A problem on bilinear maps (problem U228 in Mathematical Reflections)

Darij Grinberg

version 0.7, 30 October 2012

[not proofread]

Problem

Let L/K be a separable algebraic extension of fields.

Let V, W and U be L-vector spaces. Then, V, W and U canonically become K-vector spaces.

Let $h : V \times W \to U$ be a K-bilinear map (not necessarily an L-bilinear map). Assume that

$$h(xa, xb) = x^2 h(a, b) \qquad \text{for every } x \in L,\ a \in V \text{ and } b \in W. \tag{1}$$

Then, prove that h is L-bilinear.

Solutions of the problem

We give three solutions to the above problem.

First solution

First solution of the problem:

We first prove the following lemma:

> **Lemma 1.** Under the conditions of the problem, let $x \in L$, $a \in V$ and $b \in W$ be arbitrary. Let $\alpha = h(a, b)$ and $\beta = h(a, xb) - xh(a, b)$. Then, every positive $n \in \mathbb{N}$ satisfies
>
> $$h(a, x^n b) = x^n \alpha + n x^{n-1} \beta. \tag{2}$$

Proof of Lemma 1. Let us prove that (2) holds for every positive $n \in \mathbb{N}$. We will prove this by strong induction over n:

Induction step[1]*:* Let $N \in \mathbb{N}$ be positive. Assume that (2) holds for every positive $n \in \mathbb{N}$ satisfying $n < N$. We must then prove that (2) holds for $n = N$.

The equality (2) holds for $n = 1$ (since

$$h\left(a, \underbrace{x^1}_{=x} b\right) = h(a, xb) = \underbrace{x}_{=x^1}\underbrace{h(a,b)}_{=\alpha} + \underbrace{h(a, xb) - xh(a, b)}_{\substack{=\beta = 1\beta = 1x^{1-1}\beta \\ \text{(since } 1x^{1-1} = 1x^0 = 1 \text{ and thus } 1 = 1x^{1-1}\text{)}}} = x\alpha + 1x^{1-1}\beta$$

). In other words, if $N = 1$, then (2) holds for $n = N$. Hence, if $N = 1$, the induction step is already completed. Thus, for the rest of the induction step, we can WLOG assume that $N \neq 1$. Assume this.

[1] A strong induction does not need an induction base.

Since $N \in \mathbb{N}$ is positive, but $N \neq 1$, we must have $N \geq 2$. Thus, $N-1$ lies in $\mathbb{N}$ and is positive. Consequently, (2) holds for $n = N-1$ (since we assumed that (2) holds for every $n \in \mathbb{N}$ satisfying $n < N$). In other words, $h\left(a, x^{N-1}b\right) = x^{N-1}\alpha + (N-1)\, x^{(N-1)-1}\beta$. Now,

$$h\left(xa, \underbrace{x^N}_{=xx^{N-1}}\ b\right) = h\left(xa, xx^{N-1}b\right) = x^2 h\left(a, x^{N-1}b\right) \tag{3}$$

(by (1), applied to $x^{N-1}b$ instead of b).

It is easy to show that

$$h\left(xa, x^{N-1}b\right) = x^N\alpha + (N-2)\, x^{N-1}\beta \tag{4}$$

[2].

But (1) (applied to $1+x$ and $x^{N-1}b$ instead of x and b) yields $h\left((1+x)\, a, (1+x)\, x^{N-1}b\right) =$

[2] *Proof of (4).* We have $N \geq 2$. Thus, we must be in one of the following two cases:

Case 1: We have $N = 2$.

Case 2: We have $N > 2$.

Let us first consider Case 1. In this case, $N = 2$, so that $N-1 = 1$, and thus $x^{N-1} = x^1 = x$, so that

$$\begin{aligned} h\left(xa, x^{N-1}b\right) &= h\left(xa, xb\right) = x^2 \underbrace{h\left(a, b\right)}_{=\alpha} \qquad \text{(by (1))} \\ &= x^2\alpha \end{aligned}$$

Compared with

$$\underbrace{x^N}_{\substack{=x^2 \\ \text{(since } N=2\text{)}}}\ \alpha + \underbrace{(N-2)}_{\substack{=0 \\ \text{(since } N=2\text{)}}}\ x^{N-1}\beta = x^2\alpha + 0x^{N-1}\beta = x^2\alpha,$$

this yields $h\left(xa, x^{N-1}b\right) = x^N\alpha + (N-2)\, x^{N-1}\beta$. Thus, (4) is proven in Case 1.

Now, let us consider Case 2. In this case, $N > 2$, so that $N-2$ is a positive element of $\mathbb{N}$. Consequently, (2) holds for $n = N-2$ (since we assumed that (2) holds for every $n \in \mathbb{N}$ satisfying $n < N$). In other words, $h\left(a, x^{N-2}b\right) = x^{N-2}\alpha + (N-2)\, x^{(N-2)-1}\beta$. Now,

$$\begin{aligned} h\left(xa, \underbrace{x^{N-1}}_{=xx^{N-2}}\ b\right) &= h\left(xa, xx^{N-2}b\right) = x^2 \underbrace{h\left(a, x^{N-2}b\right)}_{=x^{N-2}\alpha+(N-2)x^{(N-2)-1}\beta} \qquad \left(\text{by (1), applied to } x^{N-2}b \text{ instead of } b\right) \\ &= x^2\left(x^{N-2}\alpha + (N-2)\, x^{(N-2)-1}\beta\right) \\ &= \underbrace{x^2x^{N-2}}_{=x^{2+(N-2)}=x^N}\ \alpha + (N-2)\ \underbrace{x^2x^{(N-2)-1}}_{=x^{2+((N-2)-1)}=x^{N-1}}\ \beta = x^N\alpha + (N-2)\, x^{N-1}\beta. \end{aligned}$$

Thus, (4) is proven in Case 2.

Thus, in each of the two cases 1 and 2, we have shown that (4) holds. Since Cases 1 and 2 are the only two possible cases, this shows that (4) always holds, qed.

$(1+x)^2 h\left(a, x^{N-1}b\right)$. Since

$$
\begin{aligned}
&h\left(\underbrace{(1+x)\,a}_{=a+xa}, \underbrace{(1+x)\,x^{N-1}b}_{=x^{N-1}b+xx^{N-1}b}\right) \\
&= h\left(a+xa, x^{N-1}b + \underbrace{xx^{N-1}}_{=x^N} b\right) = h\left(a+xa, x^{N-1}b + x^N b\right) \\
&= h\left(a, x^{N-1}b\right) + h\left(a, x^N b\right) + h\left(xa, x^{N-1}b\right) + \underbrace{h\left(xa, x^N b\right)}_{\substack{=x^2 h\left(a, x^{N-1}b\right) \\ \text{(by (3))}}} \qquad \text{(since } h \text{ is } K\text{-bilinear)} \\
&= h\left(a, x^{N-1}b\right) + h\left(a, x^N b\right) + h\left(xa, x^{N-1}b\right) + x^2 h\left(a, x^{N-1}b\right) \\
&= h\left(a, x^N b\right) + h\left(a, x^{N-1}b\right) + x^2 h\left(a, x^{N-1}b\right) + h\left(xa, x^{N-1}b\right),
\end{aligned}
$$

this rewrites as

$$
h\left(a, x^N b\right) + h\left(a, x^{N-1}b\right) + x^2 h\left(a, x^{N-1}b\right) + h\left(xa, x^{N-1}b\right) = (1+x)^2 h\left(a, x^{N-1}b\right).
$$

