary concepts.

Definition-Proposition 3.1. *Let $F \subseteq K$ be an extension of fields. For a subset $U \subseteq K$, we write $F(U)$ for the intersection of all subfield of K containing F and U. Then $F(U)$ is a subfield of K containing F and U. Moreover, $F(U)$ consists of all expressions of the form $p(u_1, \ldots, u_n)q(u_1, \ldots, u_n)^{-1}$, for some subset $u_1, \ldots, u_n \in U$ and polynomials $p(x_1, \ldots, x_n), q(x_1, \ldots, x_n)$ with coefficients in F and $q(u_1, \ldots, u_n) \neq 0$. We call $F(U)$ the subfield of K generated by U over F.*

Proof. It is easy to check that $F(U)$ is a subfield of K containing F. For example, if $0 \neq a \in F(U)$, then $a \in E$ for any subfield of K containing F and U. But then $a^{-1} \in E$. Since this holds for all E, $a^{-1} \in F(U)$. That $F(U)$ is closed under addition and multiplication follows in a similar manner.

Now, Let C be the set of all elements in K of the form $p(u_1, \ldots, u_n)q(u_1, \ldots, u_n)^{-1}$, for some subset $u_1, \ldots, u_n \in U$, for some $n \geq 1$, and polynomials $p(x_1, \ldots, x_n), q(x_1, \ldots, x_n)$ with coefficients in F and $q(u_1, \ldots, u_n) \neq 0$. We claim C is a a subfield of K containing F and U. Suppose this were true. Then, by definition, $F(U) \subseteq C$. On the other hand, since $U \subseteq F(U)$ and $F(U)$ is a field, any expression in C belongs to $F(U)$, so $C \subseteq F(U)$, and hence $C = F(U)$, as required.

C clearly contains F and U. To see that it is a subfield of K, it suffices to show that C is closed under addition, multiplication and inverses. Suppose $0 \neq a \in C$ and $a = p(u_1, \ldots, u_n)q(u_1, \ldots, u_n)^{-1}$, for some subset $u_1, \ldots, u_n \in U$ and polynomials $p(x_1, \ldots, x_n), q(x_1, \ldots, x_n)$ with coefficients in F and $q(u_1, \ldots, u_n) \neq 0$. Then $p(u_1, \ldots, u_n) \neq 0$, so $a^{-1} = q(u_1, \ldots, u_n)p(u_1, \ldots, u_n)^{-1}$ belongs to C. Suppose $b = c(u'_1, \ldots, u'_m)d(u'_1, \ldots, u'_m)^{-1} \in C$. Then clearly

$$ab = \{p(u_1, \ldots, u_n)c(u'_1, \ldots, u'_m)\}\{q(u_1, \ldots, u_n)d(u'_1, \ldots, u'_m)\}^{-1}$$

and

$$a + b = \{d(u'_1, \ldots, u'_m)p(u_1, \ldots, u_n) + q(u_1, \ldots, u_n)c(u'_1, \ldots, u'_m)\}\{q(u_1, \ldots, u_n)d(u'_1, \ldots, u'_m)\}^{-1}$$

belong to C, which completes the proof. □

Remark 3.2. Suppose that $U = \{\alpha\}$ and α is the root of an irreducible polynomial of degree d with coefficients in F. Let us note that $F(\alpha)$ as described in Proposition 2.1 is the same as $F(U)$ described in Proposition 3.1. To see this, $F(\alpha)$ as described in Proposition 2.1, is a subfield of K containing F and α, and thus contains $F(U)$ as described in Proposition 3.1. On the other hand, $F(U)$ is a field containing F and α and thus contains all expressions of the form $a_0 + a_1\alpha + \cdots + a_{d-1}\alpha^{d-1}$, with the $a_i \in F$. In other words, $F(\alpha) \subseteq F(U)$, which gives what we want.

Lecture 4: Monday, August 30.

We now come to one of the most important definitions in this part of the course.

Definition 4.1. Let $F \subseteq K$ be an extension of fields. An element $\alpha \in K$ is said to be algebraic over F if there exists $h(x) \in F[x]$ with $h(\alpha) = 0$. The extension $F \subseteq K$ is an algebraic extension if every element in K is algebraic over F. In this case we also say that K is algebraic over F.

Proposition 4.2. *Let $F \subseteq K$ be an extension of fields. Suppose $\alpha \in K$ is algebraic over F. Then:*

(i) There exists a unique monic, irreducible polynomial $f(x) \in F[x]$ of least degree satisfying $f(\alpha) = 0$.
(ii) $f(x)$ divides any $g(x) \in F[x]$ satisfying $g(\alpha) = 0$.
(iii) $[F(\alpha) : F] = \deg f(x)$.

Proof. For part (i), since α is algebraic over F the set of monic polynomials in $F[x]$ having α as a root is non-empty. Thus, by the Well Ordering Principle applied to the set of degrees of such polynomials, there exists a monic polynomial $f(x) \in F[x]$ of least degree with $f(\alpha) = 0$. Clearly $f(x)$is irreducible over F, otherwise, we could write $f(x) = a(x)b(x)$ with $a(x), b(x) \in F[x]$, with $a(x), , b(x)$ having degree less than the degree of $f(x)$. But then $0 = f(\alpha) = a(\alpha)b(\alpha)$, so that either $a(\alpha) = 0$ or $b(\alpha) = 0$, contradicting the choice of $f(x)$. Thus, $f(x)$ is irreducible over F. Set d to be the degree of $f(x)$.

For (ii), suppose $g(x) \in F[x]$ and $g(\alpha) = 0$. Then the degree of $g(x)$ is greater than or equal to d. Write $g(x) = f(x)h(x) + r(x)$, with the degree of $r(x)$ less than d. Setting $x = \alpha$, we see that $r(\alpha) = 0$, which contradicts the choice of d, unless $r(x) = 0$. Thus $f(x)$ divides $g(x)$. Now suppose further $g(x)$ is also monic of degree d. Then, on the one hand, since $f(x)$ and $g(x)$ have the same degree, $h(x)$ must be a constant. On the other hand, since $f(x)$ and $g(x)$ are monic, this constant must be 1. In other words, $f(x) = g(x)$, so that $f(x)$is the unique monic polynomial of least degree in $F[x]$ having α as a root. Finally, part (iii) follows immediately from Proposition 2.1.

Definition 4.3. For $\alpha \in K$ algebraic over F, the unique monic polynomial $f(x) \in F[x]$ of least degree satisfying $f(\alpha) = 0$ is called the minimal polynomial of α over F.

Given an extension of fields, $F \subseteq K$, the following theorem gives a criterion for an element of $\alpha \in K$ to be algebraic over F.

Theorem 4.4. *Let $F \subseteq K$ be an extension of fields. Suppose $\alpha, \beta \in K$. Then:*

(i) α is algebraic over F if and only if $[F(\alpha) : F] < \infty$.
(ii) If α and β are algebraic over F, then $\alpha \pm \beta$, $\alpha\beta$ and α^{-1} are algebraic over F. Therefore, the set of elements of K that are algebraic over F form a subfield of K containing F.

Proof. For (i), we use $F(\alpha)$ in the sense of Proposition 3.1, since we do not know *apriori* that α is algebraic over F. So, assume $[F(\alpha) : F]$ is finite, say $[F(\alpha) : F] = e$. Then $1, \alpha, \ldots, \alpha^e$ are $e + 1$ elements in the e-dimensional vector space $F(\alpha)$ over F. Thus, these elements must be linearly dependent over F. Thus, there exist $a_0, a_1, \ldots, a_e \in F$, not all zero, such that $a_0 + a_2\alpha + \cdots + a_e\alpha^e = 0$. In other words, α is algebraic over F. Conversely, suppose α is algebraic over F. Then by the previous proposition, α is the root of an irreducible polynomial in $F[x]$. Thus, by Remark 3.2, $F(\alpha)$ takes the form described in Proposition 2.1. In particular, $[F(\alpha) : F] = d$, where d is the degree of the minimal polynomial of α over F.

For part (ii) suppose α, β are algebraic over and we can show $[F(\alpha, \beta) : F] < \infty$. Since $F(\alpha \pm \beta)$, $F(\alpha\beta)$ and $F(\alpha^{-1})$ are contained in $F(\alpha, \beta)$, each of these subfields are finite dimensional over F, and thus by (i) above, $\alpha \pm \beta$, $\alpha\beta$ and α^{-1} are algebraic over F. The second statement in (ii) follows immediately from this.

To see that $F(\alpha, \beta)$ is finite dimensional over F, let's first note that by Definition 3.1, $F(\alpha, \beta) = F(\alpha)(\beta)$. Since β is algebraic over F, it is clearly algebraic over $F(\alpha)$. Thus, $F(\alpha)(\beta)$ is finite dimensional over $F(\alpha)$. Let $u_1, \ldots, u_n$ be a basis for $F(\alpha)$ over F (since α is algebraic over F) and $v_1, \ldots, v_m$ be a basis for $F(\alpha)(\beta)$ over $F(\alpha)$. We will be done if we show that the finite set $\{u_iv_j\}$ spans $F(\alpha)(\beta)$ over F. Let $\gamma \in F(\alpha)(\beta)$. Then $\gamma = \sum_{j=1}^m \alpha_j v_j$, for some $\alpha_j \in F(\alpha)$. For each $1 \leq j \leq m$ we can write $\alpha_j = \sum_{i=1}^n a_{ij}u_i$, with each $a_{ij} \in F$. Substituting these expressions into the equation for γ, we have

$$\gamma = \sum_{j=1}^m \alpha_j v_j = \sum_{j=1}^m (\sum_{i=1}^n a_{ij}u_i)v_j = \sum_{j=1}^m \sum_{i=1}^n a_{ij}u_iv_j,$$

which shows that the set $\{u_iv_j\}$ spans $F(\alpha)(\beta) = F(\alpha, \beta)$, which completes the proof. □

Corollary 4.5. *Let $F \subseteq K$ be an extension of fields. If $[K : F] < \infty$, then K is algebraic over F.*

Proof. Given any $\alpha \in K$, $F \subseteq F(\alpha) \subseteq K$, and thus $[F(\alpha) : F] < \infty$. Thus, by Theorem 4.4, α is algebraic over F. □

Example 4.6. Suppose we have a field extension $F \subseteq F(\alpha)$, where α is algebraic over F. It follows from Corollary 4.5 that *every* element of $F(\alpha)$ is algebraic over F. Given $\beta \in F(\alpha)$, here's how we can find a polynomial in $F[x]$ with β as a root, using the example $\mathbb{Q} \subseteq \mathbb{Q}(\sqrt[3]{2})$. Suppose we take $\beta = 1 + 2\sqrt[3]{2} + 3\sqrt[3]{4}$. We now consider the equations:

$$\begin{aligned}\beta \cdot 1 &= 1 + 2\sqrt[3]{2} + 3\sqrt[3]{4}\\ \beta \cdot \sqrt[3]{2} &= 6 + \sqrt[3]{2} + 2\sqrt[3]{4}\\ \beta \cdot \sqrt[3]{4} &= 4 + 6\sqrt[3]{2} + \sqrt[3]{4}\end{aligned}$$

Viewing these equations as a system of equations, we may rewrite them as a matrix equation

$$\begin{bmatrix}1-\beta & 2 & 3\\ 6 & 1-\beta & 2\\ 4 & 6 & 1-\beta\end{bmatrix} \cdot \begin{bmatrix}1\\ \sqrt[3]{2}\\ \sqrt[3]{4}\end{bmatrix} = \begin{bmatrix}0\\0\\0\end{bmatrix}.$$

The vector $\begin{bmatrix}1\\ \sqrt[3]{2}\\ \sqrt[3]{4}\end{bmatrix}$ is a non-trivial solution to the system

$$\begin{bmatrix}1-\beta & 2 & 3\\ 6 & 1-\beta & 2\\ 4 & 6 & 1-\beta\end{bmatrix} \cdot \begin{bmatrix}x\\ y\\ z\end{bmatrix} = \begin{bmatrix}0\\0\\0\end{bmatrix}.$$

Thus, the determinant of $\begin{bmatrix}1-\beta & 2 & 3\\ 6 & 1-\beta & 2\\ 4 & 6 & 1-\beta\end{bmatrix}$ equals zero. This gives an equation of degree three having β as a root. In particular, β is a root of $g(x) := x^3 - 3x^2 - 33x - 89$. Note that since the $[\mathbb{Q}(\sqrt[3]{2}) : \mathbb{Q}] = 3$, a prime, $\mathbb{Q}(\beta) = \mathbb{Q}(\sqrt[3]{2})$ (see Corollary 5.3 below) so that $[\mathbb{Q}(\beta) : \mathbb{Q}] = 3$ and thus $g(x)$ is the minimal polynomial of β over $\mathbb{Q}$. □

In general, if $F \subseteq K$ is a finite extension and $v_1, \ldots, v_n \in K$ is a basis for K over F, then for any $\beta \in K$, rewriting the products βv_j as linear combinations of $v_1, \ldots, v_n$ yields a homogeneous system of linear equations as in the example above. The determinant of the coefficient matrix then gives a polynomial $g(x)$ with coefficients in F of degree n having β as a root. However, $g(x)$ need not be the minimal polynomial of β over F.

In the last paragraph of the proof of Theorem 4.4, it is also the case that the set $\{u_iv_j\}$ is linearly independent over F. This fact, together with the last paragraph of the proof of Theorem 4.4 yields the multiplicative property of the degrees of extensions, to be seen in the next lecture.

Lecture 5: Wednesday September 1.

Proposition 5.1. *Let $F \subseteq K \subseteq L$ be extensions of fields. Then $[L : F]$ is finite if and only if $[K : F]$ and $[L : K]$ are finite, in which case $[L : F] = [L : K] \cdot [K : F]$.*

Proof. Suppose $u_1, \ldots, u_n \in K$ are linearly independent over F and $v_1, \ldots, v_m \in L$ are linearly independent over K. We claim that the set $\{u_iv_j\}$ is linearly independent over F. Suppose the claim is true. Then if $[K : F]$ and $[L : K]$ are finite, and we take $\{u_i\}$ to be a basis for K over F and $\{v_j\}$ to be a basis for L over K, then the claim and the fact that the vectors $\{u_iv_j\}$ span L over F (see the proof of Theorem 4.4) show that the set $\{u_iv_j\}$ is a basis for L over F. Thus $[L : F]$ is finite and $[L : F] = [L : K] \cdot [K : F]$. Conversely, if $[L : F]$ is finite, then $[K : F]$ is finite, since K is a subspace of L, as vector space over F. Moreover, any set that spans L over F automatically spans L over K, so $[L : K]$ must also be finite. Therefore, it remains to prove the claim.

Suppose $\sum_j \sum_i a_{ij}u_iv_j = 0$, with each $a_{ij} \in F$. Then

$$0 = \sum_j \sum_i a_{ij}u_iv_j = \sum_j (\sum_i a_{ij}u_i)v_j.$$

Since each $\sum_i a_{ij}u_i$ belongs to K and $\{v_j\}$ is linearly independent over K, each $\sum_i a_{ij}u_i = 0$. Thus all $a_{ij} = 0$, since the set $\{u_i\}$ is linearly independent over F, which gives what we want. □

Corollary 5.2. *Let $F \subseteq K$ be an extension of fields.*

(i) If $[K : F]$ is a prime number, then there are no intermediate fields between F and K.
(ii) If $\alpha_1, \dots, \alpha_s \in K$ are algebraic over F, then $[F(\alpha_1, \dots, \alpha_s) : F] < \infty$.

Proof. Part (i) follows immediately from Proposition 5.1.

For (ii), we have an extension of fields $F \subseteq F(\alpha_1) \subseteq F(\alpha_1, \alpha_2) \subseteq \cdots \subseteq F(\alpha_1, \dots, \alpha_s)$. The degree of each simple extension $F(\alpha_1, \dots, \alpha_i) \subseteq F(\alpha_1, \dots, \alpha_{i+1})$ is finite. Thus, by Proposition 5.1 and induction on s, we have that

$$[F(\alpha_1, \dots, \alpha_s : F] = [F(\alpha_1) : F] \cdot [F(\alpha_1, \alpha_2) : F(\alpha_1)] \cdots [F(\alpha_1, \dots, \alpha_s) : F(\alpha_1, \dots, \alpha_{s-1})] < \infty,$$

which gives what we want. □

The next result can be considered a generalization of Proposition 5.1.

Theorem 5.3. *Let $F \subseteq K \subseteq L$ be an extension of fields. Then L is algebraic over F if and only if L is algebraic over K and K is algebraic over F.*

Proof. If L is algebraic over F, then clearly L is algebraic over K and K is algebraic over F. Now suppose K is algebraic over F and L is algebraic over K. To see that L is algebraic over F, by Theorem 4.4, it is enough to show $[F(\alpha) : F] < \infty$, for any $\alpha \in L$.

Since α is algebraic over K, let $f(x) = x^r + a_1x^{r-1} + \cdots + a_r$ be the minimal polynomial of α over K, so that each $a_j \in K$ is algebraic over F. Then by Corollary 5.2, $[F(a_1, \dots, a_r) : F] < \infty$. On the other hand, α is algebraic over $F(a_1, \dots, a_r)$ and thus $[F(a_1, \dots, a_r, \alpha) : F(a_1, \dots, a_r)] < \infty$. Thus,

$$[F(a_1, \dots, a_r, \alpha) : F] = [F(a_1, \dots, a_r, \alpha) : F(a_1, \dots, a_r)] \cdot [F(a_1, \dots, a_r) : F] < \infty.$$

Since $F \subseteq F(\alpha) \subseteq F(a_1, \dots, a_r, \alpha)$, $[F(\alpha) : F] < \infty$, and the proof is complete. □

Lecture 6: Friday September 3. Thus far, we have worked with roots of polynomials that were assumed to exist. We now want to show that given a field F and a polynomial $g(x) \in F[x]$, there exists a field K containing F and $\alpha \in K$ such that $g(\alpha) = 0$. Note that this requires not only the construction of a root of $g(x)$, but also a field containing the root. After all, a root does not exist just floating in space. The standard way of doing this, is to use the fact that $F[x]$ is a PID. The proof runs as follows. One first takes $f(x)$, a monic, irreducible factor of $g(x)$. Any root of $f(x)$ is a root of $g(x)$, so one just has to construct a field containing a root of $f(x)$. Since $f(x)$ is irreducible over F, the ideal $(f(x))$ of $F[x]$ generated by $f(x)$ is a maximal ideal, and hence $K := F[x]/(f(x))$ is a field. F is identified as a subfield of K via the map $\alpha \to \overline{\alpha}$, the residue class of α in $F[x]/(f(x))$, for all $\alpha \in F$. This map is easily seen to be an isomorphism of fields. Finally, under this identification, we have $f(\overline{x}) \equiv \overline{f(x)} \equiv 0$ in $F[x]/(f(x)) = K$. Thus, $\overline{x}$ in K is a root of $f(x)$.

The proof of the construction of K and α we give below is more explicit and is almost identical to the proof that $F(\alpha)$ is a field, when we have a known root α. And why not: After K and α have been constructed, we can then form $F(\alpha)$.

Proposition 6.1. *Let F be a field and $f(x) \in F[x]$ be a polynomial. Then there exists a field K containing F and $\alpha \in K$ such that $f(\alpha) = 0$.*

Proof. We begin by noting that we may assume that $f(x)$ is irreducible over F. Indeed, if not, let $g(x)$ be an irreducible factor of $f(x)$. If K is an extension of F containing a root α of $g(x)$, then α is also a root of $f(x)$. Thus, without loss of generality, we may assume $f(x)$ is irreducible over F. Let d denote the degree of $f(x)$, let z be another indeterminate and set K to be the set of all polynomials in z having degree less than d. Thus, a typical element in K has the form $a_0 + a_1z + \cdots + a_{d-1}z^{d-1}$ with each $a_i \in F$. Note that K is clearly closed under the usual addition of polynomials. However, K is not even a ring, since it is not closed under multiplication of polynomials. So we endow the set K with a *new* multiplicative structure that turns K into a field. What multiplication should we impose on K? Answer: The exact same multiplicative structure $F(\alpha)$ would have if we already had a root α of $f(x)$!

So, for $A = a_0 + a_1z + \cdots + a_{d-1}z^{d-1}$ and $B = b_0 + b_1z + \cdots + b_{d-1}z^{d-1}$ in K, we define $A \cdot B$ in K to be $R = c_0 + c_1z + \cdots + c_{d-1}z^{d-1}$, where R is the remainder obtained by applying the division algorithm to the polynomial product AB, i.e., in $F[z]$, $AB = h(z)f(z) + R$, where R is either zero or has degree less than d. Note that the product $A \cdot B$ in K is well defined, since the remainder obtained from the division algorithm is unique.

Now, we must demonstrate that with this multiplication, the sum and product defined above satisfy the field axioms. Most of these properties will be inherited from $F[z]$. First of all, it is clear that addition is commutative and associative, that 0 is the additive identity and that if $A \in K$, then $-A$ is the additive inverse of A. Moreover, it is clear that $A \cdot B = B \cdot A$, since multiplication of polynomials is commutative. Furthermore, $1 \in K$ is easily seen to be the multiplicative identity. The distributive property is also not to hard to see. If $A, B, C \in K$, then in $F[z]$, if we write $AB = f(z)h(z) + R(z)$ and $AC = f(z)h_1(z) + R_1(z)$, then in K, $A \cdot B = R(z)$ and $A \cdot C = R_1(z)$, thus, $A \cdot B + A \cdot C = R(z) + R_1(z)$. On the other hand, in $F[z]$,

$$A(B + C) = AB + AC = f(z)h(z) + R(z) + f(z)h_1(z) + R_1(z) = (h(z) + h_1(z)f(z) + (R(z) + R_1(z)).$$

Since the degree of $R(z) + R_1(z)$ is less than d, and remainder are unique, this shows that

$$A \cdot (B + C) = R(z) + R_1(z) = A \cdot B + A \cdot C,$$

which gives what we want.

Now, suppose A as above is not zero. Since A and $f(z)$ are relatively prime in $F[z]$, there exist $C(z), D(z)$ in $F[z]$ such that $1 = AC(z) + D(z)f(z)$ in $F[z]$. As we saw in Lecture 2, we may assume $C(z)$ has degree less than d. This yields, $AC(z) = (-D(z))f(z) + 1$. Since the remainder in the division algorithm is unique, this last equation shows that in K, $A \cdot C(z) = 1$. Thus, A has a multiplicative inverse.

To see the associative property of multiplication, let $A, B, C \in K$. In what follows, $r(x), r_1(z), r_2(z), r_3(z)$ in $F[z]$ all have degree less than d, and thus also belong to K. Write, $AB = f(z)h(z) + r(z)$, so that in K, $A \cdot B = r(z)$. Now write $r(z)C = f(z)h_1(z) + r_1(z)$. On the one hand, this give $(A \cdot B) \cdot C = r_1(z)$ in K, while on the other hand, in $F[z]$, we have

$$(AB)C = f(z)Ch(z) + Cr(z) = f(z)\{Ch(z) + h_1(z)\} + r_1(z).$$

We now write $BC = f(z)h_2(z) + r_2(z)$, so that in K, $B \cdot C = r_2(z)$. In $F[z]$, we then write

$$Ar_2(z) = f(z)h_3(z) + r_3(z).$$

On the one hand, in K this gives $A \cdot (B \cdot C) = r_3(z)$, while on the other hand , in $F[z]$ we have

$$A(BC) = Af(z)h_2(z) + Ar_2(z) = Af(z)h_2(z) + f(z)h_3(z) + r_3(z) = (Ah_2(z) + h_3(z)f(z) + r_3(z).$$

Since $A(BC) = (AB)C$, uniqueness of remainders in the division algorithm gives $r_1(z) = r_3(z)$. Therefore $(A \cdot B) \cdot C = A \cdot (B \cdot C)$ in K. in K. Thus, with the given operations, K is a field containing F. Note that if $a, b \in F$, $ab = 0 \cdot f(z) + ab$, in $F[z]$, which means that $a \cdot b$ in K equals ab in F. In other words, the multiplication in K restricts to the multiplication in F, meaning that F is a subfield of K. Moreover, if $a \in F$ and $B \in K$, then $a \cdot B = aB$, since in $F[z]$, $aB = 0 \cdot f(z) + aB$. This shows that the product of an element in F with an element in K under the new product is that same as the old.

We now note that $z \in K$ is a root of $f(x)$.

Suppose $f(x) = x^d + u_{d-1}x^{d-1} + \cdots + u_0$. We first note that in K, an i-fold product of z with itself is the same as z^i in $F[z]$, if $i < d$. This follows since in $F[z]$, $z^i = 0 \cdot f(z) + z^i$. On the other hand, for the d-fold product of z with itself in K, we note that in the polynomial ring $F[z]$ we can write

$$z^d = 1 \cdot f(z) + (-u_{d-1}z^{d-1} - \cdots - u_0).$$

Thus, in K, $z^d = -u_{d-1} \cdot z^{d-1} - \cdots - u_0$, which gives

$$0 = z^d + u_{d-1} \cdot z^{d-1} + \cdots + u_0 = f(z),$$

in K. In other words, z, as an element of K, is a root of $f(x)$. Thus, K is a field containing F containing a root of $f(x)$. We may relabel z in K as α, and upon doing so, it is not hard to see that $K = F(\alpha)$. □

We now have the following: Given a field F, and a polynomial $g(x) \in F[x]$, there exists a field K containing F and $\alpha \in K$ such that $g(\alpha) = 0$. Thus, $g(x) = (x - \alpha)h(x)$, for some $h(x) \in K[x]$. We may now repeat the process on K and $h(x)$, to find a field L containing K and $\beta \in L$ such that $h(\beta) = 0$. Thus, $g(x) = (x - \alpha)(x - \beta)t(x)$, for some $t(x) \in L[x]$. It follows that there exists a field E containing F and $\alpha_1, \ldots, \alpha_n$, not necessarily distinct, such that $g(x) = (x - \alpha_1) \cdots (x - \alpha_n)$, where n is the degree of $g(x)$. This leads to the following definition.

Definition 6.2. Let F be a field and $g(x) \in F[x]$ a polynomial of degree n. Let $F \subseteq K$ be an extension of fields. We say that $g(x)$ splits over K, if there exist $\alpha_1, \ldots, \alpha_n \in K$, such that $g(x) = (x - \alpha_1) \cdots (x - \alpha_n)$. The subfield $E := F(\alpha_1, \ldots, \alpha_n)$ is called a splitting field of $g(x)$ over F.

Lecture 7: Wednesday, September 8. Now that we can construct a field containing all of the roots of a given polynomial, we want to see that we can construct a field containing *all* of the roots of *all* of the polynomials in F.

Definition/Proposition 7.1. (i) A field Ω is said to be algebraically closed if the following equivalent conditions hold.

(a) Every non-constant polynomial in $\Omega[x]$ has a root in Ω.
(b) Every non-constant polynomial in $\Omega[x]$ splits over Ω.
(c) There are no algebraic extensions of Ω.

(ii) Given a field F, a field $\overline{F}$ containing F is an algebraic closure of F, if $\overline{F}$ is algebraically closed and algebraic over F.

Proof of (i). Suppose $f(x) \in \Omega[x]$ is a non-constant polynomial. We proceed by induction on the degree of $f(x)$. If the degree of $f(x)$ is one, there is nothing to prove. If the degree of $f(x)$ is greater than one, by (a) there exists $\alpha \in \Omega$ such that $f(\alpha) = 0$. Thus, $f(x) = (x - \alpha)g(x)$ with $g(x) \in \Omega[x]$. By induction, $g(x)$ splits over Ω, so $f(x)$ splits over Ω. Thus, (a) implies(b). Suppose (b) holds and K is an algebraic extension of Ω. Let $\alpha \in K$, and suppose $f(x)$ is the minimal polynomial of α over Ω. Then, $f(x)$ is irreducible over Ω. On the other hand, $f(x)$ splits over Ω, by (b), which gives a contradiction, unless $f(x)$ has degree one. But then, this means $\alpha \in \Omega$. Thus, $K \subseteq \Omega$ and therefore there are no algebraic extension of Ω. Finally, for (c) implies (a), take $f(x) \in \Omega[x]$. Then by Proposition 6.1, there is a field extension $\Omega \subseteq K$ and $\alpha \in K$ such that $f(\alpha) = 0$. But $\Omega(\alpha)$ is an algebraic extension. By (c), $\Omega(\alpha) = \Omega$, so $\alpha \in \Omega$, which is what we want. □

Remark 7.2. We will show later in the semester that $\mathbb{C}$ is an algebraically closed field. Since $\mathbb{C}$ is algebraic over $\mathbb{R}$, this means that $\mathbb{C}$ is an algebraic closure of $\mathbb{R}$. Every field has an algebraic closure, even $\mathbb{Z}_2$. For any field F, $F[x]$ has infinitely many irreducible polynomials, and it follows from this that any algebraically closed field must be infinite, even an algebraic closure of $\mathbb{Z}_2$. However, an easy set-theoretic argument shows that an algebraic closure of a finite (or countable) field must be countable. □

Lecture 8: Friday September 10. We now want to show that every field F has an algebraic closure. The proof below is due to E. Artin, and is a clever generalization of the proof that $F[x]/(f(x))$ is a field containing a root of the irreducible polynomial $f(x)$. The proof below implicitly uses Zorn's Lemma. In case you are not familiar with Zorn's lemma, a brief exposition has been added to the Supplementary Material folder. In particular, this exposition gives a proof showing that if R is a commutative ring and $I \subseteq R$ is an ideal, then there exists a maximal ideal $M \subseteq R$ containing I (as done in the previous class). We will use the fact that, in a commutative ring R, an ideal $M \subseteq R$ is maximal if and only if R/M is a field.

Theorem 8.1. *Let F be a field. Then F has an algebraic closure.*

Proof. For each $f(X) \in F[X]$, introduce a separate variable X_f. Let R denote the polynomial ring in the collection of variables $\{X_f \mid f \in F[X]\}$. Recall that the polynomial ring in an infinite collection of variables is just the set of polynomials in finitely many of the variables at a time. Let J denote the ideal of R generated by $\{f(X_f) \mid f \in F[X]\}$. We claim that J is a proper ideal of R. Suppose this were not the case. Then, $1 \in J$, so there exists an expression of the form

$$1 = g_1 \cdot f_1(X_{f_1}) + g_2 \cdot f_2(X_{f_2}) + \cdots + g_d \cdot f_d(X_{f_d}),$$

where, each $g_i \in R$ and $f_i(X_{f_i}) \in J$. Note, we have note written out the variables appearing in each g_i. Now, let K be a field containing a root α_i of each $f_i(X_{f_i})$. There is no problem finding K, we could simply take K

to be the splitting field of the product of the $f_i(X_{f_i})$. Now the g_i in the equation above may involve variables other than the X_{f_i}, but all together, they involve only finitely many variables. Thus, we may substitute for the variables in the equation above as follows : Set each $X_{f_i} = \alpha_i$ and all the remaining variables equal to zero. This yields, $1 = 0$, the desired contradiction. Therefore, J is a proper ideal.

Now, let $M \subseteq R$ be a maximal ideal containing J (such an ideal exists by Zorn's Lemma), and let K_1 denote the field R/M. We identify F as a subfield of K_1 in the natural way, namely, $\lambda \in F$ corresponds to the residue class in K_1 of the constant polynomial λ. Then as in the case of a single polynomial, for each $f(X) \in F[X]$, $f(\overline{X_f}) \equiv \overline{f(X_f)} \equiv 0$ in K_1. Thus, K_1 is a field containing F having the property that *every* polynomial in $F[X]$ has a root in K_1. We may now apply the construction to K_1 and obtain a field K_2 having the property that every polynomial in $K_1[X]$ has a root in K_2. It follows that every polynomial in $F[X]$ has at least two roots in K_2 (assuming its degree is at least two). We therefore obtain a tower of fields $F \subseteq K_1 \subseteq K_2 \subseteq \cdots$ whose union is a field K. Note, a union of fields need not be a field, but a *directed union* of fields will be a field. By design, every polynomial in $F[X]$ splits over K. In fact, for each K_i, every polynomial in $K_i[X]$ splits over K, since the construction starting at K_i is the same as the one starting at F. Since any polynomial in $K[X]$ belongs to $K_i[X]$ for some i, it follows that every polynomial with coefficients in K splits over K, i.e., K is an algebraically closed field.

To finish, let $\overline{F}$ denote the elements in K algebraic over F and consider $h(X) \in \overline{F}[X]$. Then $h(X)$ splits over K (since $h(X) \in K[x]$). But the roots of $h(X)$ are algebraic over $\overline{F}$ and therefore algebraic over F. Thus the roots of $h(X)$ belong to $\overline{F}$. In other words, $h(X)$ splits over $\overline{F}$, so $\overline{F}$ is algebraically closed. Since $\overline{F}$ is algebraic over F, $\overline{F}$ is an algebraic closure of F. □

Remarks 8.2. (i) The proof above uses the set theoretic fact that the union of sets always exists as a set and the fact that a directed union of fields is a field. If $K_1 \subseteq K_s \subseteq \cdots$ is a directed union of fields, this means that for all $i < j$, the binary operations on K_i are just the restrictions of the binary operations on K_j to K_i. In other words, K_i is a subfield of K_j, for $i < j$. Thus, if $\alpha, \beta \in K := \bigcup_i K_i$, then $\alpha, \beta \in K_j$, some j, and we may unambiguously apply the binary operations of K_j to α and β to get the sum, product, inverses, etc, of α and β as elements of K. In other words, $K$is a field.

(ii) Regarding the proof of Theorem 8.1. It is not difficult to see that, in fact, $K = \overline{F}$, since each K_i, and thus K, is algebraic over F. However, R. Gilmer has shown that K_1 is already equal to $\overline{F}$!

(iii) Using the notation from the proof of Theorem 8.1, strictly speaking, though $F \subseteq R$, an isomorphic copy of F is contained in $K_1 = R/M$. So that the algebraic closure constructed above is an algebraic closure of an isomorphic copy of F. To remedy this, one can proceed as follows. Suppose F_1 is an isomorphic copy of F and $\phi : F_1 \to F$ is a field isomorphism. Let $\overline{F_1}$ be an algebraic closure of F_1. Let T be a set containing F having the same cardinality as $\overline{F_1}$, and take $\tau : \overline{F_1} \to T$ a 1-1 and onto set function such that τ extends ϕ. For $a, b \in T$, define $ab := \tau(\tau^{-1}(a)\tau^{-1}(b))$, $a + b := \tau(\tau^{-1}(a) + \tau^{-1}(b))$. Then is it not difficult to check that with these operations, T is a field and τ is a field isomorphism from $\overline{F_1}$ to T extending ϕ. If we now call this new field $\overline{F}$, it follows that $\overline{F}$ is an algebraic closure of F.

Lecture 9: Monday September 13. Now that we know algebraic closures exist, we want to show that given a field F, any two algebraic closures of F are isomorphic via a field homomorphism fixing F. For this, in the next lecture we will prove a slightly more general result concerning isomorphic splitting fields. We start with a quick review and discussion of some basic facts we will use below.

Review. Let $F \subseteq K$ be an algebraic extension of fields.

(i) If $\alpha_1, \ldots, \alpha_n \in K$, then every element in $F(\alpha_1, \ldots, \alpha_n)$ is a polynomial expression in $\alpha_1, \ldots, \alpha_n$ with coefficients in F. In fact, if $[F(\alpha_1, \ldots, \alpha_i) : F(\alpha_1, \ldots, \alpha_{i-1})] = d_i$, for each $1 \leq i \leq n$, then a basis for $F(\alpha_1, \ldots, \alpha_n)$ over F is the set of monomial expressions $\alpha_1^{e_1} \cdots \alpha_n^{e_n}$ such that $1 \leq e_i < d_i$. This follows by iterating Proposition 2.1, since each $F(\alpha_1, \ldots, \alpha_i)$ is obtained by adjoining α_i to $F(\alpha_1, \ldots, \alpha_{i-1})$.

(ii) Suppose $U \subseteq K$ is a (not necessarily finite) subset. If $\beta \in F(U)$, then there are finitely many elements $u_1, \ldots, u_n \in U$ and a polynomial $h(x_1, \ldots, x_n)$ with coefficients in F such that $\beta = h(u_1, \ldots, u_n)$. To see this, we use Proposition 3.1 to find $u_1, \ldots u_n \in U$ such that $\beta = p(u_1, \ldots, u_n)q(u_1, \ldots, u_n)^{-1}$, with $p(x_1, \ldots, x_n), q(x_1, \ldots, x_n)$ polynomials with coefficients in F. Thus, $\beta \in F(u_1, \ldots, u_n)$. Therefore, the required $h(x_1, \ldots, x_n)$ exists by part (i).

(iii) From problem 7 on Homework 1, we have the following. Suppose $F \subseteq K$ and $F_0 \subseteq K_0$ are field extensions, $\sigma : F \to F_0$ is a field isomorphism and $f(x) \in F[x]$ is irreducible over F. If we set $f_0(x) = f^{\sigma}(x)$, then $f_0(x)$ is irreducible over F_0. Let $\alpha \in K$ be a root of $f(x)$ and $\alpha_0 \in K_0$ be a root of $f_0(x)$. Then there exists an isomorphism of fields $\tau : F(\alpha) \to F_0(\alpha_0)$ such that $\tau(a) = \alpha_0$ and τ restricted to F equals σ.

Proposition 9.1. *Let $F \subseteq K$ be an algebraic extension of fields. Suppose $\phi : F \to L$ is a field homomorphism with L an algebraically closed field. Then there exists $\sigma : K \to L$, a field homomorphism extending ϕ.*

Proof. The proof will use Zorn's Lemma. We will also use the fact that field homomorphisms are always one-to-one. Let $\mathcal{C}$ denote the partially ordered set $\{(E, \rho)\}$, where E is a field between F and K and $\rho : E \to L$ is a field homomorphism extending ϕ. The partial order on $\mathcal{C}$ is given as follows: $(E_1, \rho_1) \le (E_2, \rho_2)$ if $E_1 \subseteq E_2$ and ρ_2 restricted to E_1 equals ρ_1. We first note that $\mathcal{C}$ is not empty, since (F, ϕ) belongs to $\mathcal{C}$.

Let $\mathcal{E} = \{(E_i, \rho_i)\}_{i \in I}$ be a chain in $\mathcal{C}$. We must show that $\mathcal{E}$ has an upper bound in $\mathcal{C}$. Set $\tilde{E} := \bigcup_{i \in I} E_i$. Since $\mathcal{E}$ is totally ordered, an argument similar to what we have seen before shows that $\tilde{E}$ is a field, and hence an intermediate field between F and K. We define $\tilde{\rho} : \tilde{E} \to L$ as follows. Take $x \in \tilde{E}$. Then $x \in E_i$ for some $i \in I$. Set $\tilde{\rho}(x) := \rho_i(x)$. We need to see that $\tilde{\rho}$ is well defined. Suppose $x \in E_j$, for some $j \in I$. We must show $\rho_i(x) = \rho_j(x)$. But this is clear from the definition of the partial order on $\mathcal{C}$, since $\mathcal{E}$ is a chain. For example, if $E_i \subseteq E_j$, then since ρ_j restricted to E_i is ρ_i, we have $\rho_i(x) = \rho_j(x)$. Thus, $\tilde{\rho}$ is well defined. Note that by definition, $\tilde{\rho}$ extends ϕ, and therefore $(\tilde{E}, \tilde{\rho})$ belongs to $\mathcal{C}$. Finally, for $(E_i, \rho_i) \in \mathcal{E}$, we clearly have $E_i \subseteq \tilde{E}$ and the restriction of $\tilde{\rho}$ to E_i equals ρ_i, by definition of $\tilde{\rho}$, so that $(E_i, \rho_i) \le (\tilde{E}, \tilde{\rho})$. Thus, $(\tilde{E}, \tilde{\rho})$ is an upper bound for the chain $\mathcal{E}$.

