LINEAR INDEPENDENCE OF CHARACTERS

KEITH CONRAD

1. Introduction

For an abelian group G, finite or infinite, a *character* of G is a group homomorphism $\chi\colon G \to F^\times$ where F is a field. Historically, the first examples of characters occurred with $F = \mathbf{C}$ (and the concept was introduced in this generality by Dedekind), but allowing the multiplicative groups of other fields as target values is useful.

Example 1.1. A field homomorphism $K \to F$ is a character by restricting it to the nonzero elements of K (that is, using $G = K^\times$) and ignoring the additive aspect of a field homomorphism. In particular, when L/K is a field extension every element of $\mathrm{Aut}(L/K)$ is a field homomorphism $L \to L$ and therefore is a character of $L^\times$ with values in $L^\times$.

Example 1.2. For all $\alpha \in F^\times$, the map $\mathbf{Z} \to F^\times$ by $k \mapsto \alpha^k$ is a character of $\mathbf{Z}$.

Characters $G \to F^\times$ can be regarded as special functions $G \to F$ and then can be added, but the sum is no longer a character (since the sum of multiplicative maps is usually not multiplicative, and could even take the value 0). The sum is just a function $G \to F$. The functions $G \to F$ form a vector space under addition and F-scaling. We will prove that different characters $G \to F^\times$ are linearly independent as functions $G \to F$. Then we turn to three very important applications of this linear independence:

- The normal basis theorem.
- Hilbert's Theorem 90 for cyclic Galois extensions.
- Some basic ideas in Kummer theory and Artin-Schreier theory.

While we will use Galois theory to prove results about characters, in [3] and [8] linear independence of characters is used to prove the Galois correspondence. That approach to Galois theory is due to Artin [1], who I think wanted to avoid the primitive element theorem.

2. Linear independence

Theorem 2.1. *Let $\chi_1, \dots, \chi_n$ be distinct characters $G \to F^\times$. They are linearly independent: if $c_1, \dots, c_n \in F$ satisfy*

$$c_1\chi_1(g) + \cdots + c_n\chi_n(g) = 0$$

for all $g \in G$ then every c_i is 0.

Example 2.2. For distinct $\alpha_1, \dots, \alpha_n \in F^\times$, the only F-linear relation $c_1\alpha_1^k + \cdots + c_n\alpha_n^k = 0$ for all $k \in \mathbf{Z}$ is the one where every c_i is 0. This is a case of Theorem 2.1 using the characters $\chi_i\colon \mathbf{Z} \to F^\times$ such that $\chi_i(k) = \alpha_i^k$.

Proof. We argue by induction on n. The case $n = 1$ is trivial. Suppose $n \geqslant 2$ and we know every set of $n-1$ distinct characters of G is linearly independent. Assume now that $\chi_1, \dots, \chi_n$ are distinct characters of G and we have an identity

$$c_1\chi_1(g) + \cdots + c_{n-1}\chi_{n-1}(g) + c_n\chi_n(g) = 0 \tag{2.1}$$

for some coefficients c_i in F and every $g \in G$. We want to show all c_i are 0.

Since $\chi_1 \neq \chi_n$, for some $g_0 \in G$ we have $\chi_1(g_0) \neq \chi_n(g_0)$. Since (2.1) is true for all g, it is also true with $g_0 g$ in place of g:

$$c_1\chi_1(g_0)\chi_1(g) + \cdots + c_{n-1}\chi_{n-1}(g_0)\chi_{n-1}(g) + c_n\chi_n(g_0)\chi_n(g) = 0 \tag{2.2}$$

for all $g \in G$. Now multiply (2.1) by $\chi_n(g_0)$:

$$c_1\chi_n(g_0)\chi_1(g) + \cdots + c_{n-1}\chi_n(g_0)\chi_{n-1}(g) + c_n\chi_n(g_0)\chi_n(g) = 0 \tag{2.3}$$

for all $g \in G$. Subtracting (2.3) from (2.2), the last terms cancel:

$$c_1(\chi_1(g_0) - \chi_n(g_0))\chi_1(g) + \cdots + c_{n-1}(\chi_{n-1}(g_0) - \chi_n(g_0))\chi_{n-1}(g) = 0$$

for all $g \in G$. This is a linear dependence relation among the functions $\chi_1, \ldots, \chi_{n-1}$, so by induction all the coefficients $c_i(\chi_i(g_0) - \chi_n(g_0))$ are 0. In particular, $c_1(\chi_1(g_0) - \chi_n(g_0)) = 0$. Since $\chi_1(g_0) \neq \chi_n(g_0)$ we must have $c_1 = 0$. By arguing in a similar way using $\chi_2, \ldots, \chi_{n-1}$ in place of χ_1, we obtain $c_i = 0$ for $i = 1, \ldots, n-1$. Therefore (2.1) becomes $c_n\chi_n(g) = 0$ for all g, so $c_n = 0$ since χ_n has nonzero values. ∎

In the proof we did not actually use inversion in the group, so it applies to homomorphisms from semi-groups like $\mathbf{N}$ to $F^\times$.

Example 2.3. For distinct $\alpha_1, \ldots, \alpha_n \in F^\times$ then the only F-linear relation $c_1\alpha_1^k + \cdots + c_n\alpha_n^k = 0$ for all $k \in \mathbf{N}$ is the one where every c_i is 0.

3. The normal basis theorem

Let L/K be a Galois extension of degree n and $\mathrm{Gal}(L/K) = \{\sigma_1, \ldots, \sigma_n\}$. A *normal basis* for L/K is a basis consisting of a set of K-conjugate elements $\{\sigma_1(\gamma), \ldots, \sigma_n(\gamma)\}$.

Example 3.1. The usual basis $\{1, i\}$ of $\mathbf{C}/\mathbf{R}$ is not a normal basis since the terms are not $\mathbf{R}$-conjugate, while $\{i, -i\}$ is a set of $\mathbf{R}$-conjugate elements that is not a normal basis since the terms are not linearly independent over $\mathbf{R}$. The set $\{1+i, 1-i\}$ is a normal basis: its elements are $\mathbf{R}$-conjugate and they are linearly independent over $\mathbf{R}$.

Example 3.2. For a prime p, the basis $\{1, \zeta_p, \ldots, \zeta_p^{p-2}\}$ of $\mathbf{Q}(\zeta_p)/\mathbf{Q}$ is not a normal basis (1 is not $\mathbf{Q}$-conjugate to the rest of the basis), but $\{\zeta_p, \zeta_p^2, \ldots, \zeta_p^{p-1}\}$ is a normal basis.[1]

Example 3.3. For the extension $\mathbf{Q}(\sqrt{2}+\sqrt{3})/\mathbf{Q}$, the four $\mathbf{Q}$-conjugates of $\sqrt{2}+\sqrt{3}$ are not a normal basis since they add up to 0: they are linearly dependent over $\mathbf{Q}$. If $\alpha = 1+\sqrt{2}+\sqrt{3}+\sqrt{6}$ then its $\mathbf{Q}$-conjugates are

$$1+\sqrt{2}+\sqrt{3}+\sqrt{6}, \quad 1-\sqrt{2}+\sqrt{3}-\sqrt{6}, \quad 1+\sqrt{2}-\sqrt{3}-\sqrt{6}, \quad 1-\sqrt{2}-\sqrt{3}+\sqrt{6},$$

which is linearly independent over $\mathbf{Q}$ since the matrix of coefficients

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}$$

has determinant $16 \neq 0$. So the $\mathbf{Q}$-conjugates of α form a normal basis.