In other words,

$$
\begin{aligned}
h\left(a, x^N b\right) &= \underbrace{(1+x)^2 h\left(a, x^{N-1}b\right) - h\left(a, x^{N-1}b\right) - x^2 h\left(a, x^{N-1}b\right)}_{=\left((1+x)^2 - 1 - x^2\right) h\left(a, x^{N-1}b\right)} - h\left(xa, x^{N-1}b\right) \\
&= \underbrace{\left((1+x)^2 - 1\right)}_{=2x} \underbrace{h\left(a, x^{N-1}b\right)}_{=x^{N-1}\alpha + (N-1)x^{(N-1)-1}\beta} - \underbrace{h\left(xa, x^{N-1}b\right)}_{\substack{=x^N\alpha + (N-2)x^{N-1}\beta \\ \text{(by (4))}}} \\
&= 2x\left(x^{N-1}\alpha + (N-1)\,x^{(N-1)-1}\beta\right) - \left(x^N\alpha + (N-2)\,x^{N-1}\beta\right) \\
&= 2\underbrace{xx^{N-1}}_{=x^{1+(N-1)}=x^N}\alpha + 2(N-1)\underbrace{xx^{(N-1)-1}}_{=x^{1+((N-1)-1)}=x^{N-1}}\beta - x^N\alpha - (N-2)\,x^{N-1}\beta \\
&= 2x^N\alpha + 2(N-1)\,x^{N-1}\beta - x^N\alpha - (N-2)\,x^{N-1}\beta \\
&= \underbrace{\left(2x^N\alpha - x^N\alpha\right)}_{=(2-1)x^N\alpha} + \underbrace{\left(2(N-1)\,x^{N-1}\beta - (N-2)\,x^{N-1}\beta\right)}_{=(2(N-1)-(N-2))x^{N-1}\beta} \\
&= \underbrace{(2-1)}_{=1} x^N\alpha + \underbrace{\left(2(N-1) - (N-2)\right)}_{=N} x^{N-1}\beta = x^N\alpha + Nx^{N-1}\beta.
\end{aligned}
$$

In other words, (2) holds for $n = N$. This completes the induction step. Thus, the induction proof of (2) is done. In other words, Lemma 1 is proven.

Now, we will show:

Lemma 2. Under the conditions of the problem, let $x \in L$, $a \in V$ and $b \in W$ be arbitrary. Then, $h(a, xb) = xh(a, b)$.

Proof of Lemma 2. Let $\alpha = h(a, b)$ and $\beta = h(a, xb) - xh(a, b)$. According to Lemma 1, every $n \in \mathbb{N}$ satisfies (2).

Since L is an algebraic extension of K, the element $x \in L$ has a minimal polynomial over K. Let $P \in K[X]$ be this minimal polynomial. Then, P is separable (since

L is a separable extension of K, so that x is separable over K). In other words, $\gcd\left(P, P'\right) = 1$ (where P' denotes the X-derivative of the polynomial P). Hence, no root of the polynomial P is simultaneously a root of P'. Thus, x is not a root of P' (because x is a root of the polynomial P (since P is the minimal polynomial of x)). In other words, $P'\left(x\right) \neq 0$.

Since P is the minimal polynomial of x, we have $P\left(x\right) = 0$.

Since $P \in K\left[X\right]$ is a polynomial over K, we can write P in the form $P = \sum\limits_{n=0}^{M} \lambda_n X^n$ for some $M \in \mathbb{N}$ and some elements λ_0, λ_1, ..., λ_M of K. Consider this M and these elements λ_0, λ_1, ..., λ_M.

Since $P = \sum\limits_{n=0}^{M} \lambda_n X^n$, we have $P' = \sum\limits_{n=1}^{M} n\lambda_n X^{n-1}$ (by the definition of the derivative of a polynomial), so that $P'\left(x\right) = \sum\limits_{n=1}^{M} n\lambda_n x^{n-1}$.

On the other hand, $0 = P\left(x\right) = \sum\limits_{n=0}^{M} \lambda_n x^n$ (since $P = \sum\limits_{n=0}^{M} \lambda_n X^n$), so that $0b = \sum\limits_{n=0}^{M} \lambda_n x^n b$. In other words, $0 = \sum\limits_{n=0}^{M} \lambda_n x^n b$. Hence,

$$h\left(a, 0\right) = h\left(a, \sum_{n=0}^{M} \lambda_n x^n b\right) = \sum_{n=0}^{M} \lambda_n h\left(a, x^n b\right) \qquad \text{(since } h \text{ is } K\text{-bilinear)}$$

$$= \lambda_0 h\left(a, \underbrace{x^0}_{=1} b\right) + \sum_{n=1}^{M} \lambda_n \underbrace{h\left(a, x^n b\right)}_{\substack{=x^n\alpha + nx^{n-1}\beta \\ \text{(by (2))}}} = \lambda_0 \underbrace{h\left(a, b\right)}_{\substack{=\alpha = x^0\alpha \\ \text{(since } x^0=1 \text{ and} \\ \text{thus } x^0\alpha=\alpha\text{)}}} + \sum_{n=1}^{M} \underbrace{\lambda_n \left(x^n\alpha + nx^{n-1}\beta\right)}_{=\lambda_n x^n\alpha + \lambda_n n x^{n-1}\beta}$$

$$= \lambda_0 x^0\alpha + \underbrace{\sum_{n=1}^{M} \left(\lambda_n x^n\alpha + \lambda_n n x^{n-1}\beta\right)}_{=\left(\sum\limits_{n=1}^{M} \lambda_n x^n\right)\alpha + \left(\sum\limits_{n=1}^{M} \lambda_n n x^{n-1}\right)\beta} = \underbrace{\lambda_0 x^0\alpha + \left(\sum_{n=1}^{M} \lambda_n x^n\right)\alpha}_{=\left(\lambda_0 x^0 + \sum\limits_{n=1}^{M} \lambda_n x^n\right)\alpha} + \left(\sum_{n=1}^{M} \underbrace{\lambda_n n}_{=n\lambda_n} x^{n-1}\right)\beta$$

$$= \underbrace{\left(\lambda_0 x^0 + \sum_{n=1}^{M} \lambda_n x^n\right)}_{=\sum\limits_{n=0}^{M} \lambda_n x^n = 0} \alpha + \underbrace{\left(\sum_{n=1}^{M} n\lambda_n x^{n-1}\right)}_{=P'(x)} \beta = 0\alpha + P'\left(x\right) \cdot \beta = P'\left(x\right) \cdot \beta.$$

Since $h\left(a, 0\right) = 0$ (because h is K-bilinear), this becomes $0 = P'\left(x\right) \cdot \beta$. Since $P'\left(x\right) \neq 0$, this yields $0 = \beta$ (since L is a field). Now, (2) (applied to $n = 1$) yields

$$h\left(a, x^1 b\right) = \underbrace{x^1}_{=x}\alpha + 1x^{1-1} \underbrace{\beta}_{=0} = x \underbrace{\alpha}_{=h(a,b)} + \underbrace{1x^{1-1}0}_{=0} = xh\left(a, b\right).$$

Since $x^1 = x$, this simplifies to $h\left(a, xb\right) = xh\left(a, b\right)$. This proves Lemma 2.

Notice that $h\left(a, b + b'\right) = h\left(a, b\right) + h\left(a, b'\right)$ for all $a \in V$, $b \in W$ and $b' \in W$ (since h is K-bilinear). This, combined with Lemma 2, yields that the map h is L-linear in its second variable. Similarly, the map h is L-linear in its first variable. Hence, the map h is L-linear in each of its two variables, i. e., an L-bilinear map.

Second solution

Second solution of the problem:

Let us first completely forget about the problem. In particular, let us forget about the notations K, L, V, W, U and h.

We define a well-known notion from ring theory:

Definition 11. Let L be a commutative ring. Let U be an L-module. A *derivation* from L to U means a homomorphism $\delta : L \to U$ of additive groups (not a priori required to be L-linear) which satisfies

$$(\delta(xy) = \delta(x) \cdot y + x \cdot \delta(y) \qquad \text{for all } x \in L \text{ and } y \in L).$$

Let us now prove a known fact about separable algebraic field extensions:

Proposition 12. Let L/K be a separable algebraic extension of fields. Let U be a L-vector space. Let $D : L \to U$ be a derivation[3] such that $D(K) = 0$. Then, $D = 0$.

