SEPARABILITY II

KEITH CONRAD

1. Introduction

Separability of a finite field extension L/K can be described in several different ways. The original definition is that every element of L is separable over K (that is, has a separable minimal poynomial in $K[X]$). We will give here *three* descriptions of separability for a finite extension and use each of them to prove two theorems about separable extensions.

Theorem 1.1. *Let L/K be a finite extension. Then L/K is separable if and only if the trace function $\mathrm{Tr}_{L/K}\colon L \to K$ is not identically 0.*

The trace function is discussed in Appendix A.

Theorem 1.2. *Let L/K be a finite extension and write $\overline{K}$ for an algebraic closure of K. Then L/K is separable if and only if the ring $\overline{K} \otimes_K L$ has no nonzero nilpotent elements. When L/K is separable, the ring $\overline{K} \otimes_K L$ is isomorphic to $\overline{K}^{[L:K]}$.*

Example 1.3. Consider the extension $\mathbf{Q}(\sqrt{2})/\mathbf{Q}$. Since $\mathbf{Q}(\sqrt{2}) \cong \mathbf{Q}[X]/(X^2-2)$, tensoring with $\overline{\mathbf{Q}}$ gives

$$\overline{\mathbf{Q}} \otimes_{\mathbf{Q}} \mathbf{Q}(\sqrt{2}) \cong \overline{\mathbf{Q}}[X]/(X^2-2) = \overline{\mathbf{Q}}[X]/((X+\sqrt{2})(X-\sqrt{2})) \cong \overline{\mathbf{Q}} \times \overline{\mathbf{Q}},$$

which is a product of 2 copies of $\overline{\mathbf{Q}}$ (associated to the 2 roots of X^2-2) and has no nilpotent elements besides 0.

Example 1.4. Consider the extension $\mathbf{F}_2(\sqrt{u})/\mathbf{F}_2(u)$. Since $\mathbf{F}_2(\sqrt{u}) \cong \mathbf{F}_2[X]/(X^2-u)$,

$$\overline{\mathbf{F}_2(u)} \otimes_{\mathbf{F}_2(u)} \mathbf{F}_2(\sqrt{u}) \cong \overline{\mathbf{F}_2(u)}[X]/(X^2-u) = \overline{\mathbf{F}_2(u)}[X]/(X-\sqrt{u})^2,$$

which has the nonzero nilpotent element $X - \sqrt{u}$.

Theorem 1.5. *Let L/K be a finite extension. Then L is separable over K if and only if any derivation of K has a unique extension to a derivation of L.*

Derivations are discussed in Appendix B.

The proofs of Theorems 1.1 and 1.2 both use tensor products. For those two proofs, the reader should be comfortable with the fact that injectivity and surjectivity of a linear map of vector spaces can be detected after a base extension: a linear map is injective or surjective if and only if its base extension to a larger field is injective or surjective.

Each of the three theorems above will be proved and then lead in its own way to proofs of the following two theorems.

Theorem 1.6. *If $L = K(\alpha_1, \ldots, \alpha_r)$ and each α_i is separable over K then every element of L is separable over K (so L/K is separable).*

Theorem 1.7. *Let L/K be a finite extension and F be an intermediate field. If L/F and F/K are separable then L/K is separable*

Here is a brief outline of this handout. In Sections 2, 3 and 4 we will respectively prove Theorems 1.1, 1.2 and 1.5 and apply each one to prove Theorems 1.6 and 1.7. In Section 5 we will use our new viewpoints to define separability for arbitrary (possibly non-algebraic) field extensions.

2. Theorem 1.1: Traces

Review Appendix A before reading this section.

We want to show L/K is separable if and only if $\mathrm{Tr}_{L/K}\colon L \to K$ is not identically 0. The trace map is either identically 0 or it is onto, since it is K-linear with target K, so another way of putting Theorem 1.1 is that we want to show L/K is separable if and only if the trace from L to K is onto.

Proof. We might as well take K to have positive characteristic p, since in characteristic 0 all finite field extensions are separable and the trace is not identically 0: $\mathrm{Tr}_{L/K}(1) = [L:K] \neq 0$ in characteristic 0.

If L/K is separable, by the primitive element theorem we can write $L = K(\alpha)$ where α is separable over K. To show the trace is surjective for finite separable extensions, it suffices to prove surjectivity of the trace map on $K(\alpha)/K$ when K is any base field and α is separable over K.

If L/K is inseparable, then there must be some $\alpha \in L$ which is inseparable over K. Since $\mathrm{Tr}_{L/K} = \mathrm{Tr}_{K(\alpha)/K} \circ \mathrm{Tr}_{L/K(\alpha)}$, it suffices to prove the trace map on $K(\alpha)/K$ vanishes when α is inseparable over K.

For both cases of the field extension $K(\alpha)/K$ (α separable or inseparable over K), let α have minimal polynomial $\pi(X)$ in $K[X]$. Write $\pi(X) = \widetilde{\pi}(X^{p^m})$ where m is as large as possible, so $\widetilde{\pi}(X)$ is separable. Thus $\pi(X)$ is separable if and only if $m = 0$.

