CYCLOTOMIC EXTENSIONS

KEITH CONRAD

1. Introduction

For a positive integer n, an *nth root of unity* in a field is a solution to $z^n = 1$, or equivalently is a root of $T^n - 1$. There are at most n different nth roots of unity in a field since $T^n - 1$ has at most n roots in a field. A *root of unity* is an nth root of unity for some n. The only roots of unity in $\mathbf{R}$ are ± 1, while in $\mathbf{C}$ there are n different nth roots of unity for each n, namely $e^{2\pi ik/n}$ for $0 \leq k \leq n-1$ and they form a group of order n. In characteristic p there is no pth root of unity besides 1: if $x^p = 1$ in characteristic p then $0 = x^p - 1 = (x-1)^p$, so $x = 1$. That is strange, but it is a key feature of characteristic p, *e.g.*, it makes the pth power map $x \mapsto x^p$ on fields of characteristic p injective.

For a field K, an extension of the form $K(\zeta)$, where ζ is a root of unity, is called a *cyclotomic* extension of K. The term cyclotomic means "circle-dividing," which comes from the fact that the nth roots of unity in $\mathbf{C}$ divide a circle into n arcs of equal length, as in Figure 1 when $n = 7$. The important algebraic fact we will explore is that cyclotomic extensions of every field have an abelian Galois group; we will look especially at cyclotomic extensions of $\mathbf{Q}$ and finite fields. There are not many general methods known for constructing abelian extensions (that is, Galois extensions with abelian Galois group); cyclotomic extensions are essentially the only construction that works over all fields. Other constructions of abelian extensions are Kummer extensions, Artin-Schreier-Witt extensions, and Carlitz extensions, but these all require special conditions on the base field.

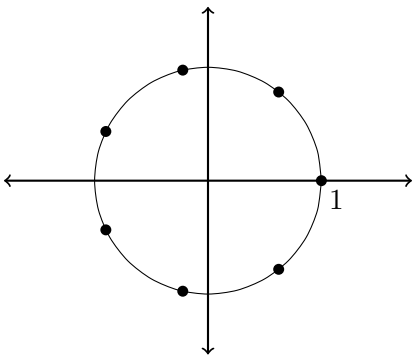


Figure 1. The 7th roots of unity.

The nth roots of unity in a field form a group under multiplication. It's obvious that this group in $\mathbf{C}$ is cyclic from the analytic formula for them, with generator $e^{2\pi i/n}$. In a general field there is no formula, but these roots of unity are still a cyclic group.

Theorem 1.1. *The group of nth roots of unity in a field is cyclic. More generally, every finite subgroup of the nonzero elements of a field is a cyclic group.*

Proof. Let F be a field and G be a finite subgroup of $F^\times$. From the general theory of abelian groups, if there are elements in G with orders n_1 and n_2 then there is an element of G with order the least common multiple $[n_1, n_2]$. Letting N be the maximal order of all the elements of G, we will show the order of every element in G divides N. If n is the order of some element in G then there is an element of G with order $[n, N] \geq N$. Since N is the maximal order we have $[n, N] \leq N$, so $[n, N] = N$, which implies n divides N (why?). Thus every element of G is a root of $T^N - 1$, which implies $|G| \leq N$ (the number of roots of a polynomial in a field is at most its degree). At the same time, since the order of each element divides the size of the group we have $N \mid |G|$. Hence $N = |G|$, which means some element of G has order $|G|$, so G is cyclic. □

Example 1.2. For each prime p, the group $(\mathbf{Z}/(p))^\times$ is cyclic by Theorem 1.1 since these are the nonzero elements in the field $\mathbf{Z}/(p)$ and they form a finite group. More generally, if F is a finite field then $F^\times$ is a cyclic group.

Watch out! Theorem 1.1 does *not* say $(\mathbf{Z}/(p^r))^\times$ is cyclic for $r > 1$, since the ring $\mathbf{Z}/(p^r)$ is not a field for $r > 1$. The theorem is simply silent about this. In fact, for other reasons, $(\mathbf{Z}/(p^r))^\times$ is cyclic for $p \neq 2$ but usually not if $p = 2$, *e.g.*, $(\mathbf{Z}/(8))^\times$ is not cyclic.

For a cyclotomic extension $K(\zeta)/K$ set n to be the order of ζ as a root of unity: the least exponent making $\zeta^n = 1$. Then $T^n - 1$ has every power of ζ as a root, so it has n different roots: $T^n - 1$ is *separable* over K. Conversely, if $T^n - 1$ is separable over K then it has n different roots, they are a group under multiplication, and Theorem 1.1 guarantees it is cyclic: there is a root of unity of order n among the nth roots of unity. Therefore when we construct cyclotomic extensions $K(\zeta)/K$ little is lost by assuming $T^n - 1$ is separable over K. That is equivalent to $T^n - 1$ being relatively prime to its derivative nT^{n-1} in $K[T]$, which is equivalent to $n \neq 0$ in K: $\mathrm{char}(K) = 0$, or $\mathrm{char}(K) = p$ and $(p, n) = 1$. We assume this is the case in all we do below.

When there are n different nth roots of unity, we denote the group of them by μ_n.[1] For instance, in $\mathbf{C}$ we have $\mu_2 = \{1, -1\}$ and $\mu_4 = \{1, -1, i, -i\}$. In $\mathbf{F}_7$, $\mu_3 = \{1, 2, 4\}$. A generator of μ_n is denoted ζ_n. That is, ζ_n denotes a root of unity of order n. **Watch out!** An nth root of unity is a solution to $z^n = 1$ but that doesn't mean it has order n. For example, 1 is an nth root of unity for every $n \geq 1$. An nth root of unity that has order n is called a *primitive* nth roots of unity ($z^n = 1$ and $z^j \neq 1$ for $j < n$). For example, -1 in $\mathbf{C}$ is a 4th root of unity but not a primitive 4th root of unity. For $a \in \mathbf{Z}$, the order of ζ_n^a is $n/(a, n)$,[2] so ζ_n^a is a primitive nth root of unity if and only if $(a, n) = 1$. Therefore when a field contains n different nth roots of unity it contains $\varphi(n)$ primitive nth roots of unity, where $\varphi(n) = |(\mathbf{Z}/(n))^\times|$. The primitive nth roots of unity are the generators of μ_n, and when $n \geq 3$ there is not a unique generator *e.g.*, if ζ_n is one generator then ζ_n^{-1} is another one), so writing ζ_n always involves making a *choice* of generator.

Any two primitive nth roots of unity in a field are powers of each other, so the field $K(\zeta_n)$ is independent of the choice of ζ_n. We will often write this field as $K(\mu_n)$: adjoining one primitive nth root of unity is the same as adjoining all nth roots of unity.

2. Embedding the Galois group

When $T^n - 1$ is separable over K, $K(\mu_n)/K$ is Galois since $K(\mu_n)$ is the splitting field of $T^n - 1$ over K. What is its Galois group?

[1] Anytime we write μ_n it is understood to contain n elements.

[2] See Theorem 3.13(3) in https://kconrad.math.uconn.edu/blurbs/grouptheory/order.pdf.

Lemma 2.1. *For $\sigma \in \mathrm{Gal}(K(\mu_n)/K)$ there is an integer $a = a_\sigma$ that is relatively prime to n such that $\sigma(\zeta) = \zeta^a$ for all $\zeta \in \mu_n$.*

Proof. Let ζ_n be a generator of μ_n (that is, a primitive nth root of unity), so $\zeta_n^n = 1$ and $\zeta_n^j \neq 1$ for $1 \leq j < n$. Since σ is multiplicative and injective, $\sigma(\zeta_n)^n = 1$ and $\sigma(\zeta_n)^j \neq 1$ for $1 \leq j < n$, so $\sigma(\zeta_n)$ is a primitive nth root of unity. This implies $\sigma(\zeta_n) = \zeta_n^a$ where $(a, n) = 1$. Each $\zeta \in \mu_n$ has the form ζ_n^k for some k, so

$$\sigma(\zeta) = \sigma(\zeta_n^k) = \sigma(\zeta_n)^k = (\zeta_n^a)^k = (\zeta_n^k)^a = \zeta^a.$$

□

The exponent a in Lemma 2.1 is well-defined modulo n: $\zeta_n^a = \zeta_n^b \Rightarrow a \equiv b \bmod n$ because ζ_n has order n. Thus we can think of $a = a_\sigma$ as an element of the group $(\mathbf{Z}/(n))^\times$.