Proof of Proposition 12. Let $x \in L$. It is easy to see that

$$D(x^n) = D(x) \cdot n x^{n-1} \qquad \text{for all positive } n \in \mathbb{N}. \tag{5}$$

[4]

Since L is an algebraic extension of K, the element $x \in L$ has a minimal polynomial over K. Let $P \in K[X]$ be this minimal polynomial. Then, P is separable (since L is a separable extension of K, so that x is separable over K). In other words, $\gcd(P, P') = 1$ (where P' denotes the X-derivative of the polynomial P). Hence, no root of the polynomial P is simultaneously a root of P'. Thus, x is not a root of P' (because x is a root of the polynomial P (since P is the minimal polynomial of x)). In other words, $P'(x) \neq 0$.

[3]The notion of a "derivation" has been defined in Definition 11. Keep in mind that a derivation isn't a priori required to be K-linear or L-linear.

[4]*Proof of (5):* We will prove (5) by induction over n:

Induction base: For $n = 1$, we have $D(x^n) = D(x^1) = D(x)$ and $D(x) \cdot n x^{n-1} = D(x) \cdot 1 \underbrace{x^{1-1}}_{=x^0=1} = D(x)$. Hence, for $n = 1$, we have $D(x^n) = D(x) = D(x) \cdot n x^{n-1}$. Thus, (5) holds for $n = 1$. This completes the induction base.

Induction step: Let m be a positive integer. Assume that (5) holds for $n = m$. We must now prove that (5) also holds for $n = m + 1$.

Since (5) holds for $n = m$, we have $D(x^m) = D(x) \cdot m x^{m-1}$. Now, since D is a derivation, we have

$$\begin{aligned} D(x^m \cdot x) &= \underbrace{D(x^m)}_{=D(x)\cdot m x^{m-1}} \cdot x + x^m \cdot D(x) = D(x) \cdot m \underbrace{x^{m-1} \cdot x}_{=x^m} + \underbrace{x^m \cdot D(x)}_{=D(x)\cdot x^m} \\ &= D(x) \cdot m x^m + D(x) \cdot x^m = D(x) \cdot \underbrace{(m x^m + x^m)}_{=(m+1)x^m} = D(x) \cdot (m+1) \underbrace{x^m}_{=x^{(m+1)-1}} \\ &= D(x) \cdot (m+1) x^{(m+1)-1}. \end{aligned}$$

Since $x^m \cdot x = x^{m+1}$, this rewrites as $D(x^{m+1}) = D(x) \cdot (m+1) x^{(m+1)-1}$. In other words, (5) also holds for $n = m + 1$. This completes the induction step. Thus, the induction proof of (5) is complete.

Since $P \in K[X]$ is a polynomial over K, we can write P in the form $P = \sum_{n=0}^{M} \lambda_n X^n$ for some $M \in \mathbb{N}$ and some elements λ_0, λ_1, ..., λ_M of K. Consider this M and these elements λ_0, λ_1, ..., λ_M.

Since $P = \sum_{n=0}^{M} \lambda_n X^n$, we have $P' = \sum_{n=1}^{M} n\lambda_n X^{n-1}$ (by the definition of the derivative of a polynomial), so that $P'(x) = \sum_{n=1}^{M} n\lambda_n x^{n-1}$.

On the other hand, from $P = \sum_{n=0}^{M} \lambda_n X^n$, we obtain $P(x) = \sum_{n=0}^{M} \lambda_n x^n$. Compared with $P(x) = 0$ (because P is the minimal polynomial of x), this yields $0 = \sum_{n=0}^{M} \lambda_n x^n$. Thus,

$$
\begin{aligned}
D(0) &= D\left(\sum_{n=0}^{M} \lambda_n x^n\right) = \sum_{n=0}^{M} D(\lambda_n x^n) \\
&\qquad \text{(since } D \text{ is a derivation and thus a homomorphism of additive groups)} \\
&= D\left(\lambda_0 \underbrace{x^0}_{=1}\right) + \sum_{n=1}^{M} \underbrace{D(\lambda_n x^n)}_{\substack{=D(\lambda_n)\cdot x^n + \lambda_n \cdot D(x^n) \\ \text{(since } D \text{ is a derivation)}}} \\
&= \underbrace{D(\lambda_0)}_{\substack{=0 \\ \text{(since } \lambda_0 \in K \text{ and thus} \\ D(\lambda_0) \in D(K) = 0)}} + \sum_{n=1}^{M} \left(\underbrace{D(\lambda_n)}_{\substack{=0 \\ \text{(since } \lambda_n \in K \text{ and thus} \\ D(\lambda_n) \in D(K) = 0)}} \cdot x^n + \lambda_n \cdot D(x^n) \right) \\
&= \sum_{n=1}^{M} \left(\underbrace{0 \cdot x^n}_{=0} + \lambda_n \cdot D(x^n) \right) = \sum_{n=1}^{M} \lambda_n \cdot \underbrace{D(x^n)}_{\substack{=D(x)\cdot n x^{n-1} \\ \text{(by (5))}}} = \sum_{n=1}^{M} \lambda_n \cdot D(x) \cdot n x^{n-1} \\
&= D(x) \cdot \underbrace{\sum_{n=1}^{M} n\lambda_n x^{n-1}}_{=P'(x)} = D(x) \cdot P'(x).
\end{aligned}
$$

Compared with $D(0) = 0$ (because D is a derivation and thus a homomorphism of additive groups), this yields $D(x) \cdot P'(x) = 0$. We can divide this equation by $P'(x)$ (this is allowed since $P'(x) \neq 0$), and thus obtain $D(x) = 0$.

Now forget that we fixed x. We thus have proven that every $x \in L$ satisfies $D(x) = 0$. In other words, $D = 0$. This proves Proposition 12.

Next, let us state a very simple fact:

Lemma 13. Let K be a commutative ring, and L a commutative K-algebra.

Let V, W and U be L-modules. Then, V, W and U canonically become K-modules.

Let $h : V \times W \to U$ be a K-bilinear map (not necessarily an L-bilinear map). Assume that

$$h(xa, xb) = x^2 h(a, b) \qquad \text{for every } x \in L,\ a \in V \text{ and } b \in W. \tag{6}$$

Then,

$$h(xa, yb) + h(ya, xb) = 2xyh(a, b) \qquad \text{for every } x \in L,\ y \in L,\ a \in V \text{ and } b \in W. \tag{7}$$

Proof of Lemma 13. Let $x \in L$, $y \in L$, $a \in V$ and $b \in W$. Applying (6) to $x + y$ instead of x, we obtain

$$h((x+y)a, (x+y)b) = (x+y)^2 h(a, b).$$

Since

$$\begin{aligned}
&h\left(\underbrace{(x+y)a}_{=xa+ya}, \underbrace{(x+y)b}_{=xb+yb}\right) \\
&= h(xa + ya, xb + yb) = \underbrace{h(xa, xb+yb)}_{\substack{=h(xa,xb)+h(xa,yb) \\ \text{(since } h \text{ is } K\text{-bilinear)}}} + \underbrace{h(ya, xb+yb)}_{\substack{=h(ya,xb)+h(ya,yb) \\ \text{(since } h \text{ is } K\text{-bilinear)}}} \\
&\qquad \text{(since } h \text{ is } K\text{-bilinear)} \\
&= \underbrace{h(xa, xb)}_{\substack{=x^2h(a,b) \\ \text{(by (6))}}} + h(xa, yb) + h(ya, xb) + \underbrace{h(ya, yb)}_{\substack{=y^2h(a,b) \\ \text{(by (6), applied} \\ \text{to } y \text{ instead of } x)}} \\
&= x^2 h(a, b) + h(xa, yb) + h(ya, xb) + y^2 h(a, b),
\end{aligned}$$

this rewrites as

$$x^2 h(a, b) + h(xa, yb) + h(ya, xb) + y^2 h(a, b) = (x+y)^2 h(a, b).$$

Subtracting $x^2 h(a, b) + y^2 h(a, b)$ from this equation, we obtain

$$\begin{aligned}
h(xa, yb) + h(ya, xb) &= (x+y)^2 h(a, b) - x^2 h(a, b) - y^2 h(a, b) \\
&= \underbrace{\left((x+y)^2 - x^2 - y^2\right)}_{=2xy} h(a, b) = 2xyh(a, b).
\end{aligned}$$

This proves Lemma 13.