Let $n = \deg \pi = p^m d$, with $d = \deg \widetilde{\pi}$. In $\overline{K}[X]$,

$$\widetilde{\pi}(X) = (X - \beta_1) \cdots (X - \beta_d)$$

for some β_i's, which are all distinct since $\widetilde{\pi}(X)$ is separable. Write $\beta_i = \gamma_i^{p^m}$, so the γ_i's are distinct. Then

$$\pi(X) = \widetilde{\pi}(X^{p^m}) = (X^{p^m} - \beta_1) \cdots (X^{p^m} - \beta_d) = (X - \gamma_1)^{p^m} \cdots (X - \gamma_d)^{p^m}.$$

Consider now the extension of scalars up to $\overline{K}$ of the trace map $\mathrm{Tr}_{K(\alpha)/K}\colon K(\alpha) \to K$:

$$\overline{\mathrm{Tr}} = \mathrm{id}_{\overline{K}} \otimes \mathrm{Tr}_{K(\alpha)/K}\colon \overline{K} \otimes_K K(\alpha) \to \overline{K} \otimes_K K \cong \overline{K}. \tag{2.1}$$

By Theorem A.8, $\overline{\mathrm{Tr}}$ is the trace map on $\overline{K} \otimes_K K(\alpha)$ as a $\overline{K}$-vector space.

Since tensoring with a field extension preserves injectivity and surjectivity of a linear map,

$$\mathrm{Tr}_{K(\alpha)/K} \text{ is onto } \iff \overline{\mathrm{Tr}} \text{ is onto}, \quad \mathrm{Tr}_{K(\alpha)/K} = 0 \iff \overline{\mathrm{Tr}} = 0.$$

Since $K(\alpha) \cong K[X]/(\pi(X))$ as K-algebras, $\overline{K} \otimes_K K(\alpha) \cong \overline{K}[X]/(\pi(X))$ as $\overline{K}$-algebras, and thus is isomorphic to the direct product of the rings $\overline{K}[X]/(X^{p^m} - \beta_i)$. Theorem A.5 tells us that the trace in (2.1) is the sum of the traces to $\overline{K}$ on each $\overline{K}[X]/(X^{p^m} - \beta_i)$. Let's look at the trace from $\overline{K}[X]/(X^{p^m} - \beta_i)$ to $\overline{K}$.

In $\overline{K}[X]$, $X^{p^m} - \beta_i = (X - \gamma_i)^{p^m}$. Then $\overline{K}[X]/(X^{p^m} - \beta_i) = \overline{K}[Y]/(Y^{p^m})$, where $Y = X - \gamma_i$. If $m = 0$, then $\overline{K}[Y]/(Y^{p^m}) = \overline{K}$, so the trace to $\overline{K}$ is the identity. If $m > 0$, any element of $\overline{K}[Y]/(Y^{p^m})$ is the sum of a constant plus a multiple of Y, which is a constant plus a nilpotent element (since Y mod Y^{p^r} is nilpotent). Any constant in $\overline{K}[Y]/(Y^{p^m})$ has

trace 0 since $p^m = 0$ in $\overline{K}$ (because $m > 0$). A nilpotent element has trace 0 (Example A.3). Thus the trace to $\overline{K}$ of any element of $\overline{K}[Y]/(Y^{p^m})$ is 0.

To summarize, when α is separable over K (*i.e.*, $m = 0$), the trace map from $K(\alpha)$ to K is onto since it is onto after extending scalars to $\overline{K}$. When α is inseparable over K (*i.e.*, $m > 0$), the trace map is identically 0 since it vanishes after extending scalars. □

Corollary 2.1. *Theorem* 1.1 *implies Theorem* 1.6.

Proof. Set $L_0 = K$, $L_1 = K(\alpha_1) = L_0(\alpha_1)$, and more generally $L_i = K(\alpha_1, \ldots, \alpha_i) = L_{i-1}(\alpha_i)$ for $i \geq 1$. So we have the tower of field extensions

$$K = L_0 \subset L_1 \subset L_2 \subset \cdots \subset L_{r-1} \subset L_r = L. \tag{2.2}$$

By transitivity of the trace,

$$\mathrm{Tr}_{L/K} = \mathrm{Tr}_{L_1/L_0} \circ \mathrm{Tr}_{L_2/L_1} \circ \cdots \circ \mathrm{Tr}_{L_r/L_{r-1}}.$$

Since α_i is separable over K and the minimal polynomial of α_i over L_{i-1} divides its minimal polynomial over K, α_i is separable over L_{i-1}. Therefore $\mathrm{Tr}_{L_{i-1}(\alpha_i)/L_{i-1}} \colon L_i \to L_{i-1}$ is onto from the proof of Theorem 1.1, so the composite map $\mathrm{Tr}_{L/K} \colon L \to K$ is onto. Therefore L/K is separable by Theorem 1.1. □

Corollary 2.2. *Theorem* 1.1 *implies Theorem* 1.7.

Proof. By Theorem 1.1 and the hypothesis of Theorem 1.7, both $\mathrm{Tr}_{L/F}$ and $\mathrm{Tr}_{F/K}$ are onto. Therefore their composite $\mathrm{Tr}_{L/K}$ is onto, so L/K is separable by Theorem 1.1. □

3. Theorem 1.2: Nilpotents

Proof. We will begin with the case of a simple extension $L = K(\alpha)$. Let $\pi(X)$ be the minimal polynomial of α over K, so $L \cong K[X]/(\pi(X))$ as K-algebras and

$$\overline{K} \otimes_K L \cong \overline{K}[X]/(\pi(X))$$

as $\overline{K}$-algebras. This ring was considered in the proof of Theorem 1.1, where we saw its structure is different when $\pi(X)$ is separable or inseparable. If $\pi(X)$ is separable in $K[X]$, then $\overline{K}[X]/(\pi(X))$ is a product of copies of the field $\overline{K}$, so it has no nonzero nilpotent elements. If $\pi(X)$ is inseparable, then $\overline{K}[X]/(\pi(X))$ is a product of copies of rings $\overline{K}[Y]/(Y^{p^m})$ with $m > 0$, which all have nonzero nilpotents.