[1] In general, the primitive nth roots of unity in the nth cyclotomic field form a normal basis over $\mathbf{Q}$ if and only if n is squarefree. See http://math.stackexchange.com/questions/87290/basis-of-primitive-nth-roots-in-a-cyclotomic-extension for a proof.

The normal basis theorem says that every finite Galois extension admits a normal basis. We will give a proof of this theorem when K is infinite following a method of Waterhouse [9]. Then we will give a proof of the normal basis theorem when L/K is a cyclic extension, which covers the case that K is finite. The proof of each case will use linear independence of characters. (It would be nice to have a proof of the normal basis theorem that treats finite and infinite fields in a uniform manner without making the details overly complicated.)

Lemma 3.4. *Let L/K be Galois, $\mathrm{Gal}(L/K) = \{\sigma_1, \ldots, \sigma_n\}$ and $e_1, \ldots, e_n$ be a K-basis of L. Then the n-tuples*

$$(\sigma_1(e_i), \sigma_2(e_i), \ldots, \sigma_n(e_i))$$

for $1 \leqslant i \leqslant n$ are an L-basis of L^n.

Example 3.5. An $\mathbf{R}$-basis of $\mathbf{C}$ is $\{1, i\}$ and $\{(1,1), (i,-i)\}$ is a $\mathbf{C}$-basis of $\mathbf{C}^2$.

Proof. Let W be the L-span of these n-tuples in L^n. We want to show $W = L^n$. If W is a proper subspace of L^n then there is a nonzero element of the L-dual space $(L^n)^\vee$ that vanishes on W. We will show, however, that each $\varphi \in (L^n)^\vee$ that is zero on W must be zero on all of L^n, so $W = L^n$.

Each $\varphi \in (L^n)^\vee$ can be viewed as the dot product with some n-tuple, say $\varphi(v) = (c_1, \ldots, c_n) \cdot v$ for all $v \in L^n$. If φ vanishes on W then

$$(c_1, \ldots, c_n) \cdot (\sigma_1(e_i), \ldots, \sigma_n(e_i)) = 0$$

for all i. Thus $\sum_{j=1}^n c_j \sigma_j(e_i) = 0$ for all i, so by taking K-linear combinations we get $\sum_{j=1}^n c_j \sigma_j(x) = 0$ for all $x \in L$. By linear independence of characters ($G = L^\times$, $F = L$) each c_j is 0, so φ is 0. ∎

Theorem 3.6. *Every finite Galois extension of an infinite field has a normal basis.*

Proof. Let L/K be the extension. Write $n = [L : K]$ and $\mathrm{Gal}(L/K) = \{\sigma_1, \ldots, \sigma_n\}$. We are seeking an $x \in L$ such that $\sigma_1(x), \ldots, \sigma_n(x)$ are linearly independent over K. Let's explore what a K-linear dependence relation on this list would look like for a fixed x in L:

$$\sum_{j=1}^n a_j \sigma_j(x) = 0$$

for some $a_1, \ldots, a_n \in K$. Applying σ_i^{-1} (arbitrary i) to both sides gives new linear relations

$$\sum_{j=1}^n a_j (\sigma_i^{-1}\sigma_j)(x) = 0.$$

Collecting these together for $i = 1, \ldots, n$ tells us that

$$\begin{pmatrix} \sigma_1^{-1}\sigma_1(x) & \cdots & \sigma_1^{-1}\sigma_n(x) \\ \vdots & \ddots & \vdots \\ \sigma_n^{-1}\sigma_1(x) & \cdots & \sigma_n^{-1}\sigma_n(x) \end{pmatrix} \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{3.1}$$

The vector on the left side is in K^n. We want an $x \in L$ such that the only solution to (3.1) is $a_1 = 0, \ldots, a_n = 0$, which can be achieved by finding an $x \in L$ such that the matrix $((\sigma_i^{-1}\sigma_j)(x))$ is invertible, *i.e.*, $\det((\sigma_i^{-1}\sigma_j)(x)) \neq 0$.

Pick a K-basis $e_1, \ldots, e_n$ of L and write an arbitrary $x \in L$ in the form $\sum_{k=1}^n b_k e_k$ with $b_k \in K$. Then

$$(\sigma_i^{-1}\sigma_j)(x) = \sum_{k=1}^n b_k((\sigma_i^{-1}\sigma_j)(e_k)).$$

Consider the polynomial

$$\Delta(X_1, \ldots, X_n) = \det\left(\sum_{k=1}^n((\sigma_i^{-1}\sigma_j)(e_k))X_k\right) \in L[X_1, \ldots, X_n].$$

Let σ_1 be the identity automorphism of L. By Lemma 3.4 (here is where linear independence of characters is used), there are $c_1, \ldots, c_n \in L$ such that

$$\sum_{k=1}^n c_k(\sigma_1(e_k), \ldots, \sigma_n(e_k)) = (1, 0, \ldots, 0).$$

Reading this off componentwise,

$$\sum_{k=1}^n c_k e_k = 1, \quad \sum_{k=1}^n c_k\sigma(e_k) = 0 \text{ for } \sigma \neq \mathrm{id}_L\,.$$

Thus

$$\sum_{k=1}^n c_k(\sigma_i^{-1}\sigma_j)(e_k) = \begin{cases} 1, & \text{if } \sigma_i = \sigma_j, \\ 0, & \text{otherwise.} \end{cases}$$

Therefore the matrix $(\sum_{k=1}^n c_k(\sigma_i^{-1}\sigma_j)(e_k))$ is I_n, so $\Delta(c_1, \ldots, c_n) = 1$. This shows the polynomial $\Delta(X_1, \ldots, X_n)$ is not the zero polynomial, although the c_i's we are using here come from L, not K.

Since K is *infinite*, there must be $b_1, \ldots, b_n \in K$ (not just in L, like $c_1, \ldots, c_n$) such that $\Delta(b_1, \ldots, b_n) \neq 0$. Then

$$\begin{aligned} 0 &\neq \Delta(b_1, \ldots, b_n) \\ &= \det\left(\sum_{k=1}^n (\sigma_i^{-1}\sigma_j)(e_k)b_k\right) \\ &= \det\left(\sigma_i^{-1}\sigma_j\left(\sum_{k=1}^n b_k e_k\right)\right) \\ &= \det(\sigma_i^{-1}\sigma_j(x)), \end{aligned}$$

where $x = \sum_{k=1}^n b_k e_k$. Thus $\{\sigma_1(x), \ldots, \sigma_n(x)\}$ is a normal basis for L/K. ∎

This theorem proves the normal basis theorem for a Galois extension L/K where K is infinite. What if K is a finite field? In the case L/K is a cyclic extension (that is, $\mathrm{Gal}(L/K)$ is a cyclic group), and the next theorem addresses that case.