The next lemma lies at the heart of our second solution:

Lemma 14. Let K be a commutative ring, and L a commutative K-algebra.

Let V, W and U be L-modules. Then, V, W and U canonically become K-modules.

Let $h : V \times W \to U$ be a K-bilinear map (not necessarily an L-bilinear map). Assume that

$$h(xa, yb) + h(ya, xb) = 2xyh(a, b) \qquad \text{for every } x \in L,\ y \in L,\ a \in V \text{ and } b \in W. \tag{8}$$

Assume that

$$\text{every derivation } D : L \to U \text{ satisfying } D(K) = 0 \text{ satisfies } D = 0. \tag{9}$$

Then:

(a) Any $x \in L$, $y \in L$, $a \in V$ and $b \in W$ satisfy $h(xa, yb) = h(ya, xb)$.

(b) Any $x \in L$, $a \in V$ and $b \in W$ satisfy $2h(a, xb) = 2xh(a, b)$.

(c) Any $x \in L$, $a \in V$ and $b \in W$ satisfy $2h(xa, b) = 2xh(a, b)$.

Proof of Lemma 14. Let $a \in V$ and $b \in W$.
Define a map $E : L \to U$ by

$$(E(x) = 2h(a, xb) - 2xh(a, b) \qquad \text{for all } x \in L).$$

We are now going to show that E is a derivation.

Since h is K-bilinear, it is easy to see that E is K-linear. In particular, E is a homomorphism of additive groups.

(b) Now, it is easy to see that any $x \in L$ and $y \in L$ satisfy

$$2yh(a, xb) - h(ya, xb) = h(a, xyb). \tag{10}$$

[5] Also, any $x \in L$ and $y \in L$ satisfy

$$\begin{aligned} 2xh(a, yb) - h(xa, yb) &= h\left(a, \underbrace{yx}_{=xy} b\right) \qquad \text{(by (10), applied to } y \text{ and } x \text{ instead of } x \text{ and } y\text{)} \\ &= h(a, xyb). \end{aligned} \tag{11}$$

Now, let $x \in L$ and $y \in L$. Adding (10) to (11), we obtain

$$2yh(a, xb) - h(ya, xb) + 2xh(a, yb) - h(xa, yb) = h(a, xyb) + h(a, xyb) = 2h(a, xyb).$$

[5] *Proof of (10):* Let $x \in L$ and $y \in L$. Applying (8) to y, 1 and xb instead of x, y and b, we obtain

$$h(ya, 1 \cdot xb) + h(1a, y \cdot xb) = 2y \cdot 1 \cdot h(a, xb).$$

Thus,

$$h(1a, y \cdot xb) = \underbrace{2y \cdot 1}_{=2y} \cdot h(a, xb) - h\left(ya, \underbrace{1 \cdot xb}_{=xb}\right) = 2yh(a, xb) - h(ya, xb).$$

In other words, $2yh(a, xb) - h(ya, xb) = h\left(\underbrace{1a}_{=a}, \underbrace{y \cdot x}_{=xy} b\right) = h(a, xyb)$. This proves (10).

Thus,

$$\begin{aligned} 2h\left(a, xyb\right) &= 2yh\left(a, xb\right) - h\left(ya, xb\right) + 2xh\left(a, yb\right) - h\left(xa, yb\right) \\ &= 2yh\left(a, xb\right) + 2xh\left(a, yb\right) - \underbrace{\left(h\left(xa, yb\right) + h\left(ya, xb\right)\right)}_{\substack{=2xyh(a,b) \\ \text{(by (8))}}} \\ &= 2yh\left(a, xb\right) + 2xh\left(a, yb\right) - 2xyh\left(a, b\right). \end{aligned} \tag{12}$$

Now, by the definition of E, we have

$$\begin{aligned} E\left(xy\right) &= \underbrace{2h\left(a, xyb\right)}_{\substack{=2yh(a,xb)+2xh(a,yb)-2xyh(a,b) \\ \text{(by (12))}}} - 2xyh\left(a, b\right) \\ &= 2yh\left(a, xb\right) + 2xh\left(a, yb\right) - 2xyh\left(a, b\right) - 2xyh\left(a, b\right) \\ &= \underbrace{2yh\left(a, xb\right) - 2xyh\left(a, b\right)}_{=(2h(a,xb)-2xh(a,b))\cdot y} + \underbrace{2xh\left(a, yb\right) - 2xyh\left(a, b\right)}_{=x\cdot(2h(a,yb)-2yh(a,b))} \\ &= \underbrace{\left(2h\left(a, xb\right) - 2xh\left(a, b\right)\right)}_{\substack{=E(x) \\ \text{(since } E(x)=2h(a,xb)-2xh(a,b) \\ \text{by the definition of } E)}} \cdot y + x \cdot \underbrace{\left(2h\left(a, yb\right) - 2yh\left(a, b\right)\right)}_{\substack{=E(y) \\ \text{(since } E(y)=2h(a,yb)-2yh(a,b) \\ \text{by the definition of } E)}} \\ &= E\left(x\right) \cdot y + x \cdot E\left(y\right). \end{aligned}$$

Now forget that we fixed x and y. We thus have proven that every $x \in L$ and $y \in L$ satisfy $E\left(xy\right) = E\left(x\right) \cdot y + x \cdot E\left(y\right)$. Combined with the fact that E is a homomorphism of additive groups, this yields that E is a derivation.

Every $x \in K$ satisfies

$$E\left(x\right) = 2\underbrace{h\left(a, xb\right)}_{\substack{=xh(a,b) \\ \text{(since } h \text{ is } K\text{-bilinear)}}} - 2xh\left(a, b\right) = 2xh\left(a, b\right) - 2xh\left(a, b\right) = 0.$$

In other words, $E\left(K\right) = 0$. Hence, (9) (applied to $D = E$) yields $E = 0$. Thus, every $x \in L$ satisfies $E\left(x\right) = 0$. Since $E\left(x\right) = 2h\left(a, xb\right) - 2xh\left(a, b\right)$ (by the definition of E), this rewrites as follows: Every $x \in L$ satisfies $2h\left(a, xb\right) - 2xh\left(a, b\right) = 0$. In other words,

$$\text{every } x \in L \text{ satisfies } 2h\left(a, xb\right) = 2xh\left(a, b\right).$$

This proves Lemma 14 **(b)**.

(c) A similar argument, but with the roles of the left and the right variable interchanged, proves Lemma 14 **(c)**.

(a) Now, let $x \in L$, $y \in L$, $a \in V$ and $b \in W$. Then, (8) yields

$$\begin{aligned} &h\left(xa, yb\right) + h\left(ya, xb\right) \\ &= 2xyh\left(a, b\right) = x \cdot \underbrace{2yh\left(a, b\right)}_{\substack{=2h(ya,b) \\ \text{(since Lemma 14 (c) (applied to } y \text{ instead of } x) \\ \text{yields } 2h(ya,b)=2yh(a,b))}} \\ &= x \cdot 2h\left(ya, b\right) = 2xh\left(ya, b\right) = 2h\left(ya, xb\right) \\ &\qquad \text{(since Lemma 14 (b) (applied to } ya \text{ instead of } a) \text{ yields } 2h\left(ya, xb\right) = 2xh\left(ya, b\right)). \end{aligned}$$

Subtracting $h(ya, xb)$ from this, we obtain $h(xa, yb) = h(ya, xb)$. This proves Lemma 14 **(a)**. Thus, Lemma 14 is proven.