Now we consider the structure of $\overline{K} \otimes_K L$ when L/K is any finite extension.

First assume L/K is separable. By the primitive element theorem, we can write $L = K(\alpha)$ and α is separable over K. By the first paragraph of the proof, $\overline{K} \otimes_K L \cong \overline{K}^{[L:K]}$ since $\pi(X)$ has distinct linear factors in $\overline{K}$.

If L/K is inseparable, then some $\alpha \in L$ is inseparable over K. Tensoring the inclusion map $K(\alpha) \hookrightarrow L$ up to $\overline{K}$, we have an inclusion

$$\overline{K} \otimes_K K(\alpha) \hookrightarrow \overline{K} \otimes_K L.$$

The ring $\overline{K} \otimes_K K(\alpha)$ has a nonzero nilpotent element by the first paragraph of the proof, so $\overline{K} \otimes_K L$ does as well. □

Corollary 3.1. *The proof of Theorem* 1.2 *implies Theorem* 1.6.

Proof. Make a tower of intermediate extensions in L/K as in (2.2). Note $\overline{K}$ is an algebraic closure of every field L_i in the tower. Since

$$\overline{K} \otimes_K L \cong (\overline{K} \otimes_K L_1) \otimes_{L_1} L$$

and $L_1 = K(\alpha_1)$ with α_1 separable over K, the proof of Theorem 1.2 implies

$$\overline{K} \otimes_K L_1 \cong \overline{K}^{[L_1:K]}$$

as $\overline{K}$-algebras. Therefore

$$\overline{K} \otimes_K L \cong \overline{K}^{[L_1:K]} \otimes_{L_1} L \cong (\overline{K} \otimes_{L_1} L)^{[L_1:K]}.$$

Since $L = L_1(\alpha_2, \ldots, \alpha_r)$ with each α_i separable over L_1, we can run through the same computation for $\overline{K} \otimes_{L_1} L$ as we did for $\overline{K} \otimes_K L$, and we get $\overline{K} \otimes_{L_1} L \cong (\overline{K} \otimes_{L_2} L)^{[L_2:L_1]}$, so

$$\overline{K} \otimes_K L \cong (\overline{K} \otimes_{L_2} L)^{[L_2:L_1][L_1:K]} = (\overline{K} \otimes_{L_2} L)^{[L_2:K]}.$$

Repeating this enough, in the end we get

$$\overline{K} \otimes_K L \cong (\overline{K} \otimes_{L_r} L)^{[L_r:K]} \cong \overline{K}^{[L:K]}.$$

□

Corollary 3.2. *The proof of Theorem* 1.2 *implies Theorem* 1.7.

Proof. The field $\overline{K}$ is an algebraic closure of F and L. Using Theorem 1.2,

$$\begin{aligned} \overline{K} \otimes_K L &\cong (\overline{K} \otimes_K F) \otimes_F L \\ &\cong \overline{K}^{[F:K]} \otimes_F L \quad \text{since } F/K \text{ is separable} \\ &\cong (\overline{K} \otimes_F L)^{[F:K]} \\ &\cong \overline{K}^{[L:F][F:K]} \quad \text{since } L/F \text{ is separable} \\ &\cong \overline{K}^{[L:K]}. \end{aligned}$$

Thus L/K is separable by Theorem 1.2. □

4. Theorem 1.5: Derivations

Review Appendix B before reading this section.

We prove Theorem 1.5 by starting with a theorem about extending derivations.

Theorem 4.1. *Let L/K be an extension of fields, and $\alpha \in L$ be algebraic over K. Then α is separable over K if and only if any derivation on K has a unique extension to a derivation on $K(\alpha)$.*

Proof. When $\alpha \in L$ is separable over K, Corollary B.10 shows any derivation on K extends uniquely to a derivation on $K(\alpha)$.

Now suppose $\alpha \in L$ is inseparable over K. Then $\pi'(X) = 0$, where $\pi(X)$ is the minimal polynomial of α over K. In particular $\pi'(\alpha) = 0$. We are going to use this vanishing of $\pi'(\alpha)$ to construct a nonzero derivation on $K(\alpha)$ which extends the zero derivation on K. Then the zero derivation on K has two lifts to $K(\alpha)$: the zero derivation on $K(\alpha)$ and this other derivation we will construct.

Define $Z\colon K(\alpha) \to K(\alpha)$ by $Z(f(\alpha)) = f'(\alpha)$, where $f(X) \in K[X]$. Is this well-defined? Well, if $f_1(\alpha) = f_2(\alpha)$, then $f_1(X) \equiv f_2(X) \bmod \pi(X)$, say

$$f_1(X) = f_2(X) + \pi(X)k(X).$$

Differentiating both sides with respect to X,

$$f_1'(X) = f_2'(X) + \pi(X)k'(X) + \pi'(X)k(X).$$

Evaluating both sides at α yields $f_1'(\alpha) = f_2'(\alpha)$ since $\pi'(\alpha) = 0$. So $Z\colon K(\alpha) \to K(\alpha)$ is well-defined.

It is left to the reader to check Z is a derivation on $K(\alpha)$. This derivation kills K, but $Z(\alpha) = 1$, so Z extends the zero derivation on K while not being the zero derivation itself. □

The reader can check more generally that when α is inseparable over K and $\beta \in K(\alpha)$ is arbitrary the map $f(\alpha) \mapsto f'(\alpha)\beta$ is a derivation on $K(\alpha)$ that extends the zero derivation on K and sends α to β. So there are many extensions of the zero derivation on K to $K(\alpha)$: one for each element of $K(\alpha)$.