Theorem 3.7. *Every finite cyclic extension L/K has a normal basis.*

Proof. Let $\mathrm{Gal}(L/K) = \langle\sigma\rangle = \{1, \sigma, \ldots, \sigma^{n-1}\}$. The generator σ is a K-linear map from L to L. We will treat L as a module over $K[X]$ by letting X act on L as σ, so $f(X)\alpha = f(\sigma)\alpha$ for all $f(X) \in K[X]$. The polynomial $X^n - 1$ annihilates all of L since $(\sigma^n - \mathrm{id}_L)(\alpha) =$

$\sigma^n(\alpha) - \alpha = \alpha - \alpha = 0$. No lower-degree polynomial in $K[X]$ has this property, since if $c_0 + c_1X + \cdots + c_{n-1}X^{n-1} \in K[X]$ kills all of L then

$$c_0\alpha + c_1\sigma(\alpha) + \cdots + c_{n-1}\sigma^{n-1}(\alpha) = 0$$

for all $\alpha \in L$, and by linear independence of characters every c_i is 0. The annihilator ideal

$$\mathrm{Ann}_{K[X]}(L) = \{f(X) \in K[X] : f(\sigma) \equiv 0 \text{ on } L\}$$

is principal, so it must be $(X^n - 1)$. From the general theory of finitely generated modules over a PID, the annihilator ideal of the module equals the annihilator ideal of *some* particular element: there is some $\alpha_0 \in L$ such that $\mathrm{Ann}_{K[X]}(\alpha_0) = (X^n - 1)$. That means $f(\sigma)\alpha_0 = 0$ if and only if $(X^n - 1)|f(X)$. Therefore the K-linear function $K[X] \to L$ by $f(X) \mapsto f(\sigma)\alpha_0$ has kernel $(X^n - 1)$, so it induces a K-linear embedding $K[X]/(X^n - 1) \hookrightarrow L$. Since $K[X]/(X^n - 1)$ and L both have K-dimension n, our embedding is surjective too. Thus the image of a K-basis of $K[X]/(X^n - 1)$ is a K-basis of L. The K-basis $1, X, \ldots, X^{n-1}$ of $K[X]/(X^n - 1)$ is sent to $\alpha_0, \sigma(\alpha_0), \ldots, \sigma^{n-1}(\alpha_0)$, so this is a normal basis of L/K. ∎

Using a normal basis and the trace function, let's see how to write down a primitive element for *every* intermediate extension in a finite Galois extension.

Theorem 3.8. *Let L/K be Galois with Galois group G and let $\{\sigma(\alpha) : \sigma \in G\}$ be a normal basis.*

For a subgroup $H \subset G$, $L^H = K(\alpha_H)$ where $\alpha_H = \sum_{\tau \in H} \tau(\alpha) = \mathrm{Tr}_{L/L^H}(\alpha)$.

If $N \lhd G$ then a normal basis for $K(\alpha_N)/K$ is $\{\overline{\tau}(\alpha_N) : \overline{\tau} \in G/N\}$.

Proof. Since $\alpha_H \in L^H$, $K(\alpha_H) \subset L^H$. So to show $K(\alpha_H) = L^H$ it is enough to show $[K(\alpha_H) : K] = [L^H : K]$. The degree $[K(\alpha_H) : K]$ equals the size of the Galois orbit $\{\sigma(\alpha_H) : \sigma \in G\}$. Using the orbit-stabilizer formula, the size of this orbit is the index of the stabilizer of α_H in G. So we ask: for which $\sigma \in G$ is $\sigma(\alpha_H) = \alpha_H$? This holds if $\sigma \in H$, and conversely if $\sigma(\alpha_H) = \alpha_H$ then

$$\begin{aligned} 0 &= \sigma(\alpha_H) - \alpha_H \\ &= \sum_{\tau \in H} (\sigma\tau)(\alpha) - \sum_{\tau \in H} \tau(\alpha), \end{aligned}$$

so by the meaning of a normal basis we must have $\sigma\tau = \tau'$ for some τ and τ' in H. Therefore $\sigma = \tau'\tau^{-1} \in H$, so the stabilizer of α_H is H, which means $|\{\sigma(\alpha_H) : \sigma \in G\}| = [G : H] = [L^H : K]$.

Now suppose $N \lhd G$. We want to show $\{\overline{\tau}(\alpha_N) : \overline{\tau} \in G/N\}$ is a normal basis for $K(\alpha_N)/K$. For $\tau \in G$, $\overline{\tau}(\alpha_N) = \tau(\alpha_N)$. Let $\tau_1, \ldots, \tau_m$ be coset representatives in G for G/N, so

$$\{\overline{\tau}(\alpha_N) : \overline{\tau} \in G/N\} = \{\tau_1(\alpha_N), \ldots, \tau_m(\alpha_N)\}.$$

We want to show this set is linearly independent over K. A linear dependence relation can be written as

$$0 = \sum_{i=1}^m c_i\tau_i(\alpha_N) = \sum_{i=1}^m c_i\tau_i \left(\sum_{\sigma \in N} \sigma(\alpha) \right) = \sum_{i=1}^m \sum_{\sigma \in N} c_i\tau_i(\sigma(\alpha)) = \sum_{i=1}^m \sum_{\sigma \in N} c_i(\tau_i\sigma)(\alpha).$$

Since the τ_i's are coset representatives for N in G, the set of all products $\tau_i\sigma$ is all of G, so the last double sum is a single sum over all elements of G. Therefore the coefficients are all 0 from our normal basis for L/K. ∎

Remark 3.9. The normal basis theorem doesn't make sense for an infinite Galois extension: in an infinite Galois extension, each particular element has only finitely many conjugates over the base field, so no element can have its conjugates be a basis for the field extension when the degree of the extension is infinite.

However, there is another way to think about the normal basis theorem that does generalize to infinite-degree extensions. For a finite Galois extension L/K with Galois group G, the group G acts as K-linear maps $L \to L$. The set $\mathrm{Map}(G, K)$ of functions $f\colon G \to K$ is also a K-vector space on which G acts (by $(\sigma f)(\tau) := f(\sigma^{-1}\tau)$) as K-linear maps. Using a normal basis $\{\sigma(\alpha) : \sigma \in G\}$ for L/K sets up a bijection $\mathrm{Map}(G, K) \to L$ (namely $f \mapsto \sum_{\sigma\in G} f(\sigma)\sigma(\alpha)$) that is an isomorphism of K-vector spaces that respects the G-actions on both sides. Conversely, such an isomorphism respecting the G-actions leads to a normal basis of L/K. (Take for α the element of L that is identified with the function $G \to K$ that is 1 at the identity and 0 elsewhere.) So the normal basis theorem for L/K is equivalent to L being isomorphic to $\mathrm{Map}(G, K)$ as K-vector spaces by an isomorphism respecting the G-action. This viewpoint generalizes to possibly infinite Galois extensions L/K with $\mathrm{Map}(G, K)$ being replaced by the space $C(G, K)$ of continuous functions $G \to K$ where G has the Krull topology and K has the discrete topology [7].