Example 2.2. The primitive 7th roots of unity are the 7th roots of unity besides 1, and they are all roots of $(T^7 - 1)/(T - 1) = T^6 + T^5 + T^4 + T^3 + T^2 + T + 1$. This polynomial is irreducible over $\mathbf{Q}$ because it becomes Eisenstein at 7 when we replace T with $T + 1$:

$$\frac{(T+1)^7 - 1}{(T+1) - 1} = T^6 + 7T^5 + 21T^4 + 35T^3 + 35T^2 + 21T + 7.$$

This implies, for instance, that ζ_7 and ζ_7^2 have the same minimal polynomial over $\mathbf{Q}$. Since, moreover, $\mathbf{Q}(\zeta_7) = \mathbf{Q}(\zeta_7^2)$, there is an automorphism $\sigma \in \mathrm{Gal}(\mathbf{Q}(\mu_7)/\mathbf{Q})$ with the effect $\sigma(\zeta) = \zeta^2$ for all $\zeta \in \mu_7$. It is **FALSE** that $\sigma(\alpha) = \alpha^2$ for all $\alpha \in \mathbf{Q}(\mu_7)$, since squaring is not additive in characteristic 0. **ONLY** on the 7th roots of unity is σ being described as a power map. Elsewhere σ is determined from additivity and multiplicativity, *e.g.*,

$$\sigma(4\zeta_7^5 - 11\zeta_7 + 9) = 4(\zeta_7^2)^5 - 11\zeta_7^2 + 9 = 4\zeta_7^3 - 11\zeta_7^2 + 9.$$

Theorem 2.3. *The mapping*

$$\mathrm{Gal}(K(\mu_n)/K) \to (\mathbf{Z}/(n))^\times$$

where $\sigma \mapsto a_\sigma \bmod n$, from $\sigma(\zeta) = \zeta^{a_\sigma}$ for all $\zeta \in \mu_n$, is an injective group homomorphism.

Proof. Pick σ and τ in $\mathrm{Gal}(K(\mu_n)/K)$. For a primitive nth root of unity ζ_n,

$$(\sigma\tau)(\zeta_n) = \sigma(\tau(\zeta_n)) = \sigma(\zeta_n^{a_\tau}) = \sigma(\zeta_n)^{a_\tau} = (\zeta_n^{a_\sigma})^{a_\tau} = \zeta_n^{a_\sigma a_\tau}.$$

Also $(\sigma\tau)(\zeta_n) = \zeta_n^{a_{\sigma\tau}}$, so $\zeta_n^{a_{\sigma\tau}} = \zeta_n^{a_\sigma a_\tau}$. Since ζ_n has order n, $a_{\sigma\tau} \equiv a_\sigma a_\tau \bmod n$. This shows $\sigma \mapsto a_\sigma \bmod n$ is a homomorphism from $\mathrm{Gal}(K(\mu_n)/K)$ to $(\mathbf{Z}/(n))^\times$.

When σ is in the kernel, $a_\sigma \equiv 1 \bmod n$, so $\sigma(\zeta_n) = \zeta_n$. Also σ fixes all the elements of K, so σ is the identity on $K(\zeta_n) = K(\mu_n)$, so σ is the identity in $\mathrm{Gal}(K(\mu_n)/K)$. □

Since $(\mathbf{Z}/(n))^\times$ is abelian, the embedded subgroup $\mathrm{Gal}(K(\mu_n)/K)$ is abelian. We have proved that *cyclotomic extensions are always abelian.* Whenever we view $\mathrm{Gal}(K(\mu_n)/K)$ in $(\mathbf{Z}/(n))^\times$, it will always be understood to be by the embedding in Theorem 2.3.

Example 2.4. Complex conjugation is an automorphism of $\mathbf{Q}(\mu_n)/\mathbf{Q}$ with order 2. Under the embedding of $\mathrm{Gal}(\mathbf{Q}(\mu_n)/\mathbf{Q})$ into $(\mathbf{Z}/(n))^\times$, complex conjugation corresponds to $-1 \bmod n$ since $\overline{\zeta} = \zeta^{-1}$ for every root of unity ζ.

The embedding of $\mathrm{Gal}(K(\mu_n)/K)$ into $(\mathbf{Z}/(n))^\times$ may not be surjective; that depends on K. For instance, if $K = \mathbf{R}$ and $n \geq 3$ then $K(\mu_n)/K = \mathbf{C}/\mathbf{R}$ is a quadratic extension. The nontrivial $\mathbf{R}$-automorphism of $\mathbf{C}$ is complex conjugation, whose effect on roots of unity in

$\mathbf{C}$ is to invert them: $\overline{\zeta} = \zeta^{-1}$. Therefore the embedding $\mathrm{Gal}(\mathbf{C}/\mathbf{R}) \hookrightarrow (\mathbf{Z}/(n))^\times$ for $n \geq 3$ has image $\{\pm 1 \bmod n\}$, which is smaller than $(\mathbf{Z}/(n))^\times$ unless $n = 2$, 3, 4, or 6.

We will figure out the image in Theorem 2.3 in two important examples: $K = \mathbf{Q}$ and $K = \mathbf{F}_p$. In the first case the embedding $\mathrm{Gal}(K(\mu_n)/K) \to (\mathbf{Z}/(n))^\times$ is surjective, while in the second case the embedding need not be surjective but we can still describe the image.

Theorem 2.5. *The embedding* $\mathrm{Gal}(\mathbf{Q}(\mu_n)/\mathbf{Q}) \hookrightarrow (\mathbf{Z}/(n))^\times$ *is an isomorphism.*

Proof. Let ζ_n be a primitive nth root of unity over $\mathbf{Q}$. The size of $\mathrm{Gal}(\mathbf{Q}(\mu_n)/\mathbf{Q}) = \mathrm{Gal}(\mathbf{Q}(\zeta_n)/\mathbf{Q})$ is the number of $\mathbf{Q}$-conjugates of ζ_n, which is at most $\varphi(n) = |(\mathbf{Z}/(n))^\times|$. To prove $\mathrm{Gal}(\mathbf{Q}(\mu_n)/\mathbf{Q}) \hookrightarrow (\mathbf{Z}/(n))^\times$ is a surjection we will show for all $a \in \mathbf{Z}$ such that $(a, n) = 1$ that ζ_n and ζ_n^a are $\mathbf{Q}$-conjugate: their minimal polynomials over $\mathbf{Q}$ agree.

Since ζ_n^a only depends on $a \bmod n$, we can take $a > 0$, and in fact $a > 1$. Write $a = p_1 p_2 \cdots p_r$ as a product of primes p_i, each not dividing n (some p_i's could coincide). To show ζ_n and ζ_n^a have the same minimal polynomial over $\mathbf{Q}$, it suffices to show for each prime p not dividing n that every primitive nth root of unity and its pth power have the same minimal polynomial over $\mathbf{Q}$, since then the successive pairs of primitive nth roots of unity

$$\zeta_n, \quad \zeta_n^{p_1}, \quad \zeta_n^{p_1 p_2}, \quad \zeta_n^{p_1 p_2 p_3}, \quad \dots, \quad \zeta_n^{p_1 p_2 \cdots p_r} = \zeta_n^a$$

have the same minimal polynomial over $\mathbf{Q}$ since each is a prime power of the previous one.

For an arbitrary primitive nth root of unity ζ_n over $\mathbf{Q}$, assume ζ_n and ζ_n^p are *not* $\mathbf{Q}$-conjugate for some prime p not dividing n. We will get a contradiction. Let $f(T)$ be the minimal polynomial of ζ_n over $\mathbf{Q}$ and $g(T)$ be the minimal polynomial of ζ_n^p over $\mathbf{Q}$, so $g(T) \neq f(T)$. The polynomials $f(T)$ and $g(T)$ are in $\mathbf{Z}[T]$ since they both divide $T^n - 1$ and every monic factor of $T^n - 1$ in $\mathbf{Q}[T]$ is in $\mathbf{Z}[T]$ by Gauss' lemma.

Every nth root of unity is a root of $T^n - 1$, so $f(T)$ and $g(T)$ each divide $T^n - 1$ in $\mathbf{Q}[T]$. They are different monic irreducibles, so $T^n - 1 = f(T)g(T)h(T)$ for a monic $h(T) \in \mathbf{Q}[T]$. By Gauss' lemma, $h(T) \in \mathbf{Z}[T]$. Reducing this equation modulo p,

$$T^n - \overline{1} = \overline{f}(T)\overline{g}(T)\overline{h}(T) \tag{2.1}$$

in $\mathbf{F}_p[T]$. The polynomial $T^n - \overline{1}$ is separable in $\mathbf{F}_p[T]$ since p doesn't divide n, so (2.1) tells us that $\overline{f}(T)$ and $\overline{g}(T)$ are relatively prime in $\mathbf{F}_p[T]$. Because $f(T)$ and $g(T)$ are monic, their reductions $\overline{f}(T)$ and $\overline{g}(T)$ have the same degrees as $f(T)$ and $g(T)$, so the reductions are nonconstant.

