SEPARABILITY

KEITH CONRAD

1. Introduction

Let K be a field. We are going to look at concepts related to K that fall under the label "separable".

Definition 1.1. A nonzero polynomial $f(X) \in K[X]$ is called *separable* when it has distinct roots in a splitting field over K. That is, each root of $f(X)$ has multiplicity 1. If $f(X)$ has a multiple root then $f(X)$ is called *inseparable*.

Example 1.2. In $\mathbf{R}[X]$, the polynomial $X^2 - X$ is separable since its roots are 0 and 1 and $X^3 - 2$ is separable since there are 3 different cube roots of 2 in the complex numbers. In $\mathbf{F}_3[X]$ the polynomial $X^3 - 2$ is inseparable because $X^3 - 2 = (X + 1)^3$ in $\mathbf{F}_3[X]$ so it has a triple root.

The term "separable" comes from distinctness of the roots: they are separate in the sense that there are no multiple roots.

Definition 1.3. If α is algebraic over K, it is called *separable over* K when its minimal polynomial in $K[X]$ is separable in the sense of Definition 1.1: the minimal polynomial of α in $K[X]$ has distinct roots in a splitting field over K. If the minimal polynomial of α in $K[X]$ is inseparable then α is called *inseparable over* K.

Example 1.4. The real numbers $\sqrt{2}$ and $\sqrt{3}$ are both separable over $\mathbf{Q}$, as they have minimal polynomials $X^2 - 2$ and $X^2 - 3$ in $\mathbf{Q}[X]$, which are both separable.

From Definition 1.1, checking a polynomial is separable requires building a splitting field to check the roots are distinct. But we will see in Section 2 a criterion for deciding when a polynomial is separable (that is, has no multiple roots) without having to work in a splitting field. In Section 3 we will define what it means for a field extension to be separable and then prove the primitive element theorem, an important result about separable field extensions.

2. Separable Polynomials

Although the definition of a separable polynomial in $K[X]$ involves how the polynomial factors over a splitting field, we can use differentiation on $K[X]$ to describe the separability condition without leaving $K[X]$:

Theorem 2.1. *A nonzero polynomial in $K[X]$ is separable if and only if it is relatively prime to its derivative in $K[X]$.*

Proof. Let $f(X)$ be a nonzero polynomial in $K[X]$. Suppose $f(X)$ is separable, and let α be a root of $f(X)$ (in some extension of K). Then $f(X) = (X - \alpha)h(X)$, with $h(\alpha) \neq 0$. Since $f'(\alpha) = h(\alpha) \neq 0$, α is not a root of $f'(X)$. Therefore $f(X)$ and $f'(X)$ have no common roots, so they have no common factors in $K[X]$: they are relatively prime.

Now suppose $f(X)$ is not separable, so by definition it has a repeated root (in a splitting field over K). This root is also a root of $f'(X)$: when $f(X) = (X-\alpha)^2 g(X)$, the product rule shows

$$f'(X) = (X-\alpha)^2 g'(X) + 2(X-\alpha)g(X),$$

so $f'(\alpha) = 0$. Since $f(X)$ and $f'(X)$ have α as a common root, they are both divisible by the minimal polynomial of α in $K[X]$. In particular, $f(X)$ and $f'(X)$ are not relatively prime in $K[X]$. Taking the contrapositive, if $f(X)$ and $f'(X)$ are relatively prime in $K[X]$ then $f(X)$ has no repeated root so it is separable. □

When we are given a specific $f(X)$, whether or not $f(X)$ and $f'(X)$ are relatively prime can be checked by Euclid's algorithm for polynomials.

Example 2.2. In $\mathbf{F}_3[X]$, let $f(X) = X^6 + X^5 + X^4 + 2X^3 + 2X^2 + X + 2$. Using Euclid's algorithm in $\mathbf{F}_3[X]$ on $f(X)$ and $f'(X)$,

$$\begin{aligned} f(X) &= f'(X)(2X^2 + X) + (2X^2 + 2) \\ f'(X) &= (2X^2 + 2)(X^2 + 2X + 2), \end{aligned}$$

so $(f(X), f'(X)) = 2X^2+2$ (which is the same as X^2+1 up to scaling). The greatest common divisor is nonconstant, so $f(X)$ is inseparable. In fact, $f(X) = (X^2+1)^2(X^2+X+2)$. Notice we were able to detect that $f(X)$ has a repeated root *before* we gave its factorization.

Example 2.3. In $\mathbf{Q}[X]$, let $f(X) = X^4 - 3X - 2$. Using Euclid's algorithm in $\mathbf{Q}[X]$ on $f(X)$ and $f'(X)$,

$$\begin{aligned} f(X) &= f'(X) \cdot \left(\frac{1}{4}X\right) - \frac{9}{4}X - 2 \\ f'(X) &= \left(-\frac{9}{4}X - 2\right)\left(-\frac{16}{9}X^2 + \frac{128}{81}X - \frac{1024}{729}\right) - \frac{4235}{729}. \end{aligned}$$

We have reached a remainder that is a nonzero constant, so $f(X)$ and $f'(X)$ are relatively prime. Therefore $f(X)$ is separable over $\mathbf{Q}$.

There are some important *families* of polynomials where relative primality with the derivative can be checked directly, not using Euclid's algorithm.

Example 2.4. Let $f(X) = X^n - a$ where $a \in K^\times$. The derivative of $f(X)$ is nX^{n-1}. If $n = 0$ in K (that is K has characteristic p and $p \mid n$) then $f'(X) = nX^{n-1} = 0$ and $(f(X), f'(X)) = f(X)$ is nonconstant, so $X^n - a$ is not separable. If $n \neq 0$ in K (that is, K has characteristic 0 or K has characteristic p and p doesn't divide n), then $f'(X) = nX^{n-1}$ is nonzero and $(X^n - a, nX^{n-1}) = 1$ since X doesn't divide $X^n - a$ (because $a \neq 0$). Therefore $X^n - a$ is separable in $K[X]$ if and only if $n \neq 0$ in K.

In particular, $X^n - 1$ is separable over K if and only if $n \neq 0$ in K. So in $\mathbf{F}_2[X]$, $X^3 - 1$ and $X^9 - 1$ are separable but $X^6 - 1$ is inseparable; in fact, $X^6 - 1 = (X^3 - 1)^2$ in $\mathbf{F}_2[X]$.

Example 2.5. If K has characteristic p and $a \in K$, the polynomial $X^p - X - a$ has derivative -1, a nonzero constant, so $X^p - X - a$ is separable. More generally, if $g(X^p) \in K[X^p]$ is an *arbitrary* polynomial in X^p and $c \in K^\times$ then $g(X^p) + cX$ has derivative c so $g(X^p) + cX$ is separable in $K[X]$. Thus $X^{p^d} - X$ and $X^{p^2} + aX^p + bX$ ($b \neq 0$) are separable in characteristic p.

As practice using the derivative criterion of Theorem 2.1 in proofs, we verify two properties of separable polynomials that are obvious if you think about how polynomials factor in a splitting field.

Corollary 2.6. *If $f(X) \in K[X]$ is separable and $L \supset K$ then every factor of $f(X)$ in $L[X]$ is also separable. An element in an extension of L that is separable over K is also separable over L.*

Proof. Let $g(X)$ be a factor of $f(X)$, say $f(X) = g(X)h(X)$ in $L[X]$. Since $f(X)$ is separable we can write $1 = f(X)u(X) + f'(X)v(X)$ for some polynomials $u(X)$ and $v(X)$ in $K[X]$. Then

$$\begin{aligned} 1 &= (g(X)h(X))u(X) + (g(X)h'(X) + g'(X)h(X))v(X) \\ &= g(X)(h(X)u(X) + h'(X)v(X)) + g'(X)(h(X)v(X)). \end{aligned}$$

The last expression shows a polynomial-linear combination of $g(X)$ and $g'(X)$ equals 1, so $g(X)$ is separable.