4. Hilbert's Theorem 90

In a finite extension L/K, K-conjugate elements have the same minimal polynomial, and thus the same characteristic polynomial (a power of the minimal polynomial), and thus the same norm and trace. In particular, when L/K is Galois we have $\mathrm{N}_{L/K}(\sigma(\gamma)) = \mathrm{N}(\gamma)$ and $\mathrm{Tr}_{L/K}(\sigma(\gamma)) = \mathrm{Tr}_{L/K}(\gamma)$ for all $\gamma \in L$ and $\sigma \in \mathrm{Gal}(L/K)$. So the homomorphisms $\mathrm{N}_{L/K}\colon L^\times \to K^\times$ and $\mathrm{Tr}_{L/K}\colon L \to K$ contain obvious elements in their kernels: all $\sigma(\gamma)/\gamma$ for the norm and all $\sigma(\gamma) - \gamma$ for the trace. When L/K is a cyclic extension, a theorem of Hilbert says that every element of the kernel of the norm and trace has this obvious form for some $\gamma \in L$.

Theorem 4.1 (Hilbert's Theorem 90)**.** *Let L/K be a cyclic extension and σ be a generator of the Galois group. For $\alpha \in L^\times$,*

$$\mathrm{N}_{L/K}(\alpha) = 1 \Longleftrightarrow \alpha = \frac{\sigma(\beta)}{\beta} \text{ for some } \beta \in L^\times,$$

and

$$\mathrm{Tr}_{L/K}(\alpha) = 0 \Longleftrightarrow \alpha = \sigma(\beta) - \beta \text{ for some } \beta \in L.$$

Notice the role of a choice of generator of $\mathrm{Gal}(L/K)$ in the theorem.

Proof. Only the direction ($\Rightarrow$) has to be proved. Let $n = [L : K]$.

First we treat the multiplicative version, for norms. Pick an $\alpha \in L^\times$ such that $\mathrm{N}_{L/K}(\alpha) = 1$ and consider the function $f\colon L \to L$ given by

$$f(x) = \alpha x + \alpha\sigma(\alpha)\sigma(x) + \alpha\sigma(\alpha)\sigma^2(x) + \cdots + \underbrace{\alpha\sigma(\alpha)\cdots\sigma^{n-1}(\alpha)}_{\mathrm{N}_{L/K}(\alpha)=1}\sigma^{n-1}(x).$$

This is an L-linear combination of the numbers $\{\sigma^i(x) Li = 0, \ldots, n-1\}$ and the coefficients are not all 0 (in fact none are 0). Viewing $\mathrm{Gal}(L/K)$ as characters of $L^\times$, linear independence

of characters implies $f(x) \neq 0$ for some $x \in L^\times$. For such an x,

$$\begin{aligned}
\sigma(f(x)) &= \sigma(\alpha)\sigma(x) + \sigma(\alpha)\sigma^2(\alpha)\sigma^2(x) + \cdots + \underbrace{\sigma(\alpha)\sigma^2(\alpha)\cdots\sigma^n(\alpha)}_{\sigma(1)=1}\sigma^n(x) \\
&= \frac{f(x) - \alpha x}{\alpha} + \mathrm{N}_{L/K}(\alpha)x \qquad \text{since } \sigma^n = \mathrm{id}_L \\
&= \frac{f(x)}{\alpha} + (\mathrm{N}_{L/K}(\alpha) - 1)x \\
&= \frac{f(x)}{\alpha},
\end{aligned}$$

so $\alpha = \sigma(\beta)/\beta$ where $\beta = 1/f(x)$.

For the additive version, we will need the fact that the trace map $\mathrm{Tr}_{L/K}\colon L \to K$ is not identically 0. This is true either because the trace map is not identically 0 for finite separable extensions, or because in the special case of a Galois extension we can write $\mathrm{Tr}_{L/K}(x) = \sum_{\tau\in G}\tau(x)$ so $\mathrm{Tr}_{L/K}$ is a nontrivial linear combination of characters of L and thus can't be identically 0 on L.

Pick $\alpha \in L$ with $\mathrm{Tr}_{L/K}(\alpha) = 0$, and define $f\colon L \to L$ by

$$f(x) = \alpha x + (\alpha + \sigma(\alpha))\sigma(x) + \cdots + \underbrace{(\alpha + \sigma(\alpha) + \cdots + \sigma^{n-1}(\alpha))}_{\mathrm{Tr}_{L/K}(\alpha)=0}\sigma^{n-1}(x).$$

Then

$$\begin{aligned}
\sigma(f(x)) &= \sigma(\alpha)\sigma(x) + (\sigma(\alpha) + \sigma^2(\alpha))\sigma^2(x) + \cdots + \\
&\qquad (\sigma(\alpha) + \sigma^2(\alpha) + \cdots + \sigma^n(\alpha))\sigma^n(x) \\
&= f(x) - \alpha x - \alpha\sigma(x) - \cdots - \alpha\sigma^{n-1}(x) + \\
&\qquad (\sigma(\alpha) + \sigma^2(\alpha) + \cdots + \sigma^n(\alpha))x \qquad \text{since } \sigma^n = \mathrm{id}_L \\
&= f(x) - \alpha\,\mathrm{Tr}_{L/K}(x) + \mathrm{Tr}_{L/K}(\alpha)x \\
&= f(x) - \alpha\,\mathrm{Tr}_{L/K}(x).
\end{aligned}$$

Since $\mathrm{Tr}_{L/K}(\alpha) = 0$, $\sigma(f(x)) = f(x) - \alpha\,\mathrm{Tr}_{L/K}(x)$ for all $x \in L$. Choose $x \in L$ such that $\mathrm{Tr}_{L/K}(x) \neq 0$. Then $\alpha = \sigma(\beta) - \beta$ for $\beta = -f(x)/\,\mathrm{Tr}_{L/K}(x)$ and we're done.