Since $g(\zeta_n^p) = 0$, $g(T^p)$ has ζ_n as a root, so $f(T) \mid g(T^p)$ in $\mathbf{Q}[T]$. Write $g(T^p) = f(T)k(T)$ for some monic $k(T)$ in $\mathbf{Q}[T]$. We have $k(T) \in \mathbf{Z}[T]$ by Gauss' lemma (why?). Reduce the equation $g(T^p) = f(T)k(T)$ modulo p and use the formula $\overline{g}(T^p) = \overline{g}(T)^p$ in $\mathbf{F}_p[T]$ to get

$$\overline{g}(T)^p = \overline{f}(T)\overline{k}(T)$$

in $\mathbf{F}_p[T]$. Thus every irreducible factor of $\overline{f}(T)$ in $\mathbf{F}_p[T]$ is a factor of $\overline{g}(T)$ (and there *are* irreducible factors since $\overline{f}(T)$ is nonconstant). That contradicts relative primality of $\overline{f}(T)$ and $\overline{g}(T)$ in $\mathbf{F}_p[T]$. □

Remark 2.6. The proof of Theorem 2.5 above goes back to Dedekind [2]. Its appearance in van der Waerden's *Moderne Algebra* in 1930 made it the standard proof in later books. Here is another proof of Theorem 2.5, due to Landau [5]. Let $f(T)$ be the minimal polynomial of ζ_n over $\mathbf{Q}$, so $f(T)$ is monic in $\mathbf{Z}[T]$ as explained above. We want to show when $(a, n) = 1$ that $f(\zeta_n^a) = 0$. For all integers $k \geq 1$ write $f(T^k) = f(T)Q_k(T) + R_k(T)$ in $\mathbf{Z}[T]$, where $R_k = 0$ or $\deg R_k < \deg f$. Since $R_k(T)$ is 0 or has degree less than $\deg f = [\mathbf{Q}(\zeta_n) : \mathbf{Q}]$,

$R_k(T)$ is determined by $R_k(\zeta_n) = f(\zeta_n^k)$, so $R_k(\zeta_n)$ only depends on k mod n. In particular, every $R_k(T)$ is one of $R_1(T), R_2(T), \ldots, R_n(T)$.

For prime p, $R_p(\zeta_n) = f(\zeta_n^p) = f(\zeta_n^p) - f(\zeta_n)^p$, so $R_p(T)$ is the remainder when $f(T^p) - f(T)^p$ is divided by $f(T)$. From $f(T)^p \equiv f(T^p) \bmod p$ we get $f(T^p) - f(T)^p \in p\mathbf{Z}[T]$, which implies (why?) that $R_p(T) \in p\mathbf{Z}[T]$. Let C be the largest absolute value of all coefficients in $R_1(T), R_2(T), \ldots, R_n(T)$. Since there are infinitely many primes (!), pick a prime $p > C$. The polynomial $R_p(T)$ must be 0: its coefficients are smaller in absolute value then C and are divisible by p, which exceeds C. Therefore $f(T^p) = f(T)Q_p(T)$ when $p > C$, so $f(\zeta_n^p) = 0$. This implies, by iteration, that $f(\zeta_n^k) = 0$ for the positive integers k whose prime factors all exceed C. If $(a, n) = 1$ and $a > 1$, define $k := a + n\prod_{p \le C, (p,a)=1} p$. Then $k \equiv a \bmod n$, so $(k, n) = 1$. The two terms in the sum defining k are relatively prime, so every prime factor of this k is larger than C (why?), which implies $0 = f(\zeta_n^k) = f(\zeta_n^a)$.

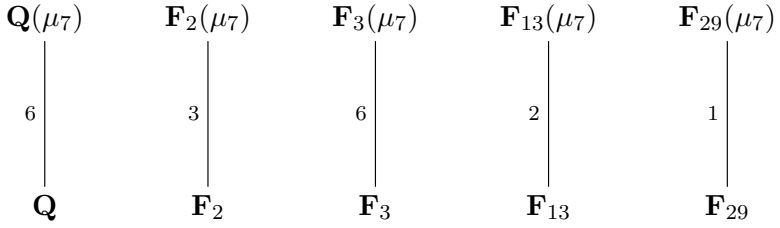
Remark 2.7. An introductory discussion of cyclotomic extensions of $\mathbf{Q}$ would not be complete without mentioning a deep theorem of Kronecker and Weber: every finite abelian extension of $\mathbf{Q}$ lies inside a cyclotomic extension of $\mathbf{Q}$. This is false for a base field K that is a finite extension of $\mathbf{Q}$ larger than $\mathbf{Q}$: some finite abelian extension of K does not lie in a cyclotomic extension of K. For instance, if $K = \mathbf{Q}(i)$ then $K(\sqrt[4]{1+i})/K$ is abelian but for no n is $K(\sqrt[4]{1+i}) \subset K(\zeta_n)$.

Now we turn to Galois groups of cyclotomic extensions of the finite field $\mathbf{F}_p$.

Theorem 2.8. *When n is not divisible by the prime p, the image of* $\mathrm{Gal}(\mathbf{F}_p(\mu_n)/\mathbf{F}_p)$ *in* $(\mathbf{Z}/(n))^\times$ *under the standard embedding is* $\langle p \bmod n \rangle$. *In particular,* $[\mathbf{F}_p(\mu_n) : \mathbf{F}_p]$ *is the order of* $p \bmod n$.

Proof. The polynomial $T^n - 1$ is separable in $\mathbf{F}_p[T]$ and the general theory of finite fields tells us $\mathrm{Gal}(\mathbf{F}_p(\mu_n)/\mathbf{F}_p)$ is generated by the pth power map $\varphi_p\colon x \mapsto x^p$ for all x in $\mathbf{F}_p(\mu_n)$. The standard embedding of $\mathrm{Gal}(\mathbf{F}_p(\mu_n)/\mathbf{F}_p)$ into $(\mathbf{Z}/(n))^\times$ associates to φ_p the congruence class $a \bmod n$ where $\varphi_p(\zeta) = \zeta^a$ for all $\zeta \in \mu_n$. Then $\zeta^p = \zeta^a$, so $a \equiv p \bmod n$. Therefore the standard embedding of $\mathrm{Gal}(\mathbf{F}_p(\mu_n)/\mathbf{F}_p)$ into $(\mathbf{Z}/(n))^\times$ turns φ_p into $p \bmod n$. Since φ_p generates the Galois group, the image of the Galois group in $(\mathbf{Z}/(n))^\times$ is $\langle p \bmod n \rangle$, so the size of the Galois group is the order of p in $(\mathbf{Z}/(n))^\times$. □

Example 2.9. The degree $[\mathbf{F}_p(\mu_7) : \mathbf{F}_p]$ is the order of $p \bmod 7$ that is 1, 2, 3, or 6 (if $p \neq 7$). The field diagram below gives some examples.



These are all consistent with how $(T^7 - 1)/(T - 1) = T^6 + T^5 + \cdots + T + 1$ factors modulo each of the primes: into irreducibles of common degree $[\mathbf{F}_p(\mu_7) : \mathbf{F}_p]$.

$$\begin{aligned} T^6+T^5+T^4+T^3+T^2+T+1 &\equiv (T^3+T+1)(T^3+T^2+1) \bmod 2 \\ &\equiv \text{irreducible mod } 3 \\ &\equiv (T^2+3T+1)(T^2+5T+1)(T^2+6T+1) \bmod 13 \\ &\equiv (T-7)(T-16)(T-20)(T-23)(T-24)(T-25) \bmod 29. \end{aligned}$$

For the cyclic group $\mathrm{Gal}(\mathbf{F}_p(\mu_n)/\mathbf{F}_p)$ to be as big as $\mathrm{Gal}(\mathbf{Q}(\mu_n)/\mathbf{Q}) \cong (\mathbf{Z}/(n))^\times$ is equivalent to saying $\langle p \bmod n\rangle = (\mathbf{Z}/(n))^\times$, so $(\mathbf{Z}/(n))^\times$ must be a cyclic group and $p \bmod n$ is a generator of it. The groups $(\mathbf{Z}/(n))^\times$ are usually not cyclic (like $n = 8$ or $n = 15$), so the standard embedding $\mathrm{Gal}(\mathbf{F}_p(\mu_n)/\mathbf{F}_p) \hookrightarrow (\mathbf{Z}/(n))^\times$ is usually not surjective.

Theorem 2.8 generalizes to all finite fields $\mathbf{F}_q$ as a base field. A proof is left to the reader.