Therefore, by Zorn's Lemma, $\mathcal{C}$ has a maximal element, (K', σ). If $K' = K$, the proof is complete. Suppose $K' \subsetneq K$. Take $\alpha \in K \backslash K'$ and let $f(X)$ be the minimal polynomial for α over K'. If we set $\sigma(K') := K'_0$, then $K'_0 \subseteq L$ is a field isomorphic to F and $f(X)$ corresponds via σ to a polynomial $f_0(X) \in K'_0[X]$. It is not too difficult to show that $f_0(x)$ is irreducible over K'_0 [3]. Since L is algebraically closed, it contains a root β of $f_0(X)$. Thus, item (iii) from the Review above shows that there exists field isomorphism ρ from $E := K'(\alpha)$ to $E_0 := F_0(\alpha_0) \subseteq L$ extending σ. Therefore, $(E, \rho) \in \mathcal{C}$ and is not equal to (K', ϕ), contradicting the maximality of (K', σ). Thus, $K' = K$ and σ is the required field homomorphism. □

Remark 9.2. The required isomorphism of algebraic closures follows almost immediately from the proposition above. We sketch the proof of this and reserve a more formal proof for the isomorphism of splitting fields in the next lecture. Let F be a field and $\overline{F_1}$ and $\overline{F_2}$ algebraic closures of F. To see that there is a field isomorphism from $\overline{F_1}$ to $\overline{F_2}$ that fixes F, we apply the previous proposition with $K = \overline{F_1}$ and $L = \overline{F_2}$. Let $\phi : F \to \overline{F_2}$ be the identity map. Then, by Proposition 9.1 there exists a field homomorphism σ from $\overline{F_1}$ to $\overline{F_2}$ extending ϕ. Since ϕ is the identity map, this means that σ fixes F. Since any field homomorphism is automatically one-to-one, we must show σ is onto. Let $\beta \in \overline{F_2}$ and write $f(x)$ for the minimal polynomial of β over F. Let d be the degree of $f(x)$. Since $\overline{F_1}$ is algebraically closed, there exist $\alpha_1, \ldots, \alpha_d \in \overline{F_1}$ such that

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

in $\overline{F_1}[x]$. Because, σ is the identity on F, we have $f^{\sigma}(x) = f(x)$. On the other hand, one can also show that $f(x) = (x - \sigma(\alpha_1)) \cdots (x - \sigma(\alpha_d)) \in \overline{F_2}[x]$. Since $\beta \in \overline{F_2}$ is a root of $f(x)$, we must have $\beta = \sigma(\alpha_i)$ for some i, showing σ is onto.

Lecture 10: Wednesday September 15. In this lecture we will see that two algebraic closures of the same field are isomorphic. To do this, we will prove a slightly stronger result for splitting fields. Splitting fields play an important role in Galois Theory.

Definition 10.1. (i) Given any set S of non-constant polynomials in $F[x]$, if we fix an algebraic closure $\overline{F}$ of F, then all of the roots of the polynomials in S belong to $\overline{F}$, and we can therefore adjoin them to F. We call the resulting field the splitting field of S over F. Note that if S is a finite set of polynomials, then the

[3] Note that $f_0(x)$ is obtained by applying σ to the coefficients of $f(x)$. If we write $f^{\sigma}(x)$ for this polynomial, then one can show that the map $\psi : K'[x] \to K'_0[x]$ defined by $\psi(g(x)) = g^{\sigma}(x)$ is an isomorphism of rings. It follows from this that $f_0(x)$ is irreducible over K'_0

splitting field of S is just the splitting field of the single polynomial obtained by taking the product of the elements in S. More generally we have, the following definition.

(ii) Let $F \subseteq K$ be an algebraic extension of fields. We say that K is a splitting field over F if K is the splitting field for some subset $S \subseteq F[x]$. In other words, if $U \subseteq \overline{F}$ is the set of roots of the polynomials in S, then $K = F(U)$. □

NOTE. Let $f(x) \in F[x]$ and suppose $c \in F$ is the leading coefficient of $f(x)$. The $f(x)$ splits over a field K if and only if $\frac{1}{c} \cdot f(x)$ splits over K. Thus, we may safely assume that, whenever we have a splitting field K for a set of polynomials $S \subseteq F[x]$, all of the polynomials in S are monic.

Examples 10.2 (i) Let $f(x) = x^2 - 2 \in \mathbb{Q}[x]$. The the roots of $f(x)$ are $\pm\sqrt{2}$. It follows that the splitting field of $f(x)$ over $\mathbb{Q}$ is $\mathbb{Q}(\pm\sqrt{2}) = \mathbb{Q}(\sqrt{2})$.

(ii) Let $g(x) = x^3 - 2 \in \mathbb{Q}[x]$, take $\sqrt[3]{2}$ to be the real cube root of 2. Let $\epsilon = e^{\frac{2\pi i}{3}}$, a *primitive cube root of* 1. Note that $\epsilon^3 = 1$ and $(\epsilon^2)^3 = 1$, for $\epsilon^2 = e^{\frac{4\pi i}{3}}$. It follows that $\sqrt[3]{2}$, $\sqrt[3]{2}\epsilon$, and $\sqrt[3]{2}\epsilon^2$ are the roots of $g(x)$. Thus $\mathbb{Q}(\sqrt[3]{2}, \sqrt[3]{2}\epsilon, \sqrt[3]{2}\epsilon^2) = \mathbb{Q}(\sqrt[3]{2}, \epsilon)$ is the splitting field of $g(x)$ over $\mathbb{Q}$.

(iii) Let $h(x) = x^n - 1$ and set $\gamma := e^{\frac{2\pi i}{n}}$. Then $1, \gamma, \gamma^2, \ldots, \gamma^{n-1}$ are distinct complex numbers and $\gamma_i^n = 1$, for all i. Thus, $1, \gamma, \ldots, \gamma^{n-1}$ are the distinct roots of $h(x)$, therefore $\mathbb{Q}(\gamma, \ldots, \gamma^{n-1}) = \mathbb{Q}(\gamma)$ is the splitting field of $h(x)$ over $\mathbb{Q}$.

(iv) Let $S \subseteq \mathbb{Q}[x]$ denote the set of polynomials $x^2 - n$, where $n \in T$, the set of positive square-free integers. Then the splitting field for S over $\mathbb{Q}$ is $\mathbb{Q}(U)$, where $U := \{\sqrt{n} \mid n \in T\}$.

Lemma 10.3. *Let F be a field with algebraic closure $\overline{F}$. Suppose $S \subseteq F[x]$ is a set of non-constant polynomials and $F \subseteq K_1, K_2 \subseteq \overline{F}$ are splitting fields for S over F. Then $K_1 = K_2$.*

Proof. This is basically clear, since any polynomial of degree n cannot have more than n roots in $\overline{F}$. More explicitly, let $K_1 = F(U)$ and $K_2 = F(V)$ be as in the definition above. Take $u \in U$ and $f(x) \in S$ with $f(u) = 0$. Suppose $f(x) \in S$ has degree n. Then over $\overline{F}$ we may write

$$(x - u)(x - u_2) \cdots (x - u_n) = f(x) = (x - v_1) \cdots (x - v_n),$$

with $u_j \in U$ and $v_j \in V$. Upon setting $x = u$, the left hand side of the displayed equation equals zero, so the right hand side equals zero as well. This shows that $u = v_j$ for some j. Thus $u \in V$. Therefore $U \subseteq V$. By symmetry, $V \subseteq U$, so $V = U$, and thus $K_1 = K_2$. □

Theorem 10.4. *Let $F \subseteq K$ be an algebraic extension. Assume K is contained in the algebraic closure $\overline{F}$ of F. The following are equivalent:*

(i) K is a splitting field over F.
(ii) If $f(x) \in F[x]$ is irreducible and has a root in K, then $f(x)$ splits over K.
(iii) If $\sigma : K \to \overline{F}$ is a field homomorphism fixing F, then $\sigma(K) = K$.

Remark 10.5. Note that the theorem above shows that if K is the splitting field for the set of polynomials $S \subseteq F[x]$, and $f(x) \in F[x]$ is irreducible, then even if $f(x)$ is *not* in S, $f(x)$ splits over K, as long as it has at least one root in K. □

Proof of Theorem 10.4. We start with (i) implies (iii). Assume $K = F(U)$ is a splitting field for the set $S \subseteq F[x]$ and suppose $\sigma : K \to \overline{F}$ is a field homomorphism fixing F. If we show that $\sigma(K) \subseteq \overline{F}$ is a splitting field for S over F, then $\sigma(K) = K$, by the previous Lemma, which is what we want. Let $f(x) \in S$ and write $f(x) = x^n + c_1x^{n-1} + \cdots + c_n$, with each $c_j \in F$. Then there exist $u_1, \ldots u_n \in U$ such that $f(x) = (x - u_1) \cdots (x - u_n)$. It follows that

$$\begin{aligned}
c_1 &= -(u_1 + \cdots + u_n)\\
c_2 &= \sum_{i<j} u_i u_j\\
&\vdots\\
c_n &= (-1)^n u_1 \cdots u_n.
\end{aligned}$$

Applying σ to these equations, we get

$$\begin{aligned}
c_1 &= -(\sigma(u_1) + \cdots + \sigma(u_n))\\
c_2 &= \sum_{i<j} \sigma(u_i)\sigma(u_j)\\
&\vdots\\
c_n &= (-1)^\sigma \sigma(u_1) \cdots \sigma(u_n).
\end{aligned}$$

Thus, in $\overline{F}[x]$, $f(x) = (x - \sigma(u_1)) \cdots (x - \sigma(u_n))$. This shows that $F(\sigma(U))$ is a splitting field for S over F. Since σ fixes F, we have $\sigma(K) = F(\sigma(U))$, showing that $\sigma(K)$ is a splitting field for S over F, which is what we want.

To see that (iii) implies (ii), suppose $f(x) \in F[x]$ is irreducible over F and has a root $\alpha \in K$. Let $\alpha' \in \overline{F}$ be any other root of $f(x)$. By Proposition 2.1, there exists a field isomorphism $\phi : F(\alpha) \to F(\alpha')$ fixing F, such that $\phi(\alpha) = \alpha'$. We may regard ϕ as a field homomorphism from $F(\alpha)$ to $\overline{F}$. By Proposition 9.1, we may extend ϕ to a field homomorphism $\sigma : K \to \overline{F}$. By assumption, $\sigma(K) = K$. In particular, $\alpha' = \phi(\alpha) \in K$. Thus, K contains all of the roots of $f(x)$, so $f(x)$ splits over K.

Finally, if we assume (ii), let $\beta \in K$ and write $f_\beta(x)$ for the minimal polynomial of β over F. Let U_β, denote the set of roots of $f_\beta(x)$, so that by assumption, U_β belong to K. If we set $U := \bigcup_\beta U_\beta$, as β ranges over $K\backslash F$, we clearly have that $K = F(U)$ is a splitting field for the set of polynomials $\{f_\beta(x)\}_\beta$, which completes the proof. □

Corollary 10.6. *Let F be a field and $S \subseteq F[x]$ be a set of non-constant polynomials. Assume that K_1 and K_2 are two splitting fields for S over F. Then there exists a field isomorphism $\sigma : K_1 \to K_2$ that fixes F. In particular, if $\overline{F_1}$ and $\overline{F_2}$ are algebraic closure of F, then there exists a field isomorphism $\sigma : \overline{F_1} \to \overline{F_2}$ fixing F.*

Proof. If $K_1 \neq K_2$, it follows from Lemma 10.3 that K_1 and K_2 are contained in two different algebraic closures of F, say $\overline{F_1}$ and $\overline{F_2}$ respectively. Let $\phi : F \to F \subseteq \overline{F_2}$ be the identity map. By Proposition 9.1, there exists a field homomorphism $\sigma : K \to \overline{F_2}$ extending ϕ. Thus, σ fixes F. The same proof of (i) implies(iii) from Theorem 10.4 shows that $\sigma(K_1) \subseteq \overline{F_2}$ is a splitting field of S over F. By Lemma 10.3, $\sigma(K_1) = K_2$. In other words, σ is the required isomorphism. The second statement follows immediately from the first, since $\overline{F_1}$ and $\overline{F_2}$ are splitting fields over F of the set of all non-constant polynomials in $F[x]$. □

Example 10.7. Theory versus computation: Consider the polynomial $f(x) = x^3 - 3x + 1$, which is irreducible over $\mathbb{Q}$. Let $\alpha \in \mathbb{R}$ be a real root. Then it is easy to check that $\alpha^2 - 2$ is also a root of $f(x)$. Since $f(x)$ has two roots in $\mathbb{Q}(\alpha)$, its third root also belongs to $\mathbb{Q}(\alpha)$. Thus, $\mathbb{Q}(\alpha)$ is the splitting field for $f(x)$ over $\mathbb{Q}$. Now consider $\beta := 1 + \alpha$. If we let $g(x)$ denote the minimal polynomial of β over $\mathbb{Q}$, then, by Theorem 10.4 above $g(x)$ splits over $\mathbb{Q}(\alpha)$. Arguing as in Example 4.6, it is not hard to show that $g(x) = x^3 - 3x^2 - 3$. Since β is a root of $g(x)$, $x - \beta$ divides $g(x)$. Upon doing so, we get $g(x) = (x - \beta)(x^2 + (\beta - 3)x + \beta^2 - 3\beta)$. Using the quadratic formula, the other two roots of $g(x)$ are

$$\frac{-(\beta - 3) \pm \sqrt{(\beta - 3)^2 - 4(\beta^2 - 3\beta)}}{2}.$$

These numbers belong to $\mathbb{Q}(\alpha)$ and thus can be written as $\mathbb{Q}$-linear combinations of $1, \alpha, \alpha^2$. How can we find such linear combinations? Clearly since $\beta = 1 + \alpha$, the difficulty lies in writing $\sqrt{(\beta - 3)^2 - 4(\beta^2 - 3\beta)}$ as a

$\mathbb{Q}$-linear combination of $1, \alpha, \alpha^2$. So, we wish to find $\gamma = a + b\alpha + c\alpha^2$ such that $\gamma^2 = (\beta - 3)^2 - 4(\beta^2 - 3\beta)$. Using the fact that $\alpha^3 - 3\alpha + 1 = 0$, one can show that

$$\gamma^2 = (a^2 - 2bc) + (2ab + 6bc - c^2)\alpha + (b^2 + 3c^2 + 2ac)\alpha^2.$$

On the other hand, using that $\beta = 1 + \alpha$, an easy calculation yields

$$(\beta - 3)^2 - 4(\beta^2 - 3\beta) = 12 + (-3)\alpha^2.$$

Thus finding γ is equivalent to solving the system of equations

$$\begin{aligned} 12 &= a^2 - 2bc \\ 0 &= 2ab + 6bc - c^2 \\ -3 &= b^2 + 3c^2 + 2ac \end{aligned}$$

over $\mathbb{Q}$. Computationally, this is decidedly a non-trivial problem! On the other hand, this system has two solutions over $\mathbb{Q}$, since $g(x)$ has three roots in $\mathbb{Q}(\alpha)$. □

Lecture 11: Friday September 17. The second type of field extension we wish to discuss is a *separable* field extension.

Definition 11.1. Let F be a field and $f(x) \in F[x]$ be irreducible over F. We say that $f(x)$ is a separable polynomial, if it has distinct roots in $\overline{F}$. The element $\alpha \in \overline{F}$ is separable, if its minimal polynomial is a separable polynomial. The algebraic extension $F \subseteq K$ is a separable extension, if every element in K is separable over F. □

The Galois Correspondence Theorem, which we will see in later lectures, applies to finite extensions $F \subseteq K$ that are both splitting fields and separable. Such extensions are called a Galois extensions. There are a number of important properties of separable extensions that one must prove, in order to state the Galois correspondence theorem in full generality. This is because, separability is not automatic when F has positive characteristic, though, as we will see below, every algebraic extension is separable when F has characteristic zero.

Proposition 11.2. *Let F be a field and $f(x) \in F[x]$ be a non-constant polynomial. Then:*

(i) $f(x)$ has distinct roots in $\overline{F}$ if and only $f'(x) \neq 0$ and $f(x)$ and $f'(x)$ have no common, non-constant, factor in $F[x]$.
(ii) If $f(x)$ is irreducible, then $f(x)$ is separable if and only if $f'(x) \neq 0$.
(iii) If F has characteristic zero, equivalently, $\mathbb{Q} \subseteq F$, then every algebraic extension of F is a separable extension.

Proof Part (i) is just problem 8 on Homework 1, stated in a different way.

For part (ii), since $f(x)$ is irreducible, and $f'(x)$ has degree less than the degree of $f(x)$, the GCD of $f(x)$ and $f'(x)$ is 1, unless $f'(x) = 0$. Thus, by (i), $f(x)$ has distinct roots if and only if $f'(x) \neq 0$.

Part (iii) follows immediately from the definitions and part (ii), since if F has characteristic zero, and $f(x)$ is not a constant, then $f'(x) \neq 0$. □

Example 11.3. This example shows how an irreducible polynomial over a field of positive characteristic can fail to be separable. Suppose $F := \mathbb{Z}_3(y)$, the rational function field in one variable over $\mathbb{Z}_3$. Then $f(x) = x^3 - y$ is irreducible, since y is not a cube in $\mathbb{Z}_3(y)$. (Check this!) Let $\alpha \in \overline{F}$ be a root of $f(x)$, so that $\alpha^3 = y$. Now suppose $\beta \in \overline{F}$ is also a root of $f(x)$. Then $\beta^3 = y = \alpha^3$, and hence $0 = \alpha^3 - \beta^3 = (\alpha - \beta)^3$, and thus $\alpha - \beta = 0$, so $\alpha = \beta$. It follows that α is the only root of $f(x)$, and thus a root of multiplicity three, i.e., over $\overline{F}$, $f(x) = (x - \alpha)^3$. Note also, that $f'(x) = 0$. □

Finite separable extensions $F \subseteq K$ have a very nice property - namely that every such extension is a simple extension, i.e., $K = F(\alpha)$, for some $\alpha \in K$. The element α is called a primitive element for the extension $F \subseteq K$.

Primitive Element Theorem 11.4. *Suppose that $F \subseteq K$ is a finite extension of fields and K is separable over F. Then there exists $\alpha \in K$ such that $K = F(\alpha)$. In particular, if F has characteristic zero, then every finite extension of F is a simple extension, i.e., admits a primitive element.*

Proof. We must consider two cases. The first case is that F is a finite field. Thus K is also a finite field and upon writing K^* for $K\backslash\{0\}$, we have that K^* is a finite abelian group under multiplication. We will see later in the semester that this means we can write

$$K^* \cong C_{n_1} \times C_{n_2} \times \cdots \times C_{n_d},$$

where each C_j is a cyclic group (written multiplicatively) of order n_j and $n_1|n_2|\cdots|n_d$. It follows that every $\alpha \in K^*$ satisfies $\alpha^{n_d} = 1$. Thus, in K there are $n_1 \cdot n_2 \cdots n_d$ roots of the polynomial $x^{n_d} - 1$. This is a contradiction unless $d = 1$, i.e., K^* is cyclic. Thus, there exists $\alpha \in K$ such that every non-zero element in K (including those in F!) is of the form α^r, for some $r \in \mathbb{Z}$. We clearly have $K = F(\alpha)$.

Suppose that F is an infinite field. Since K is finite over F, we can write $K = F(u_1, \ldots, u_n)$, with each $u_i \in K$ and separable over F. The proof is by induction on n, with the base case $n = 1$ being trivial. Suppose $n > 1$. Then $K = F(u_1, \ldots, u_{n-1})(u_n)$. By induction there exists $v \in F(u_1, \ldots, u_{n-1})$ such that $F(u_1, \ldots, u_{n-1}) = F(v)$. Thus,

$$K = F(u_1, \ldots, u_{n-1})(u_n) = F(v)(u_n) = F(v, u_n).$$

This shows that the theorem reduces to the case $n = 2$. So we start again with $K = F(u, v)$, with $u, v \in K$, separable over F. We lso assume that K is not equal to $F(u)$ or $F(v)$, for then there is nothing to prove. Let $f(x)$ be the minimal polynomial of u over F and $g(x)$ be the minimal polynomial of v over F. Thus, $f(x)$ and $g(x)$ have distinct roots in $\overline{F}$. We will show that there exist infinitely many $0 \neq \lambda \in F$ such that $K = F(u + \lambda v)$. For $\alpha := u + \lambda v$, with λ a non-zero element of F, note that if $v \in F(\alpha)$, then $u = \alpha - \lambda v \in F(\alpha)$, so $K = F(\alpha)$. Thus, it is enough to find $0 \neq \lambda \in F$ such that $v \in F(\alpha)$, for $\alpha = u + \lambda v$.

For what choices of λ, other than $\lambda = 0$, can v fail to belong to $F(\alpha)$? Fix such a $\lambda \neq 0$, so that $\alpha = u + \lambda v$ is also fixed. Since $v \notin F(\alpha)$, the minimal polynomial $p(x)$ of v over $F(\alpha)$ has to have degree greater than one. Note that $p(x)$ divides $g(x)$. Consider the polynomial $h(x) := f(\alpha - \lambda x)$. Then $h(v) = 0$. Thus $p(x)$ also divides $h(x)$. Thus, in $\overline{F}$, $g(x)$ and $h(x)$ have a common root other than v, say v_0. Since $g(x)$ is separable over F, $v \neq v_0$. Since $0 = h(v_0) = f(\alpha - \lambda v_0)$, $\alpha - \lambda v_0 =: u_0$, for u_0 a root of $f(x)$. Notice that $v_0 \neq v$ implies $u_0 \neq u$. Thus $\lambda v_0 = \alpha - u_0 = (u + \lambda v) - u_0$. Solving for λ, we obtain

$$\lambda = \frac{u - u_0}{v_0 - v}.$$

In other words, if $F(\alpha) \neq K$, then $\lambda = \frac{u-u_0}{v_0-v}$, for roots u, u_0 of $f(x)$ and roots v, v_0 of $g(x)$. Since there are only finitely many such combinations, there are only finitely many $\lambda \in F$ for which $F(u + \lambda v) \neq K$. Thus, we have infinitely many choices of (non-zero) $\lambda \in F$ so that $K = F(\alpha)$, with $\alpha = u + \lambda v$. Finally, the last statement follows immediately from Proposition 11.2 and what we have just shown. □

Remark 11.5. Note that the existence of a primitive element in the case that F is finite did not use the separability assumption. It turns out that *every* finite extension of a finite field is automatically a separable extension (and also a splitting field).

Example 11.6. Here is the standard example (modulo some details) of a finite extension that is not a simple extension. Note that by what we have shown above, the base field F must be neither finite, nor have characteristic zero. Take $F := \mathbb{Z}_p(x^p, y^p)$ and $K := \mathbb{Z}_p(x, y)$, for indeterminates x and y over $\mathbb{Z}_p$, so that $K = F(x, y)$. Since $x^p \in F$ and $y^p \in F$, $[K : F] < \infty$. It can be shown that $[K : F] = p^2$. Suppose K were a simple extension of F. Then there would exist $\alpha \in K$ such that $K = F(\alpha)$. Thus, the minimal polynomial for α over F would have degree p^2. However, $\alpha^p \in F$ (Check this!), which show that α satisfies a polynomial of degree p or less over F, a contradiction. Therefore, K is not a simple extension of F.

Lecture 12: Monday September 20. Let K be a field. The the set $\text{Aut}(K)$ of automorphisms of $K$is easily seen to be a group under composition. This group can be small or large, e.g., when $K = \mathbb{Q}$, $\text{Aut}(K)$ is just the identity, while if $K = \mathbb{C}$, $\text{Aut}(K)$ is an uncountable group. We are mainly interested in the case that K is a finite extension of a field F and the automorphisms in question fix F. Note that since the set of automorphisms of a field K is a group under composition, the automorphisms of K fixing F is a subgroup, since a composition of automorphisms fixing F, fixes F, and the inverse of an automorphism fixing F also fixes F. Thus $\text{Gal}(K/F)$ is a group. This give rise to the following definition.

Definition 12.1. Given the field extension $F \subseteq K$, the Galois group of K over F, denoted $\mathrm{Gal}(K/F)$ is the group of automorphisms of K leaving F fixed.

We make a few comments concerning $\mathrm{Gal}(K/F)$ when K is a finite extension of F. Note, that in this case, $K = F(\alpha_1, \dots, \alpha_n)$ for some $\alpha_i \in K$.

(a) If $\gamma \in K$, then (as we have seen before), we can write $\gamma = h(\alpha_1, \dots, \alpha_n)$, where $h(x_1, \dots, x_n)$ is a polynomial with coefficients in F. Thus, for $\sigma \in \mathrm{Gal}(K/F)$,

$$\sigma(\gamma) = \sigma(h(\alpha_1, \dots, \alpha_n)) = h(\sigma(\alpha_1), \dots, \sigma(\alpha_n)),$$

since σ fixes the elements of F. Thus, σ is determined by the values $\sigma(\alpha_1), \dots, \sigma(\alpha_n)$.

(b) $\mathrm{Gal}(K/F)$ is a finite group. To see this, let $f_i(x)$ be the minimal polynomial of α_i over F. Then for any $\sigma \in \mathrm{Gal}(K/F)$, $\sigma(\alpha_i)$ is a root of $f_i(x)$ (by Proposition 2.1). Since each $f_i(x)$ has finitely many roots and there are only finitely many α_i, there can only be finitely many values $\sigma(\alpha_i)$ as σ varies over $\mathrm{Gal}(K/F)$ and i runs from 1 to n. By the observation (a), $\mathrm{Gal}(K/F)$ must be finite.

(c) A nontrivial fact is that the order of $\mathrm{Gal}(K/F)$ is less than or equal to $[K : F]$. We have basically proven this in the case K is separable over F, and in particular, when F has characteristic zero. Indeed, by the Primitive Element Theorem $K = F(\alpha)$. Now $|\mathrm{Gal}(K/F)| \leq [K : F]$ follows from Proposition 2.1 part (v) .

Let us now calculate a Galois group. The crucial ingredient to this calculation is Problem 7 in Homework 1.

Example 12.2. Consider the extension $\mathbb{Q} \subseteq \mathbb{Q}(\sqrt{2}, \sqrt{3})$. We wish to show that $G := \mathrm{Gal}(\mathbb{Q}(\sqrt{2}, \sqrt{3})$ is isomorphic to $\mathbb{Z}_2 \times \mathbb{Z}_2$. Note that if $\sigma \in G$, then $\sigma(\sqrt{2}) = \pm\sqrt{2}$ and $\sigma(\sqrt{3}) = \pm\sqrt{3}$. We will show that these four possibilities occur. For this, we first note that $x^2 - 2$ is irreducible over $\mathbb{Q}(\sqrt{3})$ and $x^2 - 3$ is irreducible over $\mathbb{Q}(\sqrt{2})$. This, follows, for example, since $\sqrt{3} \notin \mathbb{Q}(\sqrt{2})$. To see this, suppose $(a + b\sqrt{2})^2 = \sqrt{3}$. This leads to the equations $a^2 + 2b^2 = 3$ and $2ab = 0$, which has no solutions in $\mathbb{Q}$.

Now, let id denote the identity automorphism on $\mathbb{Q}(\sqrt{2}, \sqrt{3})$. Set $F_1 := \mathbb{Q}(\sqrt{2})$. Since $x^2 - 3$ is irreducible over F_1, there exist two automorphisms fixing F_1, $id : F_1(\sqrt{3}) \to F_1(\sqrt{3})$ which is the identity, i.e., takes $\sqrt{3}$ to $\sqrt{3}$ and $\sigma_2 : F_1(\sqrt{3}) \to F(\sqrt{3})$ taking $\sqrt{3}$ to $-\sqrt{3}$. This follows either from Proposition 2.1 or Problem 7. Since $F_1(\sqrt{3}) = \mathbb{Q}(\sqrt{2}, \sqrt{3})$ we have two elements in G. Similarly, if we set $F_2 := \mathbb{Q}(\sqrt{3})$, then $x^2 - 2$ is irreducible over F_2, and we have an automorphism $\sigma_3 : F_2(\sqrt{2}) \to F_2(\sqrt{2})$ taking $\sqrt{2}$ to $-\sqrt{2}$ that fixes F_2. This gives a third element of G. Finally, Let $\phi : F_1 \to F_1$ be the automorphism taking $\sqrt{2} \to -\sqrt{2}$, which exists by Proposition 2.1. Note that for $f(x) := x^2 - 3$, $f^{\phi}(x) = f(x)$, so we apply Problem 7 to get an isomorphism $\sigma_4 : F_2(\sqrt{3}) \to F_2(\sqrt{3})$ extending ϕ that takes $\sqrt{3}$ to $-\sqrt{3}$. In other words, σ_4 is an automorphism of $\mathbb{Q}(\sqrt{2}, \sqrt{3})$ which fixes $\mathbb{Q}$ and takes $\sqrt{2}$ to $-\sqrt{2}$ and $\sqrt{3}$ to $-\sqrt{3}$.

Let us now look at the group relations. All of the elements of G are determined by their effect on $\sqrt{2}$ and $\sqrt{3}$. So for example, $\sigma_2^2(\sqrt{2}) = \sigma_2(\sigma_2(\sqrt{2})) = \sigma_2(\sqrt{2}) = \sqrt{2}$ and $\sigma_2^2(\sqrt{3}) = \sigma_2(\sigma_2(\sqrt{3})) = \sigma_2(-\sqrt{3}) = - - \sqrt{3} = \sqrt{3}$. In other words, $\sigma_2^2 = id$. Similarly, the other non-identity elements of G also have order two. By elementary group theory, this shows G is isomorphic to $\mathbb{Z}_2 \times \mathbb{Z}_2$, since $\mathbb{Z}_2 \times \mathbb{Z}_2$ and $\mathbb{Z}_4$ are the only groups of order four. Alternately, one can make a group table for G, by calculating all products. For example, to calculate $\sigma_3\sigma_4$, we see: $\sigma_3\sigma_4(\sqrt{2}) = \sigma_3(-\sqrt{2}) = \sqrt{2}$ and $\sigma_3\sigma_4(\sqrt{3}) = \sigma_3(-\sqrt{3}) = -\sqrt{3}$. Thus, $\sigma_3\sigma_4 = \sigma_2$. The other relations can be obtained similarly. □

Lecture 13: Wednesday September 22. We calculate two more Galois groups.

Example 13.1. Consider $f(x) = x^3 + x + 1 \in \mathbb{Z}_2[x]$. Notice that $f(x)$ does not have a root in $\mathbb{Z}_2$, and thus $f(x)$ is irreducible over $\mathbb{Z}_2$, since it has degree three. Let $\alpha \in \overline{\mathbb{Z}_2}$ be a root of $f(x)$ and consider $K = \mathbb{Z}_2(\alpha)$. Thus $[K : \mathbb{Z}_2] = 3$, so $|K| = 8$. Now, since $\mathbb{Z}_2$ has characteristic two and $\alpha^3 + \alpha + 1 = 0$, we have,

$$\begin{aligned} 0 &= (\alpha^3 + \alpha + 1)^2 \\ 0 &= (\alpha^3)^2 + \alpha^2 + 1 \\ 0 &= (\alpha^2)^3 + \alpha^2 + 1 \end{aligned}$$

which shows that α^2 is a root of $f(x)$. Similarly, we can see that α^4 is also a root of $f(x)$. Notice that these are distinct roots, since for example, if $\alpha = \alpha^4$, then $\alpha^3 = 1$ and this contradicts the fact that $f(x)$ is the minimal polynomial of α over $\mathbb{Z}_2$. Notice that this shows that K is the splitting field of $f(x)$ over $\mathbb{Z}_2$.

By Proposition 2.1, at most three automorphisms of K fixing $\mathbb{Z}_2$, one of which is the identity automorphism. Again, by Proposition 2.1, there is an automorphism σ of K taking α to α^2. Notice that

$$\sigma^2(\alpha) = \sigma(\alpha^2) = \sigma(\alpha)^2 = \alpha^4,$$

so that σ^2 take α to the third root of $f(x)$. This already tells us that $\mathrm{Gal}(K/F) = \langle\sigma\rangle \cong \mathbb{Z}_3$. However direct calculation shows that $\sigma^3(\alpha) = \alpha^8$. Since K^* is a group of order seven, $\alpha^7 = 1$, and thus $\alpha^8 = \alpha$, showing that $\sigma^3 = id$, which also shows $\mathrm{Gal}(K/F) \cong \mathbb{Z}_3$. □

Example 13.2. Let us consider the Galois group of the splitting field of $x^3 - 2$ over $\mathbb{Q}$. We start by letting $\sqrt[3]{2}$ denote the real cube root of 2 and ϵ denote a (fixed) primitive cube root of unity (e.g., $\epsilon = e^{\frac{2\pi i}{3}}$). Then $\epsilon \neq \epsilon^2$ and $(\epsilon)^3 = (e^2)^3 = 1$. It follows that $\sqrt[3]{2}$, $\sqrt[3]{2}\epsilon$, and $\sqrt[3]{2}\epsilon^2$ are the three distinct roots of $x^3 - 2$. Note also, that any field containing $\sqrt[3]{2}$ and ϵ contains $\sqrt[3]{2}$, $\sqrt[3]{2}\epsilon$, $\sqrt[3]{2}\epsilon^2$, while on the other hand, any field containing $\sqrt[3]{2}$, $\sqrt[3]{2}\epsilon$, $\sqrt[3]{2}\epsilon^2$ contains $\sqrt[3]{2}$ and ϵ. Thus, we set $K := \mathbb{Q}(\sqrt[3]{2}, \epsilon) = \mathbb{Q}(\sqrt[3]{2}, \sqrt[3]{2}\epsilon, \sqrt[3]{2}\epsilon^2)$, the splitting field of $x^3 - 2$ over $\mathbb{Q}$.

We first note that $x^3 - 2$ and $x^2 + x + 1$ are irreducible over $\mathbb{Q}$, since neither has a root in $\mathbb{Q}$. It follows that $[\mathbb{Q}(\sqrt[3]{2}) : \mathbb{Q}] = 3$ and $[\mathbb{Q}(\epsilon) : \mathbb{Q}] = 2$. Moreover, $x^3 - 2$ remains irreducible over $\mathbb{Q}(\epsilon)$ and $x^2 + x + 1$ is irreducible over each of $L_0 := \mathbb{Q}(\sqrt[3]{2}), L_1 := \mathbb{Q}(\sqrt[3]{2}\epsilon)$, and $L_2 := \mathbb{Q}(\sqrt[3]{2}\epsilon^2)$. One can either verify these statements directly, or use HW 1 Problem 3 part (iv), since for example, if $L. = \mathbb{Q}(\epsilon)$ and $M = \mathbb{Q}(\sqrt[3]{2})$, then $LM = \mathbb{Q}(\epsilon, \sqrt[3]{2})$, which then gives $[\mathbb{Q}(\epsilon, \sqrt[3]{2}) : \mathbb{Q}] = 2 \cdot 3 = 6$. This implies that $[\mathbb{Q}(\epsilon, \sqrt[3]{2}) : \mathbb{Q}(\epsilon)] = 2$ (for example), which is equivalent to saying that $x^3 - 2$ is irreducible over $\mathbb{Q}(\epsilon)$. These facts will enable us to apply Problem 7 from HW 1 in a number of different ways.

To start, let us note that if we take $F = F_0 = \mathbb{Q}$ and $f(x) = f_0(x) = x^3 - 2$, and σ the identity automorphism in Problem 7, then we get three isomorphisms : (i) $\sigma_0 : L_0 \to L_0$, (ii) $\sigma_1 : L_0 \to L_1$, and (iii) $\sigma_2 : L_0 \to L_2$, for which $\sigma_0(\sqrt[3]{2}) = \sqrt[3]{2}$, (so σ_0 is the identity map), $\sigma_1(\sqrt[3]{2}) = \sqrt[3]{2}\epsilon$, and $\sigma_2(\sqrt[3]{2}) = \sqrt[3]{2}\epsilon^2$.

We now show that each σ_i can be extended in two ways to automorphisms of K. To see this, we first note that for any $0 \leq i \leq 2$, $K = L_i(\epsilon) = L_i(\epsilon^2)$. (Check this.) We first extend σ_i as follows : In Problem 7, take $F = L_0$ and $F_0 = L_i$ and $\sigma = \sigma_i$. We also take $f(x) = f_0(x) = x^2 + x + 1$. Note that $x^2 + x + 1$ has coefficients in $\mathbb{Q}$, so that any isomorphism applied to the coefficients of $x^2 + x + 1$ must fix those coefficients. We first take $\alpha = \epsilon = \alpha_0$. Then, by Problem 7 there exist isomorphisms $\hat{\sigma_i} : L_0(\epsilon) \to L_i(\epsilon)$ extending σ_i and with $\hat{\sigma_i}(\epsilon) = \epsilon$. Since $L_0(\epsilon) = L_i(\epsilon) = K$, we obtain three automorphisms of K, namely, $\hat{\sigma_0}, \hat{\sigma_1}$, and $\hat{\sigma_2}$.

On the other hand, starting again with the same data, if we take $\alpha = \epsilon$ and $\alpha_0 = \epsilon^2$, then we may apply Problem 7 again to extend each σ_i to an isomorphism $\tilde{\sigma_i} : L_0(\epsilon) \to L_i(\epsilon^2)$ taking ϵ to ϵ^2. Again, each $\tilde{\sigma_i}$ is an automorphism of K, so we obtain six automorphisms of K altogether.

In terms of the roots, the following list shows how each automorphism permutes the roots of $X^3 - 2$:

(i) $\hat{\sigma_0} : \sqrt[3]{2} \to \sqrt[3]{2},\ \sqrt[3]{2}\epsilon \to \sqrt[3]{2}\epsilon,\ \sqrt[3]{2}\epsilon^2 \to \sqrt[3]{2}\epsilon^2$.
(ii) $\hat{\sigma_1} : \sqrt[3]{2} \to \sqrt[3]{2}\epsilon,\ \sqrt[3]{2}\epsilon \to \sqrt[3]{2}\epsilon^2,\ \sqrt[3]{2}\epsilon^2 \to \sqrt[3]{2}$ (since $\hat{\sigma_1}(\sqrt[3]{2}\epsilon) = \hat{\sigma_1}(\sqrt[3]{2}) \cdot \hat{\sigma_1}(\epsilon) = \sqrt[3]{2}\epsilon \cdot \epsilon = \sqrt[3]{2}\epsilon^2$).
(iii) $\hat{\sigma_2} : \sqrt[3]{2} \to \sqrt[3]{2}\epsilon^2,\ \sqrt[3]{2}\epsilon \to \sqrt[3]{2},\ \sqrt[3]{2}\epsilon^2 \to \sqrt[3]{2}\epsilon$.
(iv) $\tilde{\sigma_0} : \sqrt[3]{2} \to \sqrt[3]{2},\ \sqrt[3]{2}\epsilon \to \sqrt[3]{2}\epsilon^2,\ \sqrt[3]{2}\epsilon^2 \to \sqrt[3]{2}\epsilon$.
(v) $\tilde{\sigma_1} : \sqrt[3]{2} \to \sqrt[3]{2}\epsilon,\ \sqrt[3]{2}\epsilon \to \sqrt[3]{2},\ \sqrt[3]{2}\epsilon^2 \to \sqrt[3]{2}\epsilon^2$.
(vi) $\tilde{\sigma_2} : \sqrt[3]{2} \to \sqrt[3]{2}\epsilon^2,\ \sqrt[3]{2}\epsilon \to \sqrt[3]{2}\epsilon,\ \sqrt[3]{2}\epsilon^2 \to \sqrt[3]{2}$.

How do we know that we have all of the automorphisms of K? Since $K = \mathbb{Q}(\sqrt[3]{2}, \sqrt[3]{2}\epsilon, \sqrt[3]{2}\epsilon^2)$, every element in K is a polynomial expression in $\sqrt[3]{2}, \sqrt[3]{2}\epsilon, \sqrt[3]{2}\epsilon^2$ with coefficients in $\mathbb{Q}$. Thus, any automorphism of K is determined by its effect on $\sqrt[3]{2}, \sqrt[3]{2}\epsilon, \sqrt[3]{2}\epsilon^2$. Since any automorphism of K permutes $\sqrt[3]{2}, \sqrt[3]{2}\epsilon, \sqrt[3]{2}\epsilon^2$ (they are roots of the same polynomial with coefficients in $\mathbb{Q}$), there cannot be more automorphisms of K than permutations of $\sqrt[3]{2}, \sqrt[3]{2}\epsilon, \sqrt[3]{2}\epsilon^2$. Alternately, since $\mathbb{Q}$ has characteristic zero, the extension is a simple extension, and thus, by Proposition 2.1, the order of the Galois group is less than or equal to the degree of the extension. Since the extension has degree 6, there cannot be any more automorphisms of K fixing $\mathbb{Q}$ than the ones we have accounted for. Thus, the Galois group of of the splitting field of $x^3 - 2$ is the group of automorphisms $\sigma_0, \dots, \sigma_5$ above. This is a non-abelian group of order 6, and therefore is isomorphic to the symmetric group S_3. We can also see that G is isomorphic to S_3 because every permutation of the three roots appears as an automorphism of K. □

Remark 13.3. Let F be a field and $f(x) \in F[x]$ a polynomial with n distinct roots in $\overline{F}$. If we let K denote the splitting field of $f(x)$ over F, then any element in $\mathrm{Gal}(K/F)$ permutes the roots of $f(x)$ and thus can be identified with an element of S_n, the symmetric group on n letters. In other words, $\mathrm{Gal}(K/F)$ is isomorphic to a subgroup of S_n. On the other hand, it is important to note that not every permutation of roots gives rise to an automorphism of K fixing F. For example, since $\mathbb{Q}$ has characteristic zero, the field extension given in Example 10.2 is a simple extension. If we let $\alpha \in \mathbb{Q}(\sqrt{2}, \sqrt{3})$ be a primitive element, e.g., $\alpha = \sqrt{2} + \sqrt{3}$, then since $\mathbb{Q}(\sqrt{2}, \sqrt{3})$ is a splitting field over $\mathbb{Q}$, $f(x)$, the minimal polynomial of α over $\mathbb{Q}$, splits over $\mathbb{Q}(\sqrt{2}, \sqrt{3})$. It follows that $\mathbb{Q}(\alpha) = \mathbb{Q}(\alpha, \alpha_2, \alpha_3, \alpha_4)$, where the α_i are the distinct roots of $f(x)$. Since the Galois group of this extension is $\mathbb{Z}_2 \times \mathbb{Z}_2$, not every permutation of the roots of $f(x)$ gives rise to an automorphism of its splitting field, otherwise the Galois group in this case would be S_4.