We need a lemma to put inseparable extensions into a convenient form for our derivation constructions later.

Lemma 4.2. *Let L/K be a finite inseparable field extension. Then there is an $\alpha \in L$ and intermediate field F such that $L = F(\alpha)$ and α is inseparable over F.*

Proof. Inseparable field extensions only occur in positive characteristic. Let p be the characteristic of K. Necessarily $[L : K] > 1$. Since L/K is inseparable, there is some $\beta \in L$ that is inseparable over K.

Write $L = K(\alpha_1, \ldots, \alpha_r)$. We will show by contradiction that some α_i has to be inseparable over K. Assume every α_i is separable over K. Then we can treat L/K as a succession of simple field extensions as in (2.2), where $L_i = L_{i-1}(\alpha_i)$ with α_i separable over L_{i-1}. By Theorem 4.1, any derivation on L_{i-1} extends to a derivation on L_i, so any derivation on K extends to a derivation on L. Moreover, this extended derivation on L is unique.[1] To show that, consider two derivations D and D' on L that are equal on K. Since $L_1 = K(\alpha_1)$ and α_1 is separable over K, the proof of Corollary B.10 tells us that D and D' both send L_1 to L_1 and are equal on L_1. Now using L_1 in place of K, D and D' being equal on L_1 implies they are equal on L_2 since $L_2 = L_1(\alpha_2)$ and α_2 is separable over L_1. We can keep going like this until we get $D = D'$ on $L_r = L$. As a special case of this uniqueness, the only derivation on L which vanishes on K is the zero derivation on L.

Now replace K as base field with $K(\beta)$, over which the α_i's are of course still separable. Then any derivation on $K(\beta)$ extends uniquely to a derivation on L. But in the proof of Theorem 4.1 we saw there is a nonzero derivation Z on $K(\beta)$ that vanishes on K, and an extension of that to a derivation on L is[2] nonzero on L and is zero on K. We have a contradiction of the uniqueness of extensions, so in any set of field generators $\{\alpha_1, \ldots, \alpha_r\}$, some α_i must be inseparable in K.

Choose a generating set $\{\alpha_1, \ldots, \alpha_r\}$ with as few inseparable elements as possible. At least one α_i is inseparable over K and we may assume that α_r is one of them. Set $\alpha = \alpha_r$ and $F = K(\alpha_1, \ldots, \alpha_{r-1})$ (so $F = K$ if $r = 1$). Then $L = F(\alpha)$. We will show by contradiction that α must be inseparable over F, which is the point of the lemma.

[1]This is not automatic from the uniqueness in Theorem 4.1 because we need to rule out the possibility that a derivation on L might not send L_i back to L_i.

[2]It is crucial that we know in advance that all derivations on $K(\beta)$ really do extend to L, so we're not just talking hypothetically about an extended derivation. A nonzero derivation on a field sometimes has no extensions to a particular larger field. See Example B.9.

Suppose α is separable over F. Then α is separable over the larger field $F(\alpha^p)$ since its minimal polynomial over $F(\alpha^p)$ divides its minimal polynomial over F. Since α is a root of $X^p - \alpha^p \in F(\alpha^p)[X]$, its (separable) minimal polynomial in $F(\alpha^p)[X]$ is a factor of this, so that polynomial must be $X - \alpha$. Therefore $\alpha \in F(\alpha^p)$. Taking p^k-th powers for any $k \geq 0$, $\alpha^{p^k} \in F(\alpha^{p^{k+1}})$, so

$$F(\alpha^{p^k}) \subset F(\alpha^{p^{k+1}}).$$

The reverse inclusion is obvious, so $F(\alpha^{p^k}) = F(\alpha^{p^{k+1}})$ for all $k \geq 0$. Therefore

$$L = F(\alpha) = F(\alpha^{p^k}) = K(\alpha_1, \ldots, \alpha_{r-1}, \alpha_r^{p^k})$$

for any $k \geq 0$. We can pick k so that α^{p^k} is separable over K (why?). Then the generating set $\{\alpha_1, \ldots, \alpha_{r-1}, \alpha_r^{p^k}\}$ has one less inseparable element among the field generators. This contradicts the choice of generators to have as few members in the list as possible that are inseparable over K, so α has to be inseparable over F. □

Now we turn to a proof of Theorem 1.5.

Proof. Assume L/K is separable, so by the primitive element theorem $L = K(\alpha)$ where α is separable over K. Any derivation on K can be extended (using Theorem 4.1) uniquely to a derivation on L.

If L/K is inseparable, then Lemma 4.2 lets us write $L = F(\alpha)$ with α inseparable over F, and $F \supset K$. Then, by a construction used in the proof of Theorem 4.1, $f(\alpha) \mapsto f'(\alpha)$ with $f(X) \in F[X]$ is a nonzero derivation on L which is zero on F, and thus also zero on the smaller field K. This shows the zero derivation on K has a nonzero extension (and thus two extensions) to a derivation on L. □

Corollary 4.3. *The proof of Theorem* 1.5 *implies Theorem* 1.6*.*

Proof. Again we consider the tower of field extensions (2.2). Since $L_i = L_{i-1}(\alpha_i)$ and α_i is separable over L_{i-1}, the proof of Theorem 1.5 shows any derivation on L_{i-1} extends uniquely to a derivation on L_i. Therefore any derivation on $K = L_0$ can be extended step-by-step through the tower (2.2) to a derivation on $L_r = L$. By the argument in the proof of Lemma 4.2, this derivation on L is unique. □