Theorem 2.10. *Let $\mathbf{F}_q$ be a finite field with prime power order q. When n is relatively prime to q, the image of $\mathrm{Gal}(\mathbf{F}_q(\mu_n)/\mathbf{F}_q)$ in $(\mathbf{Z}/(n))^\times$ is $\langle q \bmod n\rangle$. In particular, $[\mathbf{F}_q(\mu_n) : \mathbf{F}_q]$ is the order of $q \bmod n$.*

3. Composites and Intersections of Cyclotomic Extensions

When $T^n - 1$ and $T^m - 1$ are separable over K, the composite field $K(\mu_m)K(\mu_n)$ equals $K(\mu_{[m,n]})$. Indeed, both $K(\mu_m)$ and $K(\mu_n)$ lie in $K(\mu_{[m,n]})$, so their composite does too. For the reverse inclusion, a primitive root of unity of order $[m, n]$ can be obtained by multiplying *suitable* mth and nth roots of unity (why?), so $\mu_{[m,n]} \subset K(\mu_m)K(\mu_n)$, which implies $K(\mu_{[m,n]}) \subset K(\mu_m)K(\mu_n)$. Therefore $K(\mu_m)K(\mu_n) = K(\mu_{[m,n]})$.

It is natural to guess that a counterpart of $K(\mu_m)K(\mu_n) = K(\mu_{[m,n]})$ for intersections is $K(\mu_m) \cap K(\mu_n) = K(\mu_{(m,n)})$. The inclusion $\supset$ is easy, but the other inclusion can be **FALSE**! It's possible for m and n to be relatively prime and $K(\mu_m) \cap K(\mu_n)$ to be larger than $K(\mu_1) = K$. Matt Emerton pointed out to me the following simple example.

Example 3.1. If $K = \mathbf{Q}(\sqrt{3})$ then $K(\zeta_4) = K(i) = \mathbf{Q}(\sqrt{3}, i) = \mathbf{Q}(\sqrt{3}, \sqrt{-3}) = K(\zeta_3)$ because $\zeta_3 = (-1 + \sqrt{-3})/2$. Since $K(\zeta_4)$ and $K(\zeta_3)$ are equal and larger than K itself, their intersection is larger than $K(\zeta_{(4,3)}) = K$.

Here are some more examples.

Example 3.2. For $n \geq 3$, complex conjugation on $\mathbf{Q}(\zeta_n)$ is an automorphism of order 2. Its fixed field (the real numbers in $\mathbf{Q}(\zeta_n)$) is denoted $\mathbf{Q}(\zeta_n)^+$. Show as an exercise that $\mathbf{Q}(\zeta_n)^+ = \mathbf{Q}(\zeta_n + \zeta_n^{-1})$ and $[\mathbf{Q}(\zeta_n) : \mathbf{Q}(\zeta_n)^+] = 2$. For each pair of relatively prime integers m and n both at least 3, one field K such that $K(\mu_m) \cap K(\mu_n) \neq K$ is $K = \mathbf{Q}(\zeta_{mn})^+$. Since $[\mathbf{Q}(\zeta_{mn}) : K] = 2$ and $K \subset \mathbf{R}$, $K(\zeta_m) = \mathbf{Q}(\zeta_{mn})$ and $K(\zeta_n) = \mathbf{Q}(\zeta_{mn})$. Thus $K(\zeta_m) \cap K(\zeta_n) = \mathbf{Q}(\zeta_{mn})$ is larger than K. Taking $m = 4$ and $n = 3$ we get Example 3.1: $\mathbf{Q}(\zeta_{12})^+ = \mathbf{Q}(\sqrt{3})$.

Example 3.3. Using Theorem 2.8,

$$\mathbf{F}_3(\mu_5) \cap \mathbf{F}_3(\mu_7) = \mathbf{F}_{3^4} \cap \mathbf{F}_{3^6} = \mathbf{F}_{3^2} \neq \mathbf{F}_3.$$

This weirdness does not happen when the base field is the rational numbers.

Theorem 3.4. *For all positive integers m and n, $\mathbf{Q}(\mu_m) \cap \mathbf{Q}(\mu_n) = \mathbf{Q}(\mu_{(m,n)})$; in particular, if $(m, n) = 1$ then $\mathbf{Q}(\mu_m) \cap \mathbf{Q}(\mu_n) = \mathbf{Q}$.*

Proof. By Theorem 2.5, $[\mathbf{Q}(\mu_N) : \mathbf{Q}] = |(\mathbf{Z}/(N))^\times| = \varphi(N)$ for all positive integers N. There is a formula for $\varphi(N)$ in terms of the prime factors of N:

$$\varphi(N) = N \prod_{p|N} \left(1 - \frac{1}{p}\right). \tag{3.1}$$

Since $\mathbf{Q}(\mu_d) \subset \mathbf{Q}(\mu_m)$ when $d \mid m$, we have $\mathbf{Q}(\mu_{(m,n)}) \subset \mathbf{Q}(\mu_m) \cap \mathbf{Q}(\mu_n)$. To show this containment is an equality we will show $\mathbf{Q}(\mu_m) \cap \mathbf{Q}(\mu_n)$ and $\mathbf{Q}(\mu_{(m,n)})$ have the same degree over $\mathbf{Q}$.

For finite Galois extensions L_1/K and L_2/K in a common field, we have $[L_1L_2 : K] = [L_1 : K][L_2 : K]/[L_1 \cap L_2 : K]$. The composite field $\mathbf{Q}(\mu_m)\mathbf{Q}(\mu_n)$ is $\mathbf{Q}(\mu_{[m,n]})$, so

$$[\mathbf{Q}(\mu_{[m,n]}) : \mathbf{Q}] = [\mathbf{Q}(\mu_m)\mathbf{Q}(\mu_n) : \mathbf{Q}] = \frac{[\mathbf{Q}(\mu_m) : \mathbf{Q}][\mathbf{Q}(\mu_n) : \mathbf{Q}]}{[\mathbf{Q}(\mu_m) \cap \mathbf{Q}(\mu_n) : \mathbf{Q}]}.$$

Replacing each $[\mathbf{Q}(\mu_N) : \mathbf{Q}]$ on the right side with $\varphi(N)$,

$$[\mathbf{Q}(\mu_m) \cap \mathbf{Q}(\mu_n) : \mathbf{Q}] = \frac{\varphi(m)\varphi(n)}{\varphi([m,n])}. \tag{3.2}$$

Using (3.1), (3.2) becomes

$$[\mathbf{Q}(\mu_m) \cap \mathbf{Q}(\mu_n) : \mathbf{Q}] = \frac{m\prod_{p|m}(1-1/p) \cdot n\prod_{p|n}(1-1/p)}{[m,n]\prod_{p|[m,n]}(1-1/p)}.$$

Since $[m,n](m,n) = mn$, the ratio $mn/[m,n]$ is (m,n). The prime factors of $[m,n]$ are those dividing either m or n, so the ratio of products over primes is the product of $1-1/p$ over all primes dividing both m and n, which means the prime factors of (m,n). Therefore

$$[\mathbf{Q}(\mu_m) \cap \mathbf{Q}(\mu_n) : \mathbf{Q}] = (m,n) \prod_{p|(m,n)} \left(1 - \frac{1}{p}\right) = \varphi((m,n)),$$

which is $[\mathbf{Q}(\mu_{(m,n)}) : \mathbf{Q}]$, so $\mathbf{Q}(\mu_m) \cap \mathbf{Q}(\mu_n)$ has the same degree over $\mathbf{Q}$ as $\mathbf{Q}(\mu_{(m,n)})$, hence the fields are equal since we already saw one is a subfield of the other. □

4. Coincidences of Cyclotomic Fields over $\mathbf{Q}$

Knowing the degree of cyclotomic extensions of $\mathbf{Q}$ lets us determine when two cyclotomic fields can coincide. For example, $\mathbf{Q}(\zeta_3) = \mathbf{Q}(\zeta_6)$ since $-\zeta_3$ has order 6. More simply, $\mathbf{Q}(\zeta_1) = \mathbf{Q}(\zeta_2) = \mathbf{Q}$ since $\zeta_1 = 1$ and $\zeta_2 = -1$. Here is the general result in this direction.

Theorem 4.1. *Let m and n be positive integers.*

1. *The number of roots of unity in $\mathbf{Q}(\mu_m)$ is $[2,m]$.*
2. *If $m \neq n$ then $\mathbf{Q}(\mu_m) = \mathbf{Q}(\mu_n)$ if and only if $\{m,n\} = \{k, 2k\}$ as sets for an odd k.*

Proof. 1) Our argument is adapted from [1, p. 158].

