THE ARTIN-SCHREIER THEOREM

KEITH CONRAD

1. Introduction

The algebraic closure of $\mathbf{R}$ is $\mathbf{C}$, which is a finite extension. Are there other fields which are not algebraically closed but have an algebraic closure that is a finite extension? Yes. An example is the field of real algebraic numbers. Since complex conjugation is a field automorphism fixing $\mathbf{Q}$, and the real and imaginary parts of a complex number can be computed using field operations and complex conjugation, a complex number $a + bi$ is algebraic over $\mathbf{Q}$ if and only if a and b are algebraic over $\mathbf{Q}$ (meaning a and b are real algebraic numbers), so the algebraic closure of the real algebraic numbers is obtained by adjoining i. This example is not too different from $\mathbf{R}$, in an algebraic sense. Is there an example that does not look like the reals, such as a field of positive characteristic or a field whose algebraic closure is a finite extension of degree greater than 2?

Amazingly, it turns out that if a field F is not algebraically closed but its algebraic closure C is a finite extension, then in a sense which will be made precise below, F looks like the real numbers. For example, F must have characteristic 0 and $C = F(i)$. This is a theorem of Artin and Schreier. Proofs of the Artin-Schreier theorem can be found in [5, Theorem 11.14] and [6, Corollary 9.3, Chapter VI], although in both cases the theorem is proved only after the development of some general theory that is useful for more than just the Artin-Schreier theorem. In [7] there is an elementary proof (based on the original one, as all known proofs are) under the hypothesis that F has characteristic 0; we essentially reproduce much of that proof below, but we add on some extra details to prove in an elementary manner that F must have characteristic 0.

The prerequisites are a knowledge of basic field theory and Galois theory of finite extensions, including Kummer extensions and Artin-Schreier extensions. Applications of the Artin–Schreier theorem to the Galois theory of infinite extensions will be mentioned after the proof.

2. Some lemmas

Our proof of the Artin-Schreier theorem requires two lemmas.

Lemma 2.1. *Let F be a field of characteristic $\ell > 0$ and $a \in F$. If $a \notin F^\ell$, then $X^{\ell^m} - a$ is irreducible in $F[X]$ for every $m \geq 1$.*

Proof. We will prove the contrapositive: if $X^{\ell^m} - a$ is reducible in $F[X]$ for some $m \geq 1$, then $a \in F^\ell$.

Let $X^{\ell^m} - a = f(X)g(X)$, where f, g are monic in $F[X]$ with positive degree. Let E be an extension field of F containing a root b of $X^{\ell^m} - a$, so in $E[X]$

$$X^{\ell^m} - a = X^{\ell^m} - b^{\ell^m} = (X - b)^{\ell^m}.$$

Since $E[X]$ has unique factorization and f and g are monic, $f(X) = (X-b)^r$ where $0 < r < \ell^m$. Let $r = \ell^t s$ with s not divisible by ℓ (clearly $t < m$). Then

$$f(X) = (X^{\ell^t} - b^{\ell^t})^s,$$

so the coefficient of $X^{\ell^t(s-1)}$ is $-sb^{\ell^t}$. This lies in F, so (since s is nonzero in F) $b^{\ell^t} \in F$. Thus

$$a = (b^{\ell^t})^{\ell^{m-t}} \in F^{\ell^{m-t}} \subset F^\ell.$$

□

Remark 2.2. For a prime number ℓ and a field of arbitrary characteristic, the irreducibility of $X^\ell - a$ over that field implies irreducibility of $X^{\ell^m} - a$ for all $m \geq 1$ except when $\ell = 2$: X^2+4 is irreducible in $\mathbf{Q}[X]$ but $X^4+4 = (X^2+2X+2)(X^2-2X+2)$. See [6, pp. 297–298] for more information about this.