While Lemma 14 is enough to easily solve the problem in all characteristics but 2, characteristic 2 doesn't allow the solution to be completed that easily. Let us prove another field-theoretical proposition:

> **Proposition 15.** Let p be a prime number. Let L/K be a separable algebraic extension of fields of characteristic p. Then, every element of L is a K-linear combination of p-th powers of elements of L.

To prove this, we need a known fact from algebra:

> **Lemma 16.** If B/A is a field extension, and Q and R are two polynomials in $A[X]$, then the greatest common divisor of the polynomials Q and R in $A[X]$ is identical with the greatest common divisor of the polynomials Q and R in $B[X]$.

The proof of Lemma 16 becomes trivial once we notice that the Euclidean algorithm (for computing the greatest common divisor of two polynomials) does not depend on the field in which the polynomials lie in, so that computing the greatest common divisor of the polynomials Q and R in $A[X]$ using the Euclidean algorithm yields the same result as computing the greatest common divisor of the polynomials Q and R in $B[X]$ using the Euclidean algorithm. The details of this proof are left to the reader, unless he already knows them.

Proof of Proposition 15. Let S denote the set of all K-linear combinations of p-th powers of elements of L.

First, let us show that S is a K-subalgebra of L (and thus a subfield of L).

We defined S as the set of all K-linear combinations of p-th powers of elements of L. In other words, S is the K-linear span of the set of all p-th powers of elements of L. Thus, S is a K-vector subspace of L.

It is very easy to show that S is also a K-subalgebra of L (since the product of p-th powers is a p-th power). Since L is a field and L/K is an algebraic extension, this yields that S is a subfield of L and a field extension of K (because whenever L/K is an algebraic extension of fields, every K-subalgebra of L is a subfield of L).

Now, let $x \in L$ be arbitrary. Since L is an algebraic extension of K, the element $x \in L$ has a minimal polynomial over K. Let $P \in K[X]$ be this minimal polynomial. Then, P is separable (since L is a separable extension of K, so that x is separable over K). Thus, P has no multiple roots over any field extension of K. In particular, P has no multiple roots over S (since S is a field extension of K).

Since $x^p \in S$, the polynomial $X^p - x^p$ is well-defined in the polynomial ring $S[X]$.

Now, Lemma 16 (applied to $A = S$, $B = L$, $Q = P$ and $R = X^p - x^p$) yields that the greatest common divisor of the polynomials P and $X^p - x^p$ in $S[X]$ is identical with the greatest common divisor of the polynomials P and $X^p - x^p$ in $L[X]$. Thus, we can denote the greatest common divisor of the polynomials P and $X^p - x^p$ by $\gcd(P, X^p - x^p)$ without having to worry about whether it is taken in $S[X]$ or in $L[X]$.

The polynomial P has no multiple roots over S. Thus, every divisor of P in $S[X]$ also has no multiple roots over S (because if a polynomial has no multiple roots, then

every divisor of this polynomial must also have no multiple roots). This yields that $\gcd(P, X^p - x^p)$ has no multiple roots over S (because $\gcd(P, X^p - x^p)$ is a divisor of P in $S[X]$).

But $\gcd(P, X^p - x^p) \mid X^p - x^p = (X - x)^p$ (since the characteristic of L is p). Thus, the polynomial $\gcd(P, X^p - x^p)$ is a divisor of the polynomial $(X - x)^p$. Since all divisors of the polynomial $(X - x)^p$ have the form $(X - x)^k$ with $k \in \{0, 1, ..., p\}$ (because $L[X]$ is a unique factorization domain, and the linear polynomial $X - x$ is irreducible), this yields that the polynomial $\gcd(P, X^p - x^p)$ has the form $(X - x)^k$ with $k \in \{0, 1, ..., p\}$. In other words, there exists some $k \in \{0, 1, ..., p\}$ such that $\gcd(P, X^p - x^p) = (X - x)^k$. Consider this k.

Since $\gcd(P, X^p - x^p)$ has no multiple roots over S, we thus conclude that $(X - x)^k$ has no multiple roots over S (since $\gcd(P, X^p - x^p) = (X - x)^k$).

Since $P(x) = 0$, we have $X - x \mid P$ in $L[X]$. Since $(X^p - x^p)(x) = x^p - x^p = 0$, we have $X - x \mid X^p - x^p$ in $L[X]$.

Since $X - x \mid P$ and $X - x \mid X^p - x^p$, the polynomial $X - x$ must be a common divisor of the polynomials P and $X^p - x^p$. Hence, $X - x$ must divide the greatest common divisor of the polynomials P and $X^p - x^p$. In other words, $X - x \mid \gcd(P, X^p - x^p) = (X - x)^k$. Thus, $(X - x)^k(x) = 0$. Hence, $k \neq 0$ (because otherwise, we would have $k = 0$, so that $(X - x)^k(x) = \underbrace{(X - x)^0}_{=1}(x) = 1(x) = 1 \neq 0$ would contradict $(X - x)^k(x) = 0$).

If we had $k \geq 2$, then the polynomial $(X - x)^k$ would have multiple roots over S (namely, the root x would appear k times), contradicting the fact that $(X - x)^k$ has no multiple roots over S. Thus, we cannot have $k \geq 2$. In other words, we have $k \leq 1$. Combined with $k \neq 0$, this yields $k = 1$. Thus, $(X - x)^k = (X - x)^1 = X - x$. So we have $X - x = (X - x)^k = \gcd(P, X^p - x^p) \in S[X]$ and thus $x \in S$.

Now forget that we fixed x. We thus have proven that $x \in S$ for every $x \in L$. In other words, $L \subseteq S$. Hence,

$$L \subseteq S = (\text{set of all } K\text{-linear combinations of } p\text{-th powers of elements of } L).$$

Hence, every element of L is a K-linear combination of p-th powers of elements of L. This proves Proposition 15.

We can now finally start solving the problem. Let K, L, V, W, U and h be as defined in the problem.

By the conditions of the problem,

$$h(xa, xb) = x^2 h(a, b) \qquad \text{for every } x \in L,\ a \in V \text{ and } b \in W.$$

Hence, Lemma 13 yields that

$$h(xa, yb) + h(ya, xb) = 2xyh(a, b) \qquad \text{for every } x \in L,\ y \in L,\ a \in V \text{ and } b \in W. \tag{13}$$

Proposition 12 yields that

$$\text{every derivation } D : L \to U \text{ satisfying } D(K) = 0 \text{ satisfies } D = 0. \tag{14}$$

Since (13) and (14) hold, we can apply Lemma 14 **(a)**, and conclude that

$$\text{any } x \in L,\ y \in L,\ a \in V \text{ and } b \in W \text{ satisfy } h(xa, yb) = h(ya, xb). \tag{15}$$

Now, we can easily see that

$$\text{any } x \in L,\ a \in V \text{ and } b \in W \text{ satisfy } h\left(x^2 a, b\right) = x^2 h\left(a, b\right). \tag{16}$$

[6]

Now, let $x \in L$, $a \in V$ and $b \in W$ be arbitrary. We will prove that

$$h\left(xa, b\right) = xh\left(a, b\right). \tag{17}$$

Proof of (17): We distinguish between two cases:
Case 1: We have $\operatorname{char} K \neq 2$.
Case 2: We have $\operatorname{char} K = 2$.

First, let us consider Case 1. In this case, $\operatorname{char} K \neq 2$. Since L is a field extension of K, we have $\operatorname{char} L = \operatorname{char} K \neq 2$.