As an illustration of different techniques, let's reprove the additive version using the normal basis theorem. Let $\{\sigma^i(\gamma)\}_{i=0}^{n-1}$ be a normal basis for L/K. For $\alpha \in L$, write

$$\alpha = \sum_{i=0}^{n-1} c_i\sigma^i(\gamma)$$

with $c_i \in K$. Since $\mathrm{Tr}_{L/K}(\sigma^i(\gamma)) = \mathrm{Tr}_{L/K}(\gamma)$ for all i,

$$\mathrm{Tr}_{L/K}(\alpha) = \left(\sum_{i=0}^{n-1} c_i\right)\mathrm{Tr}_{L/K}(\gamma).$$

Here $c_1, \ldots, c_n$ vary with α, but the term $\mathrm{Tr}_{L/K}(\gamma)$ does not depend on α. If $\mathrm{Tr}_{L/K}(\gamma) = 0$ then $\mathrm{Tr}_{L/K}(\alpha) = 0$ for all $\alpha \in L$. But $\mathrm{Tr}_{L/K}$ is not identically zero, so we must have

$\mathrm{Tr}_{L/K}(\gamma) \neq 0$. Therefore if $\mathrm{Tr}_{L/K}(\alpha) = 0$ we get $\sum_{i=0}^{n-1} c_i = 0$, so

$$\begin{aligned} \alpha &= c_0\gamma + c_1\sigma(\gamma) + \cdots + c_{n-1}\sigma^{n-1}(\gamma) \\ &= -\left(\sum_{i=1}^{n-1} c_i\right)\gamma + \sum_{i=1}^{n-1} c_i\sigma^i(\gamma) \\ &= \sum_{i=1}^{n-1} c_i(\sigma^i(\gamma) - \gamma). \end{aligned}$$

Letting $\beta_i = \gamma + \sigma(\gamma) + \cdots + \sigma^{i-1}(\gamma)$ for $i = 1, \ldots, n-1$, we have $\sigma^i(\gamma) - \gamma = \sigma(\beta_i) - \beta_i$, so $\alpha = \sigma(\beta) - \beta$ for $\beta = \sum_{i=1}^{n-1} c_i\beta_i$. ∎

Theorem 90 was the 90th theorem in Hilbert's Zahlbericht, his 1897 report to the German Mathematical Society on algebraic number theory (an English translation is available [6]). This theorem has a generalization (due to Speiser in 1919) to non-cyclic Galois extensions, but it is not about the kernel of the norm or trace anymore. Instead it is about the vanishing of certain Galois cohomology groups, which in the case of cyclic extensions takes on the form of Theorem 4.1 above. The label "Theorem 90" in the literature often refers to these vanishing theorems in Galois cohomology. (For a norm-interpretation of Theorem 90 in the case of biquadratic extensions $K(\sqrt{a}, \sqrt{b})/K$, see [4].)

A cute application of Theorem 90 is the classification of Pythagorean triples, which are integral solutions to $a^2 + b^2 = c^2$. If $a^2 + b^2 = c^2$ in $\mathbf{Z}^+$ then $(\frac{a}{c})^2 + (\frac{b}{c})^2 = 1$ in $\mathbf{Q}$, which means $\mathrm{N}_{\mathbf{Q}(i)/\mathbf{Q}}(\frac{a}{c} + \frac{b}{c}i) = 1$. Now apply Theorem 90 to the cyclic extension $\mathbf{Q}(i)/\mathbf{Q}$: the element $\frac{a}{c} + \frac{b}{c}i$ has norm 1, so $a/c + (b/c)i = \overline{\beta}/\beta$ for some nonzero $\beta \in \mathbf{Q}(i)$. Clearing a common denominator, we can assume $\beta = m + ni$ has integral, rather than rational, real and imaginary parts. Then

$$\frac{a}{c} + \frac{b}{c}i = \frac{m - ni}{m + ni} = \frac{m^2 - n^2 - 2mni}{m^2 + n^2},$$

and comparing both sides leads to 2-parameter formulas for a, b, and c in terms of m and n. For further details, see [5].

5. Kummer and Artin-Schreier extensions

When K does not have characteristic 2, every quadratic extension of K takes the form $K(\sqrt{a})$ for some non-square $a \in K^\times$. There is a way to generalize this explicit description to cyclic extensions of every degree n provided the base field contains a primitive nth root of unity.

Theorem 5.1 (Kummer)**.** *Let K be a field containing a primitive nth root of unity. The cyclic extensions of K with degree dividing n are precisely the extensions $K(\sqrt[n]{a})$ for some $a \in K^\times$.*

Here and below, $\sqrt[n]{a}$ is notation for some number whose nth power is a. There is no canonical choice of such an nth root (when $n > 1$).

Proof. In a nutshell, the idea in the proof is that we can keep track of how the Galois group behaves by using the nth roots of unity in K, and those form a cyclic group of size n.

First we will show for all $a \in K^\times$ that $K(\sqrt[n]{a})/K$ is cyclic with degree dividing n. Let ζ_n be a primitive nth root of unity in K. From one root $\sqrt[n]{a}$ of $X^n - a$ we get a full set of roots

as $\{\zeta_n^i \sqrt[n]{a} : 0 \leqslant i \leqslant n-1\}$, and these are all in $K(\sqrt[n]{a})$. Thus $K(\sqrt[n]{a})/K$ is Galois. Let $G = \mathrm{Gal}(K(\sqrt[n]{a})/K)$. We are going to write down an injective homomorphism $G \to \mu_n$, so G is cyclic of order dividing n.

For $\sigma \in G$, $\sigma(\sqrt[n]{a}) = \zeta\sqrt[n]{a}$ for some $\zeta \in \mu_n$. Define $\chi_a \colon G \to \mu_n$ by

$$\chi_a(\sigma) = \zeta = \frac{\sigma(\sqrt[n]{a})}{\sqrt[n]{a}}.$$

This ratio is independent of the choice of nth root of a. Every other choice of nth root of a has the form $\zeta\sqrt[n]{a}$ for some $\zeta \in \mu_n \subset K$, and

$$\frac{\sigma(\zeta\sqrt[n]{a})}{\zeta\sqrt[n]{a}} = \frac{\zeta\sigma(\sqrt[n]{a})}{\zeta\sqrt[n]{a}} = \frac{\sigma(\sqrt[n]{a})}{\sqrt[n]{a}}.$$

Because χ_a's definition is independent of the choice of nth root of a that is used, we can show χ_a is a homomorphism: for σ and τ in G,

$$\chi_a(\sigma\tau) = \frac{\sigma(\tau(\sqrt[n]{a}))}{\sqrt[n]{a}} = \frac{\sigma(\tau(\sqrt[n]{a}))}{\tau(\sqrt[n]{a})}\frac{\tau(\sqrt[n]{a})}{\sqrt[n]{a}} = \chi_a(\sigma)\chi_a(\tau)$$

since $\tau(\sqrt[n]{a})$ is an nth root of a. Easily χ_a has a trivial kernel, so χ_a embeds G in μ_n and therefore G is cyclic of order dividing n.

Now we handle the more interesting converse direction. Suppose L is a cyclic extension of K with degree dividing n, say $[L : K] = m$ and $m|n$. We want to show $L = K(\sqrt[n]{a})$ for some $a \in K^\times$.