Over a field of characteristic ℓ, the polynomial $X^\ell - X - a$ is a common replacement for $X^\ell - a$ when doing Galois theory, so it's natural to ask about an analogue of Lemma 2.1 for this polynomial: does irreducibility of $X^\ell - X - a$ in characteristic ℓ imply irreducibility of $(X^\ell - X)^{\circ m} - a$ for all $m \geq 1$, where $\circ m$ means "m iterates"? (For comparison, $X^{\ell^m} - a = (X^\ell)^{\circ m} - a$.) The answer is no. For example, $(X^2 - X)^{\circ 3} - 1$ is reducible over $\mathbf{F}_2$ and $(X^3 - X)^{\circ 2} - 1$ is reducible over $\mathbf{F}_3$. An interesting question about degrees of irreducible factors of $(X^p - X)^{\circ m} - 1$ over $\mathbf{F}_p$ is in https://mathoverflow.net/questions/167613.

Lemma 2.3. *Let F be a field in which -1 is not a square (so F does not have characteristic 2), and every element of $F(i)$ is a square in $F(i)$, where $i^2 = -1$. Then every finite sum of squares in F is again a square in F and F has characteristic 0.*

Proof. Without loss of generality we deal with two squares. Let $a, b \in F$. There are $c, d \in F$ such that $a + bi = (c + di)^2$. Thus $a = c^2 - d^2$ and $b = 2cd$, so $a^2 + b^2 = (c^2 + d^2)^2$.

If F has positive characteristic, then -1 is a finite sum of 1's in F (why?). Obviously 1 is a square in F, so therefore the sum, -1, is a square in F, a contradiction. So F must have characteristic 0! □

3. Statement and proof

Theorem 3.1 (Artin-Schreier)**.** *Let C be algebraically closed with F a subfield such that $1 < [C : F] < \infty$. Then $C = F(i)$ where $i^2 = -1$, and F has characteristic 0. Moreover, for $a \in F^\times$, exactly one of a or $-a$ is a square in F, and every finite sum of nonzero squares in F is again a nonzero square in F.*[1]

Proof. First we show C/F is Galois. Since C is algebraically closed, C/F is clearly normal. To show that C/F is separable (hence Galois), assume F has characteristic $\ell > 0$. We will show that $F = F^\ell$, which implies F is a perfect field and therefore the finite extension C/F is separable. Well, if there is some a in F that is not in F^ℓ then Lemma 2.1 implies $X^{\ell^m} - a$ is irreducible in $F[X]$ for all $m \geq 1$. Thus F has algebraic extensions of arbitrarily large degree, a contradiction.

Now we prove $[C : F] = 2$. Let $G = \mathrm{Gal}(C/F)$, so $[C : F] = |G|$. If $|G| > 2$, then $|G|$ is divisible by an odd prime or 4, hence G has a subgroup whose size is an odd prime or 4 (a finite group has a subgroup of size equal to each prime power factor of the size of the group), so C has a subfield K containing F such that $[C : K]$ equals an odd prime or 4.

[1] We exclude the *empty sum*, which is 0.

Using K in place of F, to show $[C : F] = 2$ it suffices to prove $[C : F]$ can't be an odd prime or 4.

Assume $[C : F] = p$, a prime, so $G = \mathrm{Gal}(C/F)$ is cyclic. Let σ be a generator of G, so for $x \in C$, $x \in F$ if and only if $\sigma(x) = x$. We want to show $p = 2$.

First we show F (equivalently, C) does not have characteristic p. Assume the contrary, *i.e.*, F has characteristic p. Since C/F is cyclic of degree p in characteristic p, Artin-Schreier theory tells us that $C = F(\alpha)$ where α is a root of some polynomial $X^p - X - a \in F[X]$. Since $\{1, \alpha, \ldots, \alpha^{p-1}\}$ is an F-basis of C, we can write each b in C as

$$b = b_0 + b_1\alpha + \cdots + b_{p-1}\alpha^{p-1},$$

with the b_j in F. Then

$$\begin{aligned} b^p - b &= \sum_{j=0}^{p-1}(b_j\alpha^j)^p - \sum_{j=0}^{p-1} b_j\alpha^j \\ &= \sum_{j=0}^{p-1}(b_j^p\alpha^{pj} - b_j\alpha^j) \\ &= \sum_{j=0}^{p-1} b_j^p(\alpha + a)^j - b_j\alpha^j \\ &= ((b_{p-1})^p - b_{p-1})\alpha^{p-1} + \text{ lower degree terms in } \alpha. \end{aligned}$$