Now, applying (16) to $x+1$ instead of x, we obtain $h\left(\left(x+1\right)^2 a, b\right) = \left(x+1\right)^2 h\left(a, b\right)$. Thus,

$$\begin{aligned}
\left(x+1\right)^2 h\left(a, b\right) &= h\left(\underbrace{\left(x+1\right)^2}_{=x^2+2x+1} a, b\right) = h\left(\underbrace{\left(x^2+2x+1\right)}_{=x^2a+2xa+a} a, b\right) = h\left(x^2 a + 2xa + a, b\right) \\
&= \underbrace{h\left(x^2 a, b\right)}_{\substack{=x^2 h(a,b) \\ \text{(by (16))}}} + 2h\left(xa, b\right) + h\left(a, b\right) \qquad \text{(since } h \text{ is } K\text{-bilinear)} \\
&= x^2 h\left(a, b\right) + 2h\left(xa, b\right) + h\left(a, b\right),
\end{aligned}$$

so that

$$2h\left(xa, b\right) = \left(x+1\right)^2 h\left(a, b\right) - x^2 h\left(a, b\right) - h\left(a, b\right) = \underbrace{\left(\left(x+1\right)^2 - x^2 - 1\right)}_{=2x} h\left(a, b\right) = 2xh\left(a, b\right).$$

Since 2 is invertible in L (because $\operatorname{char} L \neq 2$), we can divide this equality by 2, and obtain $h\left(xa, b\right) = xh\left(a, b\right)$. This proves (17) in Case 1.

Now, we will consider Case 2. In this case, $\operatorname{char} K = 2$. Since L is a field extension of K, we have $\operatorname{char} L = \operatorname{char} K = 2$. Thus, L/K is a separable algebraic extension of fields of characteristic 2. Hence, x is a K-linear combination of 2-nd powers of elements of L (since Proposition 15 (applied to $p = 2$) shows that every element of L is a K-linear combination of 2-nd powers of elements of L). In other words, there exists an $N \in \mathbb{N}$, some N elements λ_1, λ_2, ..., λ_N of K and some N elements α_1, α_2, ..., α_N of L such that $x = \sum\limits_{i \in \{1,2,\ldots,N\}} \lambda_i \alpha_i^2$. Consider this N, these λ_1, λ_2, ..., λ_N and these α_1,

[6] *Proof of (16):* Let $x \in L$, $a \in V$ and $b \in W$. Applying (15) to 1, xa and x instead of x, a and y, we obtain $h\left(1 \cdot xa, xb\right) = h\left(xxa, 1 \cdot b\right)$. Thus,

$$h\left(\underbrace{x^2}_{=xx} a, \underbrace{b}_{=1\cdot b}\right) = h\left(xxa, 1 \cdot b\right) = h\left(\underbrace{1 \cdot xa}_{=xa}, xb\right) = h\left(xa, xb\right) = x^2 h\left(a, b\right)$$

(by (1)). This proves (16).

α_2, ..., α_N. Since $x = \sum_{i \in \{1,2,...,N\}} \lambda_i \alpha_i^2$, we have

$$h(xa, b) = h\left(\sum_{i \in \{1,2,...,N\}} \lambda_i \alpha_i^2 a, b \right) = \sum_{i \in \{1,2,...,N\}} \lambda_i \underbrace{h\left(\alpha_i^2 a, b\right)}_{\substack{= \alpha_i^2 h(a,b) \\ \text{(by (16), applied to} \\ \alpha_i \text{ instead of } x\text{)}}}$$

$$\text{(since } h \text{ is } K\text{-bilinear)}$$

$$= \underbrace{\sum_{i \in \{1,2,...,N\}} \lambda_i \alpha_i^2 h(a, b)}_{= x} = x h(a, b).$$

This proves (17) in Case 2.

Thus, (17) is proven in each of the cases 1 and 2. Since these two cases cover all possibilities, this yields that (17) always holds.

Now, forget that we fixed x, a and b. We thus have proven (17) for all $x \in L$, $a \in V$ and $b \in W$. Combined with the fact that $h(a + a', b) = h(a, b) + h(a', b)$ for all $a \in V$, $a' \in V$ and $b \in W$ (since h is K-bilinear), this yields the map h is L-linear in its first variable. Similarly, the map h is L-linear in its second variable. Hence, the map h is L-linear in each of its two variables, i. e., an L-bilinear map. This solves the problem.

Third solution

Third solution of the problem:

The following solution is the least elementary (it requires tensor products of bimodules over rings), but results in the strongest generalization of the problem. We are going to use Lemmas 13 and 14 from the previous solution.

Let us first completely forget about the problem. In particular, let us forget about the notations K, L, V, W, U and h.

First, we define the notion of derivations from rings to bimodules:

> **Definition 21.** Let L be a ring. Let U be an (L, L)-bimodule (i. e., an abelian group endowed with both a left and a right action of L which satisfy $(xa) y = x(ay)$ for all $a \in U$, $x \in L$ and $y \in L$). A *derivation* from L to U means a homomorphism $\delta : L \to U$ of additive groups (not a priori required to be L-linear) which satisfies
>
> $$(\delta(xy) = \delta(x) \cdot y + x \cdot \delta(y) \qquad \text{for all } x \in L \text{ and } y \in L).$$

Of course, when L is a commutative ring, then any L-module U canonically becomes an (L, L)-bimodule[7], and the notion of a "derivation from L to the (L, L)-bimodule U"

[7] In fact, if U is an L-module, then we can interpret the action of L on U both as a left action and as a right action, and these two actions satisfy $(xa) y = x(ay)$ for all $a \in U$, $x \in L$ and $y \in L$ (because $(xa) y = y(xa) = \underbrace{(yx)}_{\substack{= xy \\ \text{(since } L \text{ is} \\ \text{commutative)}}} a = (xy) a = x(ya) = x(ay)$), so that U becomes an (L, L)-bimodule this way.

as defined in Definition 21 becomes equivalent to the notion of "derivation from L to the L-module U" as defined in Definition 11. Hence, Definition 21 can be considered a generalization of Definition 11.

Next, we define the notion of *separable algebras*. There are several equivalent definitions of this notion; we take the following one:

Definition 22. Let K be a commutative ring. Let L be a K-algebra (not necessarily commutative). Let $\varepsilon_{K,L}$ be the additive group homomorphism $L \otimes_K L \to L$ which satisfies

$$(\varepsilon_{K,L}(a \otimes_K b) = ab \qquad \text{for all } a \in L \text{ and } b \in L).$$

(This $\varepsilon_{K,L}$ is well-defined due to well-known properties of tensor products, and is easily seen to be an (L, L)-bimodule homomorphism.) We say that the K-algebra L is *separable* if there exists an element $e \in L \otimes_K L$ satisfying $\varepsilon_{K,L}(e) = 1$ and ($ae = ea$ for all $a \in L$). (Here, ae and ea are computed in the (K, K)-bimodule $L \otimes_K L$.)

This notion of separability is not literally a generalization of separable field extensions, but is closely related due to the following theorem:

Theorem 23. Let L/K be a separable **finite** extension of fields. Then, L is a separable K-algebra.

Note that Theorem 23 would not be true if we would drop the word "finite". Indeed, a result by Villamayor and Zelinsky yields that every separable K-algebra over a field K must be finite-dimensional as a K-vector space.

We don't need the full force of Theorem 23, but only the following particular case:

Proposition 24. Let L/K be a separable algebraic extension of fields. Let $x \in L$. Then, $K[x]$ is a separable K-algebra.

Proposition 24 is an obvious consequence of Theorem 23. Theorem 23, however, is also a trivial corollary of Proposition 24 using the primitive element theorem. We are not going to elaborate on this, since we are not going to need Theorem 23. Let us give a self-contained proof of Proposition 24 without recurrence to Theorem 23:

Alternative proof of Proposition 24. Let $S = K[x]$. Since x is algebraic over K (because L is an algebraic extension of K), this S is a field.

Since L is an algebraic extension of K, the element $x \in L$ has a minimal polynomial over K. Let $P \in K[X]$ be this minimal polynomial. Then, P is separable (since L is a separable extension of K, so that x is separable over K). In other words, P has no multiple roots over any field extension of K. In particular, P has no multiple roots over S (since S is a field extension of K).

Since P is a minimal polynomial, it is clear that P is monic. In other words, the leading coefficient of P is 1. Let $M = \deg P$.