Suppose α is in an extension of L and it is separable over K. Since its minimal polynomial in $L[X]$ divides its minimal polynomial in $K[X]$, separability of α over K implies separability of α over L by what we just showed about separable polynomials. □

Corollary 2.7. *Let K and L be fields. If $\sigma\colon K \to L$ is a field embedding, then a polynomial $f(X) \in K[X]$ is separable if and only if $(\sigma f)[X] \in L[X]$ is separable.*

Proof. Assume $f(X)$ is separable, so by Theorem 2.1 we can write

$$f(X)u(X) + f'(X)v(X) = 1$$

for some $u(X)$ and $v(X)$ in $K[X]$. Applying σ to coefficients is a ring embedding $K[X] \to L[X]$ and $\sigma(f') = (\sigma f)'$, so

$$(\sigma f)(X)(\sigma u)(X) + (\sigma f)'(X)(\sigma v)(X) = 1.$$

Therefore $(\sigma f)(X)$ and its derivative are relatively prime in $L[X]$, so $(\sigma f)(X)$ is separable.

Now assume $f(X)$ is inseparable, so some nonconstant $d(X)$ in $K[X]$ divides $f(X)$ and $f'(X)$ in $K[X]$. Then $(\sigma d)(X)$ is nonconstant and divides $(\sigma f)(X)$ and $\sigma(f'(X)) = (\sigma f)'(X)$ in $L[X]$, so $(\sigma f)(X)$ and its derivative are not relatively prime and thus $(\sigma f)(X)$ is inseparable. □

For *irreducible* polynomials, Theorem 2.1 can be refined to a separability criterion that is much simpler to check than relative primality with the derivative:

Theorem 2.8. *For every field K, an irreducible polynomial in $K[X]$ is separable if and only if its derivative is not 0 in $K[X]$. In particular, when K has characteristic 0 every irreducible in $K[X]$ is separable and when K has characteristic p, an irreducible in $K[X]$ is separable if and only if it is not a polynomial in X^p.*

Proof. Let $\pi(X)$ be irreducible in $K[X]$. Separability is equivalent to $(\pi(X), \pi'(X)) = 1$ by Theorem 2.1. If $\pi(X)$ and $\pi'(X)$ are not relatively prime, then $\pi(X) \mid \pi'(X)$ since $\pi(X)$ is irreducible. Taking the derivative drops degrees, so having $\pi'(X)$ be divisible by $\pi(X)$ forces $\pi'(X) = 0$. Conversely, if $\pi'(X) = 0$ then $(\pi(X), \pi'(X)) = \pi(X)$ is nonconstant, so $\pi(X)$ is inseparable by Theorem 2.1. Thus separability of $\pi(X)$ is equivalent to $\pi'(X) \neq 0$.

When K has characteristic 0, every irreducible in $K[X]$ has nonzero derivative since every nonconstant polynomial has nonzero derivative. So all irreducibles in $K[X]$ are separable.

Now suppose K has characteristic p. If there is an irreducible $\pi(X) \in K[X]$ that is not separable, then $\pi'(X) = 0$. Writing $\pi(X) = X^n + c_{n-1}X^{n-1} + \cdots + c_1X + c_0$, the condition $\pi'(X) = 0$ means $ic_i = 0$ in K for $0 \leq i \leq n$ (taking $c_n = 1$). This implies $p \mid i$ whenever $c_i \neq 0$, so the only nonzero terms in $\pi(X)$ occur in degrees divisible by p. In particular, $n = \deg \pi(X)$ is a multiple of p, say $n = pm$. Write each exponent of a nonzero term in $\pi(X)$ as a multiple of p:

$$\pi(X) = X^{pm} + c_{p(m-1)}X^{p(m-1)} + \cdots + c_pX^p + c_0 = g(X^p),$$

where $g(X) \in K[X]$. So $\pi(X) \in K[X^p]$. Conversely, if $\pi(X) = g(X^p)$ is a polynomial in X^p then $\pi'(X) = g'(X^p)pX^{p-1} = 0$, so $\pi(X)$ is inseparable if it is irreducible in $K[X]$. □

Example 2.9. Every irreducible in $\mathbf{Q}[X]$ is separable since $\mathbf{Q}$ has characteristic 0.

Example 2.10. Let $K = \mathbf{F}_3(u)$ be a rational function field over $\mathbf{F}_3$. The polynomial $X^7 + u^2X^5 + u \in K[X]$ is irreducible (by Eisenstein) and separable since its derivative is $X^6 + 2u^2X^4$, which is nonzero.

Theorem 2.8 tells us the phenomenon of *irreducible* polynomials being inseparable (that is, admitting repeated roots) is a purely characteristic p phenomenon.

Example 2.11. Let p be prime and $K = \mathbf{F}_p(u)$ be a rational function field over $\mathbf{F}_p$. The polynomial $X^p - u \in K[X]$ is irreducible since it is Eisenstein at u. Because $(X^p - u)' = 0$, Theorem 2.8 says $X^p - u$ is inseparable. We can also check inseparability of $X^p - u$ directly from the definition, as follows. If α is a root of $X^p - u$ (in some extension of $K = \mathbf{F}_p(u)$), then $\alpha^p = u$, so $X^p - u = X^p - \alpha^p = (X - \alpha)^p$. Thus $X^p - u$ is irreducible with degree p and has only one root, with multiplicity $p > 1$.

Example 2.12. Let $K = \mathbf{F}_3(u)$. The polynomial $X^6 + uX^3 + u \in K[X]$ is irreducible (by Eisenstein) and it is a polynomial in X^3 in characteristic 3, so it is not separable. If r and s are roots of $X^2 + uX + u$, then

$$X^2 + uX + u = (X - r)(X - s) \Longrightarrow X^6 + uX^3 + u = (X^3 - r)(X^3 - s).$$

and the two factors $X^3 - r$ and $X^3 - s$ each have only one root (of multiplicity 3), so $X^6 + uX^3 + u$ has only two roots (each of multiplicity 3).

The roots of an inseparable irreducible polynomial in characteristic p all have multiplicity a common power of p. We won't be using this. Further details about this are in Appendix A.

3. Separable Extensions

We have defined what it means for a polynomial and an element of a field extension to be separable. Now we will define what it means for a field extension to be separable.

Definition 3.1. A finite extension L/K is called *separable* if every element of L is separable over K. When L/K is not separable, it is called *inseparable*.

Example 3.2. Every finite extension of $\mathbf{Q}$ is separable since all irreducible polynomials in $\mathbf{Q}[X]$ are separable.

Example 3.3. Let $K = \mathbf{F}_p(u)$ be a rational function field over $\mathbf{F}_p$ and $L = K(\alpha)$, where α is a root of $X^p - u$. Since $X^p - u$ is inseparable irreducible, α is inseparable over K so the extension L/K is inseparable: it contains an element that is inseparable over K.

Remark 3.4. A finite extension L/K is inseparable when at least one element of L is inseparable over K. This does *not* mean all elements are inseparable over K. The distinction between separable and inseparable field extensions is like that between abelian and nonabelian groups: in an abelian group every pair of elements commutes, but in a nonabelian group only some (but not all) elements don't commute.

The rest of this section is devoted to proving the following two important theorems about finite separable extensions.

Theorem 3.5. *Let L/K be a finite extension and write $L = K(\alpha_1, \ldots, \alpha_r)$. Then L/K is separable if and only if each α_i is separable over K.*

The usefulness of Theorem 3.5 is that it gives a practical way to check a finite extension L/K is separable: rather than show every element of L is separable over K it suffices to show there is a set of field generators for L/K that are each separable over K.