Lemma 4.4. *Let L/K be a finite extension and F be an intermediate extension such that F/K is separable. Then any derivation $F \to L$ which sends K to K has values in F.*

Proof. Pick $\alpha \in F$, so α is separable over K. Now use Corollary B.10 to see the derivation $F \to L$ sends α to an element of $K(\alpha) \subset F$. □

Corollary 4.5. *Theorem* 1.5 *implies Theorem* 1.7*.*

Proof. To prove L/K is separable, we want to show any derivation on K has a unique extension to a derivation on L. Since F/K is separable, a derivation on K extends to a derivation on F. Since L/F is separable, a derivation on F extends to a derivation on L. For uniqueness, let D_1 and D_2 be derivations on L which extend the same derivation on K. Since $D_1(K) \subset K$ and $D_2(K) \subset K$, we have $D_1(F) \subset F$ and $D_2(F) \subset F$ by Lemma 4.4. Then $D_1 = D_2$ on F since F/K is separable, and $D_1 = D_2$ on L since L/F is separable. □

5. Separability for infinite extensions

When L/K is an algebraic extension of possibly infinite degree, here is the way separability is defined.

Definition 5.1. An algebraic extension L/K is called *separable* if every finite subextension of L/K is separable. Equivalently, L/K is separable when every element of L is separable over K.

This definition makes no sense if L/K is not an algebraic extension since a non-algebraic extension is not the union of its finite subextensions. To deal with non-algebraic extensions, we can considering using the three new characterizations of separability for finite extensions in Theorems 1.1, 1.2, and 1.5, if they would make sense for infinite extensions.

Theorem 1.1 has a problem in the infinite-degree case: there is no natural trace map. However, the conditions in Theorems 1.2 and 1.5 both make sense for a general L/K. (In the case of Theorem 1.2, we have to drop the specification of $\overline{K} \otimes_K L$ as a product of copies of $\overline{K}$, and just leave the statement about the tensor product having no nonzero nilpotent elements.) It is left to the reader to check for an infinite algebraic extension L/K that the conditions of Theorems 1.2 and 1.5 match Definition 5.1.

The conditions in Theorems 1.2 and 1.5 both make sense if L/K is not algebraic, so they could each potentially be used to define separability of a completely arbitrary field extension. But there is a problem: for transcendental (that is, non-algebraic) extensions the conditions in Theorems 1.2 and 1.5 are no longer equivalent. Indeed, take $L = K(u)$, with u transcendental over K. Then $\overline{K} \otimes_K L = \overline{K}(u)$ is a field, so the condition in Theorem 1.2 is satisfied. However, the zero derivation on K has more than one extension to $K(u)$: the zero derivation on $K(u)$ and differentiation with respect to u on $K(u)$.

Since the conditions in Theorems 1.2 and 1.5 do not describe the same kinds of extensions in general, which one should be used to define separability? The answer is to use the condition in Theorem 1.2.

Definition 5.2. A commutative ring with no nonzero nilpotent elements is called *reduced*.

A domain is reduced, but a more worthwhile example is a product of domains, like $\mathbf{F}_3 \times \mathbf{Q}[X]$, which is not a domain but is reduced.

Definition 5.3. An arbitrary field extension L/K is called *separable* when the ring $\overline{K} \otimes_K L$ is reduced.

Using this definition, in characteristic 0 all field extensions are separable. In characteristic p, any purely transcendental extension is separable. The condition in Theorem 1.5, that derivations on the base field admit unique extensions to a larger field, characterizes not separable field extensions in general, but separable algebraic field extensions.

A condition equivalent to that in Definition 5.3 is that $F \otimes_K L$ is reduced as F runs over the finite extensions of K.

The condition that $\overline{K} \otimes_K L$ is reduced makes sense not just for field extensions L/K, but for any commutative K-algebra. Define an arbitrary commutative K-algebra A to be separable when the ring $\overline{K} \otimes_K A$ is reduced. This condition is equivalent to $A \otimes_K F$ being reduced for every finite extension field F/K.

Example 5.4. Let $A = K[X]/(f(X))$ for any nonconstant $f(X) \in K[X]$. The polynomial $f(X)$ need not be irreducible, so A might not be a field. It is a separable K-algebra precisely when $f(X)$ is a separable polynomial in $K[X]$.

When $[A : K]$ is finite, an analogue of Theorem 1.1 can be proved: A is a separable K-algebra if and only if the trace pairing $\langle x, y\rangle = \mathrm{Tr}_{A/K}(xy)$ from $A \times A$ to K is non-degenerate.

Appendix A. Traces

Let A be a finite-dimensional commutative K-algebra (with identity), such as a finite extension field of K or the product ring K^n or even a mixture of the two: a product of finite extensions of K. To any $a \in A$ we associate the K-linear map $m_a \colon A \to A$ which is left multiplication by a:

$$x \mapsto ax.$$

For $a, b \in A$ and $\alpha \in K$, $m_{a+b} = m_a + m_b$ and $m_{\alpha a} = \alpha m_a$, so m_a is a K-linear map.

Definition A.1. For a finite-dimensional K-algebra A, the *trace* of $a \in A$ is the trace of m_a.

That is, the trace of a is $\mathrm{tr}(m_a) \in K$, usually written as $\mathrm{Tr}_{A/K}(a)$, so $\mathrm{Tr}_{A/K} \colon A \to K$. The trace from A to K is K-linear, hence identically zero or surjective since K is a one-dimensional K-vector space.