Since P is the minimal polynomial of x, we have $P(x) = 0$. Thus, $X - x \mid P$ in $S[X]$. Hence, $\dfrac{P}{X-x}$ is a well-defined element of $S[X]$. Denote this element $\dfrac{P}{X-x}$ by Q. Then, $P = Q \cdot (X - x)$.

It is easy to see that $Q(x) \neq 0$. [8] Also, $Q = \dfrac{P}{X-x}$ by the definition of Q. Thus, $\deg Q = \deg \dfrac{P}{X-x} = \underbrace{\deg P}_{=M} - \underbrace{\deg (X-x)}_{=1} = M-1$.

Since $P \in K[X]$ is a polynomial over K with degree $\deg P = M$, we can write P in the form $P = \sum\limits_{n=0}^{M} a_n X^n$ for some elements a_0, a_1, ..., a_M of K satisfying $a_M \neq 0$. Consider these elements a_0, a_1, ..., a_M. Then, clearly, a_M is the leading coefficient of P, so that $a_M = 1$ (because the leading coefficient of P is 1).

Since $P = \sum\limits_{n=0}^{M} a_n X^n$, we have $P(x) = \sum\limits_{n=0}^{M} a_n x^n = \underbrace{a_M}_{=1} x^M + \sum\limits_{n=0}^{M-1} a_n x^n = x^M + \sum\limits_{n=0}^{M-1} a_n x^n$. Thus,

$$x^M = \underbrace{P(x)}_{=0} - \sum_{n=0}^{M-1} \underbrace{a_n x^n}_{=x^n a_n} = -\sum_{n=0}^{M-1} x^n a_n. \tag{18}$$

Since $Q \in S[X]$ is a polynomial over S with degree $\deg Q = M-1$, we can write Q in the form $Q = \sum\limits_{n=0}^{M-1} b_n X^n$ for some elements b_0, b_1, ..., b_{M-1} of S. Consider these elements b_0, b_1, ..., b_{M-1}. Also, define two elements b_{-1} and b_M of S by $b_{-1} = 0$ and $b_M = 0$.

Now,

$$\sum_{n=0}^{M} a_n X^n = P = \underbrace{Q}_{=\sum\limits_{n=0}^{M-1} b_n X^n} \cdot (X-x) = \left(\sum_{n=0}^{M-1} b_n X^n \right) \cdot (X-x) = \sum_{n=0}^{M-1} b_n \underbrace{X^n \cdot X}_{=X^{n+1}} - \sum_{n=0}^{M-1} \underbrace{b_n X^n \cdot x}_{=b_n x X^n}$$

$$= \sum_{n=0}^{M-1} b_n X^{n+1} - \sum_{n=0}^{M-1} b_n x X^n = \sum_{n=1}^{M} b_{n-1} \underbrace{X^{n-1+1}}_{=X^n} - \sum_{n=0}^{M-1} b_n x X^n$$

(here, we substituted n for $n+1$ in the first sum)

$$= \sum_{n=1}^{M} b_{n-1} X^n - \sum_{n=0}^{M-1} b_n x X^n.$$

[8] *Proof.* Assume the opposite. Thus, $Q(x) = 0$, so that $X - x \mid Q$ in $S[X]$. Hence, $(X-x)^2 \mid Q \cdot (X-x) = P$ in $S[X]$. The polynomial P must therefore have a multiple root over S (namely, the root x appears at least twice). But this contradicts the fact that P has no multiple roots over S. This contradiction shows that our assumption was wrong. Hence, $Q(x) \neq 0$, qed.

Compared with

$$\sum_{n=0}^{M}\left(b_{n-1}-b_n x\right)X^n = \underbrace{\sum_{n=0}^{M} b_{n-1}X^n}_{=b_{-1}X^0+\sum\limits_{n=1}^{M} b_{n-1}X^n} - \underbrace{\sum_{n=0}^{M} b_n x X^n}_{=\sum\limits_{n=0}^{M-1} b_n x X^n + b_M x X^M}$$

$$= \left(\underbrace{b_{-1}}_{=0}X^0 + \sum_{n=1}^{M} b_{n-1}X^n\right) - \left(\sum_{n=0}^{M-1} b_n x X^n + \underbrace{b_M}_{=0} x X^M\right)$$

$$= \left(\underbrace{0X^0}_{=0} + \sum_{n=1}^{M} b_{n-1}X^n\right) - \left(\sum_{n=0}^{M-1} b_n x X^n + \underbrace{0xX^M}_{=0}\right) = \sum_{n=1}^{M} b_{n-1}X^n - \sum_{n=0}^{M-1} b_n x X^n,$$

this yields $\sum\limits_{n=0}^{M} a_n X^n = \sum\limits_{n=0}^{M}\left(b_{n-1}-b_n x\right)X^n$. Comparing coefficients on both sides of this equation, we obtain that

$$a_n = b_{n-1} - b_n x \qquad \text{for every } n \in \{0,1,...,M\}. \tag{19}$$

Applying (19) to $n = M$, we obtain $a_M = b_{M-1} - \underbrace{b_M}_{=0} x = b_{M-1} - \underbrace{0x}_{=0} = b_{M-1}$. Thus, $b_{M-1} = a_M = 1$.

Now, define an element $f \in S \otimes_K S$ by $f = \sum\limits_{n=0}^{M-1} x^n \otimes_K b_n$. Then,

$$\varepsilon_{K,S}(f) = \varepsilon_{K,S}\left(\sum_{n=0}^{M-1} x^n \otimes_K b_n\right) = \sum_{n=0}^{M-1} \underbrace{\varepsilon_{K,S}\left(x^n \otimes_K b_n\right)}_{\substack{=x^n b_n \\ \text{(by the definition of } \varepsilon_{K,S}\text{)}}}$$

(since $\varepsilon_{K,S}$ is a (K,K)-bimodule homomorphism)

$$= \sum_{n=0}^{M-1} \underbrace{x_n b^n}_{=b^n x_n} = \sum_{n=0}^{M-1} b_n x^n = Q(x) \qquad \left(\text{since } Q = \sum_{n=0}^{M-1} b_n X^n \text{ and thus } Q(x) = \sum_{n=0}^{M-1} b_n x^n\right). \tag{20}$$

Now, since $f = \sum\limits_{n=0}^{M-1} x^n \otimes_K b_n$, we have

$$xf = x \sum_{n=0}^{M-1} x^n \otimes_K b_n = \sum_{n=0}^{M-1} \underbrace{xx^n}_{=x^{n+1}} \otimes_K b_n = \sum_{n=0}^{M-1} x^{n+1} \otimes_K b_n$$

$$= \sum_{n=1}^{M} \underbrace{x^{n-1+1}}_{=x^n} \otimes_K b_{n-1} \qquad \text{(here, we substituted } n-1 \text{ for } n\text{)}$$

$$= \sum_{n=1}^{M} x^n \otimes_K b_{n-1} = \sum_{n=0}^{M} x^n \otimes_K b_{n-1}$$

$$\left(\text{since } \sum_{n=0}^{M} x^n \otimes_K b_{n-1} = x^0 \otimes_K \underbrace{b_{-1}}_{=0} + \sum_{n=1}^{M} x^n \otimes_K b_{n-1} = \underbrace{x^0 \otimes_K 0}_{=0} + \sum_{n=1}^{M} x^n \otimes_K b_{n-1} = \sum_{n=1}^{M} x^n \otimes_K b_{n-1} \right)$$

$$= \sum_{n=0}^{M-1} x^n \otimes_K b_{n-1} + \underbrace{x^M}_{\substack{=-\sum_{n=0}^{M-1} x^n a_n \\ \text{(by (18))}}} \otimes_K \underbrace{b_{M-1}}_{=1} = \sum_{n=0}^{M-1} x^n \otimes_K b_{n-1} + \left(-\sum_{n=0}^{M-1} x^n a_n \right) \otimes_K 1$$

$$= \sum_{n=0}^{M-1} x^n \otimes_K b_{n-1} - \sum_{n=0}^{M-1} \underbrace{x^n a_n \otimes_K 1}_{\substack{=x^n \otimes_K a_n 1 \\ \text{(since } a_n \in K\text{)}}} = \sum_{n=0}^{M-1} x^n \otimes_K b_{n-1} - \sum_{n=0}^{M-1} x^n \otimes_K a_n 1$$

$$= \sum_{n=0}^{M-1} x^n \otimes_K \left(b_{n-1} - \underbrace{a_n 1}_{\substack{=a_n = b_{n-1} - b_n x \\ \text{(by (19))}}} \right) = \sum_{n=0}^{M-1} x^n \otimes_K \underbrace{\left(b_{n-1} - \left(b_{n-1} - b_n x \right) \right)}_{=b_n x}$$

$$= \sum_{n=0}^{M-1} x^n \otimes_K b_n x = \underbrace{\left(\sum_{n=0}^{M-1} x^n \otimes_K b_n \right)}_{=f} x = fx.$$