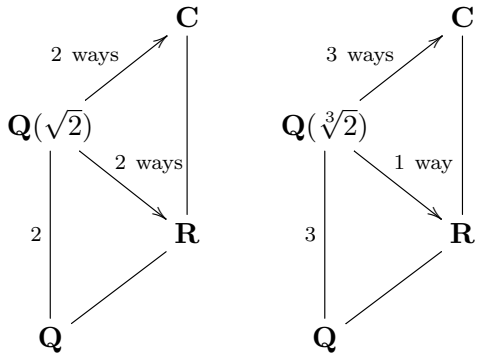
Theorem 3.6 (Primitive Element Theorem)**.** *Every finite separable extension of K has the form $K(\gamma)$ for some γ.*

When K has characteristic 0, all of its finite extensions are separable, so the primitive element theorem says every finite extension field of K has the form $K(\gamma)$ for some γ.

Example 3.7. The field $\mathbf{Q}(\sqrt{2}, \sqrt{3})$ is separable over $\mathbf{Q}$ and it equals $\mathbf{Q}(\sqrt{2} + \sqrt{3})$.

There are finite extensions that do not admit a primitive element [1, Example 2, p. 595], but such examples are not of immediate interest to us and we don't discuss them.

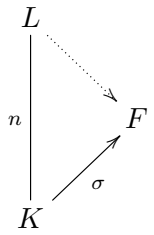
Our proofs of Theorems 3.5 and 3.6 will both depend on a connection between separable extensions and counting field embeddings (Theorem 3.8 below). To see there should be such a connection, here are two examples.



To embed the field $\mathbf{Q}(\sqrt{2})$ into $\mathbf{R}$, there are two ways this can be done: send $\sqrt{2}$ to itself or send it to $-\sqrt{2}$. That there are two embeddings is related to the fact that $X^2 - 2$ has two different roots in $\mathbf{R}$. Similarly, there are two embeddings of $\mathbf{Q}(\sqrt{2})$ into $\mathbf{C}$. If we try to embed $\mathbf{Q}(\sqrt[3]{2})$ into $\mathbf{R}$, there is only one way to do this since there is only one real cube root of 2. Enlarging our target field to $\mathbf{C}$ provides us with 3 different cube roots of 2 (one is real, two are non-real), so $\mathbf{Q}(\sqrt[3]{2})$ has 3 different embeddings into the complex numbers (determined by sending $\sqrt[3]{2}$ to each of the 3 cube roots of 2 in $\mathbf{C}$). The number of embeddings $\mathbf{Q}(\sqrt[3]{2}) \to \mathbf{C}$ is 3, but we had to make the target field large enough (target field $\mathbf{R}$ was too small). The number of embeddings of $\mathbf{Q}(\sqrt{2})$ and $\mathbf{Q}(\sqrt[3]{2})$ into $\mathbf{R}$ and $\mathbf{C}$ is related to the number of different roots of $X^2 - 2$ and $X^3 - 2$ in $\mathbf{R}$ and $\mathbf{C}$. That the

number of roots equals the degree of each polynomial when they split completely is related to the polynomials being *separable*.

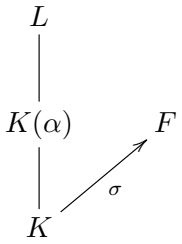
Theorem 3.8. *Let L/K be a finite extension of fields with $[L : K] = n$ and $\sigma\colon K \to F$ be a field embedding.*



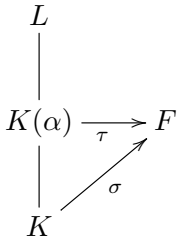
(a) *The number of extensions of σ to an embedding $L \to F$ is at most n.*
(b) *If L/K is inseparable then the number of extensions of σ to an embedding $L \to F$ is less than n.*
(c) *If L/K is separable then then there is a field $F' \supset F$ such that the number of extensions of σ to an embedding $L \to F'$ is equal to n.*

Notice parts a and b give the field degree $[L : K]$ as an upper bound on the number of extensions of σ, while part c says that if we make the target field large enough the number of extensions of σ matches the field degree. This is exactly what happens with $\mathbf{Q}(\sqrt[3]{2})/\mathbf{Q}$, where 3 embeddings are possible into $\mathbf{C}$ but not into $\mathbf{R}$.

Proof. a) We will argue by induction on $n = [L : K]$. If $n = 1$ then $L = K$ and the result is clear. Now suppose $n > 1$. Pick $\alpha \in L$ with $\alpha \notin K$. Our field diagram looks like the following.



To bound the number of extensions of σ to an embedding of L into F, we first bound the number of extensions of σ to an embedding $\tau\colon K(\alpha) \to F$ and then bound the number of extensions of each such τ to an embedding $L \to F$.



From the proof that two splitting fields of a polynomial are isomorphic, the number of τ's extending σ is the number of roots in F of $(\sigma\pi)(X)$, where $\pi(X)$ is the minimal polynomial of α in $K[X]$. The number of these roots is *at most* the degree of $(\sigma\pi)(X)$, which equals

$\deg \pi = [K(\alpha) : K]$. (This upper bound could be strict for two reasons: $(\sigma\pi)(X)$ might not split in $F[X]$ or it could split but be inseparable; the second case will be relevant in (b).)

Once we have extended σ to some τ on $K(\alpha)$, we count how many ways τ extends to L. As in the proof that splitting fields are isomorphic, the trick is to consider $K(\alpha)$ as the new base field, with τ playing the role of σ. Since $\alpha \notin K$, $[L : K(\alpha)] < [L : K]$, so by induction on the field degree the number of extensions of $\tau\colon K(\alpha) \to F$ to an embedding of L into F is at most $[L : K(\alpha)]$. Multiplying the upper bounds on the number of extensions of σ to $K(\alpha)$ and the number of further extensions up to L, the number of extensions of σ to L is at most

$$[L : K(\alpha)][K(\alpha) : K] = [L : K],$$

so by induction we're done.

b) When L/K is inseparable, some $\alpha \in L$ is inseparable over K. Running through the first part of the proof of (a) with this α, its minimal polyomial $\pi(X)$ in $K[X]$ is inseparable, so $(\sigma\pi)(X)$ is inseparable in $F[X]$ (Corollary 2.7). This inseparability forces the number of extensions of σ to $K(\alpha)$ to be *less* than $[K(\alpha) : K] = \deg \pi$. By (a), the number of extensions up to L of a field embedding $K(\alpha) \to F$ is at most $[L : K(\alpha)]$, so the number of extensions of σ to L is strictly less than

$$[L : K(\alpha)][K(\alpha) : K] = [L : K].$$

c) Write $L = K(\alpha_1, \ldots, \alpha_r)$, with each α_i separable over K. We want to construct a field $F' \supset F$ such that $\sigma\colon K \to F$ has $[L : K]$ extensions to embeddings of L into F'. We will argue in a similar way to (a), but replacing F with some larger F' will let the upper bound on the number of embeddings in the proof of (a) be reached.

For $1 \leq i \leq r$, let $\pi_i(X)$ be the minimal polynomial of α_i in $K[X]$, so each $\pi_i(X)$ is separable. Take for F'/F an extension in which each $(\sigma\pi_i)(X) \in F[X]$ splits. We will show there are $[L : K]$ extensions of σ to an embedding of L into F'. If $[L : K] = 1$ then the result we want is clear, so we may suppose $L \neq K$.

Some α_i is not in K, without loss of generality say $\alpha_1 \notin K$. By the same reasoning as in the proof of (a), the number of extensions of σ to an embedding $K(\alpha_1) \to F'$ is the number of roots of $(\sigma\pi_1)(X)$ in F'. The polynomial $(\sigma\pi_1)(X)$ is separable in $F[X]$ (Corollary 2.7) and splits in $F'[X]$ by the definition of F'. Therefore σ has $[K(\alpha_1) : K]$ extensions to embeddings $K(\alpha_1) \to F'$. If $L = K(\alpha_1)$ we are done, so we may suppose $L \neq K(\alpha_1)$.