Write σ for a generator of $\mathrm{Gal}(L/K)$ and $\zeta_m := \zeta_n^{n/m}$ for a primitive mth root of unity in K. Since $\zeta_m \in K$,

$$\mathrm{N}_{L/K}(\zeta_m) = \zeta_m^{[L:K]} = \zeta_m^m = 1,$$

so by the multiplicative version of Theorem 90 we can write $\zeta_m = \sigma(\beta)/\beta$ for some $\beta \in L^\times$. Since $\sigma(\beta) = \zeta_m\beta$, $\sigma^i(\beta) = \zeta_m^i\beta$ for all i. Therefore β has m different K-conjugates in L, so by Galois theory the minimal polynomial of β over K has degree m, which is $[L : K]$, and thus $L = K(\beta)$. Raising both sides of the equation $\sigma(\beta) = \zeta_m\beta$ to the mth power, $\sigma(\beta^m) = \beta^m$. Therefore β^m is fixed by $\langle\sigma\rangle = \mathrm{Gal}(L/K)$, so $\beta^m \in K^\times$. As $m|n$, we have $\beta^n \in K^\times$ too. Set $a = \beta^n$. Then β is an nth root of an element of $K^\times$, so $L = K(\beta) = K(\sqrt[n]{a})$.

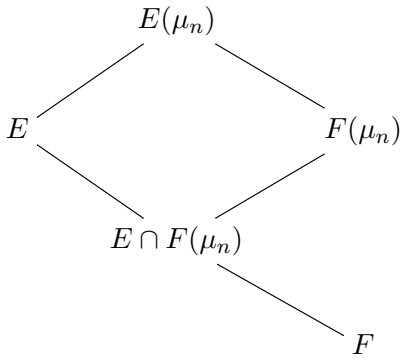
Here's another proof of the converse direction, using linear independence of characters. We use the notation σ and ζ_m from the previous paragraph. Let $f \colon L \to L$ by $f(x) = \sum_{i=0}^{m-1} \zeta_m^i\sigma^i(x)$. In ζ_m^i and $\sigma^i(x)$, i only matters modulo m. For some x_0 we have $f(x_0) \neq 0$ by linear independence of characters, so

$$\sigma(f(x_0)) = \sum_{i=0}^{m-1} \zeta_m^i\sigma^{i+1}(x_0) = \sum_{i=1}^{m} \zeta_m^{i-1}\sigma^i(x_0) = \zeta_m^{-1}f(x_0).$$

This is the analogue of the formula $\sigma(\beta) = \zeta_m\beta$ in the previous proof, and we now repeat the rest of the argument as it was given above. ∎

Corollary 5.2. *Let E/F be a cyclic extension of degree n. If $n \neq 0$ in F then $E \subset F(\mu_n, \sqrt[n]{\gamma})$ for some $\gamma \in F(\mu_n)$.*

Proof. Since $X^n - 1$ is separable in $F[X]$, there is a full group μ_n of nth roots of unity in an extension of F. Consider the composite field $E \cdot F(\mu_n) = E(\mu_n)$.



From Galois theory, $E(\mu_n)/F(\mu_n)$ is Galois and its Galois group is isomorphic to a $\mathrm{Gal}(E/E \cap F(\mu_n)) \subset \mathrm{Gal}(E/F)$. Therefore $E(\mu_n)/F(\mu_n)$ is cyclic of degree dividing n. By Theorem 5.1 with $K = F(\mu_n)$, $E(\mu_n) = K(\sqrt[n]{\gamma})$ for some $\gamma \in K$. ∎

Generally we can't expect γ in Corollary 5.2 to lie in the base field F.

Example 5.3. Take $F = \mathbf{Q}$ and $E = \mathbf{Q}(\alpha)$ where $\alpha^3 - 3\alpha - 1 = 0$. The cubic extension $\mathbf{Q}(\alpha)/\mathbf{Q}$ is cyclic, with the $\mathbf{Q}$-conjugates of α being α, $-\alpha^2 + 2$, and $\alpha^2 - \alpha - 2$. We can't write $E = \mathbf{Q}(\sqrt[3]{r})$ for rational r, since E has 3 real embeddings but the $\mathbf{Q}$-conjugates of a pure cube root are not all real. Letting ω be a nontrivial cube root of unity, Corollary 5.2 says $E \subset \mathbf{Q}(\omega, \sqrt[3]{\gamma})$ for some $\gamma \in \mathbf{Q}(\omega)$. To find γ is a matter of using the last paragraph of the proof of Theorem 5.1. Since $\mathbf{Q}(\omega)$ and E have relatively prime degree over $\mathbf{Q}$, α has degree 3 over $\mathbf{Q}(\omega)$. Letting σ be the element of $\mathrm{Gal}(\mathbf{Q}(\omega)(\alpha)/\mathbf{Q}(\omega))$ determined by $\sigma(\alpha) = -\alpha^2 + 2$, we seek a nonzero sum of the form

$$\begin{aligned} \beta &= \sum_{i=0}^{3-1} \omega^i \sigma^i(x) \\ &= x + \omega\sigma(x) + \omega^2\sigma^2(x) \\ &= x + \omega\sigma(x) + (-1-\omega)\sigma^2(x) \end{aligned}$$

for some $x \in E$. Taking $x = \alpha$ gives $\beta = (2+4\omega)+(2+\omega)\alpha-(1+2\omega)\alpha^2$, which is not 0 since its coefficients in the $\mathbf{Q}(\omega)$-basis $\{1, \alpha, \alpha^2\}$ are not all 0. Thus theory says $E \subset \mathbf{Q}(\omega)(\beta)$ and $\beta^3 \in \mathbf{Q}(\omega)$. An explicit calculation shows that in fact $\beta^3 = -27\omega = (-3)^3\omega$, so $E \subset \mathbf{Q}(\omega)(\sqrt[3]{\omega}) = \mathbf{Q}(\sqrt[3]{\omega})$. Since ω is a nontrivial cube root of unity, $\mathbf{Q}(\sqrt[3]{\omega}) = \mathbf{Q}(\zeta_9)$.

Definition 5.4. A *Kummer extension* is an extension $K(\sqrt[n]{a})/K$ where K is a field containing a primitive nth root of unity and $a \in K^\times$. Equivalently, Kummer extensions are cyclic extensions where the base field has a root of unity of order equal to the degree of the extension.

The study and applications of Kummer extensions is called Kummer theory. All quadratic extensions outside of characteristic 2 are Kummer extensions, but quadratic Galois extensions in characteristic 2 are not Kummer extensions; there is no root of unity of order 2 in characteristic 2 since $-1 = 1$.

Theorem 5.5. *With notation as in Theorem 5.1, $[K(\sqrt[n]{a}) : K]$ equals the order of a in $K^\times/(K^\times)^n$.*

Proof. Set $\alpha = \sqrt[n]{a}$ and $\mathrm{Gal}(K(\alpha)/K) = \langle\sigma\rangle$, so σ has order dividing n. We will show for $t \in \mathbf{Z}$ that $a^t \in (K^\times)^n$ if and only if $\sigma^t = \mathrm{id}_{K(\alpha)}$, and the theorem is a consequence of this since $[K(\alpha) : K] = |\mathrm{Gal}(K(\alpha)/K)|$. Write $\sigma(\alpha) = \zeta\alpha$ for some nth root of unity ζ. Then $\sigma^i(\alpha) = \zeta^i\alpha$ for all i.