Example A.2. Since m_1 is the identity function, $\mathrm{Tr}_{A/K}(1) = [A : K]$.

Example A.3. Suppose $a \in A$ is nilpotent: $a^r = 0$ for some $r \geq 1$. Then $m_a^r = O$, so m_a is a nilpotent linear transformation. Thus its eigenvalues are all 0, so $\mathrm{Tr}_{A/K}(a) = 0$.

We now consider a finite-dimensional L-algebra A with K a subfield of L such that $[L : K] < \infty$. We have finite-dimensional algebras A/L, A/K, and L/K. The next theorem is called the transitivity of the trace.

Theorem A.4. *In the above notation,* $\mathrm{Tr}_{A/K} = \mathrm{Tr}_{L/K} \circ \mathrm{Tr}_{A/L}$. *In particular, if* $a \in L$, *then* $\mathrm{Tr}_{A/K}(a) = [A : L]\mathrm{Tr}_{L/K}(a)$.

Proof. Let $(e_1, \ldots, e_m)$ be an ordered L-basis of A and $(f_1, \ldots, f_n)$ be an ordered K-basis of L. Thus as an ordered K-basis of A we can use

$$(e_1 f_1, \ldots, e_1 f_n; \ldots; e_m f_1, \ldots, e_m f_n).$$

For $a \in A$, let

$$ae_j = \sum_{i=1}^{m} c_{ij} e_i, \quad c_{ij} f_s = \sum_{r=1}^{n} b_{ijrs} f_r,$$

for $c_{ij} \in L$ and $b_{ijrs} \in K$. Thus $a(e_j f_s) = \sum_i \sum_r b_{ijrs} e_i f_r$. So

$$[m_a]_{A/L} = (c_{ij}), \quad [m_{c_{ij}}]_{L/K} = (b_{ijrs}), \quad [m_a]_{A/K} = ([m_{c_{ij}}]_{L/K}).$$

Thus

$$\begin{aligned} \mathrm{Tr}_{L/K}(\mathrm{Tr}_{A/L}(a)) &= \mathrm{Tr}_{L/K}(\sum_i c_{ii}) \\ &= \sum_i \mathrm{Tr}_{L/K}(c_{ii}) \\ &= \sum_i \sum_r b_{iirr} \\ &= \mathrm{Tr}_{A/K}(a). \end{aligned}$$

□

Theorem A.5. *Let A and B be finite-dimensional K-algebras. For (a, b) in the product ring $A \times B$, $\mathrm{Tr}_{(A\times B)/K}(a, b) = \mathrm{Tr}_{A/K}(a) + \mathrm{Tr}_{B/K}(b)$.*

Proof. Let $e_1, \ldots, e_m$ be a K-basis of A and $f_1, \ldots, f_n$ be a K-basis of B. In $A \times B$, $e_i \cdot f_j = 0$. Therefore the matrix for multiplication by (a, b), with respect to the K-basis $\{e_1, \ldots, e_m, f_1, \ldots, f_n\}$, is a block-diagonal matrix $\left(\begin{smallmatrix} [m_a] & 0 \\ 0 & [m_b] \end{smallmatrix}\right)$, whose trace is $\mathrm{Tr}_{A/K}(a) + \mathrm{Tr}_{B/K}(b)$. □

Theorem A.6. *Let A be a finite-dimensional K-algebra, L/K be a field extension, and $B = L \otimes_K A$ be the base extension of A to an L-algebra. For $a \in A$, $\mathrm{Tr}_{B/L}(1 \otimes a) = \mathrm{Tr}_{A/K}(a)$.*

Proof. Let $e_1, \ldots, e_n$ be a K-basis of A. Write $ae_j = \sum_{i=1}^n c_{ij} e_i$, so the matrix for m_a in this basis is (c_{ij}).

The tensors $1 \otimes e_1, \ldots, 1 \otimes e_n$ are an L-basis of B, and we have

$$(1 \otimes a)(1 \otimes e_j) = 1 \otimes ae_j = \sum_{i=1}^n c_{ij}(1 \otimes e_i),$$

so the matrix for $m_{1\otimes a}$ on B is the same as the matrix for m_a on A. Thus $\mathrm{Tr}_{A/K}(a) = \mathrm{Tr}_{B/L}(1 \otimes a)$. □

Remark A.7. Because $m_{1\otimes a}$ and m_a have the same matrix representation, not only are their traces the same but their characteristic polynomials are the same.

Theorem A.8. *Let A be a finite-dimensional K-algebra. For any field extension L/K, the base extension by K of the trace map $A \to K$ is the trace map $L \otimes_K A \to L$. That is, the function $\mathrm{id} \otimes \mathrm{Tr}_{A/K} \colon L \otimes_K A \to L$ which sends an elementary tensor $x \otimes a$ to $x\mathrm{Tr}_{A/K}(a)$ is the trace map $\mathrm{Tr}_{(L\otimes_K A)/L}$.*

Proof. We want to show $\mathrm{Tr}_{(L\otimes_K A)/L}(t) = (\mathrm{id} \otimes \mathrm{Tr}_{A/K})(t)$ for all $t \in L \otimes_K A$. The elementary tensors additively span $L \otimes_K A$ so it suffices to check equality when $t = x \otimes a$ for $x \in K$ and $a \in A$. This means we need to check $\mathrm{Tr}_{(L\otimes_K A)/K}(x \otimes a) = x\mathrm{Tr}_{A/K}(a)$.