Now, as in the inductive step of (a), we get ready to take $K(\alpha_1)$ as our new base field. Write $L = K(\alpha_1)(\alpha_2, \ldots, \alpha_r)$ and pick an embedding $\tau\colon K(\alpha_1) \to F'$ extending σ. (We just saw there are $[K(\alpha_1) : K]$ choices for τ.) For $i = 2, \ldots, r$, let $m_i(X)$ be the minimal polynomial of α_i in $K(\alpha_1)[X]$, so $m_i(X) \mid \pi_i(X)$. Therefore $(\tau m_i)(X)$ divides $(\tau\pi_i)(X) = (\sigma\pi_i)(X)$. (Here we need that $\pi_i(X)$ has coefficients in K.) Since $(\sigma\pi_i)(X)$ is separable and splits in $F'[X]$, the same is true of its factor $(\tau m_i)(X)$ (Corollary 2.6).

Thus $L/K(\alpha_1)$ and the embedding $\tau\colon K(\alpha_1) \to F'$ have similar properties to L/K and the embedding $\sigma\colon K \to F'$: the extension has field generators α_i that are separable over $K(\alpha_1)$ and each $(\tau m_i)(X)$ splits in $F'[X]$. One property that is not the same is the field degree: $[L : K(\alpha_1)] < [L : K]$. Since the degree is smaller, by induction we can say $\tau\colon K(\alpha_1) \to F'$ extends in $[L : K(\alpha_1)]$ ways to an embedding $L \to F'$. So the number of extensions of $\sigma\colon K \to F'$ to L is

$$[L : K(\alpha_1)][K(\alpha_1) : K] = [L : K].$$

□

Now we are ready to prove Theorems 3.5 and 3.6.

Proof. (of Theorem 3.5) If L/K is separable then each α_i is separable over K by the definition of separable extensions. Conversely, assume each α_i is separable over K. We want to show every element of L is separable over K.

The proof of Theorem 3.8c does not use the full strength of the hypothesis there that L/K is separable, but *only* that there is a set of field generators $\alpha_1, \ldots, \alpha_r$ for L/K that are each separable over K. In that proof, the set of field generators was never enlarged later on, so the proof of Theorem 3.8c applies to our current extension L/K, using $F = K$ and $\sigma = \mathrm{id}_K$, and it shows L admits $[L : K]$ embeddings into *some* field extension of K. This property is not true of inseparable extensions of K (Theorem 3.8b) so L/K must be separable, *i.e.*, every element of L must be separable over K. □

The proof of Theorem 3.6 will have two cases: K finite and K infinite. The next lemma will be used in the finite case.

Lemma 3.9. *If F is a finite field, the group $F^\times$ is cyclic.*

Proof. Let $q = |F|$, so $|F^\times| = q - 1$. We want to show some element of $F^\times$ has order $q-1$. We will appeal to a property of finite abelian groups (having in mind the group $F^\times$): the order of every element in a finite abelian group divides the maximum order of all the elements.

Let m be the maximum order of the elements of $F^\times$. Since the order of every element of $F^\times$ divides m, each $\alpha \in F^\times$ satisfies $\alpha^m = 1$, so the polynomial $X^m - 1$ has all $q - 1$ elements of $F^\times$ as roots. The number of roots of a polynomial over a field is bounded above by its degree, so $q - 1 \leq m$. At the same time, $m \mid (q - 1)$ by Lagrange's theorem, so we must have $m = q - 1$. This means there is an element of $F^\times$ with order $q - 1 = |F^\times|$, so $F^\times$ is cyclic. □

Note Lemma 3.9 is nonconstructive. This is even true in the special case $F = \mathbf{Z}/(p)$. To this day, there is no algorithm to find a generator for $(\mathbf{Z}/(p))^\times$ that runs substantially faster than just trying $2, 3, \ldots$ in succession until a generator is stumbled upon.

Proof. (of Primitive Element Theorem) If K is a finite field then every finite extension L is a finite field. Therefore $L^\times$ is cyclic by Lemma 3.9. Letting γ be a generator of $L^\times$, we have $L^\times = \langle \gamma \rangle$, so $L = K(\gamma)$.

Now consider the case when K is infinite. A finite separable extension of K has the form $K(\alpha_1, \ldots, \alpha_r)$ where each α_i is separable over K. It suffices by induction on the number of field generators to show when $K(\alpha, \beta)/K$ is separable that $K(\alpha, \beta) = K(\gamma)$ for some γ.

Let $L = K(\alpha, \beta)$ and $n = [L : K]$. Recall that a K-homomorphism is a homomorphism of extensions of K that fixes K pointwise. Since L/K is separable, Theorem 3.8c (using $F = K$ and $\sigma = \mathrm{id}_K$) tells us there is a field extension F'/K such that the number of K-homomorphisms $L \to F'$ is n. Pick $c \in K$. If $K(\alpha + c\beta) \neq L$ then $[K(\alpha + c\beta) : K] < [L : K]$. We will show there are only finitely many such c, so on account of K being infinite there is a $c \in K$ such that $K(\alpha + c\beta) = L$. (In practice, $c = 1$ often works: usually $K(\alpha, \beta) = K(\alpha + \beta)$.)

The degree $[K(\alpha + c\beta) : K]$ is an upper bound on the number of K-homomorphisms $K(\alpha + c\beta) \to F'$. Since there are n K-homomorphisms $L \to F'$, if $[K(\alpha + c\beta) : K] < [L : K] = n$ then there are different K-homomorphisms $L \to F'$, say σ and τ, which are equal on $K(\alpha + c\beta)$. Therefore $\sigma \neq \tau$ on L but $\sigma(\alpha + c\beta) = \tau(\alpha + c\beta)$. Then

$\sigma(\alpha)+c\sigma(\beta) = \tau(\alpha)+c\tau(\beta)$. If $\sigma(\beta) = \tau(\beta)$ then we get $\sigma(\alpha) = \tau(\alpha)$, so $\sigma = \tau$ as functions on $K(\alpha, \beta) = L$, which isn't true. Hence $\sigma(\beta) \neq \tau(\beta)$, so we can solve for c:

$$c = \frac{\tau(\alpha) - \sigma(\alpha)}{\sigma(\beta) - \tau(\beta)}.$$

There are only finitely many σ and τ, so only finitely many such c. □

Here are three uses of Theorem 3.5.

Corollary 3.10. *If α is separable over K then every element of $K(\alpha)$ is separable over K.*

Proof. This is the special case $r = 1$ of Theorem 3.5. □

Corollary 3.11. *If $f(X) \in K[X]$ is separable then a splitting field for f over K is separable over K.*

Proof. Let L/K be a splitting field for f over K. Then $L = K(\gamma_1, \ldots, \gamma_n)$ where the γ_i's are all roots of $f(X)$. Therefore the γ_i's are separable over K, so L/K is a separable extension. □

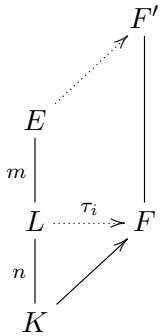
Corollary 3.12. *If the finite extension L/K contains intermediate fields E_1 and E_2 that are both separable over K then their composite E_1E_2 is separable over K. In particular, the set of elements of L that are separable over K forms a subfield of L.*

Proof. Write $E_1 = K(\gamma_1)$ and $E_2 = K(\gamma_2)$ by the primitive element theorem. Then $E_1E_2 = K(\gamma_1, \gamma_2)$, which is separable over K by Theorem 3.5. If α and β are separable over K then we can take $E_1 = K(\alpha)$ and $E_2 = K(\beta)$ to see $K(\alpha, \beta)$ is separable over K. This extension contains $\alpha \pm \beta$, $\alpha\beta$, and $1/\alpha$ if $\alpha \neq 0$, so separability of elements is preserved under field operations. □

Theorem 3.13. *If $E/L/K$ is a tower of finite extensions, then E/K is separable if and only if E/L and L/K are separable.*

Proof. If E/K is separable then every element of E is separable over K, so every element of E is separable over L (Corollary 2.6) and every element of L (which is a subset of E) is separable over K. Thus E/L and L/K are separable.