Since C is algebraically closed, we can choose $b \in C$ such that $b^p - b = a\alpha^{p-1}$, so comparing the coefficient of α^{p-1} on both sides tells us that $b_{p-1} \in F$ is a root of $X^p - X - a$. However, $X^p - X - a$ is irreducible in $F[X]$ (it's the minimal polynomial of α), so we have a contradiction. Thus F does *not* have characteristic $p = [C : F]$.

Therefore, since C is algebraically closed of characteristic different from p, C contains a root of unity of order p; call it ζ. Since $[F(\zeta) : F] \leq p - 1$ and $[C : F] = p$, we must have $[F(\zeta) : F] = 1$, so $\zeta \in F$. That means C/F is cyclic of degree p with F containing a root of unity of order p, so by Kummer theory $C = F(\gamma)$ where $\gamma^p \in F$.

We use γ to show $p = 2$. Choose $\beta \in C$ such that $\beta^p = \gamma$, so $\beta^{p^2} = \gamma^p \in F$. Thus $\beta^{p^2} = \sigma(\beta^{p^2}) = (\sigma(\beta))^{p^2}$, hence $\sigma(\beta) = \omega\beta$ with $\omega^{p^2} = 1$. Thus ω^p is a p-th root of unity (possibly 1; we'll see that it is not 1 next), so ω^p lies in F. If $\omega^p = 1$ then $(\sigma(\beta))^p = \beta^p$, so $\sigma(\beta^p) = \beta^p$. That forces $\beta^p \in F$, but $\beta^p = \gamma$ and γ is not in F, so we have a contradiction. Thus $\omega^p \neq 1$, so ω has order p^2 and ω^p has order p. Since $\omega^p = \sigma(\omega^p) = (\sigma(\omega))^p$, we have for some $k \in \mathbf{Z}$

$$\sigma(\omega) = \omega(\omega^p)^k = \omega^{1+pk}.$$

From the equation $\sigma(\beta) = \omega\beta$, we get

$$\beta = \sigma^p(\beta) = \omega\sigma(\omega)\cdots\sigma^{p-1}(\omega)\beta = \omega^{1+(1+pk)+\cdots+(1+pk)^{p-1}}\beta,$$

so we get the sequence of congruences

$$\begin{aligned} 1 + (1+pk) + \cdots + (1+pk)^{p-1} &\equiv 0 \bmod p^2, \\ \sum_{j=0}^{p-1} (1+jpk) &\equiv 0 \bmod p^2, \\ p + \frac{p(p-1)}{2} \cdot pk &\equiv 0 \bmod p^2, \\ 1 + \frac{p(p-1)}{2} \cdot k &\equiv 0 \bmod p. \end{aligned}$$

This last congruence shows p is not odd, so $p = 2$ and k is odd. Therefore ω has order $p^2 = 4$ and $\sigma(\omega) = \omega^{1+pk} = \omega^{1+2k} = \omega^3$, so $\sigma(\omega) \neq \omega$. Since ω has order 4, $\omega^2 = -1$, so we write ω as i. Thus if $[C : F]$ is a prime, then it equals 2, C doesn't have characteristic 2, and $i \notin F$.

If $[C : F] = 4$ then $\mathrm{Gal}(C/F)$ has size 4, hence a subgroup of size 2, so there is an intermediate field $F \subset K \subset C$ with $[C : K] = 2$. The above reasoning implies i is not in K, hence not in F. But then $F(i)$ is a subfield of C with $[C : F(i)] = 2$ and $F(i)$ contains i, a contradiction.

Hence if F is a nonalgebraically closed field whose algebraic closure C is a finite extension, then $[C : F]$ is not divisible by an odd prime or 4, so $[C : F] = 2$, F does not have characteristic 2, and i is not in F. Thus $C = F(i)$.