First we show there is a root of unity of order $[2,m]$ in $\mathbf{Q}(\mu_m)$. If m is odd then $[2,m] = 2m$ and the root of unity $-\zeta_m$ is in $\mathbf{Q}(\mu_m)$ with order $2m$. If m is even then $[2,m] = m$ and the root of unity ζ_m has order m.

If $\mathbf{Q}(\mu_m)$ contains an rth root of unity then $\mathbf{Q}(\mu_r) \subset \mathbf{Q}(\mu_m)$, and taking degrees over $\mathbf{Q}$ shows $\varphi(r) \leq \varphi(m)$. As $r \to \infty$, $\varphi(r) \to \infty$ (albeit erratically[3]) so there is a largest r

[3]This follows from showing a bound $\varphi(r) \leq B$ also bounds r from above. For prime powers p^e dividing r, $\varphi(p^e) \leq B$ too since $\varphi(p^e) \mid \varphi(r)$, so $p^{e-1}(p-1) \leq B$. Then $2^{e-1} \leq B$ and $p-1 \leq B$, so we get upper bounds on p and on e, which gives an upper bound on r by unique factorization.

satisfying $\mu_r \subset \mathbf{Q}(\mu_m)$. Since $\mu_m\mu_r = \mu_{[m,r]}$ is in $\mathbf{Q}(\mu_m)$ we have $[m,r] \leq r$, so $[m,r] = r$. Thus r is a multiple of m. Write $r = ms$. By (3.1), for all a and b in $\mathbf{Z}^+$ we have

$$\varphi(ab) = \frac{\varphi(a)\varphi(b)(a,b)}{\varphi((a,b))}, \tag{4.1}$$

so

$$\varphi(r) = \varphi(ms) = \varphi(m)\varphi(s)\frac{(m,s)}{\varphi((m,s))} \geq \varphi(m)\varphi(s).$$

Since $\mathbf{Q}(\mu_m) = \mathbf{Q}(\mu_r)$ for the maximal r, computing degrees over $\mathbf{Q}$ shows $\varphi(m) = \varphi(r) \geq \varphi(m)\varphi(s)$, so $1 \geq \varphi(s)$. Thus $\varphi(s) = 1$, so $s = 1$ or 2, so $r = m$ or $r = 2m$. That is, the number of roots of unity in $\mathbf{Q}(\mu_m)$ is either m or $2m$. If m is even then $\varphi(2m) = 2\varphi(m) > \varphi(m)$, so $r \neq 2m$. Thus when m is even the number of roots of unity in $\mathbf{Q}(\mu_m)$ is m. If m is odd then $\mathbf{Q}(\mu_m)$ contains $-\zeta_m$, which has order $2m$, so the number of roots of unity in $\mathbf{Q}(\mu_m)$ is $2m$. In general the number of roots of unity in $\mathbf{Q}(\mu_m)$ is $[2,m]$.

2) If $\mathbf{Q}(\mu_m) = \mathbf{Q}(\mu_n)$ and $m \neq n$ then counting roots of unity implies $[2,m] = [2,n]$. This becomes $m = [2,n]$ for even m (so $n = m/2$), and $2m = [2,n]$ for odd m (so $n = 2m$). □

Remark 4.2. Theorem 4.1 suggests two ways to parametrize cyclotomic extensions of $\mathbf{Q}$ without duplication: as $\mathbf{Q}(\mu_m)$ for m not twice an odd integer ($m \not\equiv 2 \bmod 4$) or for m equal to twice an odd integer ($m \equiv 2 \bmod 4$). In the first convention, $\mathbf{Q}(\mu_m)$ contains $2m$ roots of unity. The first convention, where $m \not\equiv 2 \bmod 4$, is commonly used since certain important results about these fields take on a simpler appearance with that convention.

When are there coincidences among different "real" cyclotomic fields $\mathbf{Q}(\mu_n)^+$?

Theorem 4.3. *If $\mathbf{Q}(\mu_m)^+ = \mathbf{Q}(\mu_n)^+$ then $\mathbf{Q}(\mu_m) = \mathbf{Q}(\mu_n)$ except for the cyclotomic fields $\mathbf{Q}(\mu_1)$, $\mathbf{Q}(\mu_3)$, and $\mathbf{Q}(\mu_4)$, which are different but all have $\mathbf{Q}(\mu_n)^+ = \mathbf{Q}$.*

Proof. We will break up the proof into two cases.

Case 1: $\mathbf{Q}(\mu_m)^+ = \mathbf{Q}(\mu_n)^+ = \mathbf{Q}$.

We will show that $\mathbf{Q}(\mu_n)^+ = \mathbf{Q}$ if and only if $\mathbf{Q}(\mu_n)$ is $\mathbf{Q}(\mu_1)$, $\mathbf{Q}(\mu_3)$, or $\mathbf{Q}(\mu_4)$. It's easy to see those three fields have $\mathbf{Q}(\mu_n)^+ = \mathbf{Q}$. To prove the converse, we can assume $n \geq 3$. Then $[\mathbf{Q}(\mu_n) : \mathbf{Q}(\mu_n)^+] = 2$ (Example 3.2), so $[\mathbf{Q}(\mu_n)^+ : \mathbf{Q}] = \varphi(n)/2$. If $\mathbf{Q}(\mu_n)^+ = \mathbf{Q}$ then $\varphi(n)/2 = 1$, so $\varphi(n) = 2$. The only such n are 3, 4, and 6, and $\mathbf{Q}(\mu_3) = \mathbf{Q}(\mu_6)$.

Case 2: $\mathbf{Q}(\mu_m)^+ = \mathbf{Q}(\mu_n)^+ \neq \mathbf{Q}$.

For general positive integers m and n, $\mathbf{Q}(\mu_m)^+ \cap \mathbf{Q}(\mu_n)^+ = \mathbf{Q}(\mu_{(m,n)})^+$ since both sides are the real numbers in $\mathbf{Q}(\mu_m) \cap \mathbf{Q}(\mu_n)$. (We are using Theorem 3.4 here.) So if $\mathbf{Q}(\mu_m)^+ = \mathbf{Q}(\mu_n)^+$ then

$$\mathbf{Q}(\mu_m)^+ = \mathbf{Q}(\mu_n)^+ = \mathbf{Q}(\mu_d)^+, \tag{4.2}$$

where $d = (m,n)$. By the hypothesis of Case 2, the fields in (4.2) are not $\mathbf{Q}$, so $m, n, d \geq 5$. All we'll really need is that these three numbers are at least 3.

Computing degrees over $\mathbf{Q}$ in (4.2), $\varphi(m)/2 = \varphi(d)/2$, so $\varphi(m) = \varphi(d)$. This is a strong restriction since $d \mid m$. Writing $m = dm'$, we'll see that m' is 1 or 2: using (4.1),

$$\varphi(m) = \varphi(dm') = \varphi(d)\varphi(m')\frac{(d,m')}{\varphi((d,m'))} = \varphi(m)\varphi(m')\frac{(d,m')}{\varphi((d,m'))} \geq \varphi(m)\varphi(m').$$

Thus $1 \geq \varphi(m')$, so $\varphi(m') = 1$. Thus m' is 1 or 2. Then $1 = \varphi(m')(d,m')/\varphi((d,m')) = (d,m')/\varphi((d,m'))$, so $(d,m') = \varphi((d,m'))$. The only positive integer equal to its φ-value is

1, so $(d, m') = 1$. If $m' = 1$ then $m = d$, so $\mathbf{Q}(\mu_m) = \mathbf{Q}(\mu_d)$. If $m' = 2$, then $m = dm' = 2d$ with $(d, 2) = 1$, so d is odd. Thus $\mathbf{Q}(\mu_m) = \mathbf{Q}(\mu_{2d}) = \mathbf{Q}(\mu_d)$.