Now assume E/L and L/K are separable. To show E/K is separable we will count field embeddings and then apply Theorem 3.8. The primitive element theorem will simplify matters, so we'll take advantage of it. Write $E = L(\beta)$ and $L = K(\alpha)$, with $m = [E : L]$ and $n = [L : K]$. Let F be a splitting field over K for the minimal polynomial of α in $K[X]$.



Since α is separable over K with degree n, there are n embeddings $\tau_i \colon L \to F$ that fix K. Moving up to E/L, the minimal polynomial of β in $L[X]$ is separable with degree m, so its

image in $F[X]$ under each τ_i is separable (Corollary 2.7). Let F' be an extension of F over which all of these polynomials in $F[X]$ split completely. These polynomials each have m roots in F', so each $\tau_i\colon L \to F$ has m extensions to embeddings $E \to F'$. Thus, the number of embeddings $E \to F'$ that extend the inclusion $K \to F'$ is $mn = [E : K]$. Since E/K admits as many embeddings into F' as its degree, E/K is separable by Theorem 3.8. □

Appendix A. Multiplicities for Inseparable Irreducibles

When a polynomial is inseparable, at least one of its roots has multiplicity greater than 1. The multiplicities of all the roots need not agree, *e.g.*, $X^2(X-1)^3$ has 0 as a root with multiplicity 2 and 1 as a root with multiplicity 3. This polynomial is reducible, so it is a dull example. When an inseparable polynomial is *irreducible*, which can only happen in positive characteristic, it is natural to ask how the multiplicities of different roots are related to each other. In fact, the multiplicities are all the same:

Theorem A.1. *Let $\pi(X) \in K[X]$ be irreducible, where K has characteristic $p > 0$. Write $\pi(X) = \widetilde{\pi}(X^{p^m})$ where $m \geq 0$ is as large as possible. Then $\widetilde{\pi}(X)$ is irreducible and separable in $K[X]$, and each root of $\pi(X)$ has multiplicity p^m.*

Proof. Since $\deg \pi = p^m \deg \widetilde{\pi}$, there is a largest possible m that can be used. Writing $\pi(X) = \widetilde{\pi}(X^{p^m})$, a non-trivial factorization of $\widetilde{\pi}(X)$ gives one for $\pi(X)$, so $\widetilde{\pi}(X)$ is irreducible in $K[X]$.

By the maximality of m, $\widetilde{\pi}(X)$ is not a polynomial in X^p, which means its derivative is not 0, so it must be separable (Theorem 2.8).

Factor $\widetilde{\pi}(X)$ in a splitting field over K:

$$\widetilde{\pi}(X) = c(X - \alpha_1)\cdots(X - \alpha_d),$$

where the α_i's are distinct since $\widetilde{\pi}(X)$ is separable. Then

$$\pi(X) = \widetilde{\pi}(X^{p^m}) = c(X^{p^m} - \alpha_1)\cdots(X^{p^m} - \alpha_d).$$

Write $\alpha_i = \gamma_i^{p^m}$ in a large enough field. Since the pth power map is injective in characteristic p, distinctness of the α_i's implies distinctness of the γ_i's. Therefore

$$\pi(X) = c(X^{p^m} - \gamma_1^{p^m})\cdots(X^{p^m} - \gamma_d^{p^m}) = c(X - \gamma_1)^{p^m}\cdots(X - \gamma_d)^{p^m},$$

which shows the roots of $\pi(X)$ (the γ_i's) are the p^mth roots of the roots of $\widetilde{\pi}(X)$ (the α_i's), and each root of $\pi(X)$ has multiplicity p^m. □

Example A.2. Let's work over the rational function field $\mathbf{F}_p(u)$. Let m and n be positive integers, with n not divisible by p. The polynomial $f(X) = X^{np^m} - u$ is irreducible in $\mathbf{F}_p(u)[X]$, since it is Eisenstein at u, and it is a polynomial in X^{p^m} but not in $X^{p^{m+1}}$. Therefore Theorem A.1 tells us that each root of $X^{np^m} - u$ has multiplicity p^m.

Let's try to understand this by working out the proof of Theorem A.1 in this case. Since n is not divisible by p, $X^n - u$ is separable over $\mathbf{F}_p(u)$, so over a larger field this polynomial picks up n distinct roots:

$$X^n - u = (X - \alpha_1)(X - \alpha_2)\cdots(X - \alpha_n).$$

Then

$$X^{np^m} - u = (X^{p^m} - \alpha_1)(X^{p^m} - \alpha_2)\cdots(X^{p^m} - \alpha_n),$$

so $X^{np^m} - u$ has n distinct roots: the p^mth roots of the α_i's. Each of these n roots has multiplicity p^m.

Remark A.3. A characterization of when a binomial $X^d - a$ is irreducible over a field is in [2, pp. 297–299],

Example A.4. In $\mathbf{F}_p(u)[X]$, let $f(X) = X^{3p} + (u^2+u)X^p + u$. This is irreducible in $\mathbf{F}_p(u)[X]$ since it is Eisenstein at u. Since this is a polynomial in X^p but not in X^{p^2}, its roots all have multiplicity p by Theorem A.1. The polynomial $X^3 + (u^2+u)X + u$ is separable (since it's irreducible by Eisenstein and its X-derivative is not 0), so in a suitable field this polynomial factors as

$$X^3 + (u^2+u)X + u = (X-\alpha_1)(X-\alpha_2)(X-\alpha_3)$$

where α_1, α_2, and α_3 are *distinct*. Therefore

$$X^{3p} + (u^2+u)X^p + u = (X^p-\alpha_1)(X^p-\alpha_2)(X^p-\alpha_3),$$

and each $X^p - \alpha_i$ has one root with multiplicity p, so each root of $f(X)$ has multiplicity p.

Corollary A.5. *When K has characteristic p, an irreducible polynomial in $K[X]$ with only one root in a splitting field has the form $X^{p^m} - a$ for some $m \geq 0$.*

Proof. Let $\pi(X)$ be such a polynomial, with one root α. Write $\pi(X) = \widetilde{\pi}(X^{p^m})$ with m maximal. By the proof of Theorem A.1, $\widetilde{\pi}(X)$ is separable with one root, so $\widetilde{\pi}(X) = X - \alpha$. Thus $\pi(X) = (X-\alpha)^{p^m} = X^{p^m} - a$, where $a = \alpha^{p^m}$. □

In Theorem A.1 we start with an irreducible polynomial $\pi(X)$ that is a polynomial in X^{p^m} and conclude that it's an irreducible polynomial of X^{p^m} (that is, $\pi(X) = \widetilde{\pi}(X^{p^m})$ with $\widetilde{\pi}(X)$ irreducible). We want to address a converse question: what are reasonable conditions under which an irreducible polynomial in $K[X]$, where K has characteristic p, remains irreducible when X in the polynomial is replaced by X^p, or more generally by X^{p^m} for all $m \geq 1$? This will provide us with a means of generating many inseparable irreducible polynomials in $\mathbf{F}_p(u)[X]$ with more than one root without having to rely so heavily on the Eisenstein criterion as in Examples A.2 and A.4.

Theorem A.6. *Let K be a field of characteristic p. If $f(X)$ is monic irreducible in $K[X]$ then $f(X^p)$ is irreducible in $K[X]$ or $f(X^p) = \pi(X)^p$ for a monic irreducible $\pi(X)$ in $K[X]$.*

Before proving this theorem, it's worth seeing each possible conclusion happening in a simple example.