Lecture 14: Friday September 24. We begin with two important definitions.

Definitions 14.1 Let $F \subseteq K$ be a finite extension of fields. (i) If $[K : F] = |\mathrm{Gal}(K/F)|$, we say that K is a Galois extension of F, or that K is *Galois* over F.

(ii) Let K be a field and G a group of automorphisms of K. Then the *fixed field of* G the set of elements $\alpha \in K$ such that $\sigma(\alpha) = \alpha$, for all $\sigma \in G$. It is easy to check that the fixed field of G is a subfield of K. □

Remarks 14.2. (i) Let $F \subseteq K$ be a finite extension and let F_0 be the fixed field of $\mathrm{Gal}(K/F)$. Then $F \subseteq F_0 \subseteq K$ and $\mathrm{Gal}(K/F) = \mathrm{Gal}(K/F_0)$.

(ii) The calculations in Examples 12.1, 13.1 and 13.2 show respectively that $\mathbb{Q}(\sqrt{2}, \sqrt{3})$ is Galois over $\mathbb{Q}$, the splitting field of $x^3 + x + 1$ over $\mathbb{Z}_2$ is Galois over $\mathbb{Z}_1$, and $\mathbb{Q}(\sqrt[3]{2}, \epsilon)$ is Galois over $\mathbb{Q}$. Note that in each of these cases, the larger field is a splitting field over the base field. We will see in Monday's lecture that this is an important component of the Galois property. □

The following theorem (due to E. Artin) is the key component in the proof of the Galois Correspondence Theorem.

Theorem 14.3. *Let K be any field, G a finite group of automorphisms of K and F_0 the fixed field of G. Then $[K : F_0] \leq |G|$.*

Proof. Suppose $|G| = n$ and there exists $v_1, \dots, v_{n+1} \in K$ linearly independent over F_0. Let $\sigma_1, \dots, \sigma_n$ be the elements of G, with $\sigma_1 = id$. Then the homogeneous system of linear equations over K

$$\begin{aligned} \sigma_1(v_1)x_1 + \cdots + \sigma_1(v_{n+1})x_{n+1} &= 0 \\ &\vdots \\ \sigma_n(v_1)x_1 + \cdots + \sigma_n(v_{n+1})x_{n+1} &= 0 \end{aligned}$$

has a non-trivial solution in K^{n+1}. Among all such solutions, choose one with the fewest non-zero entries. Suppose this solution, as an element of K^{n+1} has r non-zero entries. By changing the order of the v_j if necessary, we may assume that the solution is $\begin{pmatrix} a_1 \\ \vdots \\ a_r \\ 0 \\ \vdots \\ 0 \end{pmatrix}$, with each a_i a non-zero element of K. We first note that $r > 1$. Otherwise, the first equation in the system above becomes $v_1 \cdot a_1 = 0$, since $\sigma_1 = id$, and this is a contradiction. We next note any non-zero multiple of a solution to the system above is also a solution to the system above, so that we may assume $a_r = 1$. Thus, we have a solution of the form $\begin{pmatrix} a_1 \\ \vdots \\ a_{r-1} \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$. Finally, we note that not all of the a_j belong to F_0, otherwise upon setting $x_i = a_i$, the first equation would yield an equation of linear dependence on the v_j. Thus, some $a_i \notin F_0$. Again, by re-ordering if necessary, we can assume $a_1 \notin F_0$. Finally, We now have the following system of equations

$$\begin{aligned} \sigma_1(v_1)a_1 + \cdots + \sigma_1(v_{r-1})a_{r-1} + \sigma_1(v_r) \cdot 1 &= 0 \\ &\vdots \\ \sigma_n(v_1)a_1 + \cdots + \sigma_n(v_{r-1})a_{r-1} + \sigma_n(v_r) \cdot 1 &= 0. \end{aligned}$$

Now, since $a_1 \notin F_0$, a_1 is not fixed by some element of G, say $\sigma_k(a_1) \neq a_1$, so $a_1 - \sigma_k(a_1) \neq 0$. We now apply σ_k to the second system of equations to get the following:

$$\begin{aligned} \sigma_k\sigma_1(v_1)\sigma_k(a_1) + \cdots + \sigma_k\sigma_1(v_{r-1})\sigma_k(a_{r-1}) + \sigma_k\sigma_1(v_r) \cdot 1 &= 0 \\ &\vdots \\ \sigma_k\sigma_n(v_1)\sigma_k(a_1) + \cdots + \sigma_k\sigma_n(v_{r-1})\sigma_k(a_{r-1}) + \sigma_k\sigma_n(v_r) \cdot 1 &= 0. \end{aligned}$$

Now, since $G = \{\sigma_k\sigma_1, \ldots, \sigma_k\sigma_n\}$, this last system of equations is the same as the system

$$\begin{aligned} \sigma_1(v_1)\sigma_k(a_1) + \cdots + \sigma_1(v_{r-1})\sigma_k(a_{r-1}) + \sigma_1(v_r) \cdot 1 &= 0 \\ &\vdots \\ \sigma_n(v_1)\sigma_k(a_1) + \cdots + \sigma_n(v_{r-1})\sigma_k(a_{r-1}) + \sigma_n(v_r) \cdot 1 &= 0. \end{aligned}$$

Thus, $\begin{pmatrix} \sigma_k(a_1) \\ \vdots \\ \sigma_k(a_{r-1}) \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ is also a solution to the original system of equations. Subtracting this new solution to the system from the original solution above, we have that $\begin{pmatrix} a_1 - \sigma_k(a_1) \\ \vdots \\ a_{r-1} - \sigma_k(a_{r-1}) \\ \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ is a non-trivial solution with fewer than r non-zero entries, which is a contradiction. Thus, we must have $[K : F_0] \leq |G|$. □

Corollary 14.4. (i) *Suppose G is a finite group of automorphisms of the field K and F_0 its fixed field. Then K is Galois over F_0 and $G = \mathrm{Gal}(K/F_0)$.*

(ii) *A finite extension $F \subseteq K$ is Galois if and only if F is the fixed field of* $\mathrm{Gal}(K/F)$.

Proof. For part (i), by the theorem above, we have

$$[K : F_0] \leq |G| \leq |\mathrm{Gal}(K/F_0)| \leq [K : F_0].$$

Thus, $[K : F_0] = |\mathrm{Gal}(K/F_0)|$, and hence K is Galois over F_0, and $G = \mathrm{Gal}(K/F_0)$.

For part (ii) Set $G := \mathrm{Gal}(K/F)$ and F_0 to be the fixed field of G. Thus G is the Galois group of K over F_0. Suppose K is Galois over F. Then we have

$$|G| \leq [K : F_0] \leq [K : F] = |G|,$$

from which we see $[K : F_0] = [K : F]$, which implies $F = F_0$. Conversely, suppose $F = F_0$. Then by the theorem above, we have

$$[K : F] = [K : F_0] \leq |G| \leq [K : F].$$

Thus, $|G| = [K : F]$, so K is Galois over F. □

Lecture 15: Monday September 27. We can now state the Galois Correspondence Theorem.

Galois Correspondence Theorem. *Let $F \subseteq K$ be a finite extension with K Galois over F. For a subgroup $H \subseteq Gal(K/F)$, write K^H for the fixed field of H. Then there is a one-to-one correspondence between the subgroups $H \subseteq$* $\mathrm{Gal}(K/F)$ *and the intermediate fields $F \subseteq E \subseteq K$ given by $H \to K^H$ and $E \to Gal(K/E)$. Moreover, for any intermediate field E and any subgroup H of* $\mathrm{Gal}(K/F)$:

(i) $[G : H] = [K^H : F]$ and $[E : F] = [G : \mathrm{Gal}(K/E)]$.
(ii) K is Galois over E.
(iii) E is Galois over F if and only if $\mathrm{Gal}(K/E)$ is a normal subgroup of $\mathrm{Gal}(K/F)$, in which case $\mathrm{Gal}(E/F)$ is isomorphc to $\mathrm{Gal}(K/F)/\mathrm{Gal}(K/E)$.

Proof of the Galois Correspondence Theorem. We start by proving all parts of the theorem except part (iii). To see that $H \to K^H$ is a one-to-one correspondence between the subgroups H of $\mathrm{Gal}(K/F)$ and the intermediate fields $F \subseteq E \subseteq K$, with inverse $E \to \mathrm{Gal}(K/E)$, it suffices to show that $H = \mathrm{Gal}(K/K^H)$ and $E = K^{\mathrm{Gal}(K/E)}$. Let H be a subgroup of $\mathrm{Gal}(K/F)$. Then, by Corollary 14.3, $H = \mathrm{Gal}(K/K^H)$, which is what we want. Now let $F \subseteq E \subseteq K$ be an intermediate field and set $H := \mathrm{Gal}(K/E)$. Then, by Corollary 14.3, K is Galois over K^H, with Galois group H. Thus,

$$|H| = [K : K^H] \geq [K : E] \geq |\mathrm{Gal}(K/E)| = |H|.$$

Thus, $[K : K^H] = [K : E]$, so $E = K^H = K^{\mathrm{Gal}(K/E)}$, which is what we want. This establishes the required one-to-one correspondence.

Now let $F \subseteq E \subseteq K$. To see that K is Galois over E, set $H := \mathrm{Gal}(K/E)$. Then by what we have shown above, E is the fixed field of H, so by Corollary 14.3, K is Galois over E.

Let H be a subgroup of $\mathrm{Gal}(K/F)$. Then K is Galois over K^H, and H is the Galois group of K over K^H, by Corollary 14.4. Thus, $|H| = [K : K^H]$. However, because $|\mathrm{Gal}(K/F)| = |H| \cdot [\mathrm{Gal}(K/F) : H]$ and $[K : F] = [K : K^H] \cdot [H' : F]$, it follows that $[\mathrm{Gal}(K/F) : H] = [K^H : F]$, since $[K : F] = |\mathrm{Gal}(K/F)|$.

Similarly, if E is an intermediate field, then $|\mathrm{Gal}(K/E)| = [K : E]$, since K is Galois over E. However, because $|\mathrm{Gal}(K/F)| = |\mathrm{Gal}(K/E)| \cdot [\mathrm{Gal}(K/F) : K^H]$ and $[K : F] = [K : E] \cdot [E : F]$, it follows that $[G : \mathrm{Gal}(K/E)] = [E : F]$, since $[K : F] = |\mathrm{Gal}(K/F)|$.

We need an extra result before we can prove part (iii) of the Galois Correspondence Theorem. Before proceeding with this, we revisit a couple of our examples.

Examples 15.2. (a) Let $F = \mathbb{Q}$ and $K := \mathbb{Q}(\sqrt{2}, \sqrt{3})$. In Example 12.2 we calculated the elements of $\mathrm{Gal}(K/F)$ and saw that $\mathrm{Gal}(K/F)$ isomorphic to $\mathbb{Z}_2 \times \mathbb{Z}_2$. Thus, there are three eelemnts of order two in

$\mathrm{Gal}(K/F)$ and they each generate a subgroup of order two. Using the same notation as in Example 12.2, these elements act on the roots of $(x^2-2)(x^2-3z)$as follows:

$$\begin{aligned} \sigma_2 &: \sqrt{2} \to \sqrt{2}, \quad \sqrt{3} \to -\sqrt{3} \\ \sigma_3 &: \sqrt{2} \to -\sqrt{2}, \quad \sqrt{3} \to \sqrt{3} \\ \sigma_4 &: \sqrt{2} \to -\sqrt{2}, \quad \sqrt{3} \to -\sqrt{3} \end{aligned}$$

Since the product of any two non-identity elements in $\mathbb{Z}_2 \times \mathbb{Z}_2$ equals the third non-identity element, the only proper subgroups are $\mathrm{Gal}(K/F)$ are those generated by the σ_i. Let us consider the subgroup H of $\mathrm{Gal}(K/F)$ generated by σ_4. Thus, $[\mathrm{Gal}(K/F) : H] = 2$. What is K^H, the fixed field of H? Certainly, $\sigma_4(\sqrt{6}) = \sqrt{6}$, so that $\mathbb{Q}(\sqrt{6})$ is contained in the fixed field of H. By the Galois correspondence Theorem, $[\mathrm{Gal}(K/F) : H] = [K^H : \mathbb{Q}]$. Since $\mathbb{Q}(\sqrt{6}) \subseteq K^H$ and $[\mathbb{Q}(\sqrt{2} : \mathbb{Q}]. = 2$, we must have $K^H = \mathbb{Q}(\sqrt{6})$.

Here is another way to see that $K^H = \mathbb{Q}(\sqrt{6})$. A typical element in K is of the form $\beta = a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6}$, with $a, b, c, d \in \mathbb{Q}$. Suppose $\beta = \sigma_4(\beta) = a - b\sqrt{2} - c\sqrt{3} + d\sqrt{6}$. Then $a = a$, $b = -b$, $c = -c$, and $d = d$. It follows that $b = 0 = c$, so that $\beta = a + b\sqrt{6}$, so $\beta \in \mathbb{Q}(\sqrt{6})$. Moreover, $\sigma_4(\sqrt{6}) = \sqrt{6}$, so that $\mathbb{Q}(\sqrt{6})$ is contained in the fixed field of H. Thus, $\mathbb{Q}(\sqrt{6})$ is the fixed field of H. If we apply either strategy to the other subgroups of $\mathrm{Gal}(K/F)$ we get the following correspondence

$$\begin{aligned} id &\longleftrightarrow \mathbb{Q}(\sqrt{2}, \sqrt{3}) \\ \langle \sigma_2 \rangle &\longleftrightarrow \mathbb{Q}(\sqrt{2}) \\ \langle \sigma_3 \rangle &\longleftrightarrow \mathbb{Q}(\sqrt{3}) \\ \langle \sigma_4 \rangle &\longleftrightarrow \mathbb{Q}(\sqrt{6}) \\ \mathrm{Gal}(K/F) &\longleftrightarrow \mathbb{Q}. \end{aligned}$$

Note that $\mathrm{Gal}(K/F)$ corresponds to $\mathbb{Q}$ by Corollary 14.4. Note also, that $\mathrm{Gal}(K/F)$ is abelian, so that every subgroup is a normal subgroup. Each intermediate field above is easily seen to be Galois over $\mathbb{Q}$ (e.g., by Proposition 2.1), and this agrees with part (iii) of the correspondence theorem.

(b) Let $F = \mathbb{Q}$ and $K = \mathbb{Q}(\sqrt[3]{2}, \epsilon) = \mathbb{Q}(\sqrt[3]{2}, \sqrt[3]{2}\epsilon, \sqrt[3]{2}\epsilon^2)$, the splitting field of $x^3 - 2$ over $\mathbb{Q}$. As we have seen in Example 13.2, K is Galois over $\mathbb{Q}$ and $\mathrm{Gal}(K/F)$ is isomorphic to S_3. Maintaining the notation from Example 3.2. So, we have the following elements in $\mathrm{Gal}(K/F)$ of order two:

$$\begin{aligned} \tilde{\sigma_0} &: \sqrt[3]{2}\epsilon \longleftrightarrow \sqrt[3]{2}\epsilon^2, \ \sqrt[3]{2} \text{ fixed} \\ \tilde{\sigma_1} &: \sqrt[3]{2} \longleftrightarrow \sqrt[3]{2}\epsilon, \ \sqrt[3]{2}\epsilon^2 \text{ fixed} \\ \tilde{\sigma_2} &: \sqrt[3]{2} \longleftrightarrow \sqrt[3]{2}\epsilon^2, \ \sqrt[3]{2}\epsilon \text{ fixed} \end{aligned}$$

and the following elements of order three:

$$\begin{aligned} \hat{\sigma_1} &: \sqrt[3]{2} \to \sqrt[3]{2}\epsilon \to \sqrt[3]{2}\epsilon^2 \to \sqrt[3]{2} \\ \hat{\sigma_2} &: \sqrt[3]{2} \to \sqrt[3]{2}\epsilon^2 \to \sqrt[3]{2}\epsilon \to \sqrt[3]{2}. \end{aligned}$$

It is easy to check that $\hat{\sigma_2} = \hat{\sigma_1}^2$. It follows that $\langle \hat{\sigma_1} \rangle$ is a subgroup of order three, and since any subgroup of order three is generated by an element of order three, this is the only subgroup of order three. Moreover, the groups $\langle \tilde{\sigma_i} \rangle$ are the subgroups of order two. We have now accounted for all proper subgroups of $\mathrm{Gal}(K/F)$, since any subgroup containing two $\hat{\sigma_i}$ cannot have order three, it must be the whole group. Similarly, because our subgroup of order three has index two, there cannot be any proper subgroups containing it.

Now to calculate the corresponding fixed fields, set $H := \langle \tilde{\sigma_0} \rangle$. Then $\sqrt[3]{2}$ is fixed by H, so $\mathbb{Q}(\sqrt[3]{2}) \subseteq K^H$. On the other hand, by part (i) of the Galois Correspondence Theorem, we have $3 = [\mathrm{Gal}(K/F) : H] = [K^H : \mathbb{Q}]$. Since $[\mathbb{Q}(\sqrt[3]{2}) : \mathbb{Q}] = 3$, this forces $\mathbb{Q}(\sqrt[3]{2}) = K^H$. A similar argument shows $K^{\langle \tilde{\sigma_1} \rangle} = \mathbb{Q}(\sqrt[3]{2}\epsilon^2)$ and $K^{\langle \tilde{\sigma_2} \rangle} = \mathbb{Q}(\sqrt[3]{2}\epsilon)$.

Let us write $C := \langle \hat{\sigma_1} \rangle$ for the remaining proper subgroup of $\mathrm{Gal}(K/F)$ and calculate K^C. Now $\epsilon = \frac{\sqrt[3]{2}\epsilon^2}{\sqrt[3]{2}\epsilon}$, so that

$$\hat{\sigma_1}(\epsilon) = \hat{\sigma_1}(\frac{\sqrt[3]{2}\epsilon}{\sqrt[3]{2}}) = \frac{\hat{\sigma_1}(\sqrt[3]{2}\epsilon)}{\hat{\sigma_1}(\sqrt[3]{2})} = \frac{\sqrt[3]{2}\epsilon^2}{\sqrt[3]{2}\epsilon} = \epsilon.$$

Thus, $\mathbb{Q}(\epsilon) \subseteq K^C$. By part (i) of the correspondence theorem, we have $2 = [\mathrm{Gal}(K/F) : C] = [K^C : \mathbb{Q}]$. Since $[\mathbb{Q}(\epsilon) : \mathbb{Q}] = 2$, this gives $K^C = \mathbb{Q}(\epsilon)$. We now have the following correspondence between the subgroups of $\mathrm{Gal}(K/F)$ and the intermediate fields between F and K.

$$\begin{aligned} id &\longleftrightarrow K \\ \langle \tilde{\sigma_0} \rangle &\longleftrightarrow \mathbb{Q}(\sqrt[3]{2}) \\ \langle \tilde{\sigma_1} \rangle &\longleftrightarrow \mathbb{Q}(\sqrt[3]{2}\epsilon^2) \\ \langle \tilde{\sigma_2} \rangle &\longleftrightarrow \mathbb{Q}(\sqrt[3]{2}\epsilon) \\ \langle \hat{\sigma_1} \rangle &\longleftrightarrow \mathbb{Q}(\epsilon) \\ \mathrm{Gal}(K/F) &\longleftrightarrow \mathbb{Q}, \end{aligned}$$

where the last correspondence follows from Corollary 14.4. □

Lecture 16: Wednesday September 29. It remains to prove part (iii) of the theorem. For this, we need the following important characterization of Galois extensions.

Theorem 16.1. *Let $F \subseteq K$ be a finite extension of fields. The following are equivalent:*

(a) K is Galois over F.
(b) F is the fixed field of $\mathrm{Gal}(K/F)$.
(c) K is separable over F and a splitting field over F.

Proof. The equivalence of (a) and (b) follows from Corollary 14.4. Now suppose K is Galois over F. Let $\alpha \in K\backslash F$ and write $f_\alpha(x)$ for the minimal polynomial of α over f. If we show that $f_\alpha(x)$ has distinct roots, and all of these roots are in K, then α is separable over F, and hence K is separable over F, and K is the splitting field of the set of polynomials $\{f_\alpha(x) \mid \alpha \in K\backslash F\}$.

Let $\sigma_1(\alpha), \ldots, \sigma_n(\alpha)$ be the distinct elements in the set $\{\sigma(\alpha) \mid \sigma \in \mathrm{Gal}(K/F)\}$. Without loss of generality, we may assume $\sigma_1(\alpha) = \alpha$, so that if we set $g(x) := (x - \sigma_1(\alpha)) \cdots (x - \sigma_n(\alpha))$, $g(\alpha) = 0$. We claim $g(x) \in F[x]$. Suppose the claim holds. Then $f_\alpha(x)$ divides $g(x)$. Since $g(x)$ has distinct roots, $f_\alpha(x)$ has distinct roots. On the other hand, the roots of $g(x)$ belong to K, and thus the roots of $f_\alpha(x)$ belong to K, which is what we want.

To prove the claim, write $g(x) = x^n + c_1x^{n-1} + \cdots + c_n$. Then we have a system of equations

$$\begin{aligned} c_1 &= -(\sigma_1(\alpha) + \cdots + \sigma_n(\alpha)) \\ c_2 &= \sum_{1 \le i < j \le n} \sigma_i(\alpha)\sigma_j(\alpha) \\ &\vdots \\ c_n &= (-1)^n \sigma_1(\alpha) \cdots \sigma_n(\alpha) \end{aligned}$$

Take $\sigma \in \mathrm{Gal}(K/F)$. Then $\sigma\sigma_1(\alpha), \ldots, \sigma\sigma_n(\alpha)$ are distinct, and thus by definition,

$$\{\sigma_1(\alpha), \ldots, \sigma_n(\alpha)\} = \{\sigma\sigma_1(\alpha), \ldots, \sigma\sigma_n(\alpha)\}.$$

It follows from this that

$$\begin{aligned}
\sigma(c_1) &= -(\sigma\sigma_1(\alpha) + \cdots + \sigma\sigma_n(\alpha)) = -(\sigma_1(\alpha) + \cdots + \sigma_n(\alpha)) = c_1 \\
\sigma(c_2) &= \sum_{1\le i<j\le n} (\sigma\sigma_i(\alpha))(\sigma\sigma_j(\alpha)) = \sum_{1\le i<j\le n} \sigma_i(\alpha)\sigma_j(\alpha) = c_2 \\
&\vdots \\
\sigma(c_n) &= (-1)^n(\sigma\sigma_1(\alpha))\cdots(\sigma\sigma_n(\alpha)) = (-1)^n\sigma_1(\alpha)\cdots\sigma_n(\alpha) = c_n.
\end{aligned}$$

Since $\sigma \in \mathrm{Gal}(K/F)$ is arbitrary it follows that each c_j belongs to the fixed field of $\mathrm{Gal}(K/F)$. By the equivalence of (a) and (b), we have that each $c_j \in F$, i.e., $f_\alpha(x) \in F[x]$, which proves the claim.

For the converse, suppose that K is separable over F and a splitting field over F. By the Primitive Element Theorem, there exists $\alpha \in K$ such that $K = F(\alpha)$. Let $f(x)$ denote the minimal polynomial of α over K, so that $f(x)$ is irreducible over F, and suppose that the degree of $f(x)$ equals d. Then, one the one hand, $f(x)$ has d distinct roots in $\overline{F}$, since α is separable over F. On the other hand, since K is a splitting field, by Theorem 10.4, $f(x)$ splits over $K = F(\alpha)$, so that $f(x)$ has d roots in $F(\alpha)$. By Proposition, $\mathrm{Gal}(K/F) = d = [K : F]$, which shows that K is Galois over F. □

Proof of the correspondence theorem continued. To finish the proof of the Galois Correspondence Theorem, we now have to show that if $F \subseteq E \subseteq K$ is an intermediate field, and $H = \mathrm{Gal}(K/E)$, then H is a normal subgroup of $\mathrm{Gal}(K/F)$ if and only if E is Galois over F, and in this case $\mathrm{Gal}(E/F) \cong \mathrm{Gal}(K/F)/\mathrm{Gal}(K/E)$. To begin, for such an E and H, suppose H is normal in $\mathrm{Gal}(K/F)$. Set $r := [\mathrm{Gal}(K/F) : H]$, and let $H, \tau_2H, \ldots, \tau_rH$ be the distinct cosets of H, so that $id, \tau_1, \ldots, \tau_r$ are a set of distinct coset representatives for H. Suppose we show that each $(\tau_i)_{|_E} \in \mathrm{Gal}(E/F)$ and the $(\tau_i)_{|_E}$ are distinct. Then,

$$[E : F] = [\mathrm{Gal}(K/F) : H] = r \le |\mathrm{Gal}(E/F)| \le [E : F].$$

Then $[E : F] = |\mathrm{Gal}(E/F)|$ which shows that E is Galois over F. For τ_i as above, suppose $\tau_i(e) = \gamma$, for $e \in E$ and $\gamma \in K$. Then for all $\sigma \in H$,

$$\tau_i\sigma\tau_i^{-1}(\gamma) = \tau_i\sigma(e) = \tau_i(e) = \gamma.$$

Thus, γ belongs to the fixed field of $\tau_iH\tau_i^{-1} = H$, since H is normal in $\mathrm{Gal}(K/F)$. But the fixed field of H is E, so $\gamma \in E$, i.e., $\tau_i(e) \in E$. This shows each each τ_i maps E into E, and since τ_i is one-to-one (say, as an F-linear transformation), τ_i must map E surjectively onto E. Therefore, $(\tau_i)_{|_E} \in \mathrm{Gal}(E/F)$. Now suppose that $(\tau_i)_{|_E} = (\tau_j)_{|_E}$. Then $\tau_i(e) = \tau_j(e)$, for all $e \in E$. In other words, $\tau_j^{-1}\tau_i$ fixes E. This means $\tau_j^{-1}\tau_i \in H$, so that $\tau_iH = \tau_jH$, so $\tau_i = .\tau_j$. Thus, the $(\tau_i)_{|_E}$ are distinct, which complete the proof that E is Galois over F.

For the converse, let E be an intermediate field which is Galois over F and $H := \mathrm{Gal}(K/E)$. We have to show that H is a normal subgroup of $\mathrm{Gal}(K/F)$. By Theorem 15.1, E is a splitting field over F. Let $\tau \in \mathrm{Gal}(K/F)$, $\sigma \in H$ and $\beta \in E$. Suppose we can show that $\tau^{-1}\sigma\tau(\beta) = \beta$. Then $\tau^{-1}\sigma\,\tau \in \mathrm{Gal}(K/E) = H$, which shows that H is normal in $\mathrm{Gal}(K/F)$.

To see that $\tau^{-1}\sigma\tau(\beta) = \beta$, it suffices to see that $\sigma\tau(\beta) = \tau(\beta)$. Now $\tau_{|_E} : E \to K \subseteq \overline{F}$ is a field homomorphism fixing F. Since E is a splitting field over F, $\tau_{|_E}(E) = E$, by Theorem 10.4. Thus, $\tau(\beta) \in E$. Therefore, since $\sigma \in H$, $\sigma(\tau(\beta)) = \tau(\beta)$, which gives what we want.

Finally, assume H is normal in $\mathrm{Gal}(K/F)$ and $E := K^H$, so that E is Galois over F. We define a group homomorphism $\psi : \mathrm{Gal}(K/F) \to \mathrm{Gal}(E/F)$ as follows : $\psi(\tau) = \tau_{|_E}$. By the argument in the paragraph above, $\psi(\tau) \in \mathrm{Gal}(E/F)$. Since $\tau_{1|_E} \circ \tau_{2|_E} = (\tau_1 \circ \tau_2)_{|_E}$, ψ is a group homomorphism. Moreover, $\psi(\tau) = id$ if and only if τ fixes E, i.e., $\tau \in H$. Thus, H is the kernel of τ. The proof is complete once we show that τ is surjective. To see this, let $f \in \mathrm{Gal}(E/F)$. Then we can think of f as a field homomorphism from E to $\overline{F}$. By Theorem 9.1, there exists $\tau : K \to \overline{F}$ extending f. Since K is a splitting field, Theorem 10.4 gives $\tau(K) = K$, i.e., $\tau \in \mathrm{Gal}(K/F)$. Since τ extends f, $\psi(\tau) = f$, which shows that τ is surjective. Thus, by the first isomorphism theorem for groups, $\mathrm{Gal}(E/F)$ is isomorphic to $\mathrm{Gal}(K/F)/H$, as required. This completes the proof of the theorem □

Discussion 16.2. Suppose that $F \subseteq K$ is finite Galois extension and suppose further that there exists a chain of intermediate fields

$$(*) \qquad F = L_0 \subseteq L_1 \subseteq \cdots \subseteq L_{r-1} \subseteq L_r = K,$$

such that each L_{i+1} is Galois over L_i. Then, by the Galois Correspondence Theorem, we obtain a corresponding chain of subgroups

$$(**) \qquad id = \mathrm{Gal}(K/L_r) \subseteq \mathrm{Gal}(K/L_{r-1}) \subseteq \mathrm{Gal}(K/L_{n-1}) \subseteq \cdots \subseteq \mathrm{Gal}(K/L_1) \subseteq \mathrm{Gal}(K/L_0) = \mathrm{Gal}(K/F).$$

Now, for each i, we have $L_i \subseteq L_{i+1} \subseteq K$. Since K is Galois over F, it is also Galois over L_i. By hypothesis, L_{i+1} is Galois over L_i. Thus by the Galois Correspondence Theorem, $\mathrm{Gal}(K/L_{i+1})$ is a normal subgroup of $\mathrm{Gal}(K/L_i)$. Thus, the chain of subgroups in (**) is a *subnormal series* of $\mathrm{Gal}(K/F)$. Therefore, in order to obtain a subnormal series (**) in $\mathrm{Gal}(K/F)$, it suffices to find a chain of intermediate fields of the type (*).

Conversely, if we have a chain of subgroups

$$id := H_0 \subseteq H_1 \subseteq \cdots \subseteq H_{n-1} \subseteq H_r = \mathrm{Gal}(K/F),$$

such that each H_{i-1} is a normal subgroup of H_i, then we get a corresponding chain of intermediate fields

$$F = L_0 \subseteq L_1 \subseteq \cdots \subseteq L_{n-1} \subseteq L_r = K,$$

where L_i is the fixed field of H_{n-i} and each L_i is Galois over L_{i-1}. This follows since by the proof of the Galois Correspondence Theorem, $H''_{n-i} = \mathrm{Gal}(K/L_i) = H_{n-i}$ and $\mathrm{Gal}(K/L_{i-1}) = H_{n-i+1}$, so that if we apply the Galois Correspondence Theorem to the Galois extension $L_{i-1} \subseteq K$, L_i is Galois over over L_{i-1}, since $\mathrm{Gal}(K/L_i)$ is a normal subgroup of $\mathrm{Gal}(K/L_{i-1})$. □

It turns out that having a good understanding of subnormal series and their connections to intermediate fields as above is key to understanding the solvability by radicals theorem. Towards this end, we will need a fair amount of group theory. Thus, beginning with the next lecture, we will make a fresh start and focus on the study of groups, including those aspects of group theory required for the solvability by radicals theorem and its applications.

Lecture 17: Friday October 1. We begin with some elementary fact that most of you may have already seen, so in some cases only sketches of proofs are provided.

Proposition 17.1. (i) *Let G be a group and $H \subseteq G$ a subgroup. The relation $a \equiv_l b \bmod H$ given by $a \equiv_l b \bmod H$ if and only if $b^{-1}a \in H$ is a congruence relation on G.* We say that a is left congruent to b modulo H.

(ii) *For $a \in G$, aH, the left coset of H with respect to a is the corresponding equivalence class of a.* Consequently, given two left cosets aH, bH, either $aH = bH$ or $aH \cap bH = \emptyset$.

(iii) *All of the foregoing applies to the equivalence relation $a \equiv_r b \bmod H$ if and only if $ab^{-1} \in H$, which we call right congruence modulo H.*

Proof sketch. Clearly $a \equiv_l a$, for all $a \in G$. If $b^{-1}a \in H$, then $(b^{-1}a)^{-1} = a^{-1}b \in H$, which shows that $a \equiv_l b$ implies $b \equiv_l a$. If $b^{-1}a, c^{-1}b \in H$, then $c^{-1}bb^{-1}a = c^{-1}a \in H$, which shows the given equivalence relation is transitive. Thus, left congruence module H is an equivalence relation. It is now easy to check the the congruence class of a is just $aH := \{ah \mid h \in H\}$. That $aH = bH$ or $aH \cap bH = \emptyset$, follows from the standard property of congruences classes. Moreover, we have that the distinct left cosets (and the distinct right cosets) of H form a partition of G. □

Theorem 17.2. (Lagrange's Theorem) *Let G be a finite group and $H \subseteq G$ a subgroup. Them $|H|$ divides $|G|$.*

Proof. Recall that if X is a set with an equivalence relation, then the distinct equivalence classes partition X. Thus, from Proposition 17.1, we have $G = H \cup g_2H \cup \cdots \cup g_rH$, a disjoint union of the distinct left cosets of H. Note that since G is finite, there are only finitely many distinct left cosets. Multiplication on G by any g_i is a one-to-one function, which shows that $|g_iH| = |H|$, for all i. Thus, $|G| = r \cdot |H|$, where r equals the number of distinct left cosets of H. □

Note that the proof above shows that the number of distinct right cosets of H equals the number of distinct left cosets of H, which we denote by $[G:H]$, the *index* of H in G.

Remarks 17.3. (a) Recall that a $H \subseteq G$ is a *normal* subgroup if it satisfies the following equivalent conditions:

(i) $gH = Hg$, for all $g \in G$.
(ii) $gHg^{-1} = H$, for all $g \in G$.
(iii) $ghg^{-1} \in H$, fr all $g \in G$ and $g \in H$.

An easy consequence of this, in light of Proposition 17.1 above is that any subgroup of of index two is automatically normal in G. To see this, on the one hand, for any $g \in G$, G is a disjoint union $G = H \cup gH$, while on the other hand, G is the disjoint union $H \cup Hg$, which shows $gH = Hg$.

(b) Moreover, if H is not normal in G, then not only may the equality $gH = Hg$ fail, for some $g \in G$, Hg is not even a left coset. For if it were, say, $Hg = g_0H$, then $g \in g_0H$, which implies $gH = g_0H = Hg$, a contradiction. See the example below. □

Example 17.4. Take $G = S_3$, the set of one-to-one onto functions from the set $X = \{1, 2, 3\}$ to itself. This is a nonabelian group of order six under composition. If we define $\sigma, \tau \in S_n$ by $\sigma(1) = 2, \sigma(2) = 3, \sigma(3) = 1$ and $\tau(1) = 2, \tau(2) = 1, \tau(3) = 3$, then it is not difficult to check the following:

(i) $S_3 = \{id, \sigma, \sigma^2, \tau, \sigma\tau, \sigma^2\tau\}$.
(ii) $\sigma^3 = id = \tau^2$.
(iii) $\tau\sigma = \sigma^2\tau$ and $\tau\sigma^2 = \sigma\tau$.

With these identities in mind, consider the subgroups $H := \{id, \sigma, \sigma^2\}$ and $K := \{id, \tau\}$. Then we have the following left cosets of H:

$$H = \{id, \sigma, \sigma^2\} \quad \text{and} \quad \tau H = \{\tau, \tau\sigma, \tau\sigma^2\} = \{\tau, \sigma^2\tau, \sigma\tau\}.$$

Since these sets partition G, they must be the distinct left cosets of H in G. We also clearly have $H\tau = \tau H$, so that H, H, τ are the distinct right cosets of H in G.

The situation is a bit different for K. An easy calculation gives:

$$K = \{id, \tau\}, \quad \sigma K = \{\sigma, \sigma\tau\}, \quad \sigma^2 K = \{\sigma^2, \sigma^2\tau\},$$

which are the distinct left cosets of K in G. On the other hand, using the identities above, we can see that the distinct left cosets of K in G are:

$$K = \{id, \tau\}, \quad K\sigma = \{\sigma, \sigma^2\tau\}, \quad K\sigma^2 = \{\sigma^2, \sigma\tau\}.$$

Note that not only are corresponding left and right cosets different, e.g., $\sigma K \neq K\sigma$, but none of the right cosets are left cosets at all. □

Reminder. When $N \subseteq G$ is a normal subgroup, then the set of left (respectively, right) cosets of N in G form a group, called the *factor group of G by N*. We denote this group by G/N and recall that, by definition, $(aN) \cdot (bN) = abN$ in G/N. To see that G/N is a groups, one notes that the group axioms in G/N are inherited from those in G and thus the only issue is whether or not the coset multiplication is well defined. To see this, suppose $aN = a'N$ and $bN = b'N$, for $a, a', b, b' \in G$. Then, $a^{-1}a' \in N$ and $b - 1b' \in N$. In order to see that $abN = a'b'N$, we have to show $(ab)^{-1}a'b' \in N$. We have

$$(ab)^{-1}a'b' = b^{-1}a^{-1}a'b' = \{b^{-1}(a^{-1}a')b\}\{b^{-1}b'\},$$

which belongs to N, since $b^{-1}(a^{-1}a')b \in N$ (as N is normal, and $a^{-1}a' \in N$) and $b^{-1}b' \in N$. □

Lecture 18: Monday, October 4. We start with the following familiar notion. Given groups G, G', the function $\phi : G \to G'$ is a *group homomorphism* if $\phi(ab) = \phi(a)\phi(b)$, for all $a, b \in G$. □

The following proposition lists several properties of group homomorphisms, most of which should be familiar.

Proposition 18.1. *Let $\phi : G \to G'$ be a homomorphism of groups. Then:*

(i) $\phi(e) = e'$.
(ii) $\phi(a^{-1}) = \phi(a)^{-1}$, for all $a \in G$.
(iii) $\ker(\phi) := \{a \in G \mid \phi(a) = e'\}$ is a normal subgroup of G.

(iv) ϕ is 1-1 if and only if $\ker(\phi) = e$.
(v) If $H \subseteq G$ is a subgroup, then $\phi(H) := \text{im}(\phi)$ is a subgroup of G'.
(vi) If ϕ is surjective and H is normal in G, then $\phi(H)$ is normal in G'.
(vii) If $H' \subseteq G'$ is a (normal) subgroup of G', then $\phi^{-1}(H') := \{a \in G \mid \phi(a) \in H'\}$ is a (normal) subgroup of G'.
(viii) $G/\ker(\phi)$ is isomorphic to $\text{im}(\phi)$.
(ix) If ϕ is surjective, there is a 1-1 correspondence between the (normal) subgroups of G containing $\ker(\phi)$ and the (normal) subgroups of G'.

Sketch of Proof. Proofs of most of the items are straightforward, so we leave (i)-(v) to you. For (vi) given $g' \in G'$ and $\phi(h) \in \phi(H)$, for $h \in h$, there exists $g \in G$ such. that $\phi(g) = g'$. therefore,

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

Since H is normal in G, $g^{-1}hg \in H$, so that $\phi(g^{-1}hg) \in \phi(H)$, which shows that $\phi(H)$ is normal in G'.

For (vii), let $a, b \in \phi^{-1}(H')$, so that $\phi(a), \phi(b) \in H'$. Since H' is a subgroup, $\phi(a)\phi(b) = \phi(ab) \in H'$, so that $ab \in \phi^{-1}(H')$. Likewise, $\phi(a)^{-1} = \phi(a^{-1}) \in H'$, so that $a^{-1} \in \phi^{-1}(H)$. Thus, $\phi^{-1}(H')$ is a subgroup of G. If H' is a normal subgroup of G', then for $g \in G$ and $a \in \phi^{-1}(H)$,

$$\phi(g^{-1}ag) = \phi(g)^{-1}\phi(a)\phi(g) \in H',$$

thus $g^{-1}ag \in \phi^{-1}(H)$, so that $\phi^{-1}(H)$ is a normal subgroup of G.