Going through the previous paragraph with n in place of m, we get $\mathbf{Q}(\mu_n) = \mathbf{Q}(\mu_d)$ too, so $\mathbf{Q}(\mu_m) = \mathbf{Q}(\mu_n)$. □

Theorem 4.4. *If $E/\mathbf{Q}$ is a finite extension that contains no proper abelian extensions of $\mathbf{Q}$ then $\mathrm{Gal}(E(\mu_n)/E) \cong (\mathbf{Z}/(n))^\times$ for all $n \geq 1$, or equivalently $[E(\mu_n) : E] = \varphi(n)$.*

Proof. From Galois theory, for finite extensions L/K and F/K, $[LF : F] = [L : L \cap F]$ if L/K is Galois. Therefore $[E(\mu_n) : E] = [\mathbf{Q}(\mu_n)E : E] = [\mathbf{Q}(\mu_n) : \mathbf{Q}(\mu_n) \cap E]$. The intersection $\mathbf{Q}(\mu_n) \cap E$ is an abelian extension of $\mathbf{Q}$ since every subfield of $\mathbf{Q}(\mu_n)$ is abelian over $\mathbf{Q}$. Therefore by hypothesis $\mathbf{Q}(\mu_n) \cap E = \mathbf{Q}$, so $[E(\mu_n) : E] = [\mathbf{Q}(\mu_n) : \mathbf{Q}] = \varphi(n)$. □

Example 4.5. An extension $E/\mathbf{Q}$ of prime degree that is not Galois, such as $\mathbf{Q}(\sqrt[p]{2})/\mathbf{Q}$ for prime $p \geq 3$, satisfies the hypothesis of Theorem 4.4 so $\mathrm{Gal}(E(\mu_n)/E) \cong (\mathbf{Z}/(n))^\times$ for $n \geq 1$.

5. Cyclotomic polynomials

In the complex numbers, all primitive nth roots of unity are $\mathbf{Q}$-conjugate and therefore have a common minimal polynomial in $\mathbf{Q}[T]$. It is called the nth *cyclotomic polynomial* and is denoted $\Phi_n(T)$. Explicitly,

$$\Phi_n(T) = \prod_{\substack{1 \leq k \leq n \\ (k,n)=1}} (T - e^{2\pi ik/n}).$$

Here are the first 10 cyclotomic polynomials.

$$\begin{array}{ll} \Phi_1(T) = T - 1 & \Phi_2(T) = T + 1 \\ \Phi_3(T) = T^2 + T + 1 & \Phi_4(T) = T^2 + 1 \\ \Phi_5(T) = T^4 + T^3 + T^2 + T + 1 & \Phi_6(T) = T^2 - T + 1 \\ \Phi_7(T) = T^6 + T^5 + T^4 + T^3 + T^2 + T + 1 & \Phi_8(T) = T^4 + 1 \\ \Phi_9(T) = T^6 + T^3 + 1 & \Phi_{10}(T) = T^4 - T^3 + T^2 - T + 1 \end{array}$$

For all $n \geq 1$, $\Phi_n(T)$ is monic of degree $\varphi(n)$ in $\mathbf{Q}[T]$ and $\Phi_n(T)$ is *irreducible* over $\mathbf{Q}$. Here are some identities involving these polynomials, where p is a prime:

(1) $T^n - 1 = \prod_{d|n} \Phi_d(T)$,
(2) $\Phi_n(T) = T^{\varphi(n)}\Phi_n(1/T)$ for $n \geq 2$,
(3) $\Phi_p(T) = T^{p-1} + T^{p-2} + \cdots + T + 1$,
(4) $\Phi_{p^r}(T) = \Phi_p(T^{p^{r-1}})[= T^{(p-1)p^{r-1}} + T^{(p-2)p^{r-1}} + \cdots + T^{p^{r-1}} + 1]$,
(5) $\Phi_{2n}(T) = \Phi_n(-T)$ for odd n,
(6) $\Phi_{p_1^{r_1} \cdots p_k^{r_k}}(T) = \Phi_{p_1 \cdots p_k}(T^{p_1^{r_1-1} \cdots p_k^{r_k-1}})$, and more generally $\Phi_{mn}(T) = \Phi_n(T^m)$ if all prime factors of m divide n,
(7) if $(p, m) = 1$ then $\Phi_{p^r m}(T) = \Phi_m(T^{p^r})/\Phi_m(T^{p^{r-1}})$.

Except for the first formula, these identities can be checked by showing the right side has the correct degree and one correct root to be the cyclotomic polynomial on the left side. (Two monic irreducible polynomials with a common root are equal.) The first identity can be regarded as a recursive definition of the cyclotomic polynomials.

Example 5.1. Since $\Phi_2(T) = T+1$, we have $\Phi_8(T) = \Phi_2(T^4) = T^4+1$. Since $\Phi_3(T) = T^2+T+1$, $\Phi_6(T) = \Phi_3(-T) = T^2-T+1$ and $\Phi_{24}(T) = \Phi_6(T^4) = \Phi_3(-T^4) = T^{12}-T^4+1$.

Theorem 5.2. *For all $n \geq 1$, $\Phi_n(T) \in \mathbf{Z}[T]$.*

Proof. Let ζ_n be a primitive nth root of unity in $\mathbf{C}$, so $\Phi_n(T)$ is its minimal polynomial over $\mathbf{Q}$. Since T^n-1 vanishes at ζ_n, $\Phi_n(T)$ divides T^n-1 in $\mathbf{Q}[T]$. Since $\Phi_n(T)$ and T^n-1 are both monic and $T^n-1 \in \mathbf{Z}[T]$, by Gauss' lemma $\Phi_n(T) \in \mathbf{Z}[T]$. □

The sequence of cyclotomic polynomials is an interesting example where initial data can be misleading: the coefficients of the first 100 cyclotomic polynomials are all 0 or ± 1, but this pattern is not true in general! For instance, $\Phi_{105}(T)$ has a coefficient -2 for T^{41} and T^7 (the other coefficients are 0 and ± 1). Why does it take so long for a coefficient besides 0 and ± 1 to occur? Well, the fifth and sixth cyclotomic polynomial identities above show the nonzero coefficients of all cyclotomic polynomials are determined by the coefficients of the $\Phi_n(T)$ where n is a product of distinct odd primes. The polynomial $\Phi_p(T)$ only has coefficient 1 and it can be shown [4] that $\Phi_{pq}(T)$ only has coefficients 0 and ± 1 when p and q are different primes. Therefore each n with at most 2 odd prime factors only has coefficients among 0 and ± 1. The first positive integer that does not have at most 2 odd prime factors is $3 \cdot 5 \cdot 7 = 105 > 100$, which shows $\Phi_{105}(T)$ is the first cyclotomic polynomial that even has a chance to have a coefficient other than 0 and ± 1. By a theorem of Schur, if n has t odd prime factors then $\Phi_n(T)$ has coefficient $-(t-1)$ (thus predicting the coefficient of -2 in $\Phi_{105}(T)$). To produce large coefficients in $\Phi_n(T)$ we should give n a lot of different odd prime factors and numbers below 100 have at most 2 odd prime factors.

Cyclotomic polynomials for prime-power n, say $n = p^r$, can be written down concretely:

$$\Phi_{p^r}(T) = \frac{T^{p^r}-1}{T^{p^{r-1}}-1} = \sum_{k=0}^{p-1} T^{p^{r-1}k}.$$

Theorem 5.3. *The polynomial $\Phi_{p^r}(T+1)$ is Eisenstein with respect to p.*

Proof. The constant term of $\Phi_{p^r}(T+1)$ is

$$\Phi_{p^r}(1) = \sum_{k=0}^{p-1} 1^{p^{r-1}k} = p,$$

which is divisible by p just once. To show the non-leading coefficients are all multiples of p, we reduce the coefficients mod p. Since, in $\mathbf{F}_p[T]$, $T^{p^r}-1 = (T-1)^{p^r}$ and $T^{p^{r-1}}-1 = (T-1)^{p^{r-1}}$, we have (reducing coefficients mod p)

$$\overline{\Phi}_{p^r}(T) = \frac{T^{p^r}-1}{T^{p^{r-1}}-1} = (T-1)^{p^r-p^{r-1}} \text{ in } \mathbf{F}_p[T],$$

so

$$\overline{\Phi}_{p^r}(T+1) = T^{p^r-p^{r-1}} \text{ in } \mathbf{F}_p[T].$$

The degree of $\Phi_{p^r}(T+1)$ is p^r-p^{r-1}, so all its non-leading coefficients are 0 in $\mathbf{F}_p$, which means the coefficients as integers are multiples of p. □

Using the Eisenstein irreducibility criterion, $\Phi_{p^r}(T+1)$ is irreducible in $\mathbf{Q}[T]$, so $\Phi_{p^r}(T)$ is irreducible in $\mathbf{Q}[T]$. Therefore $[\mathbf{Q}(\mu_{p^r}) : \mathbf{Q}] = p^r - p^{r-1} = |(\mathbf{Z}/(p^r))^\times|$, so the embedding $\mathrm{Gal}(\mathbf{Q}(\mu_{p^r})/\mathbf{Q}) \hookrightarrow (\mathbf{Z}/(p^r))^\times$ is an isomorphism. This is an alternate proof of Theorem 2.5 when n is a prime power that is simpler than the proof we gave before.

Cyclotomic polynomials can be used to prove some results that don't appear to be about roots of unity in the first place. One such result is an elementary proof that for each $n > 1$ there are infinitely many primes $p \equiv 1 \bmod n$ [6, Cor. 2.11]. A second result is a proof of Wedderburn's theorem that all finite division rings are commutative [3, Thm. 13.1].