Example A.7. Suppose $K \neq K^p$ (*e.g.*, $K = \mathbf{F}_p(u)$). The first case happens when $f(X) = X - c$ for $c \in K - K^p$, since $X^p - c$ is irreducible in $K[X]$, and the second case happens when $f(X) = X - 1$ since $X^p - 1 = (X-1)^p$. If $K = K^p$ then the first case never happens, so Theorem A.6 is really only of interest when $K \neq K^p$.

Proof. (of Theorem A.6) Let $\pi(X)$ be a monic irreducible factor of $f(X^p)$. We will show $f(X^p) = \pi(X)$ or $f(X^p) = \pi(X)^p$.

Factor out the largest power of $\pi(X)$ from $f(X^p)$, say

$$f(X^p) = \pi(X)^r g(X), \tag{A.1}$$

where $r \geq 1$ and $\pi(X)$ doesn't divide $g(X)$. Differentiating both sides,

$$0 = \pi(X)^r g'(X) + r\pi(X)^{r-1}\pi'(X)g(X) = \pi(X)^{r-1}(\pi(X)g'(X) + r\pi'(X)g(X)).$$

Therfore $\pi(X)g'(X) = -r\pi'(X)g(X)$. Since $\pi(X)$ doesn't divide $g(X)$, $\pi(X)$ must divide $r\pi'(X)$, which has degree less than that of $\pi(X)$ unless it is 0. Therefore $r\pi'(X) = 0$

(obviously $\pi(X)$ can't be a factor of a polynomial of smaller degree), so either $r = 0$ in K or $\pi'(X) = 0$ in $K[X]$.

Case 1: $\pi'(X) = 0$ in $K[X]$. This condition implies $\pi(X) = \widetilde{\pi}(X^p)$ where $\widetilde{\pi}(X)$ is monic irreducible over K, so (A.1) becomes $f(X^p) = \widetilde{\pi}(X^p)^r g(X)$. This shows $g(X) = f(X^p)/\widetilde{\pi}(X^p)^r$ is a rational function of X^p. It is also a polynomial. Check as an exercise that $K[X] \cap K(X^p) = K[X^p]$[1], so $g(X)$ is a polynomial in X^p, say $g(X) = \widetilde{g}(X^p)$ for a monic $\widetilde{g}(X)$ in $K[X]$. Then (A.1) becomes $f(X^p) = \widetilde{\pi}(X^p)^r \widetilde{g}(X^p)$. Throughout this equation X appears in the context of X^p, so $f(X) = \widetilde{\pi}(X)^r \widetilde{g}(X)$. The irreducibility of $f(X)$ in $K[X]$ implies $r = 1$ and $\widetilde{g}(X) = 1$, so $f(X^p) = \pi(X)^r g(X) = \pi(X)$: $f(X^p)$ is irreducible in $K[X]$.

Case 2: $r = 0$ in K. This condition implies r is a multiple of p, say $r = ps$. So (A.1) becomes

$$f(X^p) = \pi(X)^{ps} g(X) = \pi^{\varphi}(X^p)^s g(X), \tag{A.2}$$

where $\pi^{\varphi}(X)$ is the polynomial whose coefficients are pth powers of the coefficients of $\pi(X)$. (For example, if $\pi(X) = X^2 + aX + b$ then $\pi^{\varphi}(X) = X^2 + a^p X + b^p$.) This shows $g(X) = f(X^p)/\pi^{\varphi}(X^p)^s$ is a rational function of X^p, so as in Case 1 $g(X) = \widetilde{g}(X^p)$ for a monic $\widetilde{g}(X)$ in $K[X]$. Then (A.2) becomes $f(X^p) = \pi^{\varphi}(X^p)^s \widetilde{g}(X^p)$, so $f(X) = \pi^{\varphi}(X)^s \widetilde{g}(X)$. The irreducibility of $f(X)$ in $K[X]$ implies $s = 1$ and $\widetilde{g}(X) = 1$, so (A.2) becomes $f(X^p) = \pi(X)^p$. □

Corollary A.8. *Let K be a field of characteristic p. If $f(X)$ is monic irreducible in $K[X]$ then the following conditions are equivalent:*

a) *$f(X^{p^m})$ is irreducible in $K[X]$ for all $m \geq 1$,*
b) *$f(X) \notin K^p[X]$.*

Proof. (a)⇒(b): We prove the contrapositive. If $f(X) \in K^p[X]$ then $f(X^p) \in K^p[X^p] = K[X]^p$, so $f(X^p)$ is reducible.

(b)⇒(a): For $m \geq 1$, $f(X^{p^m})$ has the same nonzero coefficients as $f(X)$, so (b) implies $f(X^{p^m}) \notin K^p[X]$. Therefore it suffices by induction to show that if $f(X)$ is monic irreducible in $K[X]$ and is not in $K^p[X]$ then $f(X^p)$ is irreducible in $K[X]$. By Theorem A.6, $f(X^p)$ is either irreducible in $K[X]$ or is a pth power in $K[X]$. Any pth power in $K[X]$ has all of its coefficients in K^p, which contradicts the hypothesis in (b). □

This corollary tells us that a monic irreducible in $K[X]$ having at least one coefficient that is not a pth power in K will stay irreducible over K when we replace X in the polynomial with X^{p^m} for all $m \geq 1$.

Here are several illustrations of Corollary A.8.

Example A.9. The polynomial $X^3 - X - 1$ is irreducible over $\mathbf{F}_3$, so $X^3 + (u^2 - 1)X + (u^3 + 2u - 1)$ is irreducible over $\mathbf{F}_3(u)$ because it is irreducible mod u (this is an analogue of the reduction mod p test used in $\mathbf{Z}[X]$). Its linear and constant coefficients are not third powers in $\mathbf{F}_3(u)$, so $X^{3^{m+1}} + (u^2 - 1)X^{3^m} + (u^3 + 2u - 1)$ is irreducible over $\mathbf{F}_3(u)$ for all $m \geq 1$ and has three distinct roots.

Example A.10. The polynomial $X^2 + uX + 1$ is irreducible over $\mathbf{F}_p(u)$ (left as an exercise, possibly requiring a separate treatment of the case $p = 2$) and one of its coefficients, u, is not a pth power in $\mathbf{F}_p(u)$, so $X^{2p^m} + uX^{p^m} + 1$ is irreducible over $\mathbf{F}_p(u)$ for all $m \geq 1$ and has two distinct roots.

[1]This doesn't depend on characteristic p. For every field F and positive integer N, $F[X] \cap F(X^N) = F[X^N]$.

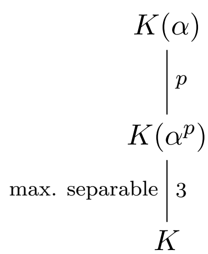
Example A.11. Let $\pi(X)$ be monic irreducible in $\mathbf{F}_p[X]$, of degree n. Then $\pi(X) + u$ is irreducible in $\mathbf{F}_p(u)[X]$ by reduction mod u, and its constant term $\pi(0) + u$ is not a pth power in $\mathbf{F}_p(u)$, so for each $m \geq 1$ the polynomial $\pi(X^{p^m}) + u$ is irreducible over $\mathbf{F}_p(u)$ and has n distinct roots having multiplicity p^m.

Examples A.2 and A.4 are special cases of Corollary A.8 with $K = \mathbf{F}_p(u)$, $f(X) = X^n - u$, and $f(X) = X^3 + (u^2 + u)X + u$.