From $C = F(i)$ and $C \neq F$, Lemma 2.3 shows that the squares in F are closed under addition and F has characteristic 0. To establish the last part of the theorem, assume a and $-a$ are both not squares in F. Then $\sqrt{a}$ and $\sqrt{-a}$ each generate the quadratic extension C/F. From $C = F(\sqrt{a}) = F(\sqrt{-a})$, the ratio $-a/a = -1$ must be a square in F, but $i \notin F$ so we have a contradiction. Thus one of a or $-a$ is a square in F, but not both since i is not in F. If $n \geq 2$ and $b_1, \ldots, b_n$ are nonzero in F, then by Lemma 2.3 the sum $b_1^2 + \cdots + b_n^2$ is a square in F. If it is zero, then $-1 = (b_2/b_1)^2 + \cdots + (b_n/b_1)^2$ is a sum of squares in F, hence -1 is a square in F, a contradiction. So $b_1^2 + \cdots + b_n^2$ is a nonzero square in F. □

Note that by defining a "positive" element of F to be a nonzero square, we see that for every nonzero $a \in F$, exactly one of a or $-a$ is positive, each finite sum of positive elements is positive, and positive elements are clearly closed under multiplication. So F admits a notion of ordering (define $a < b$ in F if $b - a$ is a nonzero square), much like the real numbers, where the positive elements are the nonzero squares.

Comparing the proof of the Artin–Schreier theorem with accounts of it in books, our only new twist (which surely is not really new, but it wasn't in the books that I checked) is the method of showing that the field has characteristic 0. The proofs in [5] and [6] follow the original proof of Artin and Schreier on this issue by studying certain cyclotomic extensions of the prime field, $\mathbf{Q}$ or $\mathbf{F}_p$, to rule out the case of prime field $\mathbf{F}_p$. Artin and Schreier first proved their theorem in 1926 under the hypothesis that F has characteristic 0 [1, Satz 4]. They indicated in a footnote that the characteristic 0 hypothesis could be removed, and they did so in a paper the following year [2, Satz 4].

4. Applications

The Artin-Schreier theorem tells us something about automorphisms of finite order in the (infinite) group of field automorphisms of the algebraic numbers. One element of finite

order is complex conjugation (restricted to the algebraic numbers), which has order 2. Let's show that *every* nonidentity automorphism of finite order has order 2.

It is no harder to deal with a more general situation, which we now do. Let K be a field, G be a finite group of automorphisms of K, and K^G the fixed field of K under G. The main theorem in Artin's development of Galois theory says K/K^G is a finite Galois extension with Galois group G. In particular, if C is algebraically closed and G is a nontrivial finite group of automorphisms of C, then $[C : C^G] = |G| > 1$, so by the Artin-Schreier theorem $|G| = 2$ and the fixed field C^G "looks like" the real numbers.

Taking G to be the finite group generated by an element of finite order in $\mathrm{Aut}(C)$, we see that the only nontrivial torsion in the automorphism group of an algebraically closed field is 2-torsion, and if there is 2-torsion then the field has characteristic 0. Two noncommuting torsion elements in $\mathrm{Aut}(C)$ have a product whose order is neither 1 nor 2 (this is just basic group theory), hence this product must have infinite order by the Artin-Schreier theorem. So for a field K, with algebraic closure $\overline{K}$, the torsion elements of $\mathrm{Aut}(\overline{K}/K) \subset \mathrm{Aut}(\overline{K})$ have order 1 or 2 and a pair of noncommuting torsion elements have a product with infinite order. In particular, each nontrivial torsion element of the automorphism group of $\overline{\mathbf{Q}}$ has order 2, and if C is an algebraically closed field of positive characteristic then $\mathrm{Aut}(C)$ is torsion-free.

With some more work (see [3] or [4, Exercise 32, Chapter VI §2]), it can be shown that $\overline{\mathbf{Q}}$ is characterized up to field isomorphism among all algebraically closed fields by the fact that its automorphism group has nontrivial torsion elements and they are all *conjugate* (this property is more important for number theory than it might appear at first glance).