For (viii), Set $K := \ker(\phi)$. We define $\psi : G/K \to \phi(G)$ as follows: $\psi(gK) = \phi(g)$. Since the map ψ is defined in terms of coset representatives, we must check that the map is independent of the chosen representative, for any $gK \in G/K$. In other words we must check that if $gK = g_0K$, then $\psi(gK) = \psi(g_0K)$. For this, if $gK = g_0K$, then $g_0^{-1}g \in K$, so that $\phi(g_0^{-1}g) = e'$. Thus, $\phi(g_0^{-1})\phi(g) = e'$, so that $\phi(g) = \phi(g_0)$, which shows that $\psi(gK) = \psi(g_0K)$, and thus ψ is well defined.

To check that ψ is an isomorphism is now straight forward.

$$\phi(gK \cdot g_0K) = \psi(gg_0K) = \phi(gg_0) = \phi(g)\phi(g_0) = \psi(gK)\psi(g_0K),$$

which shows that ψ is a group homomorphism. ψ is surjective, by definition, since given $g' \in \phi(G)$, $g' = \phi(g)$, for some $g \in G$, which shows that $\psi(gK) = \phi(g) = g'$. Finally, $\psi(gK) = e'$ if and only if $\phi(g) = e'$ if and only if $g \in K$ if and only if $gK = K$, the identity element in G/K, which shows that ψ is 1-1, and thus an isomorphism.

For (ix), set $K := \ker(\phi)$ and suppose $H \subseteq G$ is a (normal) subgroup containing K. Then $\phi(H)$ is a (normal) subgroup of G', by (v) and (vi). If H' is a (normal) subgroup of G', then $\phi^{-1}(H')$ is a (normal) subgroup of G containing K, by (vii) and the definition of K. Thus, we must show these correspondences are 1-1. Suppose H, H_0 are subgroups of G containing K and $\phi(H) = \phi(H_0)$. Take $h \in H$. Then $\phi(h) \in \phi(H) = \phi(H_0)$, so that $\phi(h) = \phi(h_0)$, for some $h_0 \in H_0$. Thus, $\phi(h_0^{-1}h) = e'$, so that $h_0^{-1}h \in K$, We can then write $h_0^{-1}h = k$, for some $k \in K$. Therefore, $h = h_0k \in H_0$, since $K \subseteq H_0$. It follows that $H \subseteq H_0$. By symmetry, $H_0 \subseteq H$, so $H = H_0$.

Now suppose that H', H_0' are subgroups of G' and $\phi^{-1}(H') = \phi^{-1}(H_0')$. Take $h' \in H'$. Since ϕ is surjective, there exists $h \in H$ with $\phi(h) = h'$. Thus, $h \in \phi^{-1}(H') = \phi^{-1}(H_0')$, which means, $h' = \phi(h) \in H_0'$. Thus, $H' \subseteq H_0'$, and by symmetry, $H_0' \subseteq H'$, showing $H' = H_0'$, which completes the proof. □

Corollary 18.2. *Let G be a group and K a normal subgroup of G. Then there is a one-to-one correspondence between the (normal) subgroups of G containing K and the (normal) subgroups of G/K.*

Proof. This follows immediately from part (ix) above. Indeed, we define $\phi : G \to G/K$ by $\phi(g) = gK$. Then ϕ is clearly a surjective group homomorphism. Moreover $\phi(g) = K$ (the identity in G/K) if and only if $gK = K$ if and only if $g \in K$. Thus, the result follows from (ix). □

What is the correspondence in Corollary 18.2? The map ϕ takes the subgroup $H \subseteq G$ to $HK/K \subseteq G/K$, which by Corollary 18.3 below makes sense since HK is a subgroup of G. Part (viii) in the Proposition 18.1 above is sometimes called the *First Isomorphism Theorem.* Two other isomorphism theorems follow quickly as corollaries.

Corollary 18.3. *Let G be a group, $H, K \subseteq G$ subgroups, with K normal in G. Then:*

(i) $HK := \{hk \mid h \in H \text{ and } k \in K\}$ is a subgroup of G and HK/K is isomorphic to $H/(H \cap K)$.

(ii) Suppose $H \subseteq K$ and H is also normal in G. Then K/H is normal in G/H and $(G/H)/(K/H)$ is isomorphic to G/K.

Proof. For (i), we first note that $HK = KH$. To see this, suppose $hk \in HK$. Then, $hkh^{-1} \in K$, since K is normal in G. Thus, $hkh^{-1} = k_0 \in K$. Therefore, $hk = k_0h \in KH$. Thus, $HK \subseteq KH$, and by symmetry, $KH \subseteq HK$, so that $HK = KH$. Note that one way of thinking about K being normal in G is that elements of K "commute" with elements of g at the expense of changing the element of K, i.e., for $k, g \in G$, $gk = k_0g$ and $kg = gk_1$, for some $k_0, k_1 \in K$. Of course this is just another way of saying that $gK = Kg$ as cosets.

Now suppose $hk \in HK$. Then

$$(hk)^{-1} = k^{-1}h^{-1} \in KH = HK.$$

If $h'k' \in HK$, then

$$(hk)(h'k') = h(kh')k' = h(h'k_0)k',$$

for some $k_0 \in K$, which shows HK is closed under the group product. Thus, HK is a subgroup of G. Define

$$\phi : HK \to H/(H \cap K)$$

by $\phi(hk) = h(H \cap K)$. To see that ϕ is well defined (since the representation hk is not unique), suppose $hk = h_0k_0$. Then, $h_0^{-1}h = k_0k^{-1} \in H \cap K$. Thus, $h_0^{-1}h \in H \cap K$, which means $h(H \cap K) = h_0(H \cap K)$, so ϕ is well-defined. Then ϕ is clearly a group homomorphism. Moreover, $\phi(he) = h(H \cap K)$, so that ϕ is surjective. In addition, $\phi(hk) = e(H \cap K)$ if and only if $h(H \cap K) = H \cap K$ if and only if $h \in H \cap K$, if and only if $h \in K$ if and only if $hk \in K$. Thus, K is the kernel of ϕ, so that part (viii) above gives HK/K is isomorphic to $H/(H \cap K)$.

For (ii), it is straightforward to check that K/H is a normal subgroup of G/H. Define

$$\psi : G/H \to G/K$$

as, $\psi(gH) = gK$. This map is well defined, since if $gH = g_0H$, then $g_0^{-1}g \in H \subseteq K$, so that $gK = g_0K$. It is easy to check that ψ is a surjective group homomorphism. Moreover, gH is in the kernel of ψ if and only if $\psi(gH) = gK = K$ if and only if $g \in K$ if and only if $gH \in K/H$, so that K/H is the kernel of ψ. By part (viii) in the proposition above, $(G/H)/(K/H)$ is isomorphic to G/K. □

Lecture 19: Wednesday October 6. We can apply the results from Proposition 18.1 to introduce a group theoretic notion that plays a central role in the solvability by radicals theorem. In fact, the group property of *solvability* obviously derives from the solvability by radicals setting.

Definition 191.1. A group G is said to be a solvable group if there exists a chain of subgroups

$$G_0 = (e) \subseteq G_1 \subseteq \cdots \subseteq G_{n-1} \subseteq G_n = G \quad (*)$$

such that each G_{i-1} is a normal subgroup of G_i and the factor group G_i/G_{i-1} is abelian, for all $1 \leq i \leq n$. The tower of groups above is called a *solvable series* for G.

Remarks 19.2. (a) Any abelian group G is a solvable group, since $(e) \subseteq G$ is a solvable series. S_3, the symmetric on three letters, is a solvable group that is not abelian. To see this, if we let τ denote the 3-cycle $(1, 2, 3)$, then $H := \langle\tau\rangle$ is a cyclic group of order three, and thus abelian. It is easy to see that H is normal in S_3. Since S_3/H is a group of order two, S_3/H is abelian. Thus, $(e) \subseteq H \subseteq S_3$ is a solvable series. It turns out that the symmetric group S_4 is also solvable, but S_n, for $n \geq 5$, is not a solvable group. It is this latter fact that is at the heart of the non-solvability by radicals of polynomials of degree five and higher.

(b) In the early part of the 20th century, the prominent English mathematician William Burnside (see `https://en.wikipedia.org/wiki/William_Burnside`) conjectured that any group of odd order must have a non-trivial normal subgroup, or equivalently, that every group of odd order is solvable. This problem stood open for almost fifty years until is was solved by group theorists Walter Feit and John Thompson in the early 1960s (see `https://en.wikipedia.org/wiki/Feit-Thompson_theorem`). The proof of the Feit-Thompson theorem was approximately 250 pages long and occupied a full issue of the Pacific Journal of Mathematics. It ushered in a new era in group theory that saw great advances and numerous papers of similar, or greater length.

Theorem 19.3. Let G be a group and $H, N \subseteq G$ subgroups such that N is normal in G.

(i) If G is solvable, then H and G/N are solvable G are solvable.

(ii) If N and G/N are solvable, then G is solvable.

(iii) If G is finite, then G is solvable if and only if G has a solvable series whose factors are cyclic of prime order.

Proof. For (i), let G_i be the terms in a solvable series (*). If $H$is a subgroup of G, we have

$$H_0 = (e) \subseteq G_1 \cap H \subseteq \cdots \subseteq G_{n-1} \cap H \subseteq G_n \cap H = H. \quad (**)$$

It is easy to see that $G_{i-1} \cap H$ is a normal subgroup of $G_i \cap H$. Moreover, the kernel of the natural map $G_i \cap H \to G_i/G_{i-1}$ is $G_{i-1} \cap H$. Consequently, we can identify $(G_i \cap H)/(G_{i-1} \cap H)$ with a subgroup of G_i/G_{i-1}, and thus each $(G_i \cap H)/(G_{i-1} \cap H)$ is abelian. Therefore, the subnormal series (**) is a solvable series.

Now, let N be a normal subgroup of G. Then, each NG_i is a subgroup containing N and thus each NG_i/N is a subgroup of G/N. We claim that the series

$$N/N \subseteq (G_1N)/N \subseteq \cdots \subseteq (NG_{n-1})/N \subseteq (NG_n)/N = G/N, \quad (***)$$

is a solvable series for G/N. For this, we first show that each NG_{i-1} is normal in NG_i. Take $n_1g_i \in NG_i$ and $n_2g_{i-1} \in NG_{i-1}$. Then

$$(n_1g_i)^{-1}n_2g_{i-1}(n_1g_i) = g_i^{-1}n_1^{-1}n_2g_{i-1}n_1g_i = g_i^{-1}n_1^{-1}n_2n_1'g_{i-1}g_i = (g_i^{-1}n_1^{-1}n_2n_1'g_i)(g_i^{-1}g_{i-1}g_i),$$

which belongs to NG_{i-1}, since N is normal in G and G_{i-1} is normal in G_i. Note, $n_1' \in N$, by our comments above, so that NG_{i-1} is normal in NG_i. Thus, by Corollary 18.3 (ii), (***) is a subnormal series in G/N. To see that (***) is a solvable series, by Corollary 18.3(ii), it suffices to shows that NG_i/NG_{i-1} is abelian. For this, it suffices to show that if $x, y \in NG_i$, then $xyx^{-1}y^{-1} \in NG_{i-1}$. Suppose $n_1g_1, n_2g_2 \in NG_i$. Then

$$(n_1g_1)(n_2g_2)(n_1g_1)^{-1}(n_2g_2)^{-1} = n_1g_1n_2g_2g_1^{-1}n_1^{-1}g_2^{-1}n_2^{-1}.$$

Now, in the last expression above, we may move all of the elements of N to the left of this expression (at the expense of changing the elements of N) to obtain

$$(n_1g_1)(n_2g_2)(n_1g_1)^{-1}(n_2g_2)^{-1} = n_0g_1g_2g_1^{-1}g_2^{-1},$$

for some $n_0 \in N$. Since G_i/G_{i-1} is abelian, $g_1g_2g_1^{-1}g_2^{-1} \in G_{i-1}$, and thus, $(n_1g_1)(n_2g_2)(n_1g_1)^{-1}(n_2g_2)^{-1}$ belongs to NG_{i-1}, which is what we want. This finishes the proof of part (i).

For part (ii), suppose

$$(e) = N_0 \subseteq N_1 \subseteq \cdots \subseteq N_{r-1} \subseteq N_r = N$$

is a solvable series for N, and

$$N/N = G_0/N \subseteq G_1/N \subseteq \cdots \subseteq G_{s-1}/N \subseteq G_s/N = G/N$$

is a solvable series for G/N. Then by Proposition 18.1 (ix) and Corollary 18.3 (ii),

$$(e) = N_1 \subseteq \cdots \subseteq N_{r-1} \subseteq N_r = N \subseteq G_1 \subseteq \cdots \subseteq G_{s-1} \subseteq G_s = G$$

is a solvable series for G, and thus G is a solvable group.

Lecture 20: Friday October 8. We continue with the proof of Theorem 19.3. Before doing so, we note that the following discussion is a consequence of Corollary 18.2.

Discussion 20.1. Suppose $H \subseteq K$ are subgroups of the group G, with K normal in H. Suppose

$$K = K_0 \subseteq K_1 \subseteq \cdots \subseteq K_r = H$$

is a sequence of subgroups satisfying K_i is normal in K_{i+1} and K_{i+1}/K_i satisfies property P (e.g., is abelian or cyclic). Then we have the following sequence in H/K

$$\{e\} = K_0/K \subseteq K_1/K_0 \subseteq \cdots \subseteq K_r/K = H/K$$

such that each K_i/K is normal in K_{i+1}/K and the quotient $(K_i/K)/(K_{i+1}/K)$ satisfies property P. Conversely, suppose we are given the sequence of subgroups

$$\{e\} = L_0 \subseteq L_1 \subseteq \cdots \subseteq L_s = H/K$$

in H/K such that L_i is normal in L_{i+1} and L_{i+1}/L_i satisfies property P. The there exists a sequence of subgroups

$$K = C_0 \subseteq C_1 \subseteq \cdots \subseteq C_s = H$$

in G such that $C_i/K = L_i$, C_i is normal in C_{i+1} and C_{i+1}/C_i satisfies property P. □

The proof of part (iii) Theorem 19.3 is comprised of repeated applications of the content of Discussion 20.1.

Proof of Theorem 19.3 (iii). Clearly if G has a solvable series whose factors are cyclic of prime order, G is a solvable group. Conversely, if G is finite, let (*) above be a solvable series. Then each factor G_i/G_{i-1} is a finite abelian group. We will use the fact that a finite abelian group A is isomorphic to a product $\mathbb{Z}_{n_1} \times \cdots \times \mathbb{Z}_{n_s}$ of cyclic groups. Note that if $A = \mathbb{Z}_{n_1} \times \cdots \times \mathbb{Z}_{n_s}$, then upon setting $A_j = \mathbb{Z}_{n_1} \times \cdots \times \mathbb{Z}_{n_j}$,

$$(e) \subseteq A_1 \subseteq \cdots \subseteq A_{s-1} \subseteq A_s = A$$

is a solvable series for A whose factors are cyclic. Thus, for each $1 \leq i \leq n$, if $G_i/G_{i-1} \cong A$, for A as above, we can insert subgroups A_j as follows:

$$G_{i-1} \subseteq A_1 \subseteq \cdots \subseteq A_{s-1} \subseteq A_s = G_i,$$

as in the discussion above, such that each term is normal in the the term following it and any factor in this chain is cyclic. Doing this for each $1 \leq i \leq n$ means we may expand the series (*) to assume that each factor is cyclic. If we now show that any cyclic group has a solvable series with factors that are cyclic of prime order, we may apply this same technique to our new solvable series, expanding it, so that the factors are cyclic of prime order. To see that a cyclic group C has a solvable series whose factors are cyclic of prime order, it is more convenient to write C using multiplicative notation, so assume $C = \langle a \rangle$, with $a^n = e$ and no smaller power of a is the identity. Write the prime factorization of n as $n = p_1^{e_1} \cdots p_r^{e_r}$. Then the following is the required solvable series for C:

$$(e) \subseteq \langle a^{p_1^{e_1} \cdots p_r^{e_r - 1}} \rangle \subseteq \langle a^{p_1^{e_1} \cdots p_r^{e_r - 2}} \rangle \subseteq \cdots \subseteq \langle a^{p_1^{e_1} \cdots p_{r-1}^{e_{r-1}}} \rangle \subseteq \langle a^{p_1^{e_1} \cdots p_{r-1}^{e_{r-1} - 1}} \rangle \subseteq \cdots \subseteq \langle a^{p_1^{e_1}} \rangle \subseteq \cdots \subseteq \langle a^{p_1} \rangle \subseteq \langle a \rangle.$$

□

Discussion 20.2. Let G be a group, and $a \in G$. If there exists $k \geq 1$ such that $a^k = e$, then we define the order of a to be the least positive n such that $a^n = e$. We will write $o(a) = n$. Suppose $a^k = e$. If we write $k = nm + r$, with $0 \leq r < n$, then

$$e = a^k = a^{mn+r} = (a^n)^m \cdot a^r = e^m \cdot a^r = a^r.$$

This is a contradiction to the choice of n, unless $r = 0$. Thus, $o(a)$ divides any integer k such that $a^k = e$. Suppose $o(a) = n$. Then it is easy to check that the subset $\{e, a, \ldots, a^{n-1}\}$ is in fact a subgroup of G, and we denote this subgroup as $\langle a \rangle$, and call it the cyclic subgroup of G generated by A. Since $|\langle a \rangle| = n$, by Lagrange's Theorem n divides $|G|$, if G is a finite group. Thus, the order of an element of a finite group divides the order of the group. □

Lecture 21: Wednesday October 13. We now would like to discuss the finite symmetric groups.

Definitions and Notation 21.1. (a) For the positive integer n, we let X_n (or just X, if n is already established) denote the set $\{1, 2, \ldots, n\}$ and S_n denote the set of 1-1 and onto functions from X_n to itself. Since any 1-1 and onto function has an inverse under composition, and composition of functions is associative, it follows that S_n is a group under composition. Moreover, since we can think of a 1-1 function from X_n to itself as a permutation of the elements of X_n - in fact one often refers to an an element of S_n as a permutation - it follows that $|S_n| = n!$. We call S_n the Symmetric Group on n objects or the Group of permutations of the set X_n.

(b) Let $\sigma \in S_n$, then $\sigma(1) = i_1, \sigma(2) = i_2, \ldots, \sigma(n) = i_n$, where $X_n = \{i_1, i_2, \ldots, i_n\}$. We will use the standard notation $\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ i_1 & i_2 & \cdots & i_n \end{pmatrix}$. Since the elements of S_n are to be regarded as functions, if we write $\sigma\tau$, for $\sigma, \tau \in S_n$, then we apply τ before σ. For example, in S_4, if $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 4 \end{pmatrix}$ and $\tau = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix}$, then $\sigma\tau = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix}$.

(c) Suppose for $1 \leq k \leq n$, there exist distinct elements $i_1, i_2, \ldots, i_k \in X_n$ with

$$\sigma(i_1) = i_2, \sigma(i_2) = i_3, \ldots, \sigma(i_{k-1}) = i_k, \sigma(i_k) = i_1 \text{ and } \sigma(j) = j \text{ for } j \notin \{i_1, \ldots, i_k\}.$$

We call σ a k-cycle and write $\sigma = (i_1, i_2, \ldots, i_k)$. In this case, it is easy to check that $o(\sigma) = k$. A 2-cycle is often referred to as a transposition. Note that in (b) above $\sigma = (1, 3), \tau = (1, 2, 3, 4)$ and $\sigma\tau = (1, 2)(3, 4)$. □

Example 21.2. Let us look closely at S_3. Consider $\tau := (1, 2, 3)$ and $\sigma := (1, 2)$. It follows that $\tau^2 = (1, 3, 2)$, $\tau^3 = id = \sigma^2$. Thus, id, τ, τ^2, σ are four distinct elements in S_3. We also have: $\sigma\tau = (2, 3)$ and $\sigma\tau^2 = (1, 3)$. Thus, $\{id, \tau, \tau^2, \sigma, \sigma\tau, \sigma\tau^2\}$ are the six distinct elements of S_3. So, S_3 has three elements of order two and two elements of order three. There is also the relation $\tau\sigma = \sigma\tau^2$ that enables one to write any product whatsoever of elements from S_e in the form of one of the six elements we have just listed. Note that if we multiply this last equation by τ we get $\tau^2\sigma = \tau\sigma\tau^2 = \sigma\tau^2\tau^2 = \sigma\tau$. For example what group element is $\tau^{14}\sigma^7\tau^{10}\sigma\tau^2$? Using the identities we have established, we have

$$\begin{aligned} \tau^{14}\sigma^7\tau^{10}\sigma\tau^2 &= \tau^2\sigma\tau\sigma\tau^2 \\ &= \sigma\tau\tau\sigma\tau^2 \\ &= \sigma\tau^2\sigma\tau^2 \\ &= \sigma\sigma\tau\tau^2 \\ &= id. \end{aligned}$$

What about normal subgroups of S_3? $K := \langle\tau\rangle = \{id, \tau, \tau^2\}$ is a subgroup of order three, and thus has index two, and so is a normal subgroup. On the other hand $H := \langle\sigma\rangle = \{id, \sigma\}$ is a subgroup that is not normal. For example, $\tau H = \{\tau, \tau\sigma\} = \{\tau, \sigma\tau^2\}$, while $H\tau = \{\tau, \sigma\tau\}$. It is now easy to check that $K, H, H_2 := \langle\sigma\tau\rangle$, and $H_3 := \langle\sigma\tau^2\rangle$ are the only proper subgroups of S_3 and that K is the only proper normal subgroup. □

Imagine a set of cards numbered 1 through n lined up on in order a long table. Professor Cauchy comes up to the table and rearranges the cards in any way he likes. Can Galois achieve the same rearrangement simply by interchanging two cards at a time? Intuitively, this seems possible. The following fundamental proposition proves this.

Proposition 21.3. *Every element in S_n can be written as the product of disjoint cycles. Moreover, every element in S_n can be written as a product of (not necessarily disjoint) transpositions.*

Proof. We say that the two cycles $(i_1 \ldots, i_k)$ and $(j_1, \ldots, j_k)$ are disjoint if the sets $\{i_1, \ldots, i_k\}$ and $\{j_1, \ldots, j_r\}$ are disjoint. Now, let $\sigma \in S_n$. We define an equivalence relation, that depends upon σ, on X_n by $i \sim j$ if and only if $j = \sigma^s(i)$ for some $s \geq 0$. Clearly $i \sim i$. If $j = \sigma^s(i)$, Then if $o(\sigma) = r$, choose c such that $s + c = r$. Then $\sigma^c(j) = i$. Thus, $i \sim j$ implies $j \sim i$.. Finally, if $j = \sigma^s(i)$ and $k = \sigma^t(j)$, then $k = \sigma^{s+t}(i)$ which gives the transitive property of the relation, $\sim$, which is therefore an equivalence relation.

Now fix $i \in X_n$ and assume $\sigma(i) \neq i$. Note that since $\sigma^n = id$ for some n, there is a least positive integer d_1 such that $\sigma^{d_1}(i) = i$. So that if $e > d_1$, we may write $e = d_1h + r$, with $r < d_1$, so that $\sigma^e(i) = \sigma^r(i)$. Thus, the equivalence class of i is the set $\{i, \sigma(i), \ldots, \sigma^{d_1-1}(i)\}$ and $\tau_1 = (i, \sigma(i), \ldots, \sigma^{d_1-1}(i))$ is a d_1-cycle corresponding to this class. Now, if X' is any other equivalence class consisting of more than one element, then for $j \in X'$, we can write $X' = \{j, \sigma(j), \ldots, \sigma^{d_2-1}(j)\}$, for some $d_2 > 1$, and thus X' corresponds to the d_2-cycle $\tau_2 = (j, \sigma(j), \ldots, \sigma^{d_2-1}(j))$. Continuing in this way, we can represent each equivalence class X_h, with $d_h > 1$ elements, as a d_h-cycle. Those classes that have a single element correspond to elements fixed by σ. It now follows that if $X_1, \ldots, X_e$ are the distinct equivalence classes with $d_h > 1$ elements, for $1 \leq h \leq e$, and τ_h is the corresponding d_h-cycle, then $\sigma = \tau_1\tau_2 \cdots \tau_e$. Since the equivalence classes are disjoint, we have written σ as a product of disjoint cycles.

To see that any $\sigma \in S_n$ is a product of transpositions, by what we have just shown, it suffices to show that any cycle is a product of transpositions. Suppose $\tau = (i_1, i_2, \dots, i_k)$. Then it is easy to check that

$$\tau = (i_1, i_k)(i_1, i_{k-1}) \cdots (i_1, i_4)(i_1, i_3)(i_1, i_2). \quad \square$$

Question. Returning to the long table with the set of cards numbered 1 through n. After Professor Cauchy has made the initial rearrangement, can Galois also make the same arrangement by interchanging two *adjacent* cards at a time?

Remark 21.4. We make a couple of comments concerning the previous proposition. The first, is that since disjoint cycles commute, when we decompose $\sigma \in S_n$ as a product of disjoint cycles, we get the same element of S_n if we interchange the cycles in the given decomposition. However, that decomposition is unique up to permutation of the terms in the decomposition, since the equivalences classes of an equivalence relation are uniquely determined by the set and the given equivalence relation. The second is that when we write $\sigma \in S_n$ as a product of transpositions, this decomposition is not unique, e.g., $(1,3) = (1,2)(2,3)(1,2)$, but its *parity* is unique. In other words, we cannot write σ as a product of an even number of transpositions in one decomposition and then write σ as a product of an odd number of transpositions in another decomposition. This is a **fundamental fact** about the elements of S_n.

Definition-Theorem 21.5. *An element $\sigma \in S_n$ is said to be an even permutation if it can be written as a product of an even number of transpositions. Otherwise σ is said to be an odd permutation. This definition is well defined.*

Proof. We use permutation matrices to show that the definitions of even and odd permutations are well defined. Fix $\sigma \in S_n$. Let A_σ be the matrix obtained from I_n, the $\times n$ identity matrix, by interchanging its rows according to σ In other words, if $R_1, R_2, \dots R_n$ are the rows of I_n, we let A_σ denote the matrix whose rows are $R_{\sigma(1)}, R_{\sigma(2)}, \dots, R_{\sigma(n)}$. Thus, the ith row of A_σ has 1 in the $\sigma(i)$th column, zero elsewhere. For example, if $\sigma \in S_4$ and $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 3 & 1 \end{pmatrix}$, then $A_\sigma = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$. Now, it is easy to see that $A_\sigma \cdot \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} a_{\sigma(1)} \\ a_{\sigma(2)} \\ \vdots \\ a_{\sigma(n)} \end{pmatrix}$. Keeping this in mind, let us set $b_i := a_{\sigma(i)}$, so that $\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = A_\sigma \cdot \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$. For $\tau \in S_n$, we have $A_\tau \cdot \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} b_{\tau(1)} \\ b_{\tau(2)} \\ \vdots \\ b_{\tau(n)} \end{pmatrix}$. However, the equation $b_j = a_{\sigma(j)}$ for all j means that $b_{\tau(i)} = a_{\sigma(\tau(i))}$. In other words, $A_\tau \cdot \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} a_{\sigma(\tau(i))} \\ a_{\sigma(\tau(2))} \\ \vdots \\ a_{\sigma(\tau(n))} \end{pmatrix}$. And thus,

$$A_\tau A_\sigma \cdot \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} a_{\sigma(\tau(i))} \\ a_{\sigma(\tau(2))} \\ \vdots \\ a_{\sigma(\tau(n))} \end{pmatrix} = A_{\sigma\tau} \cdot \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}.$$

Since this holds for all $\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$ (in say, $\mathbb{Q}^n$), we have $A_\tau A_\sigma = A_{\sigma\tau}$.

Now suppose $\sigma \in S_n$ and $\sigma = \tau_1\tau_2\cdots\tau_r$, where each τ_i is a transposition. Then

$$\det(A_\sigma) = \det(A_{\tau_1\cdots\tau_r}) = \det(A_{\tau_r}\cdots A_{\tau_1}) = \det(A_{\tau_r})\cdots\det(A_{\tau_1}) = (-1)^r,$$

since clearly each $\det(A_{\tau_i}) = -1$. If we can also write σ as a product of s transpositions, then the same argument shows that $\det(A_\sigma) = (-1)^s$, and thus $(-1)^r = (-1)^s$. It follows that either both r and s are even, or both r and s are odd. Therefore, no permutation in S_n can be written both as a product of an even number of transpositions and an odd number of transpositions. □

Clearly the inverse of an even permutation is an even permutation and the product of two even permutations is an even permutation. Thus, the set of even permutations forms a distinguished subgroup of S_n.

Definition-Proposition 21.6. The set of even permutations in S_n is subgroup, denoted A_n, and called the alternating group on n objects. $[S_n : A_n] = 2$, and therefore A_n is a normal subgroup of S_n.

Proof. We just have to show $[S_n : A_n] = 2$. Set $\sigma := (1, 2)$. Then $\tau \in S_n$ is an even permutation if and only if $\sigma\tau$ is an odd permutation. Thus, the function $f : A_n \to \{\text{Odd Permutations}\}$ defined by $f(\tau) = \sigma\tau$ is 1-1 and onto, since if γ is an odd permutation $\sigma\gamma$ is even and therefore $f(\sigma\gamma) = \gamma$. Since every element of S_n is either even or odd, it follows that half of the elements of S_n are even permutations, and thus $[S_n : A_n] = 2$. □

Lecture 22: Friday October 15. We now come to a theorem that plays an important role in the theory of solvability by radicals. Recall that a group G is said to be a *simple* group if it has no proper normal subgroups, i.e., its only normal subgroups are $\{e\}$ and G.

Remark 22.1. (i) The only finite abelian simple groups are the cyclic groups of primes order. Indeed, if G is a finite abelian group, every proper subgroup is a normal subgroup, so that G can only be simple if it has no proper subgroups. Since a finite abelian group is a product of cyclic groups, a finite simple abelian group must be a cyclic group, and if it has not proper subgroups, it must be cyclic of prime order.

(ii) There are no non-abelian simple groups of order less than 60. Up to isomorphism, A_5 is the only simple group of order 60.

Theorem 22.2. *For $n \geq 5$, A_n is a simple group.*

Proof. The proof proceeds in four steps.

Step 1. We first note that A_n is the subgroup of S_n generated by the set of all 3-cycles. To see this, if (u, v, w) is a 3-cycle, then $(u, v, w) = (u, w)(u, v)$, so that every 3-cycle belongs to A_n. Conversely, every element of A_n is a product of permutations of the form $(u, v)(s, t)$ or $(u, v)(u, s)$, for distinct elements $u, v, w, s \in X_n$. But, $(u, v)(s, t) = (u, s, v)(u, s, t)$ and $(u, v)(u, s) = (u, s, v)$, which shows that A_n is the subgroup of S_n generated by the set of 3-cycles.

Step 2. Fix $1 \leq a \neq b \leq n$. Then A_n is generated by the 3-cycles of the form (a, b, c) with $c \in X_n\backslash\{a, b\}$. To see this, let us note that any 3-cycle is of the form: $(a, b, u), (a, u, b), (a, u, v), (b, u, v), or(u, v, w)$, for a, b, u, v, w distinct elements of X_n. However, direct calculation shows that

$$\begin{aligned}(a, u, b) &= (a, b, u)^2\\ (a, u, v) &= (a, b, v)(a, b, u)^2\\ (b, u, v) &= (a, b, v)^2(a, b, u)\\ (u, v, w) &= (a, b, u)^2(a, b, w)(a, b, v)^2(a, b, u),\end{aligned}$$

which gives what we want.

Step 3. If N is a normal subgroup of A_n and N contains a 3-cycle, then $N = A_n$. To see this, let $(a, b, u) \in N$ be a 3-cycle. Then, by Step 1, it suffices to show that $(a, b, c) \in N$, for all $c \in X_n\backslash\{a, b\}$. However, using that $(a, b)(u, c)$ is its own inverse, we have

$$\{(a, b)(u, c)\}(a, b, u)^2\{(a, b)(u, c)\}^{-1} = (a, b)(u, c)(a, b, u)^2(a, b)(u, c) = (a, b, c),$$

which belongs to N since N is normal in A_n.

Step 4. If N is a normal subgroup of A_n, then N contains a 3-cycle. This is the hardest step, and requires consideration of several cases, where we analyze the decomposition of $\sigma \in N$ into a product of disjoint cycles.

Case (a). N contains an element σ of the form $\sigma := (i_1, \dots, i_k)\tau$, where $k \geq 4$ and τ is a product of disjoint cycles, each of which is disjoint from $(i_1, \dots, i_k)$. Set $\gamma := (i_1, i_2, i_3)$. Then $\sigma^{-1}(\gamma\sigma\gamma^{-1}) \in N$. However,

$$\sigma^{-1}(\gamma\sigma\gamma^{-1}) = \tau^{-1}(i_1, i_k, i_{k-1}, \dots, i_2)(i_1, i_2, i_3)(i_1, \dots, i_k)\tau(i_1, i_3, i_2) = (i_1, i_3, i_k),$$

which shows that N contains a 3-cycle.

Case (b). N contains $\sigma = (i_1, i_2, i_3)(i_4, i_5, i_6)\tau$, a product of disjoint cycles. Set $\gamma := (i_1, i_2, i_4)$. Then $\sigma^{-1}(\gamma\sigma\gamma^{-1}) \in N$. However,

$$\sigma^{-1}(\gamma\sigma\gamma^{-1}) = \tau^{-1}(i_4, i_6, i_5)(i_1, i_3, i_2)(i_1, i_2, i_4)(i_1, i_2, i_3)(i_4, i_5, i_6)\tau(i_1, i_4, i_2) = (i_1, i_4, i_2, i_6, i_3),$$

and thus, N contains a 5-cycle. By Case (a), N contains a 3-cycle.

Case (c). N contains $\sigma = (i_1, i_2, i_3)\tau$, where τ is a product of disjoint 2-cycles, disjoint from (i_1, i_2, i_3). Then $\sigma^2 \in N$, and since disjoint cycles commute,

$$\sigma^2 = (i_1, i_2, i_3)^2\tau^2 = (i_1, i_2, i_3)^2 = (i_1, i_3, i_2),$$

so N contains a 3-cycle.

Case (d). One of the previous three cases must hold. If not, then every element of N is a product of an even number of disjoint 2-cycles. Let $\sigma \in N$, and write $\sigma = (i_1, i_2)(i_3, i_4)\tau$ be the cycle decomposition of σ. Set $\gamma := (i_1, i_2, i_3)$, so that $\sigma^{-1}(\gamma\sigma\gamma^{-1}) \in N$. However,

$$\sigma^{-1}(\gamma\sigma\gamma^{-1}) = \tau^{-1}(i_3, i_4)(i_1, i_2)(i_1, i_2, i_3)(i_1, i_2)(i_3, i_4)\tau(i_1, i_3, i_2) = (i_1, i_3)(i_2, i_4).$$

For ease of notation, set $\alpha := (i_1, i_3)(i_2, i_4) \in N$.

Since $n \geq 5$, there exists $j \in X_n \backslash \{i_1, \dots, i_4\}$. Set $\beta := (i_1, i_3, j) \in A_n$. Then, $\alpha\beta\alpha\beta^{-1} \in N$. However,

$$\alpha\beta\alpha\beta^{-1} = (i_1, i_3)(i_2, i_4)(i_1, i_3, j)(i_1, i_3)(i_2, i_4)(i_1, j, i_3) = (i_1, i_3, j),$$

showing that N contains a 3-cycle. But this is a contradiction, because any 3-cycle is the product of two 2-cycles that are *not* disjoint.

All possibilities for cycle decompositions that can occur have been covered by the cases above, thus N must contain a 3-cycle. It follows immediately from Steps 1,2,3 that A_n cannot have a proper, normal subgroup, and therefore, A_n is a simple group. □

Corollary 22.3. *For $n \geq 5$, neither A_n nor S_n are solvable groups.*

Proof. By the previous theorem, A_n is not solvable for $n \geq 5$. By Theorem 19.3, S_n is not solvable. □

The following corollary to Theorem 22.2 illustrates a technique that is very useful when studying whether or not groups of small order (say, order less than or equal to 168) are simple.

Corollary 22.4. *Suppose $n \geq 5$. Then A_n is the only non-trivial normal subgroup of S_n.*

Proof. Let $N \subsetneq S_n$ be a non-trivial normal subgroup. Then $N \cap A_n$ is a normal subgroup of A_n and therefore, $N \cap A_n = A_n$ or $N \cap A_n = e$. In the first case, $A_n \subseteq N$. Since $[S_n : A_n] = 2$, this forces $A_n = N$, since $N \neq S_n$. Suppose $N \cap A_n = e$. Then, since N is normal in S_n, NA_n is a subgroup of S_n properly containing A_n. Since $[S_n : A_n] = 2$, we have $S_n = NA_n$. Using problem 6 from Homework 3, we have

$$n! = |S_n| = |NA_n| = \frac{|N| \cdot |A_n|}{|N \cap A_n|} = \frac{|N| \cdot \frac{n!}{2}}{1},$$

which implies that $|N| = 2$. Thus, $N = \{id, \sigma\}$, for $\sigma \in S_n$ having order two. Since N is normal, we have $\tau\sigma\tau^{-1} \in N$ for all $\tau \in S_n$ If $\tau\sigma\tau^{-1} = e$, then $\sigma = e$, so we must have $\tau\sigma\tau^{-1} = \sigma$, for all $\tau \in S_n$. In other words, $\tau\sigma = \sigma\tau$, so that σ commutes with every element of S_n. However, when $n \geq 3$, no $e \neq \sigma \in S_n$ has this property. To see this, suppose $\sigma(i) = j$, with $i \neq j \in X_n$. Take τ such that $\tau(i) = i$ and $\tau(\sigma(j)) \neq j$. Then $j = \sigma\tau(i)$ and $\tau(\sigma(j)) \neq j$, so that $\sigma\tau \neq \tau\sigma$. □

Remark 22.5. The classification of finite simple groups was mostly carried out during the latter quarter of the 20th century, with finishing touches applied during the early part of this century. Roughly speaking, any finite simple group belongs to one of three infinite families (e.g., A_n, $n \geq 5$, is one such family) or belongs to a list of 26 so-called *sporadic simple groups*, the latter being simple groups not part of any infinite family. Initially, the very idea that such a classification could even exist seemed improbable. Ultimately, the classification of finite simple groups was the result of a monumental effort on the part of scores of mathematicians, resulting in a proof of thousands of pages scattered among dozens of journals. The project was conceived, initiated, and organized by D. Gorenstein in the mid 1970s. Gorenstein was a leading group theorist in the last half of the 20th century, even though he did his PhD thesis at Harvard in algebraic geometry under Oscar Zariski! Major contributors to the effort included J. Thompson, W. Feit, M. Suzuki, M. Aschbacher (especially), R. Lyons, Z. Janko, R. Solomon, and many others. There are many expository (and not so expository!) articles one can find on the internet. I have uploaded to the Supplementary Materials folder in Blackboard a particularly nice article written by Ron Solomon. Aside from having a great photo of Gorenstein, the article is important, because Solomon and Aschbacher are the principal architects of the so-called *second generation proof*, which is an effort devoted to simplifying the classification proof, and putting it all into one location, the latter meaning a collection of volumes which all together comprise a complete proof.

Lecture 23: Monday October 18. We now turn to a new topic that will enable us to prove the important Sylow Theorems for finite groups. This topic concerns group actions. The basic idea is that a group G acting on a set X provides a way to multiply elements of the set X by elements of the group G. This is a very powerful concept.

Definition 23.1. Let G be a group and X a set. We say say that G acts on X if there is a function $\phi : G \times X \to X$ such that:

(i) $\phi(e, x) = x$, for all $x \in X$.
(ii) $\phi(ab, x) = \phi(a, \phi(b, x))$, for all $a, b \in G$ and $x \in X$.

We will follow the standard convention of writing ax instead of $\phi(a, x)$, for $a \in G$ and $x \in X$. Thus, the properties above become:

(i) $ex = x$, for all $x \in X$.
(ii) $(ab)x = a(bx)$, for all $a, b \in G$ and $x \in X$.

Observation 23.2. Suppose G acts on the set X. Then for any $a \in G$, the set map $f_a : X \to X$ defined by $f_a(x) = ax$, for all $x \in X$, is a one-to-one and onto function. To see this, suppose $f_a(x_1) = f_a(x_2)$, then $ax_1 = ax_2$. Multiplying by a^{-1}, we have $a^{-1}(ax_1) = a^{-1}(ax_2)$, and thus, $(a^{-1}a)x_1 = (a^{-1}a)x_2$, so $ex_1 = ex_2$, which gives $x_1 = x_2$, so f_a is one-to-one. Suppose $x \in x$. Then $a^{-1}x \in X$ and we have

$$f_a(a^{-1}x) = a(a^{-1}x) = (aa^{-1})x = ex = x,$$

which shows that f_a is onto. Therefore, f_x can be regarded as a permutation of X. In particular, when a group G acts on the set X, the elements of G permute the elements of X. We can do more, as the next important proposition shows.