Appendix B. Purely Inseparable Extensions

Corollary 3.12 says that inside a finite extension L/K, the set of elements in L that are separable over K form a subfield. Call this subfield F. It is a separable extension of K and, by its construction, it contains every separable extension of K inside L because all elements of L separable over K are inside F. We call F the maximal separable subextension of K in L. Of course if L/K is separable then $F = L$ and the concept of maximal separable subextension is dull. Let's suppose L/K is not separable, so $F \neq L$.

Example B.1. We look at the field extension from Example A.4. Let $K = \mathbf{F}_p(u)$ and $L = K(\alpha)$ where α is a root of $X^{3p} + (u^2 + u)X^p + u$. Then $[L : K] = 3p$ and L/K is not separable. Since α^p is a root of $X^3 + (u^2 + u)X + u$, $[K(\alpha^p) : K] = 3$ and we have the field diagram below.

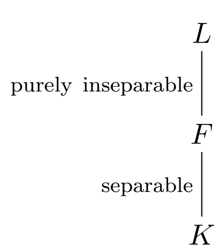


The polynomial $X^3 + (u^2 + u)X + u$ is separable, so $K(\alpha^p)/K$ is separable. This has to be the maximal separable subextension because there are no fields properly between $K(\alpha)$ and $K(\alpha^p)$.

The extension L/F is not separable in a strong sense: no element of L outside F is separable over F. Indeed, if $\alpha \in L$ is separable over F then $F(\alpha)/F$ is separable, and also F/K is separable by the definition of F, so $F(\alpha)/K$ is separable (Theorem 3.13). That implies α is separable over K, so $\alpha \in F$ by the definition of F.

Definition B.2. A field extension is called *purely inseparable* when the only elements in the top field separable over the bottom field are in the bottom field.

The relations between L, K, and the maximal separable subextension F are described in the following diagram.



Since inseparable field extensions exist only in characteristic p, purely inseparable extensions are a purely characteristic p phenomenon.

Theorem B.3. *The following conditions on a finite extension L/K in characteristic p are equivalent.*

(1) *L/K is purely inseparable.*
(2) *Each $\alpha \in L$ has minimal polynomial of the form $X^{p^m} - a$ in $K[X]$ for some $m \geq 1$ and $a \in K^\times$.*
(3) *Each $\alpha \in L$ is the root of some $X^{p^m} - a$ in $K[X]$.*

Although this says the minimal polynomials in a purely inseparable extension look like $X^{p^m} - a$, this does *not* mean all polynomials of the form $X^{p^m} - a$ are irreducible. Consider, for instance, $X^{p^m} - 1 = (X-1)^{p^m}$.

Proof. (1) $\Rightarrow$ (2): Let $\alpha \in L$ have minimal polynomial $\pi(X) \in K[X]$. Write $\pi(X) = \widetilde{\pi}(X^{p^m})$ with m maximal, so $p^m \mid \deg \pi$. Then $\widetilde{\pi}(X)$ is separable and irreducible in $K[X]$, so α^{p^m} is separable over K. Thus $\alpha^{p^m} \in K$ because L/K is purely inseparable. Set $a = \alpha^{p^m}$, so α is a root of $X^{p^m} - a \in K[X]$. That means $\pi(X) \mid (X^{p^m} - a)$. Since $\deg \pi$ is divisible by p^m, we conclude $\pi(X) = X^{p^m} - a$.

(2) $\Rightarrow$ (3): Trivial.

(3) $\Rightarrow$ (1): Pick $\alpha \in L$ that is separable over K. We want to show $\alpha \in K$. Let $\pi(X)$ be the minimal polynomial of α in $K[X]$, so $\pi(X)$ has distinct roots. By (3), α is a root of $X^{p^m} - a$ for some $a \in K$, so $\pi(X) \mid (X^{p^m} - a)$ in $K[X]$. Since $\alpha^{p^m} - a = 0$, we can factor $X^{p^m} - a$ in $L[X]$ as

$$X^{p^m} - a = X^{p^m} - \alpha^{p^m} = (X - \alpha)^{p^m}.$$

Since $\pi(X)$ is a factor of $X^{p^m} - a$, looking at this in $L[X]$ tells us $\pi(X)$ is a power of $X - \alpha$. At the same time, $\pi(X)$ is separable, so the only choice is $\pi(X) = X - \alpha$. Therefore $\alpha \in K$ because $\pi(X)$ has coefficients in K (or because it has degree 1). □

Being a root of $X^{p^m} - a \in K[X]$ is the same as having p^mth power in K. So L/K is purely inseparable if and only if each element of L has some p-power in K. This is how we will think about the purely inseparable property in the next result.

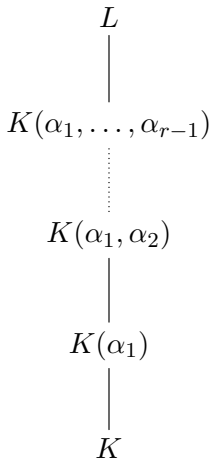
Corollary B.4. *If $L \supset E \supset K$ then L/K is purely inseparable if and only if L/E and E/K are purely inseparable.*

Proof. First suppose L/K is purely inseparable. Then all elements of L have a p-power in K, which means they have a p-power in E (use the same p-power). Therefore L/E is purely inseparable. Every element of E lies in L and thus has a p-power in K, so E/K is purely inseparable.

Now suppose L/E and E/K are purely inseparable. For each $\alpha \in L$, $\alpha^{p^m} \in E$ for some m and then $(\alpha^{p^m})^{p^n} \in K$ for some n. Since $(\alpha^{p^m})^{p^n} = \alpha^{p^{m+n}}$, all elements of L have a p-power in K, so L/K is purely inseparable. □

Corollary B.5. *A finite purely inseparable extension in characteristic p has p-power degree.*

Proof. Suppose L/K is purely inseparable. Write $L = K(\alpha_1, \ldots, \alpha_r)$. In the tower



each step is purely inseparable by Corollary B.4, and $[K(\alpha_1, \ldots, \alpha_{i+1}) : K(\alpha_1, \ldots, \alpha_i)]$ is the degree of the minimal polynomial of α_{i+1} over $K(\alpha_1, \ldots, \alpha_i)$, which has to be a power of p. Therefore $[L : K]$ is a power of p. □

It is false that all extensions of p-power degree are purely inseparable since there are many separable irreducible polynomials of p-power degree (like $X^p - X - 1$ in $\mathbf{F}_p[X]$).

Appendix C. A Weaker Primitive Element Theorem

If we give up on asking for the primitive element to be separable, then there is another approach to the Primitive Element Theorem.

Theorem C.1. *If α and β are algebraic over K and at least one of them is separable over K then there is a $\gamma \in K(\alpha, \beta)$ such that $K(\alpha, \beta) = K(\gamma)$.*

Proof. As in the usual Primitive Element Theorem, the case of finite K follows from cyclicity of $K^\times$, so we may take K to be an infinite field. We take β to be separable over K. Let $f(X)$ be the minimal polynomial of α in $K[X]$, $g(X)$ be the minimal polynomial of β in $K[X]$, and E/K be an extension in which $f(X)$ and $g(X)$ both split completely. The distinct roots of $f(X)$ in E will be denoted $\alpha_1, \ldots, \alpha_m$ (with $\alpha = \alpha_1$) and the distinct roots of $g(X)$ in E will be denoted $\beta_1, \ldots, \beta_n$ (with $\beta = \beta_1$).

For $j \neq 1$ and arbitrary i, the equation $\alpha_i + c\beta_j = \alpha_1 + c\beta_1$ has only the solution $c = (\alpha_1 - \alpha_i)/(\beta_j - \beta_1)$. Since K is infinite, there is $c \in K$ distinct from all the numbers $(\alpha_1 - \alpha_i)/(\beta_j - \beta_1)$. Let $\gamma = \alpha_1 + c\beta_1 = \alpha + c\beta$ for such c. With such a choice of c, we will show α and β are in $K(\gamma)$, so $K(\alpha, \beta) \subset K(\gamma)$. The reverse containment is obvious, so we would have $K(\alpha, \beta) = K(\gamma)$.