Suppose $\sigma^t = \mathrm{id}_{K(\alpha)}$. Then $\zeta^t = 1$, so $\sigma(\alpha)^t = \alpha^t$. Rewrite this as $\sigma(\alpha^t) = \alpha^t$. Then α^t is fixed by $\langle\sigma\rangle = \mathrm{Gal}(K(\alpha)/K)$, so $\alpha^t \in K^\times$. Then $a^t = (\alpha^t)^n \in (K^\times)^n$.

Now suppose, conversely, that $a^t \in (K^\times)^n$. Write $a^t = b^n$ with $b \in K^\times$. Then $\alpha^{nt} = b^n$, so $\alpha^t = \omega \cdot b$ for some $\omega \in \mu_n$. Since $\mu_n \subset K^\times$, $\alpha^t \in K^\times$. Raising the equation $\sigma(\alpha) = \zeta\alpha$ to the power t, $\sigma(\alpha^t) = \zeta^t\alpha^t$. Since $\alpha^t \in K^\times$, $\sigma(\alpha^t) = \alpha^t$, so $\zeta^t = 1$. Thus $\sigma^t(\alpha) = \zeta^t\alpha = \alpha$, so $\sigma^t = \mathrm{id}_{K(\alpha)}$. ∎

Corollary 5.6. *If $\mu_n \subset K$ and $a \in K^\times$, then the irreducible factors of $X^n - a$ in $K[X]$ have the form $X^t - b$, where t is the order of a in $K^\times/(K^\times)^n$ and b runs through the (n/t)th roots of a.*

Proof. Every element of $K^\times/(K^\times)^n$ has order dividing n, so $t|n$. Writing $a^t = c^n$ for some $c \in K^\times$, $a^t = (c^{n/t})^t$. Thus $a = \omega c^{n/t}$ for some $\omega \in \mu_t \subset \mu_n \subset K$. Since $\mu_n^{n/t} = \mu_t$, a is an (n/t)th power in $K^\times$.

For each root α of $X^n - a$, the proof of Theorem 5.5 shows us that its K-conjugates are $\zeta\alpha$ as ζ runs over the tth roots of unity. Therefore the minimal polynomial of α in $K[X]$ is

$$\prod_{\zeta^t=1} (X - \zeta\alpha) = X^t - \alpha^t,$$

and α^t is an (n/t)th root of a: $(\alpha^t)^{n/t} = \alpha^n = a$. ∎

When K doesn't have characteristic 2, the quadratic extensions $K(\sqrt{a})$ and $K(\sqrt{b})$ are K-isomorphic if and only if $a/b \in (K^\times)^2$. Here is a generalization to Kummer extensions.

Theorem 5.7. *Let K be a field containing a primitive nth root of unity. Write two degree-n Kummer extensions of K in the form $K(\sqrt[n]{a})$ and $K(\sqrt[n]{b})$ with a and b in $K^\times$. These extensions are K-isomorphic if and only if $a = b^r c^n$ for some r relatively prime to n and some $c \in K^\times$.*

Not every field extension $K(\sqrt[n]{a})/K$ has degree n, but only degree dividing n (extreme case: a could be an nth power in K).

Proof. Our argument is taken from [2].

We will work in a fixed algebraic closure of K, so instead of speaking about K-isomorphic fields we can speak about equal fields because the extensions $K(\sqrt[n]{a})/K$ and $K(\sqrt[n]{b})/K$ are Galois.

First we show the condition $a = b^r c^n$ for some r relatively prime to n and some c is symmetric in the roles of a and b. Write $rk + n\ell = 1$ for some integers k and ℓ. Then $a^k = b^{rk}c^{nk} = b^{1-n\ell}c^{nk} = b(c^k/b^\ell)^n$, so $b = a^k(c')^n$ where k is relatively prime to n and $c' = b^\ell/c^k \in K^\times$.

When $a = b^r c^n$ with $(r, n) = 1$ and $c \in K^\times$, let β be an nth root of b. Then $a = (\beta^r c)^n$ so there is an nth root of a in $K(\beta) = K(\sqrt[n]{b})$. Hence $K(\sqrt[n]{a}) \subset K(\sqrt[n]{b})$. Since the condition on a and b is symmetric we get the reverse inclusion too, so $K(\sqrt[n]{a}) = K(\sqrt[n]{b})$.

Conversely, suppose $K(\sqrt[n]{a}) = K(\sqrt[n]{b})$. Write α for an nth root of a and β for an nth root of b: $\alpha^n = a$ and $\beta^n = b$. Let σ be a generator of the Galois group over K, so σ has order $[K(\sqrt[n]{a}) : K]$, which we are assuming in the theorem is n. Thus $\sigma(\alpha)/\alpha$ and $\sigma(\beta)/\beta$

are primitive nth roots of unity. Write $\sigma(\beta) = \zeta\beta$ with ζ of order n, so $\sigma(\alpha) = \zeta^r\alpha$ where $(r, n) = 1$. We may take $1 \leqslant r \leqslant n-1$. Since $K(\alpha) = K(\beta)$ we can write

$$\alpha = \sum_{i=0}^{n-1} c_i\beta^i, \quad c_i \in K.$$

Applying σ,

$$\sigma(\alpha) = \sum_{i=0}^{n-1} c_i\zeta^i\beta^i.$$

Since $\sigma(\alpha) = \zeta^r\alpha$,

$$\zeta^r \sum_{i=0}^{n-1} c_i\beta^i = \sum_{i=0}^{n-1} c_i\zeta^i\beta^i.$$

Equating coefficients of powers of β on both sides, $c_i\zeta^r = c_i\zeta^i$ for $0 \leqslant i \leqslant n-1$. Therefore when $i \neq r$ we have $c_i = 0$, so $\alpha = c_r\beta^r$. Raising both sides to the nth power, $a = b^r c^n$, where $c = c_r$. ∎

Remark 5.8. The condition $a = b^r c^n$ for some r relatively prime to n and some $c \in K^\times$ is equivalent to saying a and b generate the same subgroup of $K^\times/(K^\times)^n$, so although a is not well-defined from the extension $K(\sqrt[n]{a})/K$, the subgroup generated by a in $K^\times/(K^\times)^n$ is well-defined from the extension $K(\sqrt[n]{a})/K$.

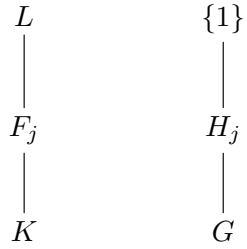
An abelian group is said to have *exponent* n if n kills every element of the group: every element has order dividing n. A subgroup and quotient group of a group with exponent n also has exponent n. A cyclic group has exponent n if and only if its size divides n.

Theorem 5.9. *Let L/K be an abelian extension with $\mu_n \subset K$. If $\mathrm{Gal}(L/K)$ has exponent n, then $L = K(\sqrt[n]{a_1}, \ldots, \sqrt[n]{a_m})$ for some a_i's in $K^\times$. Conversely, every extension $K(\sqrt[n]{a_1}, \ldots, \sqrt[n]{a_m})/K$ is abelian of exponent n.*

Proof. The group $G = \mathrm{Gal}(L/K)$ is abelian, so it is a direct product of cyclic groups, say

$$G = C_1 \times \cdots \times C_m$$

with each C_i being cyclic. Let $H_j = \prod_{i \neq j} C_i \times \{1\}$, so $G/H_j \cong C_j$. Let $F_j = L^{H_j}$ be the fixed field of H_j, so $\mathrm{Gal}(F_j/K) \cong G/H_j \cong C_j$ is a cyclic group. Thus F_j/K is a cyclic extension.