Proposition 23.3. *Assume the group G acts on the set X, and denote by S_X the group of one-to-one and onto functions from X to itself, so that if $|X| = n$, $S_X = S_n$. The function $\psi : G \to S_X$ defined by $\psi(a) = f_a$ is a group homomorphism.*

Proof. As we have seen above $f_a \in S_X$, so the function ψ makes sense. Suppose $b \in G$. We want to see that $\psi(ab) = \psi(a)\psi(b)$, in other words, the function f_{ab} should equal the function $f_a f_b$. Take $x \in X$, then

$$f_{ab}(x) = (ab)x = a(bx) = af_b(x) = f_a(f_b(x)).$$

Since $x \in X$ is arbitrary, we have $f_{ab} = f_a f_b$, which is what we want. □

We now give several examples of groups acting on set. In all of these cases it is easy to check the requirements needed in Definition .

Examples 23.4. (a) If X is any set, then clearly S_X acts on X. In particular, S_n acts on the set X_n.

(b) Let G be a group. If $X = G$, then the group multiplication gives an action of G on X In other words, G acts on itself via left (or right) multiplication. Note that in this case, the map $\psi : G \to S_X$ from Proposition is injective. To see this, suppose $a \in G$ and $f_a = id$. Therefore, $f_a(x) = x$, for all $x \in X$, i.e., $ax = x$ for all $x \in G$. Taking $x = e$, we have $a = e$. Thus, in this case, ψ is a one-to-one group homomorphism. Therefore, we have:

Cayley's Theorem. *Let G be a finite group of order n. Then G is isomorphic to a subgroup of S_n.*

(c) There is another important way that a group G can act on itself. Namely, via conjugation. So, if we set $X = G$ and take $a \in G$, $x \in X$, we define $ax := axa^{-1}$. Note that $exe^{-1} = x$, for all $x \in G$ and

$$(ab)x = (ab)x(ab)^{-1} = a(bxb^{-1})a^{-1} = a(bx),$$

so that conjugation is an action of G on itself.

(d) Let G be a group and H as subgroup. Let $X := \{g_\alpha H\}_{\alpha \in A}$ be the set of distinct left cosets of H in G. For $a \in G$ and $g_\alpha H \in X$, we define the action $a(g_\alpha H) := ag_\alpha H$. This clearly defines an action of G on X. In this case we say that *G acts on the left cosets of H via translation.* We obtain a similar action on the right cosets of H in G.

(e) Let G be a group and suppose G has a subgroup H with n elements. Then for any $a \in G$, aHa^{-1} is a subgroup of G with n elements. If we let X denote the set of all subgroups of G with n elements, then G acts on X via conjugation.

(f) Let F be a field, and $\mathrm{Gl}_n(F)$ be the group of $n \times n$ invertible matrices over F. If we let X denote the vector space F^n of column vectors of length n, then $\mathrm{Gl}_n(F)$ acts on X in the usual way via matrix multiplication: If $A \in \mathrm{Gl}_n(F)$ and $v \in X$, then $Av = A \cdot v$.

(g) Let G be a group and suppose there exists a group homomorphism $\phi : G \to \mathrm{Gl}_n(F)$. Then G acts on F^n via ϕ. In other words, for $a \in G$ and $v \in F^n$, we define $av := \phi(a) \cdot v$. This action forms the basis of the Representation Theory of Groups, since we call the map ϕ a *representation of* G.

(h) In a vein similar to the previous example, let G be a group, suppose X is a set and there is a group homomorphism $\phi : G \to S_X$. Then G acts on X via ϕ, i.e., we define $ax := \phi(a)(x)$, for all $a \in G$ and $x \in X$. It is easy to verify that this definition meets the requirements of a group action. It should now be clear that to give an action of the group G on the set X is equivalent to giving a group homomorphism $\phi : G \to S_X$.

(i) If H is a subgroup of G and G acts on X, then clearly H acts on X.

(j) Suppose the group G acts on the set X, and N is a normal subgroup of G such that $nx = x$, for all $n \in N$ and $x \in X$. Then G/N acts on X as: $(aN)x := ax$, where ax is the action of a on x. To see this action is well defined, suppose $aN = bN$, then $a = bn$, for some $n \in N$. Therefore, for all $x \in X$,

$$ax = (bn)x = a(nx) = bx,$$

so the action of aN on X is well defined. That this now satisfies the requirements of an action should be clear: (i) eN is the identity in G/N and $(eN)x = ex = x$, for all $x \in X$. If $aN, bN \in G/N$, then

$$(aNbN)x = (abN)x = (ab)x = a(bx) = a((bN)x) = aN((bN)x),$$

as required. This action of G/N on X is sometimes called the *induced action* of G/N on X.

Before developing some general properties of group actions, let us give a nice application of what we have seen so far.

Theorem 23.5. *Let G be a finite group with subgroup H. Assume that $[G : H] = p$, where p is the smallest prime dividing the order of G. Then H is a normal subgroup of G.*

Proof. By (d) above G acts on the set of left cosets of H, and consequently, by Proposition, there exists a group homomorphism from $\phi : G \to S_p$. Now suppose K is the kernel of ϕ. Then for $a \in K$, $\phi(a) = id$, which means, by definition of the action, $a(gH) = agH = gH$, for all cosets gH of H. In particular, $aH = H$, so $a \in H$. In other words, $K \subseteq H$. We claim $H = K$, which will show that H is normal in G. We have

$$|G/K| = \frac{|G|}{|K|} = \frac{|H| \cdot p}{|K|}.$$

Since G/K is isomorphic to the image of ϕ, which is a subgroup of S_p, it follows that $\frac{|H|\cdot p}{|K|}$ divides $p!$. Therefore $\frac{|H|}{|K|}$ divides $(p-1)!$. However, if $\frac{|H|}{|K|} \neq 1$, since any prime dividing $\frac{|H|}{|K|}$ divides $|G|$, such a prime must be greater that or equal to p. But clearly no such prime divides $(p-1)!$. It follows that $\frac{|H|}{|K|} = 1$, and therefore $H = K$, as required. □

Lecture 24: Wednesday October 20. We now discuss some general concepts related to group actions.

Definitions 24.1. Assume the group G acts on the set X. (a) For $x \in X$, $\{g \in G \mid gx = x\}$ is called the *stabilizer of* x. We will denote this set by G_x.

(b) Given $x \in X$, we define the *orbit of* x to be the set $\{gx \mid g \in G\}$. We denote the orbit of x by $\text{orb}(x)$.

Proposition 24.2. *Assume the group G acts on the set X.*

(i) G_x is a subgroup of G, for all $x \in X$.
(ii) X is a disjoint union of its distinct orbits.
(iii) For any $x \in X$, there is a one-to-one and onto correspondence between the elements of $\text{orb}(x)$ and the distinct left cosets of G_x in G. In particular, if either of these sets is finite, then $|\text{orb}(x)| = [G : G_x]$.
(iv) If G is finite, then $|\text{orb}(x)|$ divides $|G|$, for all $x \in X$.

Proof. For (i), fix $x \in X$. if $a, b \in G_x$, then $(ab)x = a(bx) = ax = x$, which shows that $ab \in G_x$. Moreover, $a^{-1}x = a^{-1}(ax) = (a^{-1}a)x = ex = x$, which shows $a^{-1} \in G_x$. Therefore, G_x is a subgroup.

For (ii), take $x_1, x_2 \in x$. If $y \in \text{orb}(x_1) \bigcap \text{orb}(x_2)$, then $g_1x_1 = y = g_2x_2$, for some $g_1, g_2 \in G$. Therefore, $x_1 = g_1^{-1}g_2$, which shows that $x_1 \in \text{orb}(x_1)$. Thus, $gx_1 \in \text{orb}(x_1)$, for all $g \in G$, which implies that $\text{orb}(x_1) \subseteq \text{orb}(x_2)$. By symmetry we have $\text{orb}(x_2) \subseteq \text{orb}(x_1)$ and thus, $\text{orb}(x_1) = \text{orb}(x_2)$. This shows that given any two orbits, they are either disjoint or equal. Since every element of X belongs to an orbit, we have that X is the disjoint union of its distinct orbits. Of course, we could also obtain this by defining an equivalence relation on X by saying two elements of X are equivalent if they lie in the same orbit. In which case, the distinct equivalence classes are the distinct orbits, which then gives the result.

For (iii), fix $x \in X$ and let $\{g_\alpha G_x\}_{\alpha \in A}$ be the distinct left cosets of G_x in G. Certainly, each $g_\alpha x$ belongs to $\text{orb}(x)$. Moreover, these elements are distinct, since if $g_{\alpha_1}x = g_{\alpha_2}x$, then $g_{\alpha_2}^{-1}g_{\alpha_2}x = x$, which means $g_{\alpha_2}^{-1}g_{\alpha_1} \in G_x$, and thus $g_{\alpha_1}G_x = g_{\alpha_2}G_x$. In addition, if $gx \in \text{orb}(x)$, g belongs to a coset $g_\alpha G_x$, so we may write $g = g_\alpha g_0$, for $g_0 \in G_x$. But then, $gx = g_\alpha g_0 x = g_\alpha x$. Thus, the function from distinct left cosets of G_x in G to $\text{orb}(x)$ defined by $g_\alpha G_x \to g_\alpha x$ is one-to-one and onto. The second part of (iii) follows immediately from this.

Part (iv) follows immediately from (iii) since $[G : G_x]$ divides the order of G. □

The following corollary is an immediate consequence of Proposition 24.2, but it has numerous powerful consequences.

Corollary 24.3. *Let G be a finite group acting on the finite set X. Let $orb(x_1), \ldots, orb(x_n)$ be the distinct orbits of X. Suppose that for $i = 1, \ldots, r$, $|orb(x_i)| = 1$ and for $r + 1 \leq i \leq n$, $|orb(x_i)| > 1$. Here we allow for $r = 0$ or $r = n$. Then:*

$$|X| = r + \sum_{i=r+1}^{n} |\text{orb}(x_i)| = r + \sum_{i=r+1}^{n} [G : G_{x_i}]. \qquad \square$$

Definitions and discussion 24.4. Let the group G act on itself via conjugation. If $a \in G$, then the orbit of a under this action is the set of all gag^{-1}, such that $g \in G$. In this case, we refer to the orbit of a as the *conjugacy class of* a. Any element of the conjugacy class of a is called a *conjugate* of a. We will write $C(a)$ for the conjugacy class of a. Note, that by the proposition above $|C(a)|$ divides $|G|$, if G is a finite group. Suppose $g \in G$ belongs to the stabilizer of a. Then $gag^{-1} = a$, so that $ga = ag$. In other words, $g \in G_a$ if and only if g commutes with a. In this case we write $C_G(a)$ instead of G_a and refer to $C_G(a)$ as the *centralizer of a in* G. Finally, note that $|C(a)| = 1$ if and only if $C(a) = a$ if and only if $g^{-1}ag = a$ for all $g \in G$, i..e, a commutes with every element of G. We call the set of such elements, the *center of* G, and write $Z(G)$ to denote the center of G. It is completely straight forward to check that $Z(G)$ is a *normal subgroup* of G.

The following **fundamental equation** now follows immediately from the previous corollary.

The Class Equation. *Let G be a finite group. Then*

$$|G| = |Z(G)| + \sum_i |C(a_i)| = |Z(G)| + \sum_i [G : C_G(a_i)],$$

where the sum is taken over all conjugacy classes containing more than one element. □

The following theorem shows the power of the class equation, which is really an immediate consequence of the corollary above.

Theorem 24.4. *Let G be a group, with $|G| = p^n$, where p is a prime number and $n \geq 1$. Then:*

(i) $Z(G) \neq \{e\}$.
(ii) G is a solvable group.

Proof. From the class equation we have $p^n = |Z(G)| + \sum_i |C(a_i)|$, where the sum is taken over all conjugacy classes containing more than one element. Thus, each $|C(a_i)|$ is divisible by p, since each $|C(a_i)|$ divides $|G|$. But this forces $|Z(G)|$ to be divisible by p, so that $Z(G) \neq e$.

For part (ii), we induct on n. If $n = 1$, then G is cyclic of order p and clearly solvable. Suppose $n > 1$. If G is abelian, there is nothing to prove. Otherwise by (i), $Z(G)$ is a group of order p^r, with $1 \leq r < n$. Now, $Z(G)$ is solvable since it is abelian. On the other hand, $|G/Z(G)| = p^{n-r}$, so by induction, $G/Z(G)$ is also solvable. Therefore, G is solvable, by Theorem 19.3. □

Discussion and Definition 24.5. We want to now turn to the most important theorems in the elementary aspects of finite group theory, the Sylow theorems. The first of these theorems provides the best possible converse to Lagrange's Theorem. A full converse to Lagrange's Theorem would say that if G is a finite group and the positive integer m divides $|G|$, then G has a subgroup of order m. Unfortunately, this is false. However, it is true if m is a power of a prime number. The first step along the path to the Sylow theorems is the following theorem due to Cauchy, whose celebrated proof is due to J. McKay.

Cauchy's Theorem. *Let G be a finite group and suppose p is a prime dividing $|G|$. Then G has an element of order p.*

Proof. Let X denote the set of all p-tuples $(a_1, \ldots, a_p)$ of elements of G such that $a_1 \cdots a_p = e$. Note that if $|G| = n$, then $|X| = n^{p-1}$, since we may freely choose the first $p - 1$ coordinates of an element of X, and the last coordinate is determined by the equation $a_1 \ldots a_{p-1}a_p = e$.

We now note that additive group $\mathbb{Z}$ acts on X by cyclically permuting the entries of an element of x. In other words, for $t \in \mathbb{Z}$, if $t \geq 0$, t acts on $(a_1, \ldots, a_p)$ by moving each entry forward by t positions. If $t < 0$, move each entry in reverse by $|t|$ steps.

Note that if $|t| \geq p$, we can write $t = pq + r$, with $0 \leq r < p$, and the action of r and t agree on the elements of X, since p moves forward (or backward) does not change any element of X. Writing "$\cdot$ " for the action, we then have

$$t \cdot (a_1, \ldots, a_p) = r \cdot (a_1, \ldots, a_p) = (a_{p-r}, a_{p-r+1}, \cdots, a_p, a_1, \ldots, a_{p-r-1}).$$

Note that $r \cdot (a_1, \ldots, a_p) \in X$, since if $a_1 \cdots a_{p-r-1}a_{p-r} \cdots a_p = e$, then

$$a_1 \cdots a_{p-r-1} = a_p^{-1} \cdots a_{p-r}^{-1}$$

which implies

$$a_{p-r} \cdots a_p a_1 \cdots a_{p-r-1} = e,$$

which gives $r \cdot (a_1, \ldots, a_p) \in X$. Thus, it follows at once that $\mathbb{Z}$ acts on X.

Now, any element of $p\mathbb{Z}$ fixes the elements of X, so by Example 23.4 (j) $\mathbb{Z}_p$ acts on X - and the action is the same as the $\mathbb{Z}$ action. By Proposition 24.2(iv), the order of any orbit divides $|\mathbb{Z}_p| = p$, thus, the order of any orbit, when $\mathbb{Z}_p$ acts on X, is either 1 or p. On the other hand, for $x \in X$ either all coordinates of x are the same, in which case the orbit of x has one element, or at least two coordinates in x differ, in which case the orbit of x has p elements. So, if there are r orbits consisting of one element and s orbits consisting of p elements, we have $r + sp = n^{p-1}$. Since p divides n, p must divide r. Thus, the number of orbits with a single element is a multiple of p and therefore, certainly greater than one. But each such orbit corresponds to a p-tuple of the form $(a, a, \dots, a)$ such that $a \cdots a = a^p = e$. Since $o(a)$ must divide p, we have $o(a) = p$. Therefore, G has (at least $p - 1$) elements of order p. □

Lecture 25: Friday October 22. We begin with a definition.

Definition 25.1. Suppose G is a finite group and p is a prime number dividing $|G|$. Assume that p^n is the largest power of p dividing $|G|$. A subgroup P of G having order p^n is called a Sylow p-subgroup of G. The first Sylow theorem will tell us that Sylow p-subgroups always exist. The second and third theorem relate to the number of Sylow subgroups and their relationship via conjugation.

First Sylow Theorem. *Suppose G is a finite group and p is a prime dividing $|G|$. Then G contains a Sylow p-subgroup.*

Proof. Suppose $|G| = p^n m$, where p does not divide m. We induct on $|G|$. The base case is $|G| = p$, and in this case there is nothing to prove. Now suppose the result is true for groups of order less than $p^n m$. If $n = 1$, then by Cauchy's Theorem, there exists $a \in G$ having order p, and thus $\langle a \rangle$ is a Sylow p-subgroup.

Suppose $n > 1$. If G has a proper subgroup H whose index is not divisible by p, then $|H|$ is divisible by p^n, so by induction H has a subgroup of order p^n, which is also a Sylow p-subgroup of G. Otherwise, every proper subgroup of G has index divisible by p. It follows from the class equation that $|Z(G)|$ is divisible by p. Let $a \in Z(G)$ be an element of order p. Then $\langle a \rangle$ is a normal subgroup and $|G/\langle a \rangle| = p^{n-1}m$. By induction, $G/\langle a \rangle$ has a Sylow p-subgroup. By Corollary 18.2, this Sylow subgroup is of the form $P/\langle a \rangle$, for P a subgroup of G containing $\langle a \rangle$. But now, $p^{n-1} = |P/\langle a \rangle| = \frac{1}{p} \cdot |P|$, which shows that $|P| = p^n$, which is what we want. □

Corollary 25.2. *Let G be a finite group and suppose p is a prime dividing $|G|$. Write $|G| = p^n m$, with $p \nmid m$. Then G has a subgroup of order p^i, for all $1 \le i \le n$. In fact, if P is any Sylow p-subgroup of G, then P has a subgroup of order p^i for all $1 \le i < n$.*

Proof. Sylow p-subgroups exist by the First Sylow Theorem. Thus, it suffices to prove the second statement. Let P be a Sylow p-subgroup of G. The conclusion of the corollary holds for $i = 1$ by Cauchy's Theorem. We may therefore assume $1 < i < n$. By Theorem 24.3, $Z(P) \neq \{e\}$. By Cauchy's theorem, there exists $a \in Z(P)$ having order p, so that $|P/\langle a \rangle| = p^{n-1}$. By induction (say on n), $P/\langle a \rangle$ contains a subgroup of order p^{i-1}. By Corollary 18.2, this subgroup is of the form $H/\langle a \rangle$ for a subgroup H of P containing $\langle a \rangle$. But now, $p^{i-1} = |H/\langle a \rangle| = \frac{1}{p} \cdot |H|$, which shows that $|H| = p^i$, which is what we want. □

We need the following lemma in order to prove the second and third Sylow theorems.

Lemma 25.3. *Let H be a group of order p^s, where p is prime and assume H acts on X. Let r denote the number of orbits with one element. Then $|X| \equiv r$ (mod p).*

Proof. By Corollary 24.3, $|X| = r + \sum_{i=r+1}^{n} |\text{orb}(a_i)|$, where the sum is taken over the distinct orbits with more than one element. Since each $|\text{orb}(a_i)|$ divides $|H| = p^s$, p divides each $|\text{orb}(a_i)|$. Thus, p divides $\sum_{i=r+1}^{n} |\text{orb}(a_i)|$, which gives what we want. □

Second Sylow Theorem. *Let G be a finite group and p a prime dividing $|G|$ with $|G| = p^n m$ and $p \nmid m$. Suppose that H is a subgroup of G with $|H| = p^r$, so that $1 \le r \le n$. Let P be any Sylow p-subgroup of G. Then there exists $a \in G$ such that $H \subseteq aPa^{-1}$. In particular, any subgroup of order p^r is contained in a Sylow p-subgroup and any two Sylow p-subgroups of G are conjugate.*

Proof. Let X the set of left cosets of P, so that $|X| = [G : P]$. Therefore $p \nmid |X|$. Now let H act on X via left translation, i.e., $h(aP) = haP$, for all $h \in H$ and $aP \in X$. Letting r denote the number of orbits with one element, we have by the previous lemma, that $|X| \equiv r$ (mod p). It follows that $r \neq 0$. Thus, there exists a coset aP whose orbit under the action of H consists only of aP. Therefore, $haP = aP$, for all $h \in H$. In particular, $ha \in aP$ for all $h \in H$, and therefore $h \in aPa^{-1}$, for all $h \in H$, which gives what we want.

Now, let P_1, P_2 be two Sylow p-subgroups. Then $P_1 \subseteq aP_2a^{-1}$, for some $a \in G$. Since P_1 and aP_2a^{-1} have the same number of elements, $P_1 = aP_2a^{-1}$, and the proof is complete. □

Third Sylow Theorem. *Let G be a finite group and p a prime dividing the order of G. Then the number of Sylow p-subgroups divides $|G|$ and is congruent to 1 modulo p.*

Proof. Let X be the set of Sylow p-subgroup and take $P \in X$. For the first statement, G acts on X via conjugation, and by the Second Sylow Theorem, $X = \text{orb}(P)$. Thus, $|X| = |\text{orb}(P)|$, which divides $|G|$.

For the second statement fix $P \in X$ and let P act on X via conjugation. Then $\text{orb}(P) = \{P\}$. On the other hand, suppose $Q \in X$ and $\text{orb}(Q) = Q$. Then $pQp^{-1} = Q$, for all $p \in P$, so that $pQ = Qp$, for all $p \in P$. It follows, that as subsets of G, $PQ = QP$. But this implies that PQ is a subgroup of G (by problem 6 on Homework 3). Suppose $p^n = |P| = |Q|$. Then

$$|PQ| = \frac{|P| \cdot |Q|}{|P \cap Q|} = \frac{p^{2n}}{|P \cap Q|}.$$

Since $|PQ|$ divides $|G|$ and p^n is the largest power of p dividing $|G|$, it follows that $|P \cap Q| = p^n$. In other words $P = Q$. Thus, $\{P\}$ is the only orbit of order one, and therefore by Lemma 24.3, $|X| \equiv 1 \pmod{p}$. □

Here are some examples illustrating how one can use the Sylow theorems to show that certain groups of small order are not simple. A deeper application of these techniques will be used in our proof showing that A_5 is (up to isomorphism) the only simple group of order 60.

Examples 25.4. In each of the examples below, we show that G must have a non-trivial normal subgroup.

(a) $|G| = 136 = 2^3 \cdot 17$. The number of Sylow 17-subgroups must divide 136 and be congruent to 1 mod 17. No multiple of 17 is congruent to 1 mod 17, which leaves 1, 2, 4, 8 as possible divisors. But neither 2, 4, 8 are congruent to 1 mod 17. Thus, there is only one Sylow 17-subgroup of G, which is necessarily normal in G. Therefore, G is not a simple group.

(b) $|G| = 49 = 7^2$. By Theorem 24.3, $Z(G) \neq e$, and thus $Z(G)$ is a non-trivial normal subgroup of G. Therefore, G is not simple.

(c) $|G| = 105 = 3 \cdot 5 \cdot 7$. Using Sylow's third theorem, we have:

$$\begin{aligned}&\text{Possible number of Sylow 3-subgroups} = 1, 7\\&\text{Possible number of Sylow 5-subgroups} = 1, 21\\&\text{Possible number of Sylow 7-subgroups} = 1, 15.\end{aligned}$$

In each of the cases above, if there is just one Sylow subgroup, it is normal in G, and G is not simple. Suppose that none of the numbers of Sylow p-subgroups of G equals 1. Then, among the seven Sylow 3-subgroups, the intersection of any two of them equals e, since each of these groups have prime order. Thus, there are $2 \cdot 7 = 14$ elements of order 3 in G. Similar reasoning applied to the Sylow 5-subgroups shows that there must be $4 \cdot 21 = 84$ elements of order 5 in G and likewise, $15 \cdot 6 = 90$ elements of order seven. But adding these together gives more than 105 elements. Thus, for one of the primes 3, 5, 7, there is just one Sylow subgroup, which is necessarily normal in G. Therefore G is not a simple group.

(d) $|G| = 24$. Then, $24 = 2^3 \cdot 3$. It follows from the third Sylow theorem that the number of Sylow 2-subgroups is 1 or 3. If one, we're done. Suppose there are three Sylow 2-subgroups. Let X denote the set of Sylow 2-subgroups. Then G acts on X via conjugation. By Proposition 23.3, there is a group homomorphism $\phi : G \to S_3$. Since $|S_3| = 6$, the kernel of ϕ is a non-trivial normal subgroup, so G is not a simple group.

Lecture 26: Monday October 25. The next theorem is further illustration of how to use counting techniques together with the Sylow theorems to prove significant results about finite groups.

Theorem 26.1. *If G is a simple group of order 60, then G is isomorphic to A_5.*

Proof. We seek a group homomorphism $\psi : G \to S_5$. Suppose such a homomorphism exists. Since G is simple, $\ker(\psi) = \{e\}$. Thus, G is isomorphic to a subgroup N of S_5. Since $|N| = 60$, $[S_5 : N] = 2$, which implies that N is normal in S_5. By Corollary 22.4, $N = A_5$, which is what we want. To find ψ, we apply the Sylow theorems.

We first note that $60 = 2^2 \cdot 3 \cdot 5$, so that any Sylow 2-subgroup has four elements, any Sylow 3-subgroup has three elements and any Sylow 5-subgroup have five elements. By the Third Sylow Theorem, for $p = 2, 3, 5$, the number of p-Sylow subgroups divides 60 and is congruent to 1 modulo p. Based upon this we have:

$$\begin{aligned}
&\text{Possible number of Sylow 2-subgroups} = 1, 3, 5, 15\\
&\text{Possible number of Sylow 3-subgroups} = 1, 4, 10\\
&\text{Possible number of Sylow 5-subgroups} = 1, 6
\end{aligned}$$

In each case, we can eliminate the possibility of one Sylow subgroup, for then that subgroup would be normal in G, contradicting the simplicity of G. Suppose we had four Sylow 3-subgroups. Then if G acts on the set of Sylow 3-subgroups via conjugation, by Proposition 23.2, we would have a group homomorphism $\psi : G \to S_4$. Since $|S_4| = 24$, $\ker(\psi) \neq \{e\}$. But then G would have a non-trivial normal subgroup, which is a contradiction. Thus, G must have ten Sylow 3-subgroups. Now, any two of the Sylow 3-subgroup have to intersect in e, since the intersection would is a subgroup of each. Thus, there are $10 \cdot 2 = 20$ elements in G having order 3.

Since we cannot have one Sylow 5-subgroup, there must be six of them, and since they each have prime order, the intersection of any two of them must be e. Thus, there are $6 \cdot 4 = 24$ elements of order 5 in G. We have now accounted for $20 + 24 = 44$ elements of G.

Regarding the number of Sylow 2-subgroups, we cannot have just three of them, otherwise, there would exist a group homomorphism $\psi : G \to S_3$, which would necessarily have a non-trivial kernel. If the number of Sylow 2-subgroups equals five, then we let G act on the set of Sylow 2-subgroups via conjugation. By Proposition 23.3, there exists a group homomorphism $\psi : G \to S_5$, which is what we want.

Suppose the number of Sylow 2-subgroups equals fifteen. First assume the following: Given any two such subgroups, say, P_1 and P_2, $P_1 \cap P_2 = e$. Then G would contain $15 \cdot 3 = 45$ elements of order two or four, which is a contradiction, since we already have 44 elements of order three or five. Thus, there must be at least two P_1, P_2 with $P_1 \cap P_2 \neq e$. Then $|P_1 \cap P_2| = 2$. If $e \neq x \in P_1 \cap P_2$, then since P_1 and P_2 are abelian, x commutes with every element in both P_1 and P_2. Thus, both P_1 and P_2 are contained in $C_G(x)$, the centralizer of x. Since $|P_1P_2| = 8$, and $P_1P_2 \subseteq C_G(x)$ (as sets), $|C_G(x)| \geq 8$ and $|C_G(x)|$ divides 60. Moreover, since $P_1 \subseteq C_G(x)$, 4 divides $|C_G(x)|$. Thus, $|C_G(x)| = 12$ or 20.

Suppose $|C_G(x)| = 12$. Then $[G : C_G(x)] = 5$. Thus, if we let G act on the set of left cosets of $C_G(x)$ in G via translation, by Proposition 23.3 , there exists a group homomorphism $\psi : G \to S_5$, which is what we want. If $|C_G(x)| = 20$, then, in a similar way, there would exist $\psi : G \to S_3$, which would have non-trivial kernel, so this case does not exist. Thus, if the number of Sylow 2-subgroups equals fifteen, then required map $\psi : G \to S_5$ exists and this completes the proof of the theorem. □

Remarks. (a) Let G be a simple group of order 60. The proof of the theorem above showed that the number of Sylow 3-subgroups of G equals 10 and the number of Sylow 5-subgroups equals 6. Moreover, any Sylow 3-subgroup is isomorphic to $\mathbb{Z}_3$ and any Sylow 5-subgroup is isomorphic to $\mathbb{Z}_5$. The proof of the theorem is not definitive in regards to the Sylow 2-subgroups of G. The proof shows that the number of Sylow 2-subgroups is either 5 or 15. Moreover, any such subgroup has order four. Since these subgroups are all conjugate, they are all isomorphic. So: Are the Sylow 2-subgroups of G isomorphic to $\mathbb{Z}_4$ or $\mathbb{Z}_2 \times \mathbb{Z}_2$, and how many copies of each are in G?

(b) The element x in the proof of Theorem 26.1 has order two, i.e., $x^2 = e$. Such an element is called an *involution.* We were able to make considerable headway by using $C_G(x)$, which is a *centralizer of an involution.* If you read any historical account about the Feit-Thomspon theorem, or the classification of simple groups, or results of Brauer, Suzuki, et al, leading up to these results, you will see that centralizers of involutions play a crucial part in all of these results. Since any non-abelian simple group has even order, there are plenty of involutions, and thus, plenty of centralizers of involutions. □

We are almost ready to address the issue of solvability by radicals. Before doing so, let us show how two of our major theorems - The Galois Correspondence Theorem and the first Sylow Theorem - can be used to show that $\mathbb{C}$ is the algebraic closure of $\mathbb{R}$. This is the so-called Fundamental Theorem of Algebra. We note that it is not possible to give a *purely algebraic* proof of this fact, simply because the real numbers are not a purely algebraic construct - they are, after all, equivalence classes of Cauchy sequences of rational numbers. So, we need one fact with an analytic flavor. Namely: If $f(x) \in \mathbb{R}[x]$ has odd degree, then $f(x)$ has a root in $\mathbb{R}$. This is a consequence of the mean value theorem from calculus. Thus, no polynomial of odd degree with coefficients in $\mathbb{R}$ is irreducible over $\mathbb{R}$.

The Fundamental Theorem of Algebra. *The complex number field $\mathbb{C}$ is the algebraic closure of the real number field $\mathbb{R}$. In particular, $\mathbb{C}$ is an algebraically closed field.*

Proof. Since $\mathbb{C}$ is algebraic over $\mathbb{R}$, it suffice to show that $\mathbb{C}$ is algebraically closed. For this, it suffices to show that there are no proper finite extensions of $\mathbb{C}$. We first note that if K is a finite extension of $\mathbb{R}$, then $[K : \mathbb{R}] = 2^n$, for some n. To see this, let $\mathbb{R} \subseteq K$ be a finite extension. If necessary, we may enlarge K to a finite extension K_0 of $\mathbb{R}$ that is Galois over $\mathbb{R}$. If $[K_0 : \mathbb{R}] = 2^m$, for some m, then $[K : \mathbb{R}] = 2^n$, for some n, so we may assume at the start that K is Galois over $\mathbb{R}$. Let G denote the Galois group of K over $\mathbb{R}$. If 2 divides $|G|$, let $H \subseteq G$ be a Sylow 2-subgroup of G, otherwise, let $H = \{e\}$. Then $[G : H]$ is odd. Thus, $[E : \mathbb{R}]$ is is odd, where E is the fixed field of H. (Note, if $H = \{e\}, E = K$.) Since $\mathbb{R}$ has characteristic zero, the extension $\mathbb{R} \subseteq E$ is a simple extension, say $E = \mathbb{R}(\alpha)$. But if we let $f(x)$ denote the minimal polynomial of α, its degree equals $[E : \mathbb{R}]$ which is odd, contradicting our remarks above. Thus, any finite extension of $\mathbb{R}$ has degree 2^n, for some $n \geq 1$. Of course the conclusion of the theorem tells us $n = 1$.

Now, suppose $\mathbb{C} \subseteq L$ is a finite extension of fields. As before, we may assume L is Galois over $\mathbb{C}$. Since L is also finite over $\mathbb{R}$, by what we have just seen, $[L : \mathbb{R}] = 2^n$, for some n, and since $[\mathbb{C} : \mathbb{R}] = 2$, $[Ł : \mathbb{C}] = 2^{n-1}$. If $n = 1$, $L = \mathbb{C}$, which is what we want. Suppose $n > 1$. We seek a contradiction. Let us note that there exists an intermediate field $\mathbb{C} \subseteq E \subseteq L$ with $[E : \mathbb{C}] = 2$. If $n = 2$, we take $E = L$. If $n > 2$, let G denote the Galois group of L over $\mathbb{C}$. Since $|G| = 2^{n-1}$, G has a subgroup H of index 2 (say by Corollary 25.2). If E denotes the fixed field of H, then $[E : \mathbb{C}] = 2$. We now argue that such an extension cannot exist.

Since $E = \mathbb{C}(\beta)$, for some $\beta \in E$, the minimal polynomial $h(x) \in \mathbb{C}[x]$ for β over $\mathbb{C}$ has degree two. If $h(x) = x^2 + bx + c$, with $b, c \in \mathbb{C}$, $\beta = \frac{-b \pm \sqrt{b^2-4c}}{2}$. If we show that every complex number has a square root, then $\beta \in \mathbb{C}$, which is the contradiction we sought.

There are several ways to see that any $z \in \mathbb{C}$ has a square root in $\mathbb{C}$. The easiest is to use the polar representation of any complex number. So, write $z = re^{i\theta}$, where r is a positive real number, and θ is the angle a line segment from (0,0) to z makes (going counter-clockwise) from the x-axis. Then if we set $z_0 := \sqrt{r}e^{i\frac{\theta}{2}}$, we have $z_0^2 = z$, which gives what we want. This complete the proof of the FTA. □

Corollary 26.2. *Let $f(x) \in \mathbb{R}[x]$ be an irreducible polynomial. Then the degree of $f(x)$ is less than or equal to two. Thus, any polynomial in $\mathbb{R}[x]$ can be factored (uniquely) as a product of linear and quadratic polynomials.*

Proof. It suffices to prove the first statement. Suppose $f(x) \in \mathbb{R}[x]$ is irreducible over $\mathbb{R}$ and has degree greater than one. Let α be a root of $f(x)$, so that by the Fundamental Theorem of Algebra, we may assume $\alpha \in \mathbb{C}\backslash\mathbb{R}$. Since $[\mathbb{C} : \mathbb{R}] = 2$ and $\mathbb{R}(\alpha) \subseteq \mathbb{C}$, we have $[\mathbb{R}(\alpha) : \mathbb{R}] = 2$, which implies that $f(x)$ has degree two. □

Lecture 27: Wednesday October 27. We return to the task of determining which polynomial equations are solvable by radicals. We start with two fundamental definitions.

Definitions 27.1. (a) Let $F \subseteq L$ be a finite extension of fields. We say that L is a radical extension of F if there exist $\alpha_1, \ldots, \alpha_r \in$ and positive integers $n_1, \ldots, n_r$ such that $L = F(\alpha_1, \ldots, \alpha_r)$, $\alpha_1^{n_1} \in F$ and $\alpha_i^{n_i} \in F(\alpha_1, \ldots, \alpha_i)$, for $2 \leq i \leq r$. Note that there is no assumption on the positive integers n_i. Thus, if we set $n = n_1 \cdots n_r$, then we have that $\alpha_1^n \in F$ and $\alpha_i^n \in F(\alpha_1, \ldots, \alpha_i)$, for $2 \leq i \leq r$. Therefore, in any radical extension of F, we are free to assume that a single common exponent works for each α_i.

(b) For $f(x) \in F[x]$, we say that $f(x)$ is solvable by radicals if its splitting field is contained in a radical extension of F.

Exampless 27.2. Consider $K := \mathbb{Q}(\sqrt[3]{2}, \epsilon)$, the splitting field of $f(x) = x^3 - 2$ over $\mathbb{Q}$, where ϵ is a primitive cube root of unity. Then $\epsilon^3 = 1 \in \mathbb{Q} \subseteq \mathbb{Q}(\sqrt[3]{2})$ and $(\sqrt[3]{2})^3 \in \mathbb{Q}$, so in this case the splitting field of $f(x)$ is a radical extension. In general, the splitting field of a polynomial solvable by radicals need not be a radical extension. Of course, x^3-2 is solvable by radicals, essentially by definition. On the other hand, it is not so clear that $f(x)=x^4-16x^2+4$ is solvable by radicals. However, $f(x)$ is the minimal polynomial for $\sqrt{3}+\sqrt{5}$ over $\mathbb{Q}$ (see Homework 1). Since $\sqrt{3}+\sqrt{5}\in K:=\mathbb{Q}(\sqrt{3},\sqrt{5})$, and K is Galois over $\mathbb{Q}$, $f(x)$ splits over K. Thus, the splitting field for $f(x)$ is contained in a radical extension, and in this case, K equals the splitting field of $f(x)$, since $K=\mathbb{Q}(\sqrt{3}+\sqrt{5})$, so adjoining the other roots of $f(x)$ to $\mathbb{Q}$ is redundant.

Here is the statement of the solvability by radicals theorem.

Solvability by Radicals Theorem. *Let F be a field having characteristic zero and $f(x)\in F[x]$. Let K denote the splitting field of $f(x)$. Then $f(x)$ is solvable by radicals if and only if Gal(K/F) is a solvable group.* □

Since our main goal requires that F be a field of characteristic zero, this assumption will hold in everything that follows, until further notice. Thus, if $F\subseteq K$ is a finite extension, K will automatically be a separable extension, and therefore K is Galois over F if and only if K is a splitting field over F. We will use this fact implicitly henceforth. We begin with the following key proposition.

Proposition 27.3. *Let F be a field of characteristic zero and ϵ be a primitive nth root of unity.*

(i) $F\subseteq F(\epsilon)$ is a Galois extension with abelian Galois group.

(ii) Suppose $f(x)=x^n-a\in F[x]$. Assume that $\epsilon\in F$. Let $\alpha\in\overline{F}$ be a root of $f(x)$ and set $L:=F(\alpha)$. Then L is a Galois extension of F and $\text{Gal}(L/F)$ is abelian.

Proof. For (i), since ϵ is a primitive nth root of unity, $1,\epsilon,\ldots,\epsilon^{n-1}$ are the distinct roots of $g(x)=x^n-1$. Thus, $F(\epsilon)$ is the splitting field of $g(x)$ over F, and thus $F(\epsilon)$ is Galois over F. (Note: $[F(\epsilon):F]$ is not n.) Suppose $\sigma,\tau\in\text{Gal}(F(\epsilon)/F)$. Then σ,τ must take ϵ to a root of $g(x)$, and thus $\sigma(\epsilon)=\epsilon^i$ and $\tau(\epsilon)=\epsilon^j$, for some $0\le i,j\le n-1$. Then

$$\sigma\tau(\epsilon)=\sigma(\epsilon^j)=\sigma(\epsilon)^j=\epsilon^{ij},$$

and

$$\tau\sigma(\epsilon)=\tau(\epsilon^i)=\tau(\epsilon)^j=\epsilon^{ji},$$

so that $\sigma\tau(\epsilon)=\tau\sigma(\epsilon)$. Since $L=F(\epsilon)$, it follows that $\sigma\tau=\tau\sigma$ as elements of $\text{Gal}(F(\epsilon)/F)$, and thus, $\text{Gal}(F(\epsilon)/F)$ is abelian

The proof of (ii) is essentially the same as the proof of part (i). L is clearly Galois over F since it is the splitting field of $f(x)$ over F. Suppose $\sigma,\tau\in\text{Gal}(K/F)$. Since the roots of $f(x)$ are $\{\alpha,\alpha\epsilon,\ldots,\alpha\epsilon^{n-1}\}$, $\sigma(\alpha)=\alpha\epsilon^i$ and $\tau(\alpha)=\alpha\epsilon^j$, for some $0\le i\ne j\le n-1$. Then

$$\sigma\tau(\alpha)=\sigma(\alpha\epsilon^j)=\sigma(\alpha)\sigma(\epsilon^j)=\alpha\epsilon^i\cdot\epsilon^j=\alpha\epsilon^{i+j},$$

and

$$\tau\sigma(\alpha)=\tau(\alpha\epsilon^i)=\tau(\alpha)\tau(\epsilon^i)=\alpha\epsilon^j\cdot\epsilon^i=\alpha\epsilon^{i+j},$$

so that $\sigma\tau(\alpha)=\tau\sigma(\alpha)$. Since $L=F(\alpha)$, it follows that $\sigma\tau=\tau\sigma$ as elements of $\text{Gal}(L/F)$, and thus, $\text{Gal}(L/F)$ is abelian. □

The next corollary follows immediately from the previous proposition and is clearly important for the solvability theorem.