Since cyclotomic polynomials are in $\mathbf{Z}[T]$, let's reduce them modulo p and ask how they factor. It suffices to look at $\overline{\Phi}_n(T) = \Phi_n(T) \bmod p$ when $(p, n) = 1$ since reducing the seventh algebraic identity for cyclotomic polynomials at the start of this section gives us

$$\Phi_{p^r m}(T) = \Phi_m(T)^{p^r - p^{r-1}} \bmod p \tag{5.1}$$

in $\mathbf{F}_p[T]$ when $(p, m) = 1$.

Theorem 5.4. *When the prime p does not divide n, the monic irreducible factors of $\overline{\Phi}_n(T) \in \mathbf{F}_p[T]$ are distinct and each has degree equal to the order of $p \bmod n$.*

Proof. Since $\Phi_n(T) \mid (T^n - 1)$ in $\mathbf{Z}[T]$, this divisibility relation is preserved when reducing modulo p, so $\overline{\Phi}_n(T)$ is separable in $\mathbf{F}_p[T]$ because $T^n - \overline{1}$ is separable in $\mathbf{F}_p[T]$. (Here we need $(p, n) = 1$.)

Let α be a root of $\overline{\Phi}_n(T)$ in an extension of $\mathbf{F}_p$. We will show that α inherits the expected algebraic property of being a primitive nth root of unity. Since $\overline{\Phi}_n(T) \mid T^n - \overline{1}$, from $\overline{\Phi}_n(\alpha) = 0$ we have $\alpha^n = 1$. If α were not of order n then it has some order m that properly divides n. Then α is a root of $T^m - \overline{1} = \prod_{d|m} \overline{\Phi}_d(T)$, so $\overline{\Phi}_d(\alpha) = 0$ for some d properly dividing n. Since $d \mid n$, $T^n - \overline{1}$ is divisible by $\overline{\Phi}_n(T)\overline{\Phi}_d(T)$, so α is a double root of $T^n - \overline{1}$, but $T^n - \overline{1}$ has no repeated roots. Therefore we have a contradiction, so α is a primitive nth root of unity.

Let $\pi(T)$ be an irreducible factor of $\overline{\Phi}_n(T)$ in $\mathbf{F}_p[T]$ and let α denote a root of $\pi(T)$. Then α is a primitive nth root of unity, so $\deg \pi = [\mathbf{F}_p(\alpha) : \mathbf{F}_p]$ is the order of $p \bmod n$ by Theorem 2.8. □

Example 5.5. The polynomial $\Phi_5(T) = T^4 + T^3 + T^2 + T + 1$ factors over $\mathbf{F}_p$ into irreducible factors whose degrees equal the order of $p \bmod 5$. For example, $T^4 + T^3 + T^2 + T + 1$ is irreducible in $\mathbf{F}_3[T]$ since $3 \bmod 5$ has order 4, while

$$T^4 + T^3 + T^2 + T + 1 = (T - 3)(T - 4)(T - 5)(T - 9)$$

in $\mathbf{F}_{11}[T]$ with irreducible factors of degree 1 since $11 \bmod 5$ has order 1, and

$$T^4 + T^3 + T^2 + T + 1 = (T^2 + 5T + 1)(T^2 + 15T + 1)$$

in $\mathbf{F}_{19}[T]$ with irreducible factors of degree 2 since $19 \bmod 5$ has order 2.

Example 5.6. The polynomial $\Phi_7(T) = T^6 + T^5 + T^4 + T^3 + T^2 + T + 1$ factors over $\mathbf{F}_p$ into irreducible factors whose degrees equal the order of $p \bmod 7$. For example, $2 \bmod 7$ has order 3 so $\Phi_7(T)$ factors over $\mathbf{F}_2$ into a product of irreducible cubics:

$$T^6 + T^5 + T^4 + T^3 + T^2 + T + 1 = (T^3 + T + 1)(T^3 + T^2 + 1)$$

in $\mathbf{F}_2[T]$. This explains what happened in Example 2.9: if ζ is a primitive 7th root of unity in characteristic 2, then it and ζ^3 are roots of the two different cubics on the right side: one has roots ζ, ζ^2, and ζ^4, while the other has roots ζ^3, $(\zeta^3)^2 = \zeta^6$, and $(\zeta^3)^4 = \zeta^5$.

Corollary 5.7. *The reduction $\overline{\Phi}_n(T)$ is irreducible in $\mathbf{F}_p[T]$ if and only if $(p, n) = 1$ and $p \bmod n$ is a generator of $(\mathbf{Z}/(n))^\times$.*

Proof. If $\overline{\Phi}_n(T)$ is irreducible in $\mathbf{F}_p[T]$ then $(p,n) = 1$ by (5.1), so Theorem 5.4 tells us the order of $p \bmod n$ is $\varphi(n)$: $p \bmod n$ generates $(\mathbf{Z}/(n))^\times$. Conversely, if $(p,n) = 1$ and $p \bmod n$ is a generator of $(\mathbf{Z}/(n))^\times$ then Theorem 5.4 tells us the irreducible factors of $\overline{\Phi}_n(T)$ in $\mathbf{F}_p[T]$ have degree $\varphi(n) = \deg(\overline{\Phi}_n(T))$, so $\overline{\Phi}_n(T)$ is irreducible. □

Thus many cyclotomic polynomials are examples of irreducible polynomials in $\mathbf{Z}[T]$ that factor modulo *every* prime: if $(\mathbf{Z}/(n))^\times$ is not a cyclic group then there is no generator for $(\mathbf{Z}/(n))^\times$, so Corollary 5.7 says there is no prime p such that $\Phi_n(T) \bmod p$ is irreducible. In other words, $\Phi_n(T) \bmod p$ factors for all primes p.

Example 5.8. The least n such that $(\mathbf{Z}/(n))^\times$ is non-cyclic is $n = 8$, and $\Phi_8(T) = T^4 + 1$. This polynomial is reducible mod p for all p. Here are some factorizations of $T^4 + 1 \bmod p$.

$$\begin{aligned}
\Phi_8(T) &\equiv (T+1)^4 \bmod 2,\\
\Phi_8(T) &\equiv (T^2+T+2)(T^2+2T+2) \bmod 3,\\
\Phi_8(T) &\equiv (T^2+2)(T^2+3) \bmod 5,\\
\Phi_8(T) &\equiv (T^2+3T+1)(T^2+4T+1) \bmod 7,\\
\Phi_8(T) &\equiv (T^2+3T+10)(T^2+8T+10) \bmod 11,\\
\Phi_8(T) &\equiv (T-2)(T-8)(T-9)(T-15) \bmod 17,\\
\Phi_8(T) &\equiv (T^2+6T+18)(T^2+13T+18) \bmod 19,\\
\Phi_8(T) &\equiv (T^2+5T+1)(T^2+18T+1) \bmod 23,\\
\Phi_8(T) &\equiv (T^2+12)(T^2+17) \bmod 29,\\
\Phi_8(T) &\equiv (T^2+8T+1)(T^2+23T+1) \bmod 31.
\end{aligned}$$

As an elementary application of cyclotomic polynomials, we will consider a generalization of *Mersenne primes*, which are prime numbers of the form $2^n - 1$. A necessary condition that $2^n - 1$ is prime is that n is prime: if n is composite with $n = rs$ where $r \geq 2$ and $s \geq 2$, then

$$2^n - 1 = (2^r)^s - 1 = (2^r - 1)(2^{r(s-1)} + 2^{r(s-2)} + \cdots + 2^r + 1) \tag{5.2}$$

by setting $T = 2^r$ in the identity $T^s - 1 = (T-1)(T^{s-1} + T^{s-2} + \cdots + T + 1)$. Both factors on the right in (5.2) are greater than 1, so $2^n - 1$ is composite. While $2^n - 1$ being prime implies n is prime, the converse is false, *e.g.*, $2^{11} - 1$ is composite.

In place of $2^n - 1$, could $a^n - 1$ be prime when $a \geq 3$ and $n \geq 2$? The answer is no since

$$a^n - 1 = (a-1)(a^{n-1} + \cdots + a + 1)$$

and both factors on the right are greater than 1 when $a \geq 3$. However, since $(a-1) \mid (a^n - 1)$ for all n, we should divide $a^n - 1$ by its automatic factor $a - 1$ and ask if what remains might be prime. This leads to the following generalization of the case $a = 2$ (when $a - 1 = 1$).