The polynomials $g(X)$ and $f(\gamma - cX)$ both have root β (check!), so $X - \beta$ is a common factor of $g(X)$ and $f(\gamma - cX)$ in $E[X]$. We will show it is the greatest common factor. Since β is not a double root of $g(X)$, $(X - \beta)^2$ is not a common factor of $g(X)$ and $f(\gamma - cX)$. Could another root β_j of $g(X)$ be a root of $f(\gamma - cX)$? If so then $\gamma - c\beta_j = \alpha_i$ for some i, so $\gamma = \alpha_i + c\beta_j$. By our choice of c, such equality is impossible unless $i = 1$ and $j = 1$. So in $E[X]$ the greatest common divisor of $g(X)$ and $f(\gamma - cX)$ is $X - \beta$.

Since $g(X)$ and $f(\gamma - cX)$ are in $K(\gamma)[X]$, Euclid's algorithm shows their monic greatest common divisor in $E[X]$ is also their monic greatest common divisor in $K(\gamma)[X]$. Therefore $X - \beta \in K(\gamma)[X]$, so $\beta \in K(\gamma)$. Then $\alpha = \gamma - c\beta \in K(\gamma)$, so $K(\alpha, \beta) \subset K(\gamma)$. □

Appendix D. Separability and Tensor Products

A deeper study of separability makes extensive use of tensor products. To illustrate how tensor products arise, we now reprove Corollary 3.10 using them.

Proof. Since all finite extensions in characteristic 0 are separable, we may assume K has characteristic p. Pick β in $K(\alpha)$, so $K(\beta) \subset K(\alpha)$. We want to show the minimal polynomial of β in $K[X]$ has no repeated roots. This will be proved by comparing the fields $K(\beta)$ and $K(\alpha)$ using tensor products.

We start by indicating why tensor products are a reasonable tool to use. Let $\pi_\alpha(X)$ and $\pi_\beta(X)$ be the minimal polynomials of α and β in $K[X]$. Then $K(\alpha) \cong K[X]/(\pi_\alpha(X))$ and $K(\beta) \cong K[X]/(\pi_\beta(X))$ as rings. For a field extension L/K, the tensor product $L \otimes_K K(\alpha)$ is isomorphic as a ring to $L[X]/(\pi_\alpha(X))$ and the ring structure of $L[X]/(\pi_\alpha(X))$ is related to the way $\pi_\alpha(X)$ factors in $L[X]$. We can similarly look at $L \otimes_K K(\beta) \cong L[X]/(\pi_\beta(X))$. The reason for using tensor products, rather than just contemplating the rings $L[X]/(\pi_\alpha(X))$ and $L[X]/(\pi_\beta(X))$ directly, is that it is easier to relate these rings to each other from the tensor product viewpoint because tensoring has mapping properties. (Warning: the ring $L[X]/(\pi_\alpha(X))$ is usually *not* a field: for $L/K = \mathbf{C}/\mathbf{R}$ and $\alpha = i$ we have $\mathbf{C}[X]/(X^2 + 1) \cong \mathbf{C} \times \mathbf{C}$.)

Quite generally, if we have an injective linear map $W \hookrightarrow V$ of K-vector spaces the induced L-linear map $L \otimes_K W \to L \otimes_K V$ is also injective since a subspace of a vector space is always a direct summand and the tensor product commutes with direct sums. (For comparison, given a ring extension S/R and an injective linear map $M \hookrightarrow N$ of R-modules, the base extended S-linear map $S \otimes_R M \to S \otimes_R N$ need not be injective.) Thus the natural L-linear map $L \otimes_K K(\beta) \to L \otimes_K K(\alpha)$ that base extends the inclusion map $K(\beta) \hookrightarrow K(\alpha)$ is injective. Moreover, $L \otimes_K K(\beta)$ and $L \otimes_K K(\alpha)$ are not just vector spaces, but rings, and the map from the former to the latter is a ring homomorphism. So we can think of $L \otimes_K K(\beta)$ as a subring of $L \otimes_K K(\alpha)$ in the obvious way, for every field extension L/K.

The goal now is to pick L so that the structure of $L \otimes_K K(\beta)$ is only compatible with being a subring of $L \otimes_K K(\alpha)$ if β is separable over K.

By hypothesis $\pi_\alpha(X)$ is separable. Take for L an extension of K in which both $\pi_\alpha(X)$ and $\pi_\beta(X)$ split completely, so

$$\pi_\alpha(X) = (X - \alpha_1) \cdots (X - \alpha_m) \tag{D.1}$$

with distinct α_i's and (by Theorem A.1)

$$\pi_\beta(X) = (X - \beta_1)^{p^r} \cdots (X - \beta_n)^{p^r} \tag{D.2}$$

with distinct β_j's and $r \geq 0$ in $L[X]$. The polynomial $\pi_\beta(X)$ is separable precisely when $r = 0$. We now look at the rings $L \otimes_K K(\alpha)$ and $L \otimes_K K(\beta)$.

Since $K(\alpha) \cong K[X]/(\pi_\alpha(X))$ as rings,

$$L \otimes_K K(\alpha) \cong L \otimes_K (K[X]/(\pi_\alpha(X))) \cong L[X]/(\pi_\alpha(X)),$$

where the isomorphisms are the obvious maps. In $L[X]$, $\pi_\alpha(X)$ splits completely as in (D.1). Since the roots α_i are distinct, the factors $X - \alpha_i$ are relatively prime in $L[X]$, so by the

Chinese remainder theorem

$$L[X]/(\pi_\alpha(X)) \cong L[X]/(X-\alpha_1) \times \cdots \times L[X]/(X-\alpha_m).$$

Each $L[X]/(X-\alpha_i)$ is isomorphic to L, so $L \otimes_K K(\alpha) \cong L^m$ as rings (where L^m denotes the m-fold product ring of L with itself). This is not a field, but rather is a product of as many copies of the field L as $\pi_\alpha(X)$ has roots.

Now we look at $L \otimes_K K(\beta)$. Our arguments will be similar to the treatment of $L \otimes_K K(\alpha)$. First of all, $L \otimes_K K(\beta) \cong L[X]/(\pi_\beta(X))$. Since $\pi_\beta(X)$ factors in $L[X]$ as in (D.2),

$$L \otimes_K K(\beta) \cong L[X]/(X-\beta_1)^{p^r} \times \cdots \times L[X]/(X-\beta_n)^{p^r}.$$

We have $L[X]/(X-\beta_i)^{p^r} \cong L[Y]/(Y^{p^r})$ by identifying $X-\beta_i$ with Y. So $L \otimes_K K(\beta)$ is isomorphic to the n-fold product ring of $L[Y]/(Y^{p^r})$ with itself. The inclusion of $L \otimes_K K(\beta)$ into $L \otimes_K K(\alpha)$ amounts to an embedding of $(L[Y]/(Y^{p^r}))^n$ into L^m as a subring.

There is a problem with such an embedding when $r \geq 1$. For positive r, $Y \bmod Y^{p^r}$ is a nonzero nilpotent element of $L[Y]/(Y^{p^r})$, so $(Y, \ldots, Y)$ is a nonzero nilpotent element of $(L[Y]/(Y^{p^r}))^n$. The only nilpotent element in L^m is $(0, \ldots, 0)$, so it is impossible for $(L[Y]/(Y^{p^r}))^n$ to embed as a subring of L^m unless $r = 0$, which means $\pi_\beta(X)$ has distinct roots, so β is separable over K. □

References

[1] D. Dummit and R. Foote, "Abstract Algebra," 3rd ed., Wiley, New York, 2004.
[2] S. Lang, "Algebra," revised 3rd ed., Springer-Verlag, New York, 2002.