Corollary 27.4. *Let F be a field of characteristic zero and $F\subseteq L$ a Galois extension that is also a radical extension with $L=F(\alpha_1,\ldots,\alpha_r)$, $\alpha_1^n\in F$ and $\alpha_i^n\in F(\alpha_1,\ldots,\alpha_i)$, for $2\le i\le r$. Assume that $\epsilon\in F$ is a* primitive nth root of unity. *Then* $\text{Gal}(L/F)$ *is a solvable group.*

Proof. Set $L_0:=F$, and for $1\le i\le r$, set $L_i=F(\alpha_1,\ldots,\alpha_i)$, so that $L_r=L$. Note that for each $1\le i\le r$, $\epsilon\in L_i$. Thus, by the previous proposition, $\text{Gal}(L_i/L_{i-1})$ is abelian, for each $1\le i\le r$. On the other hand, L is Galois over F and therefore Galois over L_{i-1}. Since L_i is Galois over L_{i-1}, $\text{Gal}(L/L_i)$ is a normal subgroup of $\text{Gal}(L/L_{i-1})$. Thus,

$$id=\text{Gal}(L/L_r)\subseteq\text{Gal}(L/L_{r-1})\subseteq\cdots\subseteq\text{Gal}(L/L_1)\subseteq\text{Gal}(L/F)$$

is a solvable series, since $\text{Gal}(L/L_{r-1})/\text{Gal}(L/L_r)=\text{Gal}(L_i/L_{i-1})$, by the Galois Correspondence Theorem. Therefore, $\text{Gal}(L/F)$ is a solvable group.

The next proposition shows that a radical extension can always be enlarged to a radical extension that is also a Galois extension. Thus, if $f(x) \in F[x]$ is solvable by radicals, we can assume that its splitting field is contained in an extension of F that is both a radical and a Galois extension of F.

Proposition 27.5. *Let F be a field of characteristic zero and let $F \subseteq L$ be a radical extension with $L = F(\alpha_1, \ldots, \alpha_r)$, $\alpha_1^n \in F$ and $\alpha_i^n \in F(\alpha_1, \ldots, \alpha_i)$, for $2 \leq i \leq r$. Then there exists a radical extension $F \subseteq E$ such that E contains L and E is Galois over F.*

Proof. Each α_i is algebraic over F, so we let $f_i(x)$ be the minimal polynomial of α_i over F. For each $1 \leq i \leq r$, suppose the degree of $f_i(x)$ equals e_i and let $\alpha_i := \alpha_{i,1}, \ldots, \alpha_{i,e_i}$ be the distinct roots of $f_i(x)$. We set $E_0 := F$ and for each $1 \leq i \leq r$,

$$E_i := F(\alpha_{1,1}, \ldots, \alpha_{1,e_1}, \ldots, \alpha_{i,1}, \ldots, \alpha_{i,e_i}),$$

Thus, E_i is the splitting field of $f_1(x) \cdots f_i(x)$ over F. In particular, E_r is a splitting field over F, and therefore is Galois over F. Note also that $L \subseteq E_r$. We set $E := E_r$

We will now show that E is also a radical extension of F. Upon doing so, the proof will be complete. Note that $\alpha_i^n \in E_{i-1}$, for $1 \leq i \leq r$. Fix $1 \leq i \leq r$ and $1 \leq j \leq e_i$. If we show that $\alpha_{i,j}^n \in E_{i-1}$, this will show that E is a radical extension of F. Now, by Proposition 2.1, there exists an isomorphism $\phi : F(\alpha_i) \to F(\alpha_{i,j}) \subseteq \overline{F}$ such that $\phi(\alpha_i) = \alpha_{i,j}$ and ϕ fixes F. By Theorem 9.1, there exists a field homomorphism $\sigma : E_i \to \overline{F}$ extending ϕ. Since E_{i-1} is a splitting field over F, Theorem 10.4 gives $\sigma(E_{i-1}) = E_{i-1}$. Now, there exists $a \in E_{i-1}$ such that $\alpha_i^n = a$. Applying σ we get

$$\sigma(a) = \sigma(\alpha_i^n) = \sigma(\alpha_i)^n = \alpha_{i,j}^n,$$

which gives what we want, since $\sigma(a) \in E_{i-1}$. □

Lecture 28: Friday October 29. We are almost ready to prove half of the criterion for solvability by radicals. We wish to apply Corollary 27.4 above. Unfortunately an arbitrary field F may not contain the required primitive root of unity. So this must be adjoined to F in order to apply the Corollary 27.4. Fortunately, doing so does not disturb things too greatly, as the following proposition shows.

Proposition 28.1. Suppose that F has characteristic zero, $F \subseteq K$ be a Galois extension and ϵ a primitive nth root of unity. Assume $\epsilon \notin F$. Then:

(i) K is contained in a radical extension of F if and only if $K(\epsilon)$ is contained in a radical extension of $F(\epsilon)$.
(ii) $\text{Gal}(K/F)$ is a solvable group if and only if $\text{Gal}(K(\epsilon)/F(\epsilon))$ is a solvable group.

Proof. Suppose K is contained in the radical extension L of F. Then clearly $L(\epsilon)$ is a radical extension of $F(\epsilon)$. Since $K(\epsilon) \subseteq L(\epsilon)$, this shows $K(\epsilon)$ is contained in a radical extension of $F(\epsilon)$. Conversely, suppose $K(\epsilon)$ is contained in the radical extension L_0 of $F(\epsilon)$. Since $\epsilon^n \in F$, L_0 is a radical extension of F, and it clearly contains K.

Suppose $\text{Gal}(K/F)$ is solvable. Define $\phi : \text{Gal}(K(\epsilon)/F(\epsilon) \to \text{Gal}(K/F)$ by $\phi(\sigma) = \sigma_{|_K}$. Since K is a splitting field over F, $\sigma(K) = K$, and since σ fixes $F(\epsilon)$, $\sigma_{|_K} \in \text{Gal}(K/F)$, and so ϕ is well defined. The function ϕ is clearly a group homomorphism. Moreover, if $\phi(\sigma) = id$, then σ fixes every element of K, and since by definition, σ fixes ϵ, σ fixes $K(\epsilon)$, i.e., $\sigma = id$. Therefore ϕ is one-to-one and therefore $\text{Gal}(K(\epsilon)/F(\epsilon))$ is isomorphic to a subgroup of $\text{Gal}(K/F)$. Therefore, by Proposition 19.3(i), $\text{Gal}(K(\epsilon)/F(\epsilon))$ is solvable.

Conversely, suppose that $\text{Gal}(K(\epsilon)/F(\epsilon))$ is a solvable group. We claim $\text{Gal}(K(\epsilon)/F)$ is solvable. Suppose the claim holds. Consider the extension $F \subseteq K \subseteq K(\epsilon)$. Note that $K(\epsilon)$ is Galois over F, since K is Galois over F. Indeed, whatever set of polynomials K is the splitting field for over F, $K(\epsilon)$ is the splitting field for that set of polynomials together with $x^n - 1$. Since K is Galois over F, by the Galois Correspondence Theorem $\text{Gal}(K/F)$ is a quotient group of $\text{Gal}(K(\epsilon)/F)$. If the latter group is solvable, then $\text{Gal}(K/F)$ is solvable by Proposition 19.3(i). It remains to prove that $\text{Gal}(K(\epsilon)/F)$ is solvable.

We consider the extension $F \subseteq F(\epsilon) \subseteq K(\epsilon)$. By assumption, the subgroup $\text{Gal}(K(\epsilon)/F(\epsilon))$ of $\text{Gal}(K(\epsilon)/F)$ is solvable. Moreover, since $F(\epsilon)$ is Galois over F, this subgroup is a normal subgroup of $\text{Gal}(K(\epsilon)/F)$. On the other hand, the quotient group $\mathrm{Gal}(K(\epsilon)/F)/\mathrm{Gal}(K(\epsilon)/F(\epsilon))$ is isomorphic to $\mathrm{Gal}(F(\epsilon)/F)$, which is abelian by Proposition 27.3. Thus, $\mathrm{Gal}(K(\epsilon)/F)$ is solvable by Proposition 18.1 (ii). □

Here is one half of the Solvability by Radicals Theorem.

Theorem 28.2. *Let F be a field having characteristic zero and assume $f(x) \in F[x]$ is solvable by radicals. Then Gal(K/F) is a solvable group, where K denotes the splitting field of $f(x)$.*

Proof. Let $F \subseteq K \subseteq L$, where L is a radical extension of F. If need be, by Proposition 27.5, we may enlarge L to assume that it is both a radical and Galois extension of F. Let us write $L = F(\alpha_1, \ldots, \alpha_r)$, $\alpha_1^n \in F$ and $\alpha_i^n \in F(\alpha_1, \ldots, \alpha_i)$, for $2 \leq i \leq r$. Let ϵ be a primitive nth root of unity. By Corollary 27.4, if $\epsilon \in F$, then $\mathrm{Gal}(L/F)$ is a solvable group. Since K is a splitting field over F, it is Galois over F, consequently, by the Galois Correspondence Theorem, $\mathrm{Gal}(L/K)$ is a normal subgroup of $\mathrm{Gal}(L/F)$ and the corresponding quotient group is isomorphic to $\mathrm{Gal}(K/F)$. By Proposition 18.1(i), $\mathrm{Gal}(K/F)$ is solvable.

If $\epsilon \notin F$, then $L(\epsilon)$ is a radical and Galois extension of $F(\epsilon)$, with intermediate field $K(\epsilon)$. Therefore, by what we have just shown, $\mathrm{Gal}(K(\epsilon)/F(\epsilon))$ is solvable. That $\mathrm{Gal}(K/F)$ is a solvable group follows from Proposition 28.1. □

Theorem 28.2 can be used to find polynomials that are *not* solvable by radicals. For example, it will follow from the proposition below that for $f(x) = x^5 - 4x + 2 \in \mathbb{Q}[x]$, if K is the splitting field of $f(x)$ over $\mathbb{Q}$, then $\mathrm{Gal}(K/F)$ is isomorphic to S_5, which is not a solvable group. Thus, by the theorem above, $f(x)$ is not solvable by radicals.

Proposition 28.3. *Let $f(x) \in \mathbb{Q}[x]$ be an irreducible polynomial of degree p, with p a prime number and write K for the splitting field of $f(x)$ over $\mathbb{Q}$. If $f(x)$ has exactly two non-real roots, then Gal$(K/\mathbb{Q})$ is isomorphic to S_p.*

Proof. If we let $X = \{\alpha_1, \ldots, a_p\}$ denote the set of distinct roots of $f(x)$, then any element $\sigma \in \mathrm{Gal}(\mathbb{Q}/K)$ permutes the α_i. If we think of S_p as S_X, then we may regard $\mathrm{Gal}(K/\mathbb{Q})$ as a subgroup of S_p. Let α_i be any root of $f(x)$. Then $\mathbb{Q} \subseteq \mathbb{Q}(\alpha_i) \subseteq K$, so p divides $[K : \mathbb{Q}] = |\mathrm{Gal}(K/\mathbb{Q})|$. Thus, by Cauchy's theorem, there exists an element $\tau \in \mathrm{Gal}(K/\mathbb{Q})$ of order p. Since p is prime, τ must be a p-cycle. On the other hand, if $\sigma_0 : \mathbb{C} \to \mathbb{C}$ is complex conjugation, since K is a splitting field, $\sigma_0(K) = K$, by Theorem 10.4, so that σ, the restriction of σ_0 to K, belongs to $\mathrm{Gal}(K/\mathbb{Q})$. By definition of $f(x)$ and K, σ leaves the real roots of $f(x)$ fixed and interchanges the two non-real roots of $f(x)$ Thus, as an element of S_p, $\sigma = (a, b)$ is a transposition. Now, we can write $\tau = (a, j_2, \ldots, j_p)$. It follows that for some $1 \leq i \leq p-1$, $\tau^i = (a, b, j_2', \ldots, j_p')$ and after re-indexing, we may assume $\tau^i = (1, 2, \ldots, p)$ and $\sigma = (1, 2)$. By problem 10 on Homework 3, $\mathrm{Gal}(K/\mathbb{Q}) = S_p$. □

Example 28.4. The polynomial $f(x) = x^5 - 4x + 2 \in \mathbb{Q}[x]$ is not solvable by radicals. To see this, first note that one can see (say using ideas from calculus) that $f(x)$ has exactly three real roots, and therefore, exactly two non-real roots. In order to apply Proposition 28.3, we need to see that $f(x)$ is irreducible over $\mathbb{Q}$. Assuming this, then if K denotes the splitting field of $f(x)$ over $\mathbb{Q}$, $\mathrm{Gal}(K/\mathbb{Q})$ is isomorphic to S_5, and therefore by Corollary 22.3, $\mathrm{Gal}(K/\mathbb{Q})$ is not solvable. Thus, by Theorem 28.2, $f(x)$ is not solvable by radicals. To see that $f(x)$ is irreducible over $\mathbb{Q}$, by Gauss's Lemma, it suffices to show that $f(x)$ is irreducible as an element of $\mathbb{Z}[x]$. Suppose not. Then $f(x) = g(x)h(x)$, where $g(x), h(x) \in \mathbb{Z}[x]$, and (say) $g(x)$ has degree two and $h(x)$ has degree three. Now, without loss of generality, we may assume the constant term of $g(x)$ equals 1 and the constant term of $h(x)$ equals 2. Letting overline denote mod 2, we have, on the one hand,

$$\overline{x^5} \equiv \overline{f(x)} \equiv \overline{g(x)} \cdot \overline{h(x)},$$

but, on the other hand, the constant term of $\overline{g(x)}$ is not zero mod 2, which is a contradiction. Thus, $f(x)$ is irreducible over $\mathbb{Z}$, and thus over $\mathbb{Q}$, which is what we want. □

Lecture 29: Monday November 1. The following theorem leads to the converse to Theorem 28.2 above.

Theorem 29.1. *Let $F \subseteq K$ be a finite Galois extension such that Gal(K/F) is cyclic of order p, a prime. Let $\epsilon \in F$ be a primitive pth root of unity. Then, there exists $\alpha \in K \backslash F$ such that $\alpha^p \in F$. In particular, $K = F(\alpha)$ and K is a radical extension of F.*

Proof. We offer two proofs of this important theorem, the first relying more heavily upon ideas from linear algebra. We first note that since $[K : F] = p$ is prime, there are no intermediate fields between K and F, and hence $K = F(\alpha)$, for any $\alpha \in K\backslash F$. Therefore, we seek $\alpha \in K\backslash F$ with $\alpha^p \in F$. Towards this end, we let σ be a cyclic generator for $\mathrm{Gal}(K/F)$, and note that since σ fixes F, we can regard σ as linear transformation of F-vector space K. Now, since the extension is Galois of degree p, we have $\sigma^p = id$. Thus, σ satisfies the polynomial $x^p - 1$ and therefore the minimal polynomial of σ as a linear transformation over F, divides $x^p - 1$. It follows that the minimal polynomial of σ has distinct roots, which are all in F, since $\epsilon \in F$. These roots must be a power of ϵ. Let $1 \neq \epsilon^i$ be one such root and take $\alpha \in K$ an eigenvector of ϵ^i. Note that $\alpha \notin F$, since σ fixes F. Therefore $\sigma(\alpha) = \epsilon^i\alpha$. Now, $\sigma(\alpha^p) = \sigma(\alpha)^p = (\epsilon^i\alpha)^p = \epsilon^{ip}\alpha^p = \alpha^p$. Thus, α^p is fixed by σ, and hence all powers of σ. Since $\mathrm{Gal}(K/F)$ is generated by σ, α^p is in the fixed field of $\mathrm{Gal}(K/F)$, i.e., $\alpha^p \in F$, which is what we want.

Another standard proof proceeds as follows. As before, let σ be a cyclic generator for $\mathrm{Gal}(K/F)$. We claim that as linear transformations of K over F, $1, \sigma, \ldots, \sigma^{p-1}$ are linearly independent. Suppose not. Then there exist $a_0, a_1, \ldots, a_{p-1} \in F$, not all zero, such that $T := a_0 + a_1\sigma + \cdots + a_{p-1}\sigma^{p-1}$ is the zero transformation. Then this transformation is zero on a basis for K over F. Take $\beta \in K\backslash F$, so that $1, \beta, \ldots, \beta^{p-1}$ is a basis for K over F. We now apply T to each of these basis vectors. Upon doing so, we get the following system of equations:

$$\begin{aligned} a_0 + a_1 + a_2 + \cdots + a_{p-1} &= 0 \\ a_0\beta + a_1\sigma(\beta) + a_2\sigma^2(\beta) + \cdots + a_{p-1}\sigma^{p-1}(\beta) &= 0 \\ a_0\beta^2 + a_1\sigma(\beta^2) + a_2\sigma^2(\beta^2) + \cdots + a_{p-1}\sigma^{p-1}(\beta^2) &= 0 \\ &\vdots \\ a_0\beta^{p-1} + a_1\sigma(\beta^{p-1}) + a_2\sigma_2(\beta^{p-1}) + \cdots + \sigma_{p-1}(\beta^{p-1}) &= 0 \end{aligned}$$

Writing this system of equations as a matrix equation, we get

$$A \cdot \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_{p-1} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

where the (i, j) entry of the coefficient matrix A is $A_{ij} = \sigma^{j-1}(\beta^{i-1})$, for $1 \leq i, j \leq p$. The homogeneous system of equations with cefficient matrix A has a non-trivial solution, and thus $\det(A) = 0$. On the other hand, since $A_{ij} = \sigma^{j-1}(\beta^{i-1}) = (\sigma^{j-1}(\beta))^{i-1}$, A is a Vandermonde matrix whose ith row is

$$\beta^{i-1} \ \sigma(\beta)^{i-1} \ (\sigma^2(\beta))^{i-1} \ \cdots \ (\sigma^{p-1}(\beta))^{i-1},$$

and thus, $\det(A) \neq 0$, a contradiction. It follows that $1, \sigma, \ldots, \sigma^{p-1}$ are linearly independent over F.

Therefore, the linear transformation $T_0 := 1 + \epsilon\sigma + \epsilon^2\sigma^2 + \cdots + \epsilon^{p-1}\sigma^{p-1}$ is not the zero transformation. Therefore, if we choose any basis for K over F and apply T_0 to that basis, then $T_0(v) \neq 0$, for at least one basis element v. Note, however, that since $1 + \epsilon + \cdots + \epsilon^{p-1} = 0$, $T_0(u) = 0$, for all $u \in F$. Thus, there exists $\gamma \in K\backslash F$ such that $\alpha := T_0(\gamma) \neq 0$. We have

$$\alpha = T_0(\gamma) = \gamma + \epsilon\sigma(\gamma) + \cdots + \epsilon^{p-1}\sigma^{p-1}(\gamma).$$

Applying σ to this equations we get

$$\begin{aligned} \sigma(\alpha) &= \sigma(\gamma) + \epsilon\sigma^2(\gamma) + \cdots \epsilon^{p-2}\sigma^{p-1}(\gamma) + \epsilon^{p-1}\gamma \\ &= \epsilon^{p-1} \cdot \{\gamma + \epsilon\sigma(\gamma) + \cdots + \epsilon^{p-1}\sigma^{p-1}(\gamma)\}. \\ &= \epsilon^{p-1}\alpha. \end{aligned}$$

Therefore,

$$\sigma(\alpha^p) = \sigma(\alpha)^p = (\epsilon^{p-1}\alpha)^p = \alpha^p.$$

Clearly, now, $\sigma^i(\alpha^p) = \alpha^p$, for all i which implies that α^p belongs to the fixed field of $\mathrm{Gal}(K/F)$, which is F. That is, $\alpha^p \in F$. Finally, $\alpha \notin F$, because $\sigma(\alpha) = \epsilon^{p-1}\alpha$, so that α is not fixed by σ, and the proof is complete. □

We can now prove the converse to Theorem 28.2.

Theorem 29.2. *Let F be a field of characteristic zero, $f(x) \in F[x]$ and suppose K is the splitting field of $f(x)$ over F. If Gal(K/F) is a solvable group, then $f(x)$ is solvable by radicals.*

Proof. By Proposition 19.3(iii), we may assume that $\mathrm{Gal}(K/F)$ has a solvable series, with factors that are cyclic of prime order. Set $G := \mathrm{Gal}(K/F)$ and suppose that

$$(id) = G_0 \subseteq G_1 \subseteq \cdots \subseteq G_{n-1} \subseteq G_m = G$$

is a solvable series and Gi/G_{i-1} is cyclic of order p_i, where each p_i, for $1 \leq i \leq m$ is prime. Let L_i denote the fixed field of G_i, so that we have a chain of fields extensions

$$F = L_m \subseteq L_{m-1} \subseteq \cdots \subseteq L_1 \subseteq L_0 = K.$$

Note that, by the Galois Correspondence Theorem, $G_i = \mathrm{Gal}(K/L_i)$. For each i, we have $L_i \subseteq L_{i-1} \subseteq K$, with K Galois over L_i. Moreover, L'_{i-1} is a normal subgroup of $L'_i = \mathrm{Gal}(K/L_i)$. Thus, by the Galois Correspondence Theorem, L_{i-1} is a Galois extension of L_i, and $[L_{i-1} : L_i] = [L'_i : L'_{i-1}] = p_i$.

Set $n := p_1 \cdots p_n$, and take ϵ a primitive nth root of unity. If $\epsilon \in F$, then L_i contains a primitive p_ith root of unity, and thus, by Theorem 29.1, there exists $\alpha_{i-1} \in L_{i-1}$ such that $L_{i-1} = L_i(\alpha_{i-1})$ and $\alpha_{i-1}^{p_i} \in L_i$. It follows that K is a radical extension of F, and thus, $f(x)$ is solvable by radicals.

Suppose $\epsilon \notin F$. Consider the extension $F(\epsilon) \subseteq K(\epsilon)$. By Proposition 28.1, $\mathrm{Gal}(K(\epsilon)/F(\epsilon))$ is a solvable group and in fact, we saw that $\mathrm{Gal}(K(\epsilon)/F(\epsilon)$ is isomorphic to a subgroup of $\mathrm{Gal}(K/F)$. Set $H := \mathrm{Gal}(K(\epsilon)/F(\epsilon))$, and let

$$(id) = H_0 \subseteq H_1 \subseteq \cdots \subseteq H_{r-1} \subseteq H_r = H$$

be a solvable series, such that the factor H_i/H_{i-1} is cyclic of order q_i, where q_i is prime. We claim that each $q_j \in \{p_1, \ldots, p_m\}$. Suppose the claim holds. Then $F(\epsilon)$ contains every primitive q_jth root of unity. On the other hand, $K(\epsilon)$ is the splitting field of $f(x)$ over $F(\epsilon)$, so that $K(\epsilon)$ is Galois over $F(\epsilon)$. Thus, by what we have previously shown, $K(\epsilon)$ is a radical extension of $F(\epsilon)$. Since $K(\epsilon)$ is also a radical extension of F, and $F \subseteq K \subseteq K(\epsilon)$, it follows that K is contained in a radical extension of F, and therefore $f(x)$ is solvable by radicals.

For the claim, it suffices to prove that $|G| = p_1 \cdots p_m$. If so, the same reasoning shows $|H| = q_1 \cdots q_r$, and the claim will follow since $|H|$ divides $|G|$. So, we now induct on m. If $m = 1$, $|G| = p_1$, which is what we want. Otherwise, $|G| = |G/G_{m-1}| \cdot |G_{m-1}| = p_m \cdot |G_{m-1}|$. By induction, $|G_{m-1}| = p_1 \cdots p_{m-1}$, so that $|G| = p_1 \cdots p_m \cdot p_m$, and the proof is complete. □

We now turn to the question of which finite groups arise as the Galois group of a Galois extension. The answer is any finite group is the Galois group of a finite Galois extension, if we do not specify the fields in question. In other words, given a finite group G, there exists a finite Galois extension of fields $F \subseteq K$ such that $\mathrm{Gal}(K/F) \cong G$. However, the answer is not so simple if we specify the ground field F. The classical Inverse Galois Problem asks the following. Given a finite group G, is there a finite Galois extension $\mathbb{Q} \subseteq K$, such that $\mathrm{Gal}(K/\mathbb{Q}) \cong G$? While the answer to this question is known for many groups G, the answer is not known in general. There is a positive answer for all groups, if one does not require the extension $\mathbb{Q} \subseteq K$ to be Galois. This follows from a theorem in Graph theory known as *Frucht's theorem.* As time permits, we will show that the Inverse Galois Problem (IGP) has a positive answer for abelian groups and for S_n.

We begin by giving the classical proof that any finite group is a Galois group for some Galois extension $F \subseteq K$.

Theorem 29.3. *Let E be a field a field of characteristic zero, and $x_1, \ldots, x_n$ be independent variables over E. Let $K := E(x_1, \ldots, x_n)$ be the rational function field in n-variables over E and let $s_1, \ldots, s_n$ denote the* *elementary symmetric functions in the x_i, i.e.,*

$$\begin{aligned} s_1 &= x_1 + x_2 + \cdots + x_n \\ s_2 &= \sum_{1 \le i < j \le r} x_i x_j \\ &\vdots \\ s_n &= x_1 x_2 \cdots x_n. \end{aligned}$$

Set $F := E(s_1, \ldots, s_n)$. Then K is a finite Galois extension of F with Gal$(K/F) = S_n$. Moreover, F is the field of symmetric rational functions in the x_i.

Proof. Let $\sigma \in S_n$. Then, σ gives rise to an automorphism of K by permuting the variables x_i. We will denote this automorphism by σ. To elaborate, if $f(x_1, \ldots, x_n)$ is a polynomial in $x_1, \ldots, x_n$ with coefficients in E, then $f^\sigma := f(x_{\sigma(1)}, \ldots, x_{\sigma(n)})$ denotes the polynomial obtained from f by permuting its variables according to σ. For example, if $n = 3$, $\sigma = (1, 2, 3)$ and $f(x_1, x_2, x_3) = x_1^2 x_2 x_3 + x_2^3 + x_1^{17} x_3$, then $f^\sigma = x_2^3 x_3 x_1 + x_3^3 + x_2^{17} x_1$. It is relatively straightforward to check that $(f+g)^\sigma = f^\sigma + g^\sigma$ and $(fg)^\sigma = f^\sigma g^\sigma$. Thus, σ is an automorphism of the polynomial ring $E[x_1, \ldots, x_n]$. If we now define $(\frac{f}{g})^\sigma := \frac{f^\sigma}{g^\sigma}$, then it follows that σ is an automorphism of K. Thus, we may regard S_n as a finite group of automorphisms of K. Let F_0 be the fixed field of S_n. Then, by Corollary 14.4, K is Galois over F_0 and $\text{Gal}(K/F_0) = S_n$. Clearly $F \subseteq F_0$. If we show that $[K : F] = n! = [K : F_0]$, then $F = F_0$, so that K is Galois over F and $\text{Gal}(K/F) = S_n$.

Now, on the one hand, since $F \subseteq F_0$, $[K : F] \ge [K : F_0] = n!$. On the other hand, if we let t be another independent variable, and we write $f(t) = (t - x_1)(t - x_2) \cdots (t - x_n) = t^n - s_1 t^{n-1} + \cdots + (-1)^n s_n \in F[t]$, it follows that K is the splitting field of $f(t)$ over F, which implies $[K : F] \le n!$. Thus, $[K : F] = n!$, as required. Finally, since by definition, F_0 is the field of symmetric rational functions, i.e., the rational functions that remain the same under every permutation of variables, it follows that F is the field of symmetric rational functions in $x_1, \ldots, x_n$. □

Remark 29.4. The last statement in the previous theorem can be summarized by the statement: *Every symmetric rational function is a rational function in the elementary symmetric functions.* For example, $x_1^2 + x_2^2 + x_3^2$ is a symmetric rational function (in fact, a symmetric polynomial) in three variables. Thus, it can be expressed as an element of $\mathbb{Q}(s_1, s_2, s_3)$, say. To see this, simply note that

$$x_1^2 + x_2^2 + x_3^2 = (x_1 + x_2 + x_3)^2 - 2(x_1 x_2 + x_1 x_3 + x_2 x_3) = s_1^2 - 2s_2.$$

Of course, writing other symmetric functions, like $x_1^n + x_2^n + x_3^n$ in terms of s_1, s_2, s_3 is not as easy. A final comment: A symmetric polynomial is a polynomial $f(x_1, \ldots, x_n)$ such that $f^\sigma = f$, for all $\sigma \in S_n$. One can show that *any symmetric polynomial is a polynomial in the elementary symmetric functions.*

Corollary 29.5. *Let G be a finite group. Then there exists a finite Galois extension $L \subseteq K$ such that G is isomorphic to Gal(K/L).*

Proof. Suppose $|G| = n$. Then by Cayley's Theorem, G is isomorphic to a subgroup G_0 of S_n. By the previous theorem, there exists a finite Galois extension $F \subseteq K$ such that $\text{Gal}(K/F) = S_n$. If we let L denote the fixed field of G_0, then by the Galois Correspondence Theorem, K is Galois over L and $\text{Gal}(K/L) = G_0$. Thus, $\text{Gal}(K/L) \cong G$, as required. □

Remark 29.6. It should be clear from the proof of Corollary 29.5 that whenever S_n is the Galois group of a Galois extension $F \subseteq K$, then if G has order n, G is the Galois group of a Galois extension $L \subseteq K$, for some intermediate field L between F and K. If there is time, we will also show that the Inverse Galois Problem has a positive solution for S_n. Thus, for every n, there is a finite Galois extension $\mathbb{Q} \subseteq K$ with $\text{Gal}(K/\mathbb{Q}) \cong S_n$, and hence, if G has order n, then $G \cong \text{Gal}(K/L)$, with L a finite extension of $\mathbb{Q}$, which is close to, but certainly not, a solution to the Inverse Galois Problem for G. □

Lecture 30: Wednesday November 3. Our next goal in Galois theory is to show that any finite abelian group is the Galois group of a finite Galois extension of $\mathbb{Q}$, i.e., the Inverse Galois Problem has a positive solution for finite abelian groups. For this, we will need the recurring fact that a finite abelian group is a product of cyclic groups. Rather than do this directly, which would get us to the aforementioned goal fairly quickly, we will take this opportunity to change directions and set as a short range goal the structure theorem for finitely generated modules over a Principal Ideal Domain, of which the theorem regarding finite abelian groups is a special case. This will give us a chance to start fresh with an entirely new topic, as well as provide an antidote to the depth and abstraction we have recently been dealing with.

In every thing that follows, we will be working with a commutative ring R, with multiplicative identity 1, though our main interest will be commutative rings are are an integral domain. Recall that R is an *integral domain* if whenever $ab = 0$, for $a, b \in R$, then $a = 0$ or $b = 0$. The ring of integers is certainly an integral domain, as is $F[x]$, for F a field. In fact it is not difficult to show that $R[x]$ is an integral domain if and only if R is an integral domain, which then implies that a polynomial ring in any finite number of variables over a field is an integral domain. Integral domains are especially nice, since even though we cannot divide elements in an integral domain, we have the *cancellation property*: Suppose R is an integral domain, and $ab = ac$ for $a, b, c \in R$ with $a \neq 0$. Then, $a(b - c) = 0$, so $b - c = 0$, and thus $b = c$. Thus, we have cancelled a from the equation $ab = ac$.

We now list some basic definitions and properties relevant to our immediate goal.

Definitions and Properties 30.1. Let R be a commutative ring.

(i) If $a, b \in R$ and $0 \neq a$, we say *a divides b* if $b = ac$ for some $c \in R$. In this case we write $a \mid b$.

(ii) An element $u \in R$ is a *unit* if there exists $v \in R$ such that $uv = 1$.

(iii) Two elements $a, b \in R$ are *associates* if $b = ua$, for some unit $u \in R$. Of course, $a = u^{-1}b$. Note that if a, b are non-zero elements in R, with $a \mid b$ and $b \mid a$, then a and b are associates. To see this: write $a = cb$ and $b = da$. Then $a = c(da)$. By cancellation, $cd = 1$, showing that c is a unit, and thus a and b are associates.

(iv) A non-zero, non-unit $p \in R$ is a *prime* element, if whenever p divides ab, then p divides a or p divides b. Of course prime numbers are prime elements in $\mathbb{Z}$ and irreducible polynomials in $F[x]$, F a field, are prime elements in $F[x]$.

(v) A non-zero non-unit $q \in R$ is an *irreducible* element if whenever we can write $q = ab$, with $a, b \in R$, then a or b is a unit. For rings like $\mathbb{Z}$ and $F[x]$, there is no difference between irreducible elements and prime elements. This has to do with the fact that both rings are *Unique Factorizations Domains.*

(vi) If R is an integral domain and $p \in R$ is a prime element, then p is irreducible, but the converse need not hold. To see this, suppose $p = ab$. Then $p \mid ab$, so $p \mid a$ or $p \mid b$. Suppose $p \mid a$. Then $a = pc$. Thus, $p = pcb$, so by cancellation, $cb = 1$, and thus, b is a unit, showing p is irreducible.

Consider the ring $\mathbb{Z}[\sqrt{-5}] := \{a + b\sqrt{-5} \mid a, b \in \mathbb{Z}\}$. Then it is easy to see that
$$3 \cdot 3 = 9 = (2 + \sqrt{-5}) \cdot (2 - \sqrt{-5}).$$
In Homework 5 you will show that $3, 2 + \sqrt{-5}, 2 - \sqrt{-5}$ are irreducible elements of R and that $3 \nmid 2 \pm \sqrt{-5}$, showing that 3 is not prime in R.

(vii) An *ideal* of R is a subset $I \subseteq R$ that is closed under addition, and further has the property that $ra \in I$, for all $r \in R$ and $a \in I$.

(viii) Let X be a subset of R. Then we write $\langle X \rangle$ for the ideal of R generated by X. One can define $\langle X \rangle$ to be the intersection of all ideals in R containing X. Then, it is not difficult to show that $\langle X \rangle$ is the set of all finite linear combinations of the form $r_1x_1 + \cdots + r_nx_n$ with $r_i \in R$ and $x_j \in X$. When $X = \{a\}$ is generated by a single element, then we call $\langle a \rangle$ a *principal ideal* and often just write aR for this ideal.

(ix) If $I \subseteq R$ is an ideal, then I is a subgroup of the abelian group $(R, +)$. Thus, we can form the factor group R/I. We can define multiplication in the obvious way: $(r_1 + I) \cdot (r_2 + I) := r_1r_2 + I$. It is straightforward to check that this multiplication is well defined and that R/I is a commutative ring.

(x) A ideal $P \subseteq R$ is a *prime ideal* if whenever $ab \in P$, then $a \in P$ or $b \in P$. An ideal $Q \subseteq R$ is a *maximal* ideal if whenever $J \subseteq R$ is an ideal containing Q, then $J = R$ or $J = Q$. It is easy to check that $P \subseteq R$ is a prime ideal if and only if R/P is an integral domain. One also has that $Q \subseteq R$

is a maximal ideal if and only R/Q is a field. To see this, suppose Q is a maximal ideal and take $0 \neq \overline{r} \in R/Q$. Then $r \notin Q$, so that $\langle r, Q\rangle$ is an ideal of R properly containing Q. Thus, $\langle r, Q\rangle = R$. Therefore, we can write $1 = ar + q$, for some $a \in r$ and $q \in Q$. Thus, in R/Q, we have $1 \equiv \overline{a} \cdot \overline{r}$, showing that $\overline{r}$ has a multiplicative inverse. Thus, R/Q is a field. The converse is similar.

□

We now come to an important definition.

Definition 30.2. An integral domain R is said to be a *Principal Ideal Domain* or *PID* if every ideal of R is principal. Using the division algorithm, it is easy to check that $\mathbb{Z}$ and $F[x]$, with F a field, are PIDs.

Remark 30.3 (i) Not every PID has a division algorithm. It can be shown that the ring

$$R := \mathbb{Z}[\frac{1+\sqrt{-19}}{2}] = \{a + b \cdot \frac{1+\sqrt{-19}}{2} \mid a, b \in \mathbb{Z}\}$$

is a PID, but R cannot be endowed with any sort of division algorithm. If you look up a proof of this, you will get a small taste of what goes on in the field of Algebraic Number Theory.

(ii) A polynomial ring in more than one variable is not a PID, though it does have the unique factorization property. In particular, the polynomial ring $\mathbb{Q}[x, y]$ is not a PID. Can you see why? □

Lecture 31: Friday November 5. We begin with the following important property enjoyed by PIDs.

Proposition 31.1. *Assume R is a PID. Then:*

(i) R satisfies the ascending chain condition on ideals, i.e., if $I_1 \subseteq I_2 \subseteq \cdots$ is a chain of ideals in R then there exists $n \geq 1$ such that $I_n = I_{n+1} = \cdots$.

(ii) Every non-empty collection of ideals has a maximal element.

Proof. For (i), set $J := \bigcup_{s\geq 1} I_s$. Then, as we have seen in the second application of Zorn's Lemma, J is an ideal of R. Thus, $J = aR$, for some $a \in R$. But $a \in I_n$, for some $n \geq 1$. Thus,

$$J = aR \subseteq I_n \subseteq J,$$

so $J = I_n$. It follows immediately from this that $I_n = I_{n+1} = \cdots$.

For (ii), let $\mathcal{C}$ be a non-empty collection of ideals in R. Suppose $\mathcal{C}$ does not have a maximal element. Take a_1R an ideal in $\mathcal{C}$. Since a_1R is not maximal, there exists a_2R in $\mathcal{C}$ such that $a_1R \subsetneq a_2R$. Proceeding inductively, we may construct an ascending chain of ideals from $\mathcal{C}$ $a_1R \subsetneq a_2R \subsetneq \cdots$, which contradicts (i). Thus, $\mathcal{C}$ must have a maximal element. □

In a first semester abstract algebra class you may have seen that every positive integer can be written uniquely as a product of primes or that every non-zero, non-constant polynomial can be written uniquely (up to order, and constant multiples) as a product of irreducible polynomials. Since both types of rings are PIDs, it is not too surprising that arbitrary PIDs have the unique factorization property. However, $\mathbb{Z}$ and $F[x]$ also have a division algorithm, which makes the proofs of the unique factorization property easier. An arbitrary PID need not have a division algorithm, so the proof has to rely on purely ring theoretic properties of a PID.

Theorem 31.2. *Let R be a PID. Then every non-zero, non-unit can be written uniquely as a product of irreducible elements. In particular, if $a \in R$ is a non-zero, non-unit and $a = p_1 \cdots p_s = q_1 \cdots q_t$, where each p_i and q_j is irreducible, then $s = t$ and after re-indexing, each q_i is an associate of p_i.*

Proof. The proof will proceed in three steps.

Step 1. Every non-zero, non-unit in R is a product of irreducible elements. Suppose this were false. Let $\mathcal{C}$ denote the set of principal ideals of R that are generated by non-zero, non-units that cannot be factored as a product of irreducible elements. Then $\mathcal{C}$ is non-empty, and by the preceding proposition there exists $qR \in \mathcal{C}$ a maximal element. By definition of $\mathcal{C}$, q is not an irreducible element. Thus, we may write $q = ab$, where neither a nor b is a unit. Thus, aR and bR properly contain qR. By the maximality of aR, a and b each can be written as a product of irreducible elements, and therefore q is a product of irreducible elements, which is a contradiction. Thus, $\mathcal{C}$ must be empty, and every non-zero non-unit in R is a product of irreducible elements.

Step 2. If $q \in R$ is irreducible, then q is a prime element. To see this, suppose $q \mid ab$, and $q \nmid a$, for $a, b \in R$. Then $a \notin qR$. Thus, $\langle a, q\rangle$ properly contains qR. Since R is a PID, $\langle a, q\rangle = cR$, for some $c \in R$. Since q is irreducible and $qR \subsetneq cR$, we must have $cR = R$. Thus, $1 \in \langle a, q\rangle$ so we can write $1 = r_1a + r_2q$. Multiplying this equation by b we have $b = r_1ab + r_2bq$. Since $q \mid ab$, we have $q \mid b$, which is what we want.