Since G has exponent n, each C_j has exponent n. Therefore $|C_j|$ divides n, so $F_j = K(\sqrt[n]{a_j})$ for some $a_j \in K^\times$ (Theorem 5.1). Since $H_1 \cap \cdots \cap H_m$ is trivial, $L = F_1 \cdots F_m = K(\sqrt[n]{a_1}, \ldots, \sqrt[n]{a_m})$.

Conversely, for $a_1, \ldots, a_m \in K^\times$ each $K(\sqrt[n]{a_j})/K$ is abelian with exponent n so the field $K(\sqrt[n]{a_1}, \ldots, \sqrt[n]{a_m})$ is Galois over K with its Galois group embedding into the direct product of the groups $\mathrm{Gal}(K(\sqrt[n]{a_j})/K)$, so the Galois group is abelian of exponent n. ∎

Because there are no p-th power roots of unity in characteristic p other than 1, Kummer theory can't be used to describe cyclic extensions of p-power degree in characteristic p. There is a substitute for Kummer theory in this case, found by Artin and Schreier in degree p and by Witt in p-power degree. We will focus on the simplest case: cyclic extensions of degree p. The radical polynomials $X^n - a$ are replaced with the new (Artin-Schreier) polynomials $X^p - X - a$, and the multiplicative aspects of Kummer theory become *additive*.

Theorem 5.10 (Artin-Schreier)**.** *Let K be a field of characteristic p. Then L/K is cyclic of degree p if and only if $L = K(\alpha)$ where α is a root of $X^p - X - a \in K[X]$ and this polynomial has no root in K.*

Proof. If α is a root of $X^p - X - a$ and there are no roots of this polynomial in K then $\alpha \notin K$. For each $c \in \mathbf{F}_p$ the number $\alpha + c$ is in $K(\alpha)$ and is also a root of $X^p - X - a$ since

$$\alpha^p = \alpha + a \Rightarrow (\alpha + c)^p = \alpha^p + c^p = (\alpha + a) + c = (\alpha + c) + a \Rightarrow (\alpha + c)^p - (\alpha + c) - a = 0.$$

Thus $K(\alpha)/K$ is Galois: it is a splitting field of the separable polynomial $X^p - X - a$. We have $1 < [K(\alpha) : K] \leqslant p$. Let σ be a nontrivial element of $\mathrm{Gal}(K(\alpha)/K)$, so $\sigma(\alpha) = \alpha + i$ for some $i \neq 0$ in $\mathbf{F}_p$. Then $\sigma^j(\alpha) = \alpha + ij$, so $\sigma^j(\alpha) \neq \alpha$ when $1 \leqslant j \leqslant p-1$ while $\sigma^p(\alpha) = \alpha$. Therefore σ has order p in the Galois group. Since p is an upper bound on $[K(\alpha) : K] = |\mathrm{Gal}(K(\alpha)/K)|$, the Galois group has order p and is cyclic. (In particular, $X^p - X - a$ is irreducible in $K[X]$ when it does not have a root in K.)

Now assume L/K is cyclic of degree p, say $\mathrm{Gal}(L/K) = \langle\sigma\rangle$. Since $\mathrm{Tr}_{L/K}(1) = p = 0$, by the additive version of Theorem 90 we have $1 = \sigma(\alpha) - \alpha$ for some $\alpha \in L$. Therefore $\sigma(\alpha) = \alpha + 1$, so $\sigma^i(\alpha) = \alpha + i$ for $0 \leqslant i \leqslant p-1$. Thus $\alpha, \alpha+1, \ldots, \alpha+p-1$ are all K-conjugates in L. This fills up the p possible K-conjugates of α in L, so the minimal polynomial of α in $K[X]$ is

$$\prod_{i=0}^{p-1}(X - (\alpha + i)) = \prod_{i=0}^{p-1}((X - \alpha) - i) = (X - \alpha)^p - (X - \alpha) = X^p - X - (\alpha^p - \alpha).$$

This is in $K[X]$, so $a := \alpha^p - \alpha$ lies in K. Then α is a root of $X^p - X - a$ and the roots of this polynomial aren't in K since they are $\alpha + i$ as i runs over $\mathbf{F}_p$. ∎

Let $\wp(X) = X^p - X$, which is an additive function in characteristic p. The equation $\wp(x) = a$ plays the role for cyclic extensions of degree p in characteristic p that the equation $x^n = a$ does for cyclic extensions of degree dividing n in Kummer theory. The subgroup $(K^\times)^n$ of $K^\times$ in Kummer theory is replaced by the subgroup $\wp(K) = \{c^p - c : c \in K\}$ of K. Writing $\wp^{-1}(a)$ for a solution to $x^p - x = a$ (analogue of $\sqrt[n]{a}$ as a solution of $x^n = a$), Theorem 5.10 tells us every cyclic extension of K with degree p has the form $K(\wp^{-1}(a))$ for some $a \in K$ with $a \notin \wp(K)$. There is an analogue of Theorem 5.7: $K(\wp^{-1}(a)) = K(\wp^{-1}(b))$ if and only if $a = rb + c^p - c$ for some $r \in \mathbf{F}_p^\times$ and some $c \in K$, which is equivalent to saying a and b have the same $\mathbf{F}_p$-span in the additive group $K/\wp(K)$. (Compare to Remark 5.8.) The proof of this is similar to that of Theorem 5.7.

References

[1] E. Artin, "Galois Theory," 2nd ed., Notre Dame, 1948.

[2] B. Birch, Cyclotomic fields and Kummer extensions, pp. 85–93 in: Algebraic Number Theory (J. W. S. Cassels, A. Fröhlich, ed.), Academic Press, London, 1986.

[3] D. Dummit and R. Foote, "Abstract Algebra," 3rd ed., Wiley, New York, 2004.

[4] R. Dwilewicz, J. Minca, A. Schultz, J. Swallow, *Hilbert* 90 *for biquadratic extensions*, http://arxiv.org/pdf/math.NT/0510154.
[5] N. Elkies, *Pythagorean triples and Hilbert's Theorem* 90, Amer. Math. Monthly **110** (2003), 678.
[6] D. Hilbert, "The Theory of Algebraic Number Fields," Springer-Verlag, Berlin, 1998.
[7] H. W. Lenstra, Jr., *A normal basis theorem for infinite Galois extensions*, Nederl. Akad. Wetensch. Indag. Math **47** (1985), 221–228.
[8] J. Rotman, "Advanced Modern Algebra," Prentice-Hall, Upper Saddle River, NJ, 2002.
[9] W. Waterhouse, *The normal basis theorem*, Amer. Math. Monthly **86** (1979), 212.