Pick a K-basis $e_1, \ldots, e_n$ for A and write $ae_j = \sum_{i=1}^n c_{ij} e_i$ with $c_{ij} \in K$. The elementary tensors $1 \otimes e_1, \ldots, 1 \otimes e_n$ are an L-basis of $L \otimes_K A$ and

$$(x \otimes a)(1 \otimes e_j) = x \otimes ae_j = \sum_{i=1}^n c_{ij}(x \otimes e_i) = \sum_{i=1}^n c_{ij} x(1 \otimes e_i)$$

by the definition of the L-vector space structure on $L \otimes_K A$. So the matrix for multiplication by $x \otimes a$ in the basis $\{1 \otimes e_i\}$ is $(c_{ij}x)$, which implies

$$\mathrm{Tr}_{(L\otimes_K A)/L}(x \otimes a) = \sum_{i=1}^n c_{ii}x = x \sum_{i=1}^n c_{ii} = x\mathrm{Tr}_{A/K}(a).$$

□

Appendix B. Derivations

A derivation is an abstraction of differentiation on polynomials. We want to work with derivations on fields, but polynomial rings will intervene, so we need to understand derivations on rings before we focus on fields.

Let R be a commutative ring and M be an R-module (*e.g.*, $M = R$). A *derivation* on R with values in M is a map $D \colon R \to M$ such that $D(a + b) = D(a) + D(b)$ and

$D(ab) = aD(b) + bD(a)$. Easily, by induction $D(a^n) = na^{n-1}D(a)$ for any $n \geq 1$. When $M = R$, we will speak of a derivation on R.

Example B.1. For any commutative ring A, differentiation with respect to X on $A[X]$ is a derivation on $A[X]$ ($R = M = A[X]$).

Example B.2. Let $R = A[X]$ and $M = A$ as an R-module by $f(X)a := f(0)a$. Then $D\colon R \to M$ by $D(f) = f'(0)$ is a derivation.

Example B.3. Let $D\colon R \to R$ be a derivation. For $f(X) = \sum_i a_i X^i$ in $R[X]$, set $f^D(X) = \sum_i D(a_i)X^i$. This is the application of D coefficientwise to $f(X)$. The operation $f \mapsto f^D$ is a derivation on $R[X]$ (to check the product rule, it suffices to look at monomials).

If $R = \mathbf{F}_2[u]$ and D is the usual u-derivative on $\mathbf{F}_2[u]$, then the polynomial $f(X) = (u^3 + u)X^4 + uX^3 + u^2X + 1$ in $R[X]$ has $f^D(X) = (u^2 + 1)X^4 + X^3$.

Any element of R satisfying $D(a) = 0$ is called a D-constant, or just a constant if the derivation is understood. The constants for a derivation form a subring. For instance, from the product rule, taking $a = b = 1$, we obtain $D(1) = 0$.

Example B.4. The set of all constants for X-differentiation on $K[X]$ is K when K has characteristic 0 and $K[X^p]$ when K has characteristic p.

Example B.5. If $D\colon R \to R$ is a derivation and $f \mapsto f^D$ is the corresponding derivation on $R[X]$ from Example B.3, its ring of constants is $C[X]$, where C is the constants for D.

We will generally focus on derivations from R to R, although it will be convenient to allow R-modules as the target space for derivations in Corollary B.10 below, which is used in the main text in the proofs of Theorem 1.5 and Lemma 4.4.

Example B.6. Let's check that any derivation on $K[X]$ which has the elements of K among its constants has the form $D(f) = hf'$ for some $h \in K[X]$. (When $h = 1$, this is the usual X-derivative.)

When K is among the constants of D, D is K-linear: $D(cf) = cD(f) + fD(c) = cD(f)$. Therefore D is determined by what it does to a K-basis of $K[X]$, such as the power functions X^n. By induction, $D(X^n) = nX^{n-1}D(X)$ for all $n \geq 1$. Therefore, by linearity, $D(f) = f'(X)D(X)$ for every $f \in K[X]$. Set $h = D(X)$.

Theorem B.7. *Let R be a domain with fraction field K. Any derivation $D\colon R \to K$ uniquely extends to $\widetilde{D}\colon K \to K$, given by the quotient rule: $\widetilde{D}(a/b) = (bD(a) - aD(b))/b^2$.*

Proof. Suppose there is an extension of D to a derivation on K. Then, if $x = a/b$ is in K (with $a, b \in R$), $a = bx$, so

$$D(a) = bD(x) + xD(b).$$

Therefore in K,

$$D(x) = \frac{D(a) - xD(b)}{b} = \frac{bD(a) - aD(b)}{b^2}.$$

To see, conversely, that this formula does give a derivation $\widetilde{D}$ on K, first we check it is well-defined: if $a/b = c/d$ (with b and d nonzero), then $ad = bc$, so

$$aD(d) + dD(a) = bD(c) + cD(b).$$

Therefore

$$\begin{aligned}\frac{bD(a)-aD(b)}{b^2}-\frac{dD(c)-cD(d)}{d^2} &= \frac{d^2(bD(a)-aD(b))-b^2(dD(c)-cD(d))}{b^2d^2}\\ &= \frac{bd(dD(a)-bD(c))-d^2aD(b)+b^2cD(d)}{b^2d^2}\\ &= \frac{bd(cD(b)-aD(d))-d^2aD(b)+b^2cD(d)}{b^2d^2}\\ &= \frac{(bc-ad)dD(b)-(ad-bc)bD(d)}{b^2d^2}\\ &= 0 \quad \text{since } ad=bc.\end{aligned}$$

That $\widetilde{D}$ satisfies the sum and product rules is left to the reader to check. □

Theorem B.8. *Let L/K be a finite extension of fields, and $D\colon K\to K$ be a derivation. Suppose $\alpha\in L$ is separable over K, with minimal polynomial $\pi(X)\in K[X]$. That is, $\pi(X)$ is irreducible in $K[X]$, $\pi(\alpha)=0$, and $\pi'(\alpha)\neq 0$. Then D has a unique extension from K to a derivation on the field $K(\alpha)$, and it is given by the rule*

$$D(f(\alpha)) = f^D(\alpha) - f'(\alpha)\frac{\pi^D(\alpha)}{\pi'(\alpha)} \tag{B.1}$$

for any $f(X)\in K[X]$.