As a final remark, note that we really don't need C to be algebraically closed in the Artin-Schreier theorem to conclude that $C = F(i)$. We only need that certain special polynomials in $C[X]$ have a root in C. More precisely, the properties we used of C were:

(1) If there is a subfield $F \subset C$ such that $1 < [C : F] < \infty$, then C/F is Galois.
(2) If p is prime and C has characteristic p, then $X^p - X - a$ has a root in C for all a in C. Note $X^p - X - a$ has no multiple roots in characteristic p.
(3) If p is a prime and C does not have characteristic p, then C contains a root of unity of order p.

Assuming $[C : F]$ is prime, the fourth paragraph in the proof of the Artin-Schreier theorem used properties 1 and 2 to show that $\mathrm{char}(F) = \mathrm{char}(C) \neq [C : F]$. The rest of the proof used property 3 to show that $[C : F] = 2$ and $C = F(i)$.

This observation is important because C does not have to be algebraically closed to satisfy properties 1, 2, and 3. For example, it suffices to assume only that C is *separably* closed, *i.e.*, that C has no proper separable algebraic extensions. So if a field S is a separable closure of a field F with $1 < [S : F] < \infty$, the proof of the Artin-Schreier theorem with S in the role of C yields $\mathrm{char}(S) \neq 2$ (so each element of S is a square in S since S is separably closed) and $S = F(i)$. Then Lemma 2.3 implies that $\mathrm{char}(S) = 0$, so S is algebraically closed after all! This leads to the following analogue of the Artin–Schreier theorem for separable closures.

Theorem 4.1. *Let F be a field whose separable closure is a finite proper extension (so its algebraic closure is a priori possibly an infinite extension). Then F has characteristic 0 and its algebraic (= separable) closure is the quadratic extension $F(i)$. Moreover, for $a \in F^\times$, exactly one of a or $-a$ is a square in F, and every finite sum of nonzero squares in F is again a nonzero square in F.*

The last part of this theorem, about nonzero squares, follows from $F(i)$ being algebraically closed by the last paragraph of the proof of Theorem 3.1.

Separable closures are contained in algebraic closures, so does Theorem 4.1 imply Theorem 3.1? Yes, with a little bit of extra work. Suppose F is a field with algebraic closure C and $1 < [C : F] < \infty$.

Case 1: F is separably closed. Since $F \neq C$, F is not a perfect field (for perfect fields, separable and algebraic closures match), so F has a positive characteristic ℓ with $F \neq F^\ell$, and then Lemma 2.1 tells us that there are irreducibles in $F[X]$ of arbitrarily large degree, so the algebraic closure of F is not a finite extension. This is a contradiction, so this case doesn't arise.

Case 2: F is not separably closed. The separable closure of F is a finite proper extension of F. Calling it S, Theorem 4.1 tells us that $S = F(i)$ and F has characteristic 0, which implies $C = S = F(i)$ and everything else we want in Theorem 3.1.

References

[1] E. Artin and O. Schreier, *Algebraische Konstruktion reeller Körper*, pp. 258–272 in: Artin's Collected Papers (Ed. S. Lang and J. Tate), Springer-Verlag, New York, 1965.

[2] E. Artin and O. Schreier, *Eine Kennzeichnung der reell abgeschlossenen Körper*, pp. 289–295 in: Artin's Collected Papers (Ed. S. Lang and J. Tate), Springer-Verlag, New York, 1965.

[3] R. Baer, *Die Automorphismengruppe eines algebraisch abgeschlossen Körpers der Charakteristik* 0, Math. Zeit. **117**, 1970, 7-17.

[4] N. Bourbaki, "Algebra II," Springer-Verlag, New York, 1990.

[5] N. Jacobson, "Basic Algebra II," 2nd ed., W. H. Freeman and Co., New York, 1989.

[6] S. Lang, "Algebra," 3rd revised ed., Springer-Verlag, New York, 2002.

[7] R. Guralnick and M. Miller, *Subfields of Algebraically Closed Fields*, Math. Mag. **50**, 1977, 260-261.