Theorem 5.9. *Let $a \geq 2$ in $\mathbf{Z}$. For a positive integer n, if $(a^n - 1)/(a-1)$ is prime then n is prime.*

Proof. When $n = 1$, $(a^n - 1)/(a-1) = 1$ and this is not prime. For $n \geq 2$, we will show that $(a^n - 1)/(a-1)$ being prime implies n is prime by first showing n is a prime power and then refining that to n being prime.

The identity $T^n - 1 = \prod_{d|n} \Phi_d(T)$ at $T = a$ implies

$$a^n - 1 = \prod_{d|n} \Phi_d(a) = (a-1) \prod_{\substack{d|n \\ d>1}} \Phi_d(a) \Longrightarrow \frac{a^n - 1}{a-1} = \prod_{\substack{d|n \\ d>1}} \Phi_d(a). \tag{5.3}$$

Among the integers $\Phi_d(a)$ where $d \mid n$ and $d > 1$ we have $\Phi_p(a)$ when p is a prime factor of n, and $\Phi_p(a) = a^{p-1} + \cdots + a + 1 \geq a + 1 > 1$. Thus when n has at least two distinct prime factors, the integers in the product on the right in (5.3) include at least two integers greater than 1, so $(a^n - 1)/(a - 1)$ is composite. Thus if $(a^n - 1)/(a - 1)$ is prime, then n doesn't have more than one prime factor, so n is a prime power: $n = p^r$ where p is prime and $r \geq 1$, so (5.3) becomes

$$\frac{a^n - 1}{a-1} = \frac{a^{p^r} - 1}{a-1} = \prod_{i=1}^{r} \Phi_{p^i}(a) = \Phi_p(a) \cdots \Phi_{p^r}(a).$$

When $i \geq 1$, $\Phi_{p^i}(a) = a^{p^{i-1}(p-1)} + a^{p^{i-1}(p-2)} + \cdots + a^{p^{i-1}} + 1 \geq a^{p^{i-1}} + 1 \geq a + 1 > 1$, so all factors in $\prod_{i=1}^{r} \Phi_{p^i}(a)$ exceed 1 and thus this product is composite when $r \geq 2$. Therefore if $(a^n - 1)/(a - 1)$ is prime we need $r = 1$, so $n = p^r = p$ is a prime number. □

The next theorem gives a restriction on a if $(a^p - 1)/(a - 1)$ is going to be prime for infinitely many prime exponents p and it is also proved with cyclotomic polynomials.

Theorem 5.10. *Let $a \geq 2$ in $\mathbf{Z}$. If a is a k-th power where $k \geq 2$, then $(a^p - 1)/(a - 1)$ is composite at all primes $p \nmid k$. Therefore if $(a^p - 1)/(a - 1)$ is prime for infinitely many prime exponents p, a can't be a kth power where $k \geq 2$.*

Proof. Write $a = b^k$ with $b \geq 2$. Then

$$\frac{a^p - 1}{a-1} = \frac{b^{kp} - 1}{b^k - 1} = \prod_{d|k} \frac{\Phi_d(b^p)}{\Phi_d(b)}. \tag{5.4}$$

We will show when $d \mid k$ that each ratio $\Phi_d(b^p)/\Phi_d(b)$ is an integer when $p \nmid k$. What we will actually do is show $\Phi_d(T) \mid \Phi_d(T^p)$ in $\mathbf{Z}[T]$ when $p \nmid d$ (and that includes the case when $d \mid k$ and $p \nmid k$), so setting $T = b$ gives us $\Phi_d(b) \mid \Phi_d(b^p)$ in $\mathbf{Z}$.

Let ζ be a primitive dth root of unity. Then ζ^p also is a primitive dth root of unity, as $(p, d) = 1$, so $\Phi_d(\zeta^p) = 0$. Thus $\Phi_d(T^p)$ has ζ as a root. Since $\Phi_d(T)$ is the minimal polynomial of ζ over $\mathbf{Q}$, $\Phi_d(\zeta^p) = 0 \Rightarrow \Phi_d(T) \mid \Phi_d(T^p)$ in $\mathbf{Q}[T]$, so $\Phi_d(T^p) = \Phi_d(T)A(T)$ with $A(T) \in \mathbf{Q}[T]$. What we want is $\Phi_d(T) \mid \Phi_d(T^p)$ in $\mathbf{Z}[T]$. Since $\Phi_d(T)$ is monic in $\mathbf{Z}[T]$, by the division algorithm for monic polynomials we can write $\Phi_d(T^p) = \Phi_d(T)Q(T) + R(T)$ where $Q(T)$ and $R(T)$ are in $\mathbf{Z}[T]$ and $R(T) = 0$ or $\deg R < \deg(\Phi_d)$. By the uniqueness of quotient and remainder for the division algorithm in $\mathbf{Q}[T]$, the two equations $\Phi_d(T^p) = \Phi_d(T)A(T)$ and $\Phi_d(T^p) = \Phi_d(T)Q(T) + R(T)$ imply $A(T) = Q(T) \in \mathbf{Z}[T]$ (and $R(T) = 0$), so $\Phi_d(T) \mid \Phi_d(T^p)$ in $\mathbf{Z}[T]$.

To show (5.4) implies $(a^p - 1)/(a - 1)$ is composite, we will show two of the factors $\Phi_d(b^p)/\Phi_d(b)$ in (5.4) are greater than 1. We will use the factors at $d = 1$ and $d = q$, where q is a prime factor of k. At $d = 1$,

$$\frac{\Phi_1(b^p)}{\Phi_1(b)} = \frac{b^p - 1}{b - 1} = b^{p-1} + \cdots + b + 1 \geq b + 1 > 1.$$

At $d = q$, to show $\Phi_q(b^p)/\Phi_q(b) > 1$ we look at the numerator and denominator:

$$\Phi_q(b^p) = \sum_{i=0}^{q-1} b^{pi}, \quad \Phi_q(b) = \sum_{i=0}^{q-1} b^i.$$

The terms in the sums at $i = 0$ are both 1. When $1 \leq i \leq q-1$, $b^{pi} > b^i$. Thus $\Phi_q(b^p) > \Phi_q(b)$, so $\Phi_q(b^p)/\Phi_q(b) > 1$. □

The proof shows when $a = b^k$ that the only p where $(a^p - 1)/(a - 1)$ might be prime are p dividing k, and such prime values can occur: $(4^p - 1)/(4 - 1)$ is prime at $p = 2$ and $(8^p - 1)/(8 - 1)$ is prime at $p = 3$.

By the previous two theorems, if $a \geq 2$ in $\mathbf{Z}$ and $(a^n - 1)/(a - 1)$ is prime for infinitely many positive integers n, then all such n are prime and a is not a kth power for $k \geq 2$. It is believed that the converse holds: when $a \geq 2$ is not a kth power for $k \geq 2$, such as $a = 6$ and $a = 12$ but not $a = 9$, we expect that $(a^p - 1)/(a - 1)$ is prime for infinitely many prime exponents p, but there is no a for which this is proved. Some numerical data is in the table below when $2 \leq a \leq 10$, with suitable OEIS links.[4]

a	p making $\frac{a^p-1}{a-1}$ prime	OEIS link
2	$2, 3, 5, 7, 13, 17, 19, \ldots$	https://oeis.org/A000043
3	$3, 7, 13, 71, 103, 541, 1091, \ldots$	https://oeis.org/A028491
5	$3, 7, 11, 13, 47, 127, 149, \ldots$	https://oeis.org/A004061
6	$2, 3, 7, 29, 71, 127, 271, \ldots$	https://oeis.org/A004062
7	$5, 13, 131, 149, 1699, 14221, 35201, \ldots$	https://oeis.org/A004063
10	$2, 19, 23, 317, 1031, 49081, 86453, \ldots$	https://oeis.org/A004023

References

[1] Z. I. Borevich and I. R. Shafarevich, "Number Theory," Academic Press, New York, 1966.

[2] R. Dedekind, Beweis für die Irreductibilität der Kreisteilungs-Gleichungen, *J. Reine Angew. Math.* **54** (1857), 27–30.

[3] T. Y. Lam, "A First Course in Noncommutative Rings," Springer-Verlag, New York, 1991.

[4] T. Y. Lam and K. H. Cheung, On the cyclotomic polynomial $\Phi_{pq}(T)$, *Amer. Math. Monthly* **103** (1996), 562–564.

[5] E. Landau, Über die Irreduzibilität der Kreisteilungsgleichung, *Math. Zeitschrift* **29** (1929), 462.

[6] L. Washington, "Introduction to Cyclotomic Fields," 2nd ed., Springer-Verlag, New York, 1997.

[4]The OEIS pages indicate in the Extensions row that some entries are only probable primes, such as $(7^p - 1)/6$ at $p = 35201$.