Proof. The rule (B.1) looks bizarre at first. To make it less so, we start by assuming D has an extension to $K(\alpha)$, and prove by a direct computation that it must be given by the indicated formula. For any $\beta\in K(\alpha)$, write $\beta=f(\alpha)$, where $f(X)=\sum_{i=0}^r c_iX^i$ and $c_i\in K$. Then

$$D(\beta)=D(f(\alpha))=\sum_{i=0}^{r}(D(c_i)\alpha^i+c_i(i\alpha^{i-1}D(\alpha)))=f^D(\alpha)+f'(\alpha)D(\alpha). \tag{B.2}$$

Taking $f(X)=\pi(X)$ to be the minimal polynomial of α over K, $f(\alpha)=0$, so if D has an extension to $K(\alpha)$ then (B.2) becomes

$$0=\pi^D(\alpha)+\pi'(\alpha)D(\alpha),$$

which proves (since $\pi'(\alpha)\neq 0$) that $D(\alpha)$ must be given by the formula $-\pi^D(\alpha)/\pi'(\alpha)$. Plugging this formula for $D(\alpha)$ into (B.2) shows $D(\beta)$ must be given by the formula (B.1). Since β was a general element of $K(\alpha)$, this proves D has at most one extension to a derivation on $K(\alpha)$.

Now, to show the formula (B.1) works, we start over and define

$$D(f(\alpha)) := f^D(\alpha) - f'(\alpha)\frac{\pi^D(\alpha)}{\pi'(\alpha)}.$$

We need to show this formula is well-defined.

Suppose $f_1(\alpha)=f_2(\alpha)$ for $f_1,f_2\in K[X]$. Then $f_1(X)\equiv f_2(X) \bmod \pi(X)$, say

$$f_1(X)=f_2(X)+\pi(X)k(X) \tag{B.3}$$

for some $k(X)\in K[X]$. Differentiating both sides with respect to X in the usual way,

$$f_1'(X)=f_2'(X)+\pi(X)k'(X)+\pi'(X)k(X).$$

Evaluating at $X = \alpha$,

$$f_1'(\alpha) = f_2'(\alpha) + \pi'(\alpha)k(\alpha).$$

Since $\pi'(\alpha) \neq 0$, we divide by $\pi'(\alpha)$ and multiply through by $-\pi^D(\alpha)$ to get

$$-f_1'(\alpha)\frac{\pi^D(\alpha)}{\pi'(\alpha)} = -f_2'(\alpha)\frac{\pi^D(\alpha)}{\pi'(\alpha)} - \pi^D(\alpha)k(\alpha). \tag{B.4}$$

We want to add $f_1^D(\alpha)$ to both sides. First, apply D to the coefficients in (B.3), which is a derivation on $K[X]$ (Example B.3), to get

$$f_1^D(X) = f_2^D(X) + \pi(X)k^D(X) + \pi^D(X)k(X).$$

Therefore

$$f_1^D(\alpha) = f_2^D(\alpha) + \pi^D(\alpha)k(\alpha).$$

Add this to both sides of (B.4) to get

$$\begin{aligned} f_1^D(\alpha) - f_1'(\alpha)\frac{\pi^D(\alpha)}{\pi'(\alpha)} &= f_2^D(\alpha) + \pi^D(\alpha)k(\alpha) - f_2'(\alpha)\frac{\pi^D(\alpha)}{\pi'(\alpha)} - \pi^D(\alpha)k(\alpha) \\ &= f_2^D(\alpha) - f_2'(\alpha)\frac{\pi^D(\alpha)}{\pi'(\alpha)}. \end{aligned}$$

This proves the formula for a derivation on $K(\alpha)$ is well-defined. It is left to the reader to check this really is a derivation. □

Example B.9. In contrast with Theorem B.8, consider $K = \mathbf{F}_p(u)$ and $L = K(\alpha)$ where α is a root of $X^p - u \in K[X]$. This is an inseparable irreducible polynomial over K. The u-derivative on K does not have any extension to a derivation on L. Indeed, suppose the u-derivative on K has an extension to L, and call it D. Applying D to the equation $\alpha^p = u$ gives

$$p\alpha^{p-1}D(\alpha) = D(u).$$

The left side is 0 since we're in characteristic p. The right side is 1 since D is the u-derivative on $\mathbf{F}_p(u)$. This is a contradiction, so D does not exist.

Corollary B.10. *Let L/K be a finite extension of fields. For any derivation $D\colon K \to L$ and $\alpha \in L$ which is separable over K, D has a unique extension to a derivation $K(\alpha) \to L$. If $D(K) \subset K$ then $D(K(\alpha)) \subset K(\alpha)$.*

Proof. Follow the argument in the proof of Theorem B.8, allowing derivations to have values in L rather than in $K(\alpha)$. The formula for $D(f(\alpha))$ still turns out to be the same as in (B.1). In particular, if $D(K) \subset K$ then the extension of D to a derivation on $K(\alpha)$ actually takes values in $K(\alpha)$. □