Step 3. Suppose $a \in R$ is a non-zero non-unit, and

$$a = p_1p_2\cdots p_s = q_1q_2\cdots q_t, \quad (*)$$

with each p_i, q_j irreducible in R. We prove the required uniqueness statement by induction on s. If $s = 1$, we have $p_1 = q_1\cdots q_t$, which can only hold if $t = 1$, since p_1 is irreducible. Supppose now that $s > 1$. Then by Step 2, p_1 is a prime element. Since p_1 divides $q_1\cdots q_t$, we have $p_1 \mid q_j$ for some j. After re-indexing, we may assume $j = 1$. Thus we may write $q_1 = p_1u$, for some $u \in R$. Since q_1 is irreducible and p_1 is not a unit, u must be a unit, and thus, p_1 and q_1 are associates. Cancelling p_1 from both sides of (*), we have

$$p_2p_3\cdots p_2 = (uq_2)q_3\cdots q_t.$$

It is easy to check that uq_2 is an irreducible element. Thus, by induction on s, $s = t$, and after re-indexing (if necessary), p_i and q_i are associates, for all $2 \leq i \leq s$, which gives what we want. □

We now want to define the notion of a *module* over a commutative ring. The easiest way to think of a module is a follows: A module is to a ring, as a vector space is to a field, i.e., an abelian group endowed with a scalar multiplication.

Definition 31.3 Let R be a commutative ring. A (left) R-module is an abelian group $(M, +)$ together with a scalar multiplication $* : R \times M \to M$ satisfying the following properties:

(i) $(a + b) * m = a * m + b * m$, for al $a, b \in R$ and $m \in M$.
(ii) $a * (m_1 + m_2) = a * m_1 + a * m_2$, for all $a \in R$ and $m_1, m_2 \in M$.
(iii) $(ab) * m = a * (b * m)$, for al $a, b \in R$ and $m \in M$.
(iv) $1 * m = m$, for all $m \in M$.

Henceforth, we write am instead of $a * m$. Notice that we use juxtaposition for both the product in R and the scalar product of elements of R on elements of M and $+$ for addition in R and addition in M. However, context should make it clear which operations are at work in any given expression. The following properties of an R module M are easily checked:

(i) $0m = 0$, for all $m \in M$, where the 0 on the left is 0 in R and the 0 on the right is 0 in M.
(ii) $(-a)m = -(am)$, for all $a \in R$ and $m \in M$.

Remark 31.4. Examples of R-modules include ideals of R, factor rings R/I, and direct sums $R \oplus \cdots \oplus R$. More to the point: Let G be an abelian group, written additively, so that for n a positive integer and $g \in G$, ng means $g + \cdots + g$, n times and $(-n)g$ means $-(ng)$. Then G is a $\mathbb{Z}$-module.

Lecture 32: Monday November 8. We continue with our discussion of modules over a commutative ring.

Definitions 32.1. Let R be a commutative ring and M an R-module.

(i) A subset $N \subseteq M$ is a *submodule* of M if it is closed under addition and scalar multiplication. This is easily seen to be equivalent to saying that N is an R-module in its own right under the existing operations on M.
(ii) If $N \subseteq M$ is a submodule, then the abelian group M/N inherits an R-module structure with the obvious scalar multiplication: $r * (m + N) = r * m + N$.
(iii) If $X \subseteq M$ is a subset of R, then $\langle X\rangle$, the submodule generated by X, is the intersection of all submodules of M containing X. As one might readily suspect, $\langle X\rangle$ is the set of all finite linear combinations $r_1x_1 + \cdots + r_nx_n$, with $r_j \in R$ and $x_j \in X$ (for any $n \geq 1$).
(iv) M is *finitely generated* if $M = \langle X\rangle$, for a finite subset $X \subseteq M$. If M can be generated by a one element set, we say that M is a *cyclic* R-module. In this case we often write $M = Rm$, for m the cyclic generator of M.
(v) A finitely generated R-module F is said to be a *free* R-module, if it has a basis, i.e., there exist $x_1, \ldots, x_n$ in M such that $F = \langle x_1, \ldots, x_n\rangle$ and $x_1, \ldots, x_n$ are linearly independent, i.e., given a relation $r_1x_1 + \cdots + r_nx_n = 0$, then $r_i = 0$, for all i. Just like for vector spaces, this is equivalent

to saying that every element in F can be written *uniquely* as a linear combination of $x_1, \ldots, x_n$. It turns out that any two bases for M have the same number of elements. We will say that M has *rank* n if it has a basis with n elements.

(vi) Let A, B be R-modules. A map $\phi : A \to B$ satisfying $\phi(a_1 + a_2) = \phi(a_1) + \phi(a_2)$ and $\phi(ra) = r\phi(a)$, for all $r \in R$ and $a, a_1, a_2 \in A$ is called an *R-module homomorphism.* Just like for vector spaces, the kernel of ϕ, which is $\{a \in A \mid \phi(a) = 0\}$ and the image of ϕ, which is $\phi(A)$, are submodules of A and B, respectively. We say ϕ is an isomorphism if ϕ is both 1-1 and onto, or equivalently, the kernel of ϕ is zereo and $\phi(A) = B$.

Proposition 32.2. *Let R be a commutative ring and F a finitely generated free R-module. Then any two bases for F have the same number of elements.*

Proof. To begin, let R^n denote the set of n-tuples of elements of R. Then R^n is a free R-module with standard basis $\{e_1, \ldots, e_n\}$, where e_i is the n-tuple with 1 in the ith entry, and zeros elsewhere. Suppose F is a free R-module with basis $\{x_1, \ldots, x_n\}$. Then, just like with vector spaces, the map $\phi : R^n \to F$ taking $(r_1, \ldots, r_n)$ to $r_1x_1 + \cdots + r_nx_n$ is easily seen to be an isomorphism of R-modules. If now, $\{y_1, \ldots, y_t\}$ is another basis of F, we have $F \cong R^t$. Thus, $R^n \cong R^t$, which cannot happen when R is a commutative ring, unless $n = t$. To see this, let $Q \subseteq R$ be a maximal ideal, so that $K := R/Q$ is a field. Let $Q^n \subseteq R^n$ denote the n-tuples, all of whose entries are in Q, and define Q^t likewise. Then, on the one hand, $R^n/Q^n \cong K^n$ and $R^t/Q^t \cong K^t$, while on the other hand, an isomorphism between R^n and R^t induces and isomorphism between R^n/Q^n and R^t/Q^t (since any such isomorphism maps Q^n isomorphically to Q^t). Thus, the vector spaces F^n and F^t are isomorphic. Since any two bases for a vector space over F have the same number of elements, $n = t$. Thus, any two bases for a free module over R also have the same number of elements. □

In general, unlike vector spaces over a field, a typical module over a ring that is not a field will not have a basis. And even if M is a free R-module, its submodules need not be free. However, we will see that if R is a PID, then a submodule of a finitely generated free module is, in fact, free.

Definition 32.3. Let R be a commutative ring and M and R-module. Suppose that $N_1, \ldots, N_s$ are submodules of M. The *sum* of the modules N_i is the set of all expressions $n_1 + \cdots + n_s$ such that each $n_i \in N_i$. This is a submodule of M and is denoted by $\Sigma_i N_i$ or just $N_1 + \cdots + N_s$. We say that the sum is *direct* if for all $1 \le j \le s$, $N_j \cap (\Sigma_{i \neq j} N_i) = 0$, and write $N_1 \oplus \cdots \oplus N_s$ for $\Sigma_i N_i$. If $M = N_1 \oplus \cdots \oplus N_s$, we say that M *is the direct sum of* $N_1, \ldots, N_s$. It is not difficult to show that M is the direct sum of $N_1, \ldots, N_s$ if and only if every element $m \in M$ can be written as $m = n_1 + \cdots + n_s$ for *unique* elements $n_i \in N_i$ (see Homework 5).

We will need the following proposition concerning free modules.

Proposition 32.4. *Suppose R is a commutative ring and $\phi : M \to F$ is a surjective R-module homomorphism such that F is a finitely generated free R-module. Let K denote the kernel of ϕ. Then there exists an R-module homomorphism $j : F \to M$ satisfying:*

(i) $\phi \circ j = Id_F$, the identity on F.
(ii) $F \cong j(F)$.
(iii) $M = K \oplus j(F)$.

In particular, there is a free R module that is a direct summand of M.

Proof. Just like with vector spaces, a homomorphism whose domain is a free module is determined by its values on a basis. Let $x_1, \ldots, x_n \in F$ be a basis for F. Since ϕ is surjective, for each $1 \le i \le n$, there exists $m_i \in M$ such that that $\phi(m_i) = x_i$. We define $j(x_i) = m_i$ and extend this definition to all of F, i.e., if $f \in F$, we can write f uniquely as $f = r_1x_1 + \cdots + r_nx_n$, with $r_i \in R$. We define $j(f) := r_1m_1 + \cdots + r_nm_n$. Note that this is well defined, since the r_i are unique. It is straight forward to check that j is an R-module homomorphism, which by definition, satisfies $\phi \circ j = id_F$. Moreover, the map $j : F \to j(C)$ is certainly surjective. Suppose $j(c) = 0$, for $c \in F$. Then $0 = \phi(j(c) = c$, showing that j is also injective, and thus j is an isomorphism from F to $j(F)$. In particular, $j(F)$ is a free R-submodule of M.

Suppose $m \in M$. Then

$$\phi(m - j\phi(m)) = \phi(m) - \phi(j(\phi(m))) = \phi(m) - (\phi \circ j)(\phi(m)) = \phi(m) - \phi(m) = 0.$$

Thus, $m - j(\phi(m)) \in K$, which shows $m \in K + j(F)$. Therefore, $M = K + j(F)$. Suppose $a \in K \cap j(F)$. Then $a = j(c$, for some $c \in F$. Thus, $0 = \phi(k) = \phi(j(c)) = c$, so $c = 0$. Thus $0 = j(c) = a$, showing $K \cap j(F) = 0$. Therefore, $M = K \oplus j(F)$. □

Lecture 33: Wednesday November 10. We can now state one form of the main theorem we are interested in.

Structure Theorem for finitely generated modules over a PID. *Let R be a PID and M a finitely generated R-module. Then M is a direct sum of cyclic modules.*

The next definition plays a key role in the structure theorem.

Definitions 33.1. Let R be an integral domain, M an R-module.

(i) Set $T(M) := \{m \in M \mid rm = 0, \text{ for some non-zero } r \in R\}$. It is easy to check that $T(M)$ is a submodule of M. $T(M)$ is called the *torsion submodule* of M.

(ii) M is *torsion-free* if $T(M) = 0$. M is a *torsion module* if $M = T(M)$.

If R is an integral domain, then a finitely generated free R-module F is torsion free. Indeed, if the rank of F is n, then $F \cong R^n$, which is clearly torsion-free. The following crucial ingredient to the structure theorem we seek shows that, over a PID, torsion-free modules are free. It follows that any $M \subseteq F$ is also torsion-free. The next theorem shows that we the converse to this holds over a PID.

Theorem 33.2. *Let R be a* PID and M a finitely generated R-module.

(i) If F is a finitely generated free R-module of rank r and $N \subseteq F$ is a submodule, then N is a finitely generated free R-module of rank less than or equal to r.

(ii) If M is torsion-free, then M is a free R-module.

(iii) If M can be generated by n elements, then any submodule of M can be generated by n elements or less. In particular, a submodule of a cyclic module is cyclic.

(iv) There exists a submodule $F \subseteq M$ such that F is a free R-module and $M = F \oplus T(M)$.

Proof. For (i), we induct on r. If $r = 1$, then $F = Rx$ for some $x \in R$. If $N = 0$, there is nothing to prove. Otherwise, since $N \subseteq Rx$, we consider the ideal $J := \{t \in R \mid tx \in N\}$. We then have $J = aR$, for some $a \in R$. We claim $N = R(ax)$, which is a free R-module, since F is torsion-free. Clearly $R(ax) \subseteq N$, by definition of a. On the other hand, if $n \in N$, then $n = rx$, for some $r \in J$, so $r = r'a$, for some $r' \in R$. Thus, $n = tx = (r'a)x = r'(ax)$, which shows that $N \subseteq R(ax)$, which gives what we want.

Suppose $r > 1$. Set $G := \langle x_1, \ldots, x_{r-1}\rangle$, a free module of rank $r - 1$. By induction on r, $N \cap G$ is either (0) or a free R-module of rank $r - 1$ or less. Now every element n in N can be written as $n = a_1x_1 + \cdots + a_{r-1}x_{r-1} + a_rx_r$, with each $a_j \in R$. If, for every element in N, $a_r = 0$, then $N \subseteq G$, and we are done by induction on r. Otherwise, let J denote the ideal of R generated by all of the coefficients of x_r as n varies over the elements of N. J is clearly an ideal of R. Thus, $J = aR$, for sone $a \in R$. By definition of J, there exists $n_0 \in N$ of the form $n_0 = s_1x_1 + \cdots + s_{r-1}x_{r-1} + ax_r$.

We claim $N = (N \cap G) \oplus Rn_0$. If so, then, on the one hand, Rn_0 is a free R-module of rank one. On the other hand, $M \cap G$ has a basis consisting of $r - 1$ or fewer elements. Putting these bases together gives a basis for M having no more than n elements (see Homework 5), which is what we want. For the claim, take $n = c_1x_1 + \cdots c_{r-1}x_{r-1} + c_rx_rx_n$ in N. If $c_r = 0$, then $n \in N \cap G$. Otherwise, by definition of J, $c_r = da$, for some $d \in R$. it follows that $n - dn_0 \in N \cap G$. Therefore, $n \in (N \cap G) + Rn_0$, and therefore $N = (N \cap G) + Rn_0$. On the other hand, suppose $t_1x_1 + \cdots t_{r-1}x_{r-1} = sn_0$ belongs to $(N \cap G) \cap Rn_0$. Then

$$t_1x_1 + \cdots + t_{r-1}x_{r-1} = ss_1x_1 + \cdots + ss_{r-1}x_{r-1} + sax_r.$$

Since the x_j are linearly independent, this gives $sax_r = 0$, forcing $s = 0$. Thus, $sn_0 = 0$, showing $(N \cap G) \cap Rn_0 = 0$, and thus, $N = (N \cap G) \oplus Rn_0$, as required.

For part (ii) Suppose $M = \langle x_1, \ldots, x_n\rangle$ and k is the largest integer such that k elements in the set $\{x_1, \ldots, x_n\}$ are linearly independent over R. Without loss of generality, we may assume these elements are $x_1, \ldots, x_k$. Set $F := \langle x_1, \ldots, x_k\rangle$, a free R-module. Then for each x_j with $k < j \leq n$, there exists $0 \neq a_j \in R$ such that $a_jx_j \in F$. Set $a := a_{k+1} \cdots a_n$. It follows that $aM \subseteq F$. By part (i), aM is a free R-module. However, since M is torsion-free, then map $\phi : M \to aM$ given by $\phi(m) = am$ is an isomorphism of R-modules. Thus, M is a free R-module.

For part (iii), Suppose $M = \langle x_1, \ldots, x_n \rangle$. Then we have a surjective homomorphism $\phi : R^n \to M$, which takes $(r_1, \ldots, r_n)$ to $r_1x_1 + \cdots + r_nx_n$. Let N be a submodule of M. Then $\phi^{-1}(N)$ is a submodule of R^n. By part (i), $\phi^{-1}(N)$ is a finitely generated free R-module, whose rank is less than or equal to n. Thus, N is a homomorphic image of R-module generated by n or fewer elements, and hence N is generated by n or fewer elements.

For part (iv) let's first observe that $M/T(M)$ is torsion-free. Indeed, suppose $r \cdot \overline{m} = \overline{0}$, for $\overline{m} \in M/T(M)$, and $0 \neq r \in R$. Then $rm \in T(M)$, so there exists $0 \neq r' \in R$ such that $r'(rm) = 0$. Thus, $(r'r)m = 0$. Since $r'r \neq 0$, $m \in T(m)$, so that $\overline{m} = \overline{0}$ in $M/T(M)$. Thus, $M/T(M)$ is torsion-free, and finitely generated. By part (ii), $M/T(M)$ is a finitely generated free R-module. Thus, the canonical map $M \to M/T(M)$ is a surjective map onto a finitely generated. free R-module. Since the kernel of this map is $T(M)$, by Proposition 31.7, there exists a free R-submodule $F \subseteq M$ such that $M = F \oplus T(M)$. □

Lecture 34: Friday November 12. Regarding part (iv) of Theorem 33.2, $F \subseteq M$ may not be unique, as it depends upon the map $j : M/T(M) \to M$, but its rank is unique. To see this just note that if

$$F' \oplus T(M) = M = F \oplus T(M),$$

with F' a free R-module, then modding out $T(M)$, we have $F' \cong M/T(M) \cong F$, so that F' and F are isomorphic free modules and therefore have the same rank (by Proposition 32.2). We can then define the *rank of* M to be the rank of F. Note also, that when we write, $M = F \oplus T(M)$, if $\{x_1, \ldots, x_r\}$ is a basis for F, then $F = Rx_1 \oplus \cdots \oplus Rx_r$ is a direct sum of cyclic R-modules (see Homework 5). Thus, to finish the structure theorem for finitely generated modules over a PID, we only have to show that a finitely generated torsion module is a direct sum of cyclic modules.

We begin with a definition.

Definition 34.1. If R is a commutative ring, and M an R-module. For $x \in M$ we define the *annihilator* of x to be the set $\text{ann}(x) := \{r \in R \mid rx = 0\}$. We set $\text{ann}(M) := \{r \in r \mid rx = 0, \text{ for all } x \in M\}$, the *annihilator of* M. Both annihilators are ideals of R. Note that M is torsion-free if and only if $\text{ann}(x) = 0$, for all x and M is a torsion module if and only if $\text{ann}(x) \neq 0$, for all $x \in M$.

Proposition 34.2. *Let R be a* PID *and M a finitely generated, torsion R-module.*

(i) $\text{ann}(M) = aR \neq 0$, for some $a \in R$. In particular, a annihilates M and divides any element of R annihilating M.

(ii) If $\text{ann}(M) = aR$ and p is a prime dividing a, then $M(p) := \{x \in M \mid p^jx = 0 \text{ for some } j \geq 1\}$ is a non-zero submodule with $\text{ann}(M(p)) = p^eR$, for some $e \geq 1$.

(iii) Suppose $x, z \in M$ satisfy $\text{ann}(x) = p^eR = \text{ann}(z)$, for $p \in R$ a prime and $e \geq 1$. If $\langle x \rangle \subseteq \langle z \rangle$, then $\langle x \rangle = \langle z \rangle$.

(iv) Suppose $M = \langle x \rangle$ is a cyclic module with $\text{ann}(M) = p^eR$, with $p \in R$ prime and $e \geq 1$. Then for any non-zero submodule $N \subseteq M$, we have $N = \langle p^dx \rangle$ for some $0 \leq d < e$.

(v) Suppose M is annihilated by $p \in R$, p a prime element. Then M is a direct sum of cyclic submodules.

Proof. For (i), suppose M is generated over R by $x_1, \ldots, x_n$. For each i there exists $0 \neq a_i \in R$ such $a_ix_i = 0$. If we set $a_0 := a_1 \cdots a_n$, then we have that $a_0x_i = 0$ for all i and hence $a_0x = 0$, for all $x \in M$. Note that a_0 is a no-zero element in $\text{ann}(M)$. Since R is a PID, $\text{ann}(M)$ is principal, so we write $\text{ann}(M) = aR$, for $0 \neq a \in R$. Moreover, if $bx = 0$, for all $x \in M$, then $a \mid b$ (since $b \in aR$), which gives what we want.

(ii) We first note that $M(p) \neq 0$. Write $a = pa'$. Note that $a' \notin Ra = \text{ann}(M)$, since a' is not a multiple of a, so there exists $0 \neq z \in M$ such that $a'z \neq 0$. Thus, $a'z \in M(p)$. Suppose $x, y \in M(p)$ and $p^jx = 0 = p^ky$, then $p^{j+k}(x + y) = 0$. Moreover $p^j(rx) = 0$, for all $r \in R$. Therefore $M(p)$ is a submodule of M. By Proposition 33.2, $M(p)$ is a finitely generated module. If $v_1, \ldots, v_c$ generate $M(p)$, let $e \geq 1$ be the least exponent such that $p^ev_j = 0$, for all j. Then, clearly, $p^e \in \text{ann}(M(p))$. Suppose $\text{ann}(M(p)) = cR$, for $c \in R$. Then, by Part (i), $c \mid p^e$. By the unique factorization property, $c = p^j$, for some $j \geq 1$. By definition, $j \geq e$, and thus $c \in Rp^e$. Therefore, $\text{ann}(M(p)) = Rp^e$.

For (iii), we have $x = rz$, for some $r \in R$. If $p \mid r$, then $p^{e-1}x = 0$, a contradiction (since p^{e-1} is not a multiple of p^e). Thus, $p \nmid r$, so we can write $1 = ur + vp^e$, with $u, v \in R$. We then have

$$ux = urz = z - vp^ez = z,$$

showing that $z \in \langle x \rangle$, which implies $\langle z \rangle = \langle x \rangle$.

For (iv), by Theorem 33.2 (iii), N is a cyclic module, say $N = \langle n \rangle$. Since $N \subseteq \langle x \rangle$, we have $n = rx$, for some $r \in R$. Since R has the unique factorization property, we may write $r = r_0 p^c$, for $r_0 \in R$ not divisible by p. Thus, $n = r_0 p^c x$. Since $n \neq 0$, we must have $0 \leq c < p$. On the one hand, we have $\langle n \rangle \subseteq \langle p^c x \rangle$. On the other hand, by Problem 4 on Homework 5, we may write $1 = ur_0 + vp^{e-c}$, since p does not divide r_0. Multiplying this equation by $p^c x$, and using the fact that $p^e x = 0$, we have, $p^c x = ur_0 p^c x$. Thus, $p^c x = un$, showing that $p^c x \in \langle n \rangle$, and thus $\langle p^c x \rangle \subseteq \langle n \rangle$, which gives $\langle p^c x \rangle = \langle n \rangle = N$, which is what we want.

For (v), let $n \geq 1$ be such that M can be generated by n elements and M cannot be generated by fewer than n elements. Suppose $M = \langle x_1, \ldots, x_n \rangle$. We show by induction on n that $M = \langle x_1 \rangle \oplus \cdots \oplus \langle x_n \rangle$. If $n = 1$, there is nothing to prove. Suppose $n > 1$. Set $M' := \langle x_1, \ldots, x_{n-1} \rangle$. Clearly M' cannot be generated by fewer than $n-1$ elements, otherwise these elements together with x_n would generated M, contradicting the choice of n. Thus, $M' = \langle x_1 \rangle \oplus \cdots \oplus \langle x_{n-1} \rangle$. It suffices to show $M = M' \oplus \langle x_n \rangle$. Clearly $M = M' + \langle x_n \rangle$. Suppose $z \in M' \cap \langle x_n \rangle$. If $z \neq 0$, $\langle z \rangle \subseteq \langle x_n \rangle$, so by part (iii), $\langle z \rangle = \langle x_n \rangle$, so $x_n \in \langle z \rangle \subseteq M'$, which is a contradiction - since this would imply $M = M'$. Therefore, $z = 0$ and thus, $M' \cap \langle x_n \rangle = 0$, which gives $M = M' \oplus \langle x_n \rangle$, as required. □

Remark 34.3. Let M be an R-module over the PID R. Suppose $x \in T(M)$ and let $\mathrm{ann}(x) = aR$. Then we have a surjective R-module map $\phi : R \to \langle x \rangle$ whose kernel is aR. Thus, $\langle x \rangle \cong R/aR.$, which looks exactly like the isomorphism we obtain when we have a cyclic group. When $a = p$ is prime, then $\langle x \rangle$ is the module analogue of a cyclic groups of order p: p kills its generator, and the group has no proper subgroups.

Lecture 35: Monday November 15. In Theorem 34.2 (v) we showed that if M is a finitely generated module over a PID having the property that M is annihilated by a prime element, then M is a direct sum of cyclic modules. Our next theorem extends this to modules annihilated by a power of a prime element, and is the most difficult part of the structure theorem for modules over a PID. However, we first need the following remark.

Remark 35.1. Suppose M is a finitely generated module over the PID R. By Proposition 33.2 (iii), every submodule of M is finitely generated. In this case, by Homework 5, problem 8, M also satisfies both the ascending chain condition on submodules, and the maximal condition on submodules - the latter meaning that every non-empty collection of submodules of M has a maximal element. A module satisfying these properties is said to be *Noetherian.* Our use of the maximal condition here takes the place of an induction argument that would be applied in the case of finite abelian groups or finite dimensional vector spaces over a field. An argument making use of the maximal condition is sometimes called *Noetherian induction.*

Theorem 35.2. *Let R be a PID and M a finitely generated R-module with $ann(M) = p^e R$, with $p \in R$ prime and $e \geq 1$. Then M is a direct sum of cyclic modules. In fact, there exist $x_1, \ldots, x_n \in R$ and $e = e_1 \geq \cdots \geq e_n$ such that $M = \langle x_1 \rangle \oplus \cdots \oplus \langle x_n \rangle$, with $ann(x_i) = p^{e_i} R$, for all i.*

Proof. We begin by noting that there exists $0 \neq x \in M$ such that $p^{e-1} x \neq 0$. Otherwise, p^{e-1} annihilates every x in M, and thus is divisible by p^e, which cannot happen. So we start with $0 \neq x$ such that $p^{e-1} x \neq 0$. If $M = \langle x \rangle$, we are done. If $M \neq \langle x \rangle$, we set $x_1 := x$ and claim there is a submodule M_1 such that $M = \langle x_1 \rangle \oplus M_1$. Suppose we could always find such an M_1 whenever a cyclic submodule has the same annihilator as the module. Then, taking $x_2 \in M_1$ so that $\mathrm{ann}(x_2) = \mathrm{ann}(M_1)$, either $M_1 = \langle x_2 \rangle$, and thus, $M = \langle x_1 \rangle \oplus \langle x_2 \rangle$ or there exists $M_2 \subseteq M_1$ such that $M_1 = \langle x_2 \rangle \oplus M_2$, so that $M = \langle x_1 \rangle \oplus \langle x_2 \rangle \oplus M_2$. If we apply the construction inductively, then we have a chain of submodules,

$$\langle x_1 \rangle \subseteq \langle x_1 \rangle \oplus \langle x_2 \rangle \subseteq \langle x_1 \rangle \oplus \langle x_2 \rangle \oplus \langle x_3 \rangle \subseteq \cdots .$$

Since M satisfies the ascending chain condition, this process must stop when M is a direct sum of cyclic submodules. For the statement about annihilators, first note that since $p^e R$ is in the annihilator of every element and submodule of M, the annihilator of every element and submodule divides p^e and is thus generated by a power of p. Moreover, since $M_{i+1} \subseteq M_i$, $\mathrm{ann}(M_i) \subseteq \mathrm{ann}(M_{i+1})$, and therefore, if $\mathrm{ann}(M_i) = p^{e_i} R$ and $\mathrm{ann}(M_{i+1}) = p^{e_{i+1}} R$, $e_i \geq e_{i+1}$. Since $\mathrm{ann}(x_i) = \mathrm{ann}(M_i)$, the statement concerning annihilators follows.

Thus, we must prove the following statement: If M is a finitely generated R-module with $\mathrm{ann}(M) = p^e R$ and $x \in M$ satisfies $\mathrm{ann}(x) = p^e R$, then there exists a submodule $K \subseteq M$ such that $M = \langle x \rangle \oplus K$. To see this, we will show that there exists an R-module homomorphism $\alpha : M \to Rx$ that is the identity on Rx. Suppose α exists. Set K to be the kernel of α. Let $m \in M$. Then $\alpha(m) \in Rx$, and hence $\alpha(\alpha(m)) = \alpha(m)$. Thus, $\alpha(m - \alpha(m)) = \alpha(m) - \alpha(\alpha(m)) = 0$, so that $m - \alpha(m) \in K$. Thus, $m \subseteq Rx + K$, since $\alpha(m) \in Rx$, by definition. Therefore, $M = Rx + K$. Suppose $rx \in K$ belongs to $Rx \cap K$. Then $rx = \alpha(rx) = 0$. Thus, $Rx \cap K = 0$, showing $M = Rx \oplus K$.

To find $\alpha : M \to Rx$ which is the identity on Rx, let $\mathcal{C}$ denote the collection of submodules $N \subseteq M$ containing Rx for which there exists a homomorphism $\gamma : N \to Rx$ which is the identity on Rx. Note that Rx belongs to $\mathcal{C}$ by just taking the identity map on R, so $\mathcal{C}$ is not empty. Then $\mathcal{C}$ has a maximal element, say N, together with a homomorphism $\alpha : N \to Rx$, which is the identity on Rx. We claim $N = M$. If so, then we are done. Suppose not. Take $m \in M \backslash N$ such that $pm \in N$. Then $p^e m = 0$, so that $p^{e-1}\alpha(pm) = 0$. Now, since $\alpha(pm) \in Rx$, we may write $\alpha(pm) = rx$, for some $r \in R$. Thus, $0 = p^{e-1}(rx) = (p^{e-1}r)x$, so $p^{e-1}r \in \text{ann}(x) = p^e R$. Thus, $p^{e-1}r$ is divisible by p^e, so r is divisible by p. Thus, we may write $r = r_0 p$ and therefore $\alpha(pm) = p(r_0 x)$ [4]. Set $z := r_0 x$, so that

$$\alpha(pm) = pz \in Rx.$$

We now define $\gamma : N + \langle m \rangle \to Rx$ as follows: $\gamma(n + rm) = \alpha(n) + rz$, for all $n \in N$ and $r \in R$. If γ is well defined, then the fact that γ extends α and $N + Rm$ is strictly larger that N contradicts the maximality of N. Thus, we must have $N = M$, which gives what we want.

Finally, to see that γ is well defined, suppose that $n + rm = n' + r'm$, for $r, r' \in R$ and $n, n' \in N$. Then $(r' - r)m \in N$. Set $J := \{s \in R \mid sm \in N\}$. This is a principal ideal, so $J = aR$, say. Since $p \in J$, $p \in aR$. Thus, $a \mid p$. This can only happen if a is a unit multiple of p, and hence $aR = pR$. Therefore, $r' - r = tp$, so that $r' = r + tp$. Thus, $n + rm = n' + (r + tp)m$ so that $n = n' + tpm$. Therefore

$$\begin{aligned}
\gamma(n + rm) &= \gamma((n' + tpm) + rm) \\
&= \alpha(n' + tpm) + rz \\
&= \alpha(n') + t\alpha(pm) + rz \\
&= \alpha(n') + tpz + rz \\
&= \alpha(n') + (r + tp)z \\
&= \alpha(n') + r'z \\
&= \gamma(n' + rm).
\end{aligned}$$

This shows that γ is well defined and provides the contradiction we sought. This completes the proof of the theorem. □

Lecture 36: Wednesday November 17. The next proposition shows that the structure theorem for torsion modules reduces to the case covered in Theorem 35.2.

Proposition 36.1. *Let R be a* PID *an M a finitely generated, torsion R-module. Suppose $ann(M) = aR$, and $a = p_1^{e_1} \cdots p_r^{e_r}$, for primes $p_i \in R$ and $e_i \geq 1$. Then*

(i) $M = M(p_1) \oplus \cdots \oplus M(p_r)$.
(ii) $\text{ann}(M(p_i)) = p_i^{e_i} R$.

Proof. For (i), set $s_i := \Pi_{j \neq i} p_j^{e_j}$, for $1 \leq i \leq r$. Since the GCD of the s_i equals 1, the ideal generated by the s_i is R. Thus, we may write $1 = t_1 s_1 + \cdots + t_r s_r$. For any $x \in M$, we have $x = (t_1 s_1 x) + \cdots + (t_r s_r x)$. Since $p_i^{e_i} \cdot (t_i s_i x) = 0$, each $t_i s_i x \in M(p_i)$. This shows that $M = M(p_1) + \cdots + M(p_r)$. On the other hand, by the previous proposition, each $M(p_i)$ is annihilated by a power of p_i, so we take α_i to be the least power of p_i annihilating $M(p_i)$. Then there exist $c, d \in R$ such $1 = cp_i^{\alpha_i} + du_i$, where $u_i = \Pi_{j \neq i} p_j^{\alpha_j}$. Now suppose $y \in M(p_i) \cap (\sum_{j \neq i} M(p_j))$. Then $y = (cp_i^{\alpha_i} y) + (du_i y)$. Since $y \in M(p_i)$, $cp_i^{\alpha_i} y = 0$, while on the other hand, $du_i y = 0$, since $y \in \sum_{j \neq i} M(p_j)$. Thus, $y = 0$, showing $M = M(p_1) \oplus \cdots \oplus M(p_r)$, as required. □

[4] Note: Here is where we are using the fact that $\text{ann}(x) = \text{ann}(M)$. If $\text{ann}(x) = p^c R$, with $c < x$, and the only power of p annihilating m is e, then the equation $(p^{e-1}r)x = 0$ does not tell us anything, since $p^{e-1}x$ is already 0.

For (ii), by the previous paragraph, $\mathrm{ann}(M(p_i)) = p_i^{\alpha_i}R$, where α_i to be the least power of p_i annihilating $M(p_i)$. Set $a' := p_i^{\alpha_1} \cdots p_r^{\alpha_r}$. Then a' annihilates M, since every element in M is a sum of elements from the $M(p_i)$. Thus, $a \mid a'$. It follows that each $\alpha_i \geq e_i$. Now suppose, for example, $\alpha_1 > e_1$. Then there exists a non-zero $x \in M(p_1)$ such that $p_1^{e_1}x \neq 0$. Since $M(p_1) \cap M(p_2) = 0$, $p_2^{e_2}p_1^{e_1}x \neq 0$. Continuing in this way, using the fact that $M(p_1) \cap \Sigma_{j\neq 1}M(p_j) = 0$, we have that $p_r^{e_r} \cdots p_1^{e_1}x \neq 0$, which is a contradiction, since $a = p_1^{e_1} \cdots p_r^{e_r}$ annihilates M. Thus, $e_1 = \alpha_1$, so that $\mathrm{ann}(M(p_1)) = p_1^{e_1}R$, as required. Exactly the same argument shows $\alpha_i = e_i$ for all i, which gives what we want. □

We now have all of the ingredients for the structure theorem for finitely generated torsion modules over a PID. Note, then, that this yields a structure theorem for finite abelian groups.

Structure Theorem for Torsion modules over a PID. *Let R be a* PID *and M a finitely generated torsion module over R, i.e., $M = T(M)$. Then M is a direct sum of cyclic modules. In fact, if we write ann$(M) = aR$ and $a = p_1^{e_1} \cdots p_r^{e_r}$, with each $p_i \in R$ prime and $e_i \geq 1$, then there exist integers $n_1, \ldots, n_r \geq 1$ and for each $1 \leq i \leq r$, integers $e_i = e_{i,1} \geq \cdots \geq e_{i,n_i}$ and elements $x_{i,1}, \ldots x_{i,n_i} \in M$ such that ann$(x_{i,j}) = p_i^{e_{i,j}}R$ and*

$$M = \langle x_{1,1}\rangle \oplus \cdots \oplus \langle x_{1,n_1}\rangle \oplus \cdots \oplus \langle x_{r,1}\rangle \oplus \cdots \oplus \langle x_{r,n_r}\rangle.$$

Moreover, we have

$$M \cong (R/p_1^{e_{1,1}}R) \oplus \cdots \oplus (R/p_1^{e_{1,n_1}}R) \oplus \cdots \oplus (R/p_r^{e_{r,1}}R) \oplus \cdots \oplus (R/p_r^{e_{r,n_r}}R).$$

This second decomposition may be thought of as an external direct sum.

Proof. The first statement is immediate from Proposition 35.3 and Theorem 35.2. The second statement follows from the first and the fact that the homomorphism $\phi_{i,j} : R \to \langle x_{i,j}\rangle$ defined by $\phi(r) = rx_{i,j}$ is a surjective R-module homomorphism with kernel $p_i^{e_{i,j}}R$. □

For a finite abelian group G, one says that G has *index* $a \geq 2$ if G has elements of order a and the order of every element of G is divisible by a. This is the same thing as saying that, when we view G as a $\mathbb{Z}$-module, $\mathrm{ann}(G) = a\mathbb{Z}$. Thus, the following corollary is immediate from the structure theorem above.

Corollary 36.2. Let G be a finite abelian group of index $a = p_1^{e_1} \cdots p_r^{e_r}$, where $p_1, \ldots, p_r$ are prime. Then, for each $1 \leq i \leq r$, there exists $e_i = e_{i1} \geq \cdots \geq e_{is_i}$ such that

$$G \cong (\mathbb{Z}_{p_1^{e_{1,1}}} \oplus \cdots \oplus \mathbb{Z}_{p_1^{e_{1,n_1}}}) \oplus \cdots \oplus (\mathbb{Z}_{p_r^{e_{r,1}}} \oplus \cdots \oplus \mathbb{Z}_{p_r^{e_{r,n_r}}}).$$

Remark 36.3. The collection of primes powers $\{p_i^{e_{i,j}}\}$ appearing in the structure theorem above are called the *elementary divisors* of M and are uniquely determined by M. We will sketch a uniqueness argument, focusing on the case that $R = \mathbb{Z}$, though it should be fairly clear how this applies for a general PID. Suppose M is a finite abelian group, i.e., a finitely generated torsion $\mathbb{Z}$-module and on the one hand, we have a decomposition

$$M \cong (\mathbb{Z}/p_1^{e_{1,1}}\mathbb{Z}) \oplus \cdots \oplus (\mathbb{Z}/p_1^{e_{1,n_1}}\mathbb{Z}) \oplus \cdots \oplus (\mathbb{Z}/p_r^{e_{r,1}}\mathbb{Z}) \oplus \cdots \oplus (\mathbb{Z}/p_r^{e_{r,n_r}}\mathbb{Z}),$$

where $a = p_1^{e_1} \cdots p_r^{e_r}$, with each $p_i \in \mathbb{Z}$ prime, is the order of M, i.e., the annihilator of M, while on the other hand,

$$M \cong (\mathbb{Z}/q_1^{f_{1,1}}\mathbb{Z}) \oplus \cdots \oplus (\mathbb{Z}/q_1^{f_{1,s_1}}\mathbb{Z}) \oplus \cdots \oplus (\mathbb{Z}/q_t^{f_{t,1}}\mathbb{Z}) \oplus \cdots \oplus (\mathbb{Z}/q_t^{f_{t,s_t}}\mathbb{Z}),$$

for primes $q_1, \ldots, q_t$ and positive integers $f_{i,1} \geq \cdots \geq f_{i,s_i}$. Set $b := q_1^{f_{1,1}} \cdots q_t^{f_{t,1}}$. From the second expression above, we see that $bR = \mathrm{ann}(M)$. Thus, $b\mathbb{Z} = a\mathbb{Z}$, so that a and b are associates in $\mathbb{Z}$, i.e., $a = \pm b$. Thus, by unique factorization, after a change in indices, $r = t$, $p_i = q_i$ and $e_{i,1} = f_{i,1}$ for all i. Now, for each $1 \leq i \leq r$, the subgroup $M(p_i)$ depends only on p_i and M, and not on either of the decompositions above. Thus, we may focus on a single prime and assume we have a prime $p \in \mathbb{Z}$ and a finite abelian group M such that

$$M \cong (\mathbb{Z}/p^{e_1}\mathbb{Z}) \oplus \cdots \oplus (\mathbb{Z}/p^{e_n}\mathbb{Z}) \cong (\mathbb{Z}/p^{f_1}\mathbb{Z}) \oplus \cdots \oplus (\mathbb{Z}/p^{f_s}\mathbb{Z}),$$

with $e_1 \geq \cdots \geq e_n$ and $f_1 \geq \cdots \geq f_t$. Now, thinking of M as a $\mathbb{Z}$-module, we let N denote the set of elements annihilated by p, i.e., the elements in the group having order p. This is a subgroup of M and an element in M has order p if an only if each coordinate does. Moreover, any cyclic group G has a unique subgroup of order p, if p divides G, and this subgroup consists of the $p-1$ elements of order p together with 0. Thus, we have

$$N \cong (p^{e_1-1}\mathbb{Z}/p^{e_1}\mathbb{Z}) \oplus \cdots \oplus (p^{e_n-1}\mathbb{Z}/p^{e_n}\mathbb{Z}) \cong (p^{f_1-1}\mathbb{Z}/p^{f_1}\mathbb{Z}) \oplus \cdots \oplus (p^{f_s-1}\mathbb{Z}/p^{f_s}\mathbb{Z}).$$

We now note that N can be regarded as a vector space over $\mathbb{Z}_p$. Given $\bar{c} \in \mathbb{Z}_p$ and $x \in N$, we define $\bar{c} \cdot x$ to be cx. Note that if $\bar{c} = \bar{d} \in \mathbb{Z}_p$, then $d = c + hp$, for $h \in \mathbb{Z}$. Thus, $dx = (c+hp)x = cx + hpx = cx$, so the action of $\mathbb{Z}_p$ on N is well-defined. The decompositions of N above show that $n = \dim_{\mathbb{Z}_p}(N) = s$ (since $p^{e_i-1}\mathbb{Z}/p^{e_i}\mathbb{Z} \cong \mathbb{Z}_p$), so that $n = s$. Note that induction on $e_1 + \cdots + e_n$ will give $e_1 = f_1, \ldots, e_n = f_n$. To see this, we first have to consider the factors in each decomposition of M and N consisting of $\mathbb{Z}_p$. So suppose

$$M \cong (\mathbb{Z}/p^{e_1}\mathbb{Z}) \oplus \cdots \oplus (\mathbb{Z}/p^{e_i}\mathbb{Z}) \oplus (\mathbb{Z}_p)^l \cong (\mathbb{Z}/p^{f_1}\mathbb{Z}) \oplus \cdots \oplus (\mathbb{Z}/p^{f_j}\mathbb{Z}) \oplus (\mathbb{Z}_p)^m,$$

where $e_i, f_j > 1$, $e_{i+1} = \cdots = e_n = 1 = f_{j+1} = \cdots = f_n$ and $i + l = n = j + m$. We then have

$$N \cong (p^{e_1-1}\mathbb{Z}/p^{e_1}\mathbb{Z}) \oplus \cdots \oplus (p^{e_i-1}\mathbb{Z}/p^{e_i}\mathbb{Z}) \oplus (\mathbb{Z}_p)^l \cong (p^{f_1-1}\mathbb{Z}/p^{f_1}\mathbb{Z}) \oplus \cdots \oplus (p^{f_j-1}\mathbb{Z}/p^{f_j}\mathbb{Z}) \oplus (\mathbb{Z}_p)^m.$$

Modding out N gives

$$M/N \cong (\mathbb{Z}/p^{e_1-1}\mathbb{Z}) \oplus \cdots \oplus (\mathbb{Z}/p^{e_i-1}\mathbb{Z}) \cong (\mathbb{Z}/p^{f_1-1}\mathbb{Z}) \oplus \cdots \oplus (\mathbb{Z}/p^{f_j-1}),$$

so by induction, $i = j$ and $e_1 - 1 = f_1 - 1, \ldots, e_i - 1 = f_i - 1$, and hence $e_1 = f_1, \ldots, e_i = f_i$. Since $i = j$, we have $l = m$ and the proof is complete, i.e., $n = s$ and $e_1 = f_1, \ldots, e_n = f_n$. □

Putting together Theorem 33.2 with the structure theorem for torsion modules, we obtain the full structure theorem for finitely generated modules over a PID.

Structure theorem for finitely generated modules over a PID. *Let R be a* PID *and M be a finitely generated module over R. Then their exists unique integers $s, r \geq 0, \geq 0$, unique primes $p_1, \ldots, p_r \in R$ and unique positive integers $e_{i,1} \geq \cdots \geq e_{i,n_i}$, for $1 \leq i \leq r$, such that*

$$M \cong R^s \oplus (R/p_1^{e_{1,1}}R) \oplus \cdots \oplus (R/p_1^{e_{1,n_1}}R) \oplus \cdots \oplus (R/p_r^{e_{r,1}}R) \oplus \cdots \oplus (R/p_r^{e_{r,n_r}}R).$$

Proof. By Theorem 33.2, and the remark following it, $M = F \oplus T(M)$, where F is a free R-module whose rank depends only on M. Thus, if the rank of F is s, $F \cong R^s$. By the structure theorem for torsion modules, if we write $\operatorname{ann}(T(M)) = aR$, and $a = p_1^{e_1} \cdots p_r^{e_r}$ with each $p_i \in R$ prime and $e_i \geq 1$, then

$$T(M) \cong (R/p_1^{e_{1,1}}R) \oplus \cdots \oplus (R/p_1^{e_{1,n_1}}R) \oplus \cdots \oplus (R/p_r^{e_{r,1}}R) \oplus \cdots \oplus (R/p_r^{e_{r,n_r}}R),$$

for unique $\{e_{i,j}\}$, as prescribed. □

Remark 36.4. In the statement of the full structure theorem, we have chosen to express M isomorphic to the external direct sum of cyclic modules. Of course, given Theorem 33.2 and the structure theorem for torsion modules, it is clear that M is also the internal direct sum of a free module and cyclic submodules as described in the structure theorem for finitely generated torsion modules. And of course, it can be the case that $s = 0$, in which case M is a torsion modules, or $r = 0$, in which case M is a free R-module.

Lecture 37: Friday November 19. There is another standard way to present the structure theorem for finitely generated modules over a PID, in which the decomposition of $T(M)$ takes a different form. Suppose we write

$$T(M) \cong (R/p_1^{e_{1,1}}R) \oplus \cdots \oplus (R/p_1^{e_{1,n_1}}R) \oplus \cdots \oplus (R/p_r^{e_{r,1}}R) \oplus \cdots \oplus (R/p_r^{e_{r,n_r}}R),$$

as in the structure theorem. One can use the *Chinese Remainder Theorem*, CRT for short, for PIDs, to rearrange the terms in the decomposition above. Recall, that for the integers, the CRT states that if $n, m \in \mathbb{Z}$ are relatively prime, then $\mathbb{Z}_{nm} \cong \mathbb{Z}_n \oplus \mathbb{Z}_m$.[5] Since GCDs and Bezout's Principle hold in R, given $r, s \in R$ with GCD equal to 1, we have $R/rsR \cong R/rR \oplus R/sR$. This isomorphism easily extends to finitely many elements whose GCD equals 1, i.e., elements generating R as an ideal. Thus, if we set $a_1 := p_1^{e_{1,1}} \cdots p_r^{e_{r,1}}$, we have that

$$R/p_1^{e_{1,2}}R \oplus \cdots \oplus R/p_r^{e_{r,1}}R \cong R/a_1R.$$

[5] To so this, let $\phi : \mathbb{Z} \to \mathbb{Z}_n \oplus \mathbb{Z}_m$ be the map that sends $a \in \mathbb{Z}$ to $(\bar{a}, \bar{a}) \in \mathbb{Z}_n \oplus \mathbb{Z}_m$. Since n, m are relatively prime, there exist $r \in n\mathbb{Z}$ and $s \in m\mathbb{Z}$ such that $1 = r + s$. Thus, $\bar{1} = \bar{s}$ in $\mathbb{Z}_n$ and $\bar{1} = \bar{r}$ in $\mathbb{Z}_m$. Given $(\bar{a}, \bar{b}) \in \mathbb{Z}_n \oplus \mathbb{Z}_m$, we have $\phi(sa + rb) = (\overline{sa + rb}, \overline{sa + rb}) = (\overline{sa}, \overline{rb}) = (\bar{a}, \bar{b})$, showing ϕ is surjective. It is easy to see that the kernel of ϕ is $nm\mathbb{Z}$, so the required isomorphism holds.

Now set $a_2 := p_1^{e_{1,2}} \cdots p_2^{e_{r,2}}$. We then have $a_2|a_1$ and

$$R/p_1^{e_{1,2}}R \oplus \cdots \oplus R/p_r^{e_{r,2}}R \cong R/a_2R,$$

and hence

$$R/p_1^{e_{1,2}}R \oplus \cdots \oplus R/p_r^{e_{r,1}}R \oplus R/p_1^{e_{1,2}}R \oplus \cdots \oplus R/p_r^{e_{r,2}}R \cong R/a_1R \oplus R/a_2R.$$

Continuing in this fashion, we have the existence of $a_1, \dots, a_d \in R$ such that $a_d \mid a_{d-1} \mid \cdots \mid a_1$ and

$$M \cong R^s \oplus R/a_1R \oplus R/a_2R \oplus \cdots \oplus R/a_dR.$$

The ideals $a_1R, \dots a_dR$ are called the *invariant factors of* R and are uniquely determined by M.

Application to Linear Algebra. Let V be a finite dimensional vector space over the field F and $T : V \to V$ a linear transformation. Then V becomes an $F[x]$-module under the action of T: For $p(x) \in F[x]$ and $v \in V$, one defines $p(x) \cdot v$ to be $p(T)(v)$. It is easy to check that this makes V into a finitely generated $F[x]$-module. Since $F[x]$ is a PID, we may apply the structure theorem for finitely generated modules over a PID. Let us first note that V is a torsion $F[x]$-module. For this, set $d := \dim(V)$ and take $0 \neq v \in V$. Then $v, T(v), \cdots, T^d(v)$ are linearly dependent vectors, and thus, there exist $\alpha_0, \dots, \alpha_d \in F$ such that $\alpha_0 v + \alpha_1 T(v) + \cdots + \alpha_d T^d(v) = 0$. Setting $p(x) = \alpha_0 + \alpha_1 x + \cdots + \alpha_x^d$, this yields $p(T)(v) = 0$, or in module notation, $p(x)v = 0$. Since V is a finitely generated torsion module over $F[x]$, its annihilator is non-zero. If we set $q(x)R := \mathrm{ann}(V)$, then $q(T) = 0$, since $q(T)(v) = 0$, for all $v \in V$. Not that if $p(x) \in F[x]$ and $p(T) = 0$, then $p(x)$ annihilates V and thus $q(x) \mid p(x)$. For this reason, $q(x)$ is called the *minimal polynomial of* T.

Now take $v \in V$. Then $\mathrm{ann}(v)$ is easily seen to be generated by the polynomial $q_v(x)$ of least degree such that $q_v(T)(v) = 0$, and moreover $q_v(x)$ divides any polynomial $p(x)$ such that $p(T)(v) = 0$. What is $\langle v \rangle$, the cyclic submodule of V generated by v ? By definition it is $\{f(x)v \mid f(x) \in F[x]\} = \{f(T)(v) \mid f(x) \in F[x]\}$. Let $c := \deg(q_v(x))$. Then, for any $p(x) \in F[x]$ with degree greater than or equal to c, we can write $p(x) = h(x)q_v(x) + r(x)$, with $\deg r(x) < c$, so that $p(T)(v) = h(T)q_v(T)(v) + r(T)(v) = r(T)(v)$. It follows easily from this, that $B := \{v, T(v), \dots, T^{c-1}(v)\}$ forms a basis for $\langle v \rangle$, when we think of $\langle v \rangle$ as a vector space over F. [6]. Note that by definition, $\langle v \rangle$ is invariant under T, i.e., $T_{|\langle v \rangle}$ is a linear transformation from $\langle x \rangle$ to itself. What is the matrix of $T|_{\langle v \rangle}$ with respect to the basis B? To find this, write $q_v(x) = x^c + \alpha_1 x^{c-1} + \cdots + \alpha_c$. Then we have:

$$\begin{aligned}
T(v) &= 0 \cdot v + 1 \cdot T(v) + 0 \cdot T^2(c) + 0 \cdot T^3(v) + \cdots + 0 \cdot T^{c-1}(v)\\
T(T(v)) &= 0 \cdot v + 0 \cdot T(v) + 1 \cdot T^2(v) + 0 \cdot T^3(v) + \cdots + 0 \cdot T^{c-1}(v)\\
\vdots &= \vdots\\
T(T^{c-1}(v)) &= -\alpha_c \cdot v + \alpha_{c-1} \cdot T(v) + \alpha_{c-2} \cdot T^2(v) - \cdots - \alpha_1 \cdot T^{c-1}(v),
\end{aligned}$$

the last inequality following from $q_v(T)(v) = 0$. Thus, the matrix of $T|_{\langle v \rangle}$ with respect to B is

$$\begin{pmatrix}
0 & 0 & 0 & \cdots & -a_0\\
1 & 0 & 0 & \cdots & -\alpha_1\\
0 & 1 & 0 & \cdots & -\alpha_2\\
\vdots & \vdots & \vdots & \ddots & \vdots\\
0 & 0 & 0 & \cdots & -\alpha_{c-1}
\end{pmatrix},$$

the *Companion Matrix of* $q_v(x)$, denoted $C(q_v(x))$. Note that this $c \times c$ matrix depends only on $q_v(x)$, so that given any $f(x) \in F[x]$ we may define its companion matrix $C(f(x))$ analogously.

We now state the *Rational Canonical Form Theorem* for T and note that it is really just a restatement of the structure theorem for finitely generated modules over a PID in the case that $F[x]$ is the ring, V is the module and the ring action is defined by $T : V \to V$.

[6] These vectors clearly span $\langle v \rangle$ as a vector space over F, and any non-trivial dependence relation among them would yield a polynomial $g(x)$ of degree less than c such that $g(T)(v) = 0$.

Rational Canonical Form via elementary divisors. *Let V be a finite dimensional vector space and $T : V \to V$ a linear transformation. Factor the minimal polynomial of T as $q(x) = p_1(x)^{e_1} \cdots p_r(x)^{e_r}$, with each $p_i(x)$ irreducible over F. Then V is a direct sum of cyclic subspaces. In particular, for each $1 \leq i \leq r$ there exist positive integers $e_i = e_{i,1} \geq \cdots \geq e_{i,n_i}$, and a basis $\mathcal{B}$ for V such that A, the matrix of T with respect to $\mathcal{B}$, has the the block diagonal form*

$$A = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_r \end{pmatrix},$$

and for each $1 \leq i \leq r$,

$$A_i = \begin{pmatrix} C(p_i(x)^{e_{i,1}}) & 0 & \cdots & 0 \\ 0 & C(p_i(x)^{e_{i,2}}) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & C(p_i(x)^{e_{i,n_i}}) \end{pmatrix}. \quad \square$$

We note that from the module perspective, $V = V(p_1(x)) \oplus \cdots \oplus V(p_r(x))$. Each $V(p_i(x))$ is a direct sum of cyclic subspaces that yield a companion matrix of the form $C(p_i(x)^{e_{i,j}})$, and each A_i is the matrix of $T|_{V(p_i(x))}$ with respect to the union of the cyclic bases from each of those cyclic subspaces.

Lecture 38: Monday November 22. We now return to the inverse Galois problem for finite abelian groups. Before showing that the Inverse Galois Problem has a positive answer for abelian groups, let's first see the idea behind the case when G, the group in question, is cyclic, say $G = \mathbb{Z}_n$. Using a special case of Dirichlet's theorem on primes in an arithmetic progression, one can find a prime number p such that n divides $p - 1$. We will prove this below. If we write $p - 1 = n \cdot m$, then any cyclic group of order $p - 1$ has a (unique) cyclic subgroup of order m, and if we factor out that subgroup, the quotient group is then cyclic of order n. Let ϵ is a primitive pth root of unity. We will also see that $\text{Gal}(\mathbb{Q}(\epsilon)/\mathbb{Q})$ is isomorphic to $\mathbb{Z}_p^*$, a cyclic group of order $p - 1$. Let $H \subseteq \text{Gal}(\mathbb{Q}(\epsilon)/\mathbb{Q})$ be the subgroup of order m and K be the fixed field of H. Then, since $\text{Gal}(\mathbb{Q}(\epsilon)/\mathbb{Q})$ is abelian, H is normal and K is Galois over $\mathbb{Q}$. Its Galois group over $\mathbb{Q}$ is the factor group $\text{Gal}(\mathbb{Q}(\epsilon)/\mathbb{Q})/H \cong \mathbb{Z}_n$, which gives what we want. The general case will follow along similar lines. Thus, we need to deal with the Galois group obtained by adjoining a root of unity and also the statement about primes in a particular type of arithmetic progression.

As a first step along our path, we'll make some general comments. For any positive integer n, $\mathbb{Z}_n$ is a ring. As such, not every element has a multiplicative inverse. Recall that an element in a commutative ring is said to be a *unit* precisely when it does have a multiplicative inverse. The set of units in a ring forms a group under multiplication. Now, it is not hard to see that the units in the ring $\mathbb{Z}_n$ are precisely the residue classes of the positive integers that are less than n and relatively prime to n. Let $\mathbb{Z}_n^*$ denote the multiplicative group of units in the ring $\mathbb{Z}_n$ and set $\phi(n) := |\mathbb{Z}_n^*|$. Then $\phi(n)$ is called the the *Euler phi function* or *Euler totient function*. Thus, $\phi(n)$ is the number of positive integers less than n and relatively prime to n. An important property of this function is that $\phi(nm) = \phi(n)\phi(m)$, whenever n and m are relatively prime. One way to see this is as follows. Since n and m are relatively prime, the rings $\mathbb{Z}_{nm}$ and $\mathbb{Z}_n \oplus \mathbb{Z}_m$ are isomorphic (via the Chinese Remainder Theorem). Thus the groups of units in these rings are isomorphic, i.e., $\mathbb{Z}_{nm}^* \cong (\mathbb{Z}_n \oplus \mathbb{Z}_n)^*$. However, it is easy to see that an element in a product of rings is a unit if and only if each coordinate is a unit in the corresponding factor. Thus, $(\mathbb{Z}_n \oplus \mathbb{Z}_m)^* = \mathbb{Z}_n^* \oplus \mathbb{Z}_m^*$, so we have $\mathbb{Z}_{nm}^* \cong \mathbb{Z}_n^* \oplus \mathbb{Z}_m^*$. Counting elements in these groups gives $\phi(nm) = \phi(n)\phi(m)$.

Theorem 38.1. *Let ϵ be a primitive nth root of unity. Then $\mathbb{Q} \subseteq \mathbb{Q}(\epsilon)$ is a Galois extension with Galois group isomorphic to $\mathbb{Z}_n^*$, and therefore, $[\mathbb{Q}(\epsilon) : \mathbb{Q}] = \phi(n)$. Moreover, if p is prime, then $\mathbb{Z}_p^*$ is cyclic*

Proof. Since $\mathbb{Q}(\epsilon)$ is the splitting field for $x^n - 1$ over $\mathbb{Q}$ and the extension is separable, the extension is Galois. We first show that $[\mathbb{Q}(\epsilon) : \mathbb{Q}] = \phi(n)$. To see this, set $G := \text{Gal}(\mathbb{Q}(\epsilon)/\mathbb{Q})$ and let $\sigma \in G$. Then σ is determined by its value $\sigma(\epsilon)$. On the one hand, its easy to check that $\sigma(\epsilon)$ is also a primitive nth root of unity. Since the number of primitive nth roots of unity equals $\phi(n)$, it follows that $\phi(n) \geq |G| = [\mathbb{Q}(\epsilon) : \epsilon]$. On the other hand, let $f(x) \in \mathbb{Q}(x)$ be the minimal polynomial for ϵ over $\mathbb{Q}$. We claim that if p is a prime less than n and relatively prime to n, then $f(\epsilon^p) = 0$. Suppose the claim holds. Then, since ϵ^p is also a primitive nth root of unity, $(\epsilon^p)^q = \epsilon^{pq}$ is also a root of $f(X)$ for any prime q less than n and relatively prime to n. By iterating, it follows that for any r less than n and relatively prime to n, ϵ^r is a root of $f(x)$. Thus $[\mathbb{Q}(\epsilon) : \mathbb{Q}] = \text{degree}(f(x)) \geq \phi(n)$, which gives what we want.

To prove the claim, write $x^n - 1 = f(x)g(x)$, with $g(x) \in \mathbb{Q}[x]$. Since $x^n - 1$ is a primitive polynomial, $f(x)$ and $g(x)$ have integer coefficients[7].

Since ϵ^p is also an nth root of unity, it is a root of $x^n - 1$, and thus either a root of $f(x)$ or $g(x)$. Suppose $g(\epsilon^p) = 0$. Then ϵ is a root of $g(x^p)$. Thus, $f(x)$ divides $g(x^p)$, and again, since $f(x)$ is monic, the same argument using Gauss's Lemma shows that we can write $g(x^p) = f(x)h(x)$, where $h(x)$ has integer coefficients. If we reduce the coefficients mod p, then we have an equation $g(x^p) \equiv f(x)h(x)$ holding in $\mathbb{Z}_p[x]$. Now, for any integer a, $a^p \equiv a$, modulo $p\mathbb{Z}$, so it follows that in $\mathbb{Z}_p[x]$, $g(x^p) \equiv g(x)^p$. Thus, $g(x)^p \equiv f(x)h(x)$ in $\mathbb{Z}_p[x]$. It follows from this that $f(x)$ and $g(x)$ have a common factor modulo $p\mathbb{Z}$. Since $x^n - 1 \equiv f(x)g(x)$ in $\mathbb{Z}_p[x]$, this means that in the algebraic closure of $\mathbb{Z}_p$, $x^n - 1$ has a repeated root. On the other hand, $x^n - 1$ and its derivative are relatively prime as elements of $\mathbb{Z}_p[x]$. This is because p does not divide n, so nx^{n-1} is not the zero polynomial in $\mathbb{Z}_p[x]$. Thus, $x^n - 1$ has distinct roots, a contradiction. Therefore, $g(\epsilon^p) \neq 0$, so we must have $f(\epsilon^p) = 0$, which establishes the claim and thus the equality $[\mathbb{Q}(\epsilon) : \mathbb{Q}] = \phi(n)$.

We now show that G is isomorphic to $\mathbb{Z}_n^*$. To see this, as above, we note that if $\sigma \in G$, then $\sigma(\epsilon)$ is a primitive nth root of unity, so $\sigma(\epsilon) = \epsilon^r$, for a unique positive integer r less than n and relatively prime to n. Define $\theta : G \to \mathbb{Z}_n^*$, by $\theta(\sigma) = \overline{r}$, the image of r in $\mathbb{Z}_n^*$. If $\tau \in G$ and $\tau(\epsilon) = \epsilon^s$, with s less than n and relatively prime to n, then $\tau\sigma(\epsilon) = (\epsilon^r)^s = \epsilon^{sr}$. Now, sr is relatively prime to n, but may be larger than n, so we write $sr = a \cdot n + q$, with q less than n and relatively prime to n. Thus, $\theta(\tau\sigma) = \overline{q} = \overline{sr} \in \mathbb{Z}_n^*$. But $\overline{sr} = \overline{s} \cdot \overline{r} = \theta(\tau)\theta(\sigma)$, so θ is a group homomorphism. Morover, by definition, if θ takes γ in G to $\overline{1}$, then $\gamma(\epsilon) = \epsilon$, so γ is the identity element in G, i.e., θ is one-to-one. Since G and $\mathbb{Z}_n^*$ both have order $\phi(n)$, θ must also be onto, so θ is an isomorphism. The last statement in the theorem follows since $\mathbb{Z}_p$ is a finite field, so its multiplicative group of non-zero elements is cyclic by Problem 2(v) on Homework 2. The proof of the theorem is now complete. □

Remark 38.2. Let $n \geq 1$ and ϵ be a primitive nth root of unity. It follows from the proof above that if $f(x)$ is the minimal polynomial for ϵ over $\mathbb{Q}$ then the roots of $f(x)$ are precisely the primitive nth roots of unity, i.e., $f(x) = (x - \epsilon_1) \cdots (x - \epsilon_{\phi(n)})$, where $\epsilon_1, \ldots, \epsilon_{\phi(n)}$ are the primitive nth roots of unity. Set $\Phi_n(x) := f(x)$. Then $\Phi_n(x)$ is called the nth *cyclotomic polynomial*. It also follows from the proof above that $\Phi_n(x)$ is an irreducible monic polynomial with integer coefficients. Moreover, if we let δ be any nth root of unity, then for the least d such that $\delta^d = 1$, d divides n (since the nth roots of unity are a group of order n) and δ is a primitive dth root of unity. Conversely, for any d dividing n, a primitive dth root of unity is an nth root of unity. It follows immediately from this that $x^n - 1 = \prod_{d|n} \Phi_d(x)$. Using this formula and recursion, one can compute the polynomials $\Phi_n(x)$. For example, $\Phi_1(x) = x - 1$, $\Phi_2(x) = x + 1$, $\Phi_3(x) = x^2 + x + 1$, $\Phi_4(x) = x^2 + 1$, $\Phi_5(x) = x^4 + x^3 + x^2 + x + 1$, $\Phi_6(x) = x^2 - x + 1$. If p is prime, then $\Phi_p(x) = x^{p-1} + \cdots + x + 1$.

Corollary 38.3. *Let $n_1, \ldots, n_t$ be pairwise relatively prime positive integers. For each $1 \leq i \leq t$, let ϵ_i be a primitive n_ith root of unity. Then $\epsilon := \epsilon_1 \cdots \epsilon_t$ is a primitive $n_1 \cdots n_t$ root of unity and Gal$(\mathbb{Q}(\epsilon)/\mathbb{Q})$ is isomorphic to $\mathbb{Z}_{n_1}^* \oplus \cdots \oplus \mathbb{Z}_{n_t}^*$.*

Proof. We use the following fact, whose proof we leave to the reader. Namely, that in an abelian group, if elements $x_1, \ldots, x_t$ have orders $n_1, \ldots, n_t$ that are pairwise relatively prime, then the order of $x_1 \cdots x_t$ is $n_1 \cdots n_t$. If we apply this to the group $(\mathbb{C}\backslash\{0\}, \cdot)$ and the elements $\epsilon_1, \ldots, \epsilon_t$, it follows that $\epsilon_1 \cdots \epsilon_t$ has order $n_1 \cdots n_t$. In other words, ϵ is a primitive $n_1 \cdots n_t$ root of unity. From Theorem 38.1, Gal$(\mathbb{Q}(\epsilon)/\mathbb{Q})$ is isomorphic to $(\mathbb{Z}_{n_1 \cdots n_t})^*$. But since this group is isomorphic to $\mathbb{Z}_{n_1}^* \oplus \cdots \oplus \mathbb{Z}_{n_t}^*$, the corollary follows. □

[7] To see this, first note that if $g_0(x), h_0(x) \in \mathbb{Z}[x]$ and $p \in \mathbb{Z}$ is prime, then, if p divides all of the coefficients of $g_0(x)h_0(x)$, p must divide all of the coefficients of $g_0(x)$ or all of the coefficients of $h_0(x)$ - since in $\mathbb{Z}_p[x]$, the product $\overline{g_0(x)h_0(x)} \equiv 0$, and hence either $\overline{g_0(x)} \equiv 0$ or $\overline{h_0(x)} \equiv 0$. This is one version of *Gauss's Lemma*. Now, from $x^n - 1 = g(x)h(x)$, we can write $g(x)h(x) = \frac{a}{b}g_0(x)h_0(x)$, with a, b relatively prime, $g_0(x), h_0(x) \in \mathbb{Z}[x]$, and $g_0(x), h_0(x)$ primitive polynomials. Thus, $b(x^n - 1) = ag_0(x)h_0(x)$. This forces $b = 1$, since a prime dividing b does not divide a or any of the coefficient of $g_0(x)h_0(x)$ (by Gauss's Lemma).

Lecture 39: Monday November 29. In order to show that the inverse Galois problem has a positive solution for finite abelian groups, we need the following following theorem which is a special case of the celebrated theorem of Dirichlet which states that if a and b are relatively prime positive integers, then there are infinitely many primes in the arithmetic progression $\{at + b\}_{t\geq 1}$. Theorem 39.1 will show that that for fixed n, there are infinitely many primes in the arithmetic progression $\{nt + 1\}_{t\geq 1}$, or equivalently, there are infinitely many primes congruent to 1 modulo $n\mathbb{Z}$. We will save the proof of this theorem until after we have shown our main result..

Theorem 39.1. *For a fixed positive integer n, there are infinitely many prime numbers p such that $p \equiv 1$ (mod $n\mathbb{Z}$).*

All of the pieces are now in place to prove that the Inverse Galois Problem has a positive solution for abelian groups.

Theorem 39.2. *Let G be a finite abelian group. There there exists a finite Galois extension $\mathbb{Q} \subseteq K$ such that Gal$(K/\mathbb{Q}) \cong G$.*

Proof. Since G is a finite abelian group, by the structure theorem for finite abelian groups, we may write it as a direct product of cylic groups, say $G = \mathbb{Z}_{n_1} \oplus \cdots \oplus \mathbb{Z}_{n_t}$. It follows from Theorem 39.1 that we can find t *distinct* primes $p_1, \ldots, p_t$ such that for each $1 \leq i \leq t$, p_i is congruent to 1 modulo $n_i\mathbb{Z}$. For each $1 \leq i \leq t$, there is an integer m_i such that $p_i - 1 = n_i \cdot m_i$. Then for each $1 \leq i \leq t$, there is a unique subgroup $H_i \subseteq \mathbb{Z}^*_{p_i}$ of order m_i and moreover, $\mathbb{Z}^*_{p_i}/H_i$ is isomorphic to $\mathbb{Z}_{n_i}$. Note, $\mathbb{Z}^*_{p_i}/H_i$ is a multiplicative group, while $\mathbb{Z}_{n_i}$ is an additive group. In any case, they are both cyclic groups of order n_i.

Now, for each $1 \leq i \leq t$, let ϵ_i be a primitive p_ith root of unity and write ϵ for the product $\epsilon_1 \cdots \epsilon_t$. By Corollary 38.3, $\mathrm{Gal}(\mathbb{Q}(\epsilon)/\mathbb{Q}) \cong \mathbb{Z}^*_{p_1} \oplus \cdots \oplus \mathbb{Z}^*_{p_t}$. Let $H \subseteq \mathrm{Gal}(\mathbb{Q}(\epsilon)/\mathbb{Q})$ be the subgroup corresponding to $H_1 \oplus \cdots \oplus H_t \subseteq \mathbb{Z}^*_{p_1} \oplus \cdots \oplus \mathbb{Z}^*_{p_t}$ and take K to be the fixed field of H. Then, since $\mathrm{Gal}(\mathbb{Q}(\epsilon)/\mathbb{Q})$ is abelian, H is normal, so K is a Galois extension of $\mathbb{Q}$. Moreover, $\mathrm{Gal}(K/\mathbb{Q}) = \mathrm{Gal}(\mathbb{Q}(\epsilon)/\mathbb{Q})/H$ is isomorphic to G. Thus, G is the Galois group of a Galois extension of $\mathbb{Q}$ and the proof of Theorem 39.2 is complete. □

Proof of Theorem 39.1 (J. Tate). The proof will proceed in a number of steps. The first step is interesting in its own right. It generalizes Euclid's proof that there are infinitely many prime numbers.

Step 1. Let $g(x)$ be a polynomial with integer coefficients. Then there are infinitely many prime numbers among the prime divisors of the elements in the set $\{g(0), g(1), g(2), \ldots\}$. To see this, we first reduce to the case that the constant term of $g(x)$ is 1. If the constant term of $g(X)$ is 0, the result is trivial. Suppose that $a \in \mathbb{Z}$ is the constant term of $g(x)$. Then every coefficient of $g(aX)$ is divisible by a. We can then write $g(ax) = a \cdot h(x)$, where $h(x)$ has constant term equal to 1. If we knew the set of values $\{h(0), h(1), h(2), \ldots\}$ had infinitely many prime divisors, the same would certainly hold for $g(ax)$, and thus for $g(x)$, since the values of $g(ax)$ are among those of $g(x)$. Therefore replacing $h(x)$ by $g(x)$, we may assume that the constant term of $g(x)$ is 1. Now, by way of contradiction, suppose the elements in the required set of values of $g(x)$ had just finitely many prime divisors, say $p_1, \ldots, p_r$. Consider the values of $g(t \cdot p_1 \cdots p_r)$ as $t \in \mathbb{Z}$ goes to infinity. Eventually these values are larger in absolute value than any p_i, yet are not divisble by any p_i. Thus for t sufficiently large, $g(t \cdot p_1 \cdots p_r)$ must be divisible by a prime different from any of the p_i, a contradiction. This proves Step 1.

Lecture 40: Wednesday December 1. We continue with the proof of Theorem 39.1.

Step 2. Let n be a positive integer and p be a prime number not dividing n. Then p divides $\Phi_n(c)$ for some integer c if and only if the order of the element $\bar{c}$ in $\mathbb{Z}^*_p$ equals n. To see this, suppose first that p divides $\Phi_n(c)$. Then p divides $c^n - 1$, so $(\bar{c})^n \equiv 1$ (mod $p\mathbb{Z}$). Suppose the order of $\bar{c}$ were less than n. Then for some e dividing n, $(\bar{c})^e \equiv 1$ (mod $p\mathbb{Z}$). Since $x^e - 1 \equiv \prod_{d|e} \Phi_e(x)$ modulo $p\mathbb{Z}$, it follows that $\bar{c}$ must also be a root modulo $p\mathbb{Z}$ of some $\Phi_d(x)$, with $d|n$ and $d < n$. Since $\Phi_d(x)$ and $\Phi_n(x)$ are factors of $x^n - \bar{1}$ mod $p\mathbb{Z}$, $\bar{c}$ is a multiple root of $x^n - \bar{1}$ over $\mathbb{Z}_p$. But as in the proof of Theorem 38.1, $x^n - 1$ and its derivative are relatively prime modulo $p\mathbb{Z}$, so $x^n - 1$ has distinct roots modulo $p\mathbb{Z}$. Therefore, the order of $\bar{c}$ in $\mathbb{Z}^*_p$ is n. Conversely, if the order of $\bar{c}$ in $\mathbb{Z}^*_p$ is n, then p divides $c^n - 1$. Therefore, p divides $\Phi_d(c)$, for some d dividing n. But if d were less than n, then p would divide $c^d - 1$, so the order of $\bar{c}$ in $\mathbb{Z}^*_p$ would be less than n. Thus, p divides $\Phi_n(c)$, as required. This finishes the proof of Step 2.

Step 3. For an integer $n \geq 1$ and a prime p not dividing n, $\mathbb{Z}_p^*$ has an element of order n if and only if $p \equiv 1 \pmod{n\mathbb{Z}}$. To see this, just note that since $\mathbb{Z}_p^*$ is cyclic, $\mathbb{Z}_p^*$ has an element of order n if and only if n divides $|\mathbb{Z}_p^*| = p - 1$, which happens if and only if $p \equiv 1 \pmod{n\mathbb{Z}}$.

We now finish the proof of Theorem 39.1. By Step 1, there are infinitely many primes p dividing the values $\Phi_n(c)$ as c varies over the positive integers. Of course, infinitely many of these primes do not divide n, so by Step 2, there are infinitely many primes p for which p does not divide n and for which there exists c such that $\bar{c}$ has order n in $\mathbb{Z}_p^*$. By Step 3, there are infinitely many primes congruent to 1 modulo $n\mathbb{Z}$, which is what we wanted to show. □

Examples 40.1. (i) The general construction of a Galois extension of $\mathbb{Q}$ for a given abelian group is easy, if $G = \mathbb{Z}_{n_1} \oplus \cdots \oplus \mathbb{Z}_{n_t}$ for $n_i = p_i - 1$, with p_i primes. For then, each $\mathbb{Z}_{n_i} \cong \mathbb{Z}_{p_i}^*$, so that if $n = p_1 \cdots p_r$, it follows from Corollary 38.3 that $G \cong \mathrm{Gal}(\mathbb{Q}(\epsilon)/\mathbb{Q})$, where ϵ is a primitive nth root of unity. Thus, for example, if $G = \mathbb{Z}_2 \oplus \mathbb{Z}_4 \oplus \mathbb{Z}_{18}$, and ϵ is a primitive 285th root of unity (since $285 = 3 \cdot 5 \cdot 19$), then $G \cong \mathrm{Gal}(\mathbb{Q}(\epsilon)/\mathbb{Q})$.

(ii) In general, the proof of Theorem 39.2 tells us how to construct a Galois extension over $\mathbb{Q}$ with given finite abelian Galois group. But doing so explicitly, is not alway so easy, even for cyclic groups of relatively small order. For example, suppose $G = \mathbb{Z}_8$. Then the first prime p such that $p \equiv 1 \bmod 8$ is 17. Thus, if ϵ is a primitive 17th root of unity, $\mathrm{Gal}(\mathbb{Q}(\epsilon)/\mathbb{Q}) \cong \mathbb{Z}_{17}^*$, which is a cyclic group of order 16. To find a Galois extension K of $\mathbb{Q}$ whose Galois group is isomorphic to $\mathbb{Z}_8$, we must take the fixed field of H, where $H \subseteq \mathrm{Gal}(\mathbb{Q}(\epsilon)/\mathbb{Q})$ is a subgroup of index 8. Now, $\overline{5}$ is a cyclic generator for $\mathbb{Z}_{17}^*$ (check this), and thus, the proof of Theorem 38.1 tells us that $\sigma \in \mathrm{Gal}(\mathbb{Q}(\epsilon)/\mathbb{Q})$ given by $\sigma(\epsilon) = \epsilon^5$ is a cyclic generator of $\mathrm{Gal}(\mathbb{Q}(\epsilon)/\mathbb{Q})$. Therefore σ^8 is an element of order two, and so it generates a subgroup of index 8. If we call this subgroup H, then $\gamma \in \mathbb{Q}(\epsilon)$ is in the fixed field of H if and only if $\sigma^8(\gamma) = \gamma$

Now, the minimal polynomial for ϵ over $\mathbb{Q}$ is $\Phi_{17}(x) = x^{16} + x^{15} + \cdots + x + 1$, so that $1, \epsilon, \ldots, \epsilon^{15}$ is a basis for $\mathbb{Q}(\epsilon)$ over $\mathbb{Q}$. Thus, if $\gamma \in \mathbb{Q}(\epsilon)$, we can write

$$\gamma = a_0 + a_1\epsilon + a_2\epsilon^2 + \cdots + a_{15}\epsilon^{15},$$

and therefore,

$$\sigma^8(\gamma) = a_0 + a_1\sigma^8(\epsilon) + a_2\sigma^8(\epsilon^2) + \cdots + a_{15}\sigma^8(\epsilon^{15}).$$

A somewhat tedious (though not difficult) calculation shows that

$$\begin{gathered}
\sigma^8(\epsilon) = -1 - \epsilon - \epsilon^2 - \cdots - \epsilon^{15} \\
\sigma^8(\epsilon^2) = \epsilon^{15}, \sigma^8(\epsilon^3) = e^{14}, \sigma^8(\epsilon^4) = \epsilon^{13} \\
\sigma^8(\epsilon^5) = \epsilon^{12}, \sigma^8(\epsilon^6) = \epsilon^{11}, \sigma^8(e^7) = \epsilon^{10} \\
\sigma^8(\epsilon^8) = \epsilon^9, \sigma^8(\epsilon^9) = \epsilon^8, \sigma^8(\epsilon^{10}) = \epsilon^7 \\
\sigma^8(\epsilon^{11}) = \epsilon^6, \sigma^8(\epsilon^{12}) = \epsilon^5, \sigma^8(\epsilon^{13}) = \epsilon^4 \\
\sigma^8(\epsilon^{13}) = \epsilon^4, \sigma^8(\epsilon^{14}) = \epsilon^3, \sigma^8(\epsilon^{15}) = \epsilon^2.
\end{gathered}$$

It follows that

$$\sigma^8(\gamma) = (a_0 - a_1) + (-a_1)\epsilon + (a_{15} - a_1)\epsilon^2 + (a_{14} - a_1)\epsilon^3 + \cdots + (a_2 - a_1)\epsilon^{15}.$$

Setting $\gamma = \sigma^8(\gamma)$, we obtain

$$a_0 = a_0, a_1 = 0, a_2 = a_{15}, a_3 = a_{14}, a_3 = a_{13}, \ldots, a_{14} = a_3, a_{15} = a_2.$$

Therefore,

$$\gamma = a_0 + a_2(\epsilon^2 + \epsilon^{15}) + a_3(\epsilon^3 + \epsilon^{14}) + \cdots + a_6(\epsilon^6 + \epsilon^{11}) + a_7(\epsilon^7 + \epsilon^{10}) + a_8(\epsilon^8 + \epsilon^9).$$

It follows that $K := \mathbb{Q}(\epsilon^2+\epsilon^{15}, \epsilon^3+\epsilon^{14}, \ldots, \epsilon^6+\epsilon^{11}, \epsilon^7+\epsilon^{10}, \epsilon^8+\epsilon^9)$ is the fixed field of H, and $\mathrm{Gal}(K/\mathbb{Q}) \cong \mathbb{Z}_8$, as required. We can simplify K by noting that

$$\begin{aligned}(\epsilon^2+\epsilon^{15})^2 &= \epsilon^4+2+\epsilon^{13}\\(\epsilon^3+\epsilon^{14})^2 &= \epsilon^6+2+\epsilon^{11}\\(\epsilon^5+\epsilon^{12})^2 &= \epsilon^{10}+2+\epsilon^7\\(\epsilon^4+\epsilon^{13})^2 &= \epsilon^8+2+\epsilon^9\end{aligned}$$

It follows that $K = \mathbb{Q}(\epsilon^2+\epsilon^{15}, \epsilon^3+\epsilon^{14}, \epsilon^5+\epsilon^{12})$. Can you find a simpler expression for K? □

Remark 40.2. As mentioned before, the Inverse Galois Problem is known to have a positive solution in many cases, including all solvable groups, the symmetric groups and all simple groups, except for one of the sporadic simple groups. Various other groups are known to yield a positive answer to the Inverse Galois Problem. Of course, one can fix any field F and ask if a corresponding Inverse Galois Problem over F has a positive answer for a given group G. A particular case of interest where the Inverse Galois Problem has a positive solution for $F \neq \mathbb{Q}$, and for all finite groups is when the base field is $F = \mathbb{C}(t)$, the rational function field in one variable over $\mathbb{C}$. □

Department of Mathematics, The University of Kansas, Lawrence KS 66045