So we have proven that $xf = fx$. Using this formula, it is easy to prove that

$$x^n f = f x^n \qquad \text{for every } n \in \mathbb{N}. \tag{21}$$

[9]

Now, we have

$$af = fa \qquad \text{for every } a \in S. \tag{22}$$

[9] *Proof of (21):* We will prove (21) by induction over n:

Induction base: For $n = 0$, we have $x^n = x^0 = 1$. Thus, for $n = 0$, the equation (21) is equivalent to $1f = f1$, which is trivial (since $1f = f = f1$). Thus, (21) holds for $n = 0$. This completes the induction base.

Induction step: Let $N \in \mathbb{N}$. Assume that (21) holds for $n = N$. We must now prove that (21) holds for $n = N + 1$.

[10]

Now, set $e = \dfrac{1}{Q(x)} f$. (This is allowed since $Q(x) \neq 0$.) Since $e = \dfrac{1}{Q(x)} f$, we have

$$ae = a \cdot \frac{1}{Q(x)} f = \frac{1}{Q(x)} \underbrace{af}_{\substack{=fa \\ \text{(by (22))}}} = \underbrace{\frac{1}{Q(x)} f}_{=e} a = ea$$

for every $a \in S$. Also, since $e = \dfrac{1}{Q(x)} f$, we have

$$\begin{aligned}
\varepsilon_{K,S}(e) &= \varepsilon_{K,S}\left(\frac{1}{Q(x)} f\right) \\
&= \frac{1}{Q(x)} \underbrace{\varepsilon_{K,S}(f)}_{\substack{=Q(x) \\ \text{(by (20))}}} \qquad \text{(since } \varepsilon_{K,S} \text{ is an } (L,L)\text{-bimodule homomorphism)} \\
&= \frac{1}{Q(x)} Q(x) = 1.
\end{aligned}$$

We thus have checked that $\varepsilon_{K,S}(e) = 1$ and ($ae = ea$ for all $a \in S$). Thus, there exists an element $e \in S \otimes_K S$ satisfying $\varepsilon_{K,S}(e) = 1$ and ($ae = ea$ for all $a \in S$). By Definition 22, this means that the K-algebra S is separable. Since $S = K[x]$, this yields that the K-algebra $K[x]$ is separable. Proposition 24 is proven.

Next, a rather famous property of separable algebras:

Theorem 25. Let K be a commutative ring. Let L be a separable K-algebra (not necessarily commutative). Let U be an (L,L)-bimodule. Let $d : L \to U$ be a derivation from L to U satisfying $d(K) = 0$. Then, there exists an $u \in U$ such that

$$(d(a) = au - ua \qquad \text{for every } a \in L).$$

Since (21) holds for $n = N$, we have $x^N f = f x^N$. Now,

$$\underbrace{x^{N+1}}_{=xx^N} f = x \underbrace{x^N f}_{=fx^N} = \underbrace{xf}_{=fx} x^N = f \underbrace{xx^N}_{=x^{N+1}} = f x^{N+1}.$$

Thus, (21) holds for $n = N+1$. This completes the induction step. Thus, the induction proof of (21) is complete.

[10] *Proof of (22):* Let $a \in S$. Then, $a \in S = K[x]$. Hence, there exists a polynomial $V \in K[X]$ such that $a = V(x)$. Consider this V.

Since $V \in K[X]$ is a polynomial over K, we can write V in the form $V = \sum_{n=0}^{N} v_n X^n$ for some $N \in \mathbb{N}$ and some elements v_0, v_1, ..., v_N of K. Consider this N and these elements v_0, v_1, ..., v_N. Then, $a = V(x) = \sum_{n=0}^{N} v_n x^n$ (since $V = \sum_{n=0}^{N} v_n X^n$), so that

$$af = \left(\sum_{n=0}^{N} v_n x^n\right) f = \sum_{n=0}^{N} v_n \underbrace{x^n f}_{\substack{=fx^n \\ \text{(by (21))}}} = \sum_{n=0}^{N} v_n f x^n = f \underbrace{\left(\sum_{n=0}^{N} v_n x^n\right)}_{=a} = fa.$$

This proves (22).

Proof of Theorem 25.[11] Every $x \in K$ and $y \in L$ satisfy

$$\begin{aligned} d(xy) &= \underbrace{d(x)}_{\substack{=0 \\ (\text{since } x \in K \text{ and thus} \\ d(x) \in d(K) = 0)}} \cdot y + x \cdot d(y) \qquad (\text{since } d \text{ is a derivation}) \\ &= \underbrace{0 \cdot y}_{=0} + x \cdot d(y) = x \cdot d(y). \end{aligned}$$

Combined with the fact that d is a homomorphism of additive groups, this yields that d is K-linear in the first variable. Similarly, d is K-linear in the second variable. Thus, d is K-linear in both variables. In other words, d is a (K, K)-bimodule homomorphism. Hence, the map $d \otimes_K \mathrm{id}_L : L \otimes_K L \to U \otimes_K L$ is well-defined.

The K-algebra L is separable. By Definition 22, this means that there exists an element $e \in L \otimes_K L$ satisfying $\varepsilon_{K,L}(e) = 1$ and ($ae = ea$ for all $a \in L$). Consider this e.

Since $e \in L \otimes_K L$ is a tensor, we can write it in the form $e = \sum_{i=1}^{n} t_i \otimes_K s_i$ for some $n \in \mathbb{N}$, some elements t_1, t_2, ..., t_n of L and some elements s_1, s_2, ..., s_n of L. Consider this n, these t_1, t_2, ..., t_n and these s_1, s_2, ..., s_n.

Since $e = \sum_{i=1}^{n} t_i \otimes_K s_i$, we have

$$\begin{aligned} \varepsilon_{K,L}(e) &= \varepsilon_{K,L}\left(\sum_{i=1}^{n} t_i \otimes_K s_i \right) = \sum_{i=1}^{n} \underbrace{\varepsilon_{K,L}(t_i \otimes_K s_i)}_{\substack{= t_i s_i \\ (\text{by the definition of } \varepsilon_{K,L})}} \\ &\qquad (\text{since } \varepsilon_{K,L} \text{ is a } (K,K)\text{-bimodule homomorphism}) \\ &= \sum_{i=1}^{n} t_i s_i. \end{aligned}$$

Thus, $\sum_{i=1}^{n} t_i s_i = \varepsilon_{K,L}(e) = 1$.

Let now $u = -\sum_{i=1}^{n} d(t_i)\, s_i$.

Let $a \in L$. Then, $ae = ea$ (by the condition that $ae = ea$ for all $a \in L$). Since $e = \sum_{i=1}^{n} t_i \otimes_K s_i$, this rewrites as

$$\sum_{i=1}^{n} a t_i \otimes_K s_i = \sum_{i=1}^{n} t_i \otimes_K s_i a.$$

Applying the map $d \otimes_K \mathrm{id}_L$ to this equation, we get

$$\sum_{i=1}^{n} d(a t_i) \otimes_K s_i = \sum_{i=1}^{n} d(t_i) \otimes_K s_i a. \tag{23}$$

[11] I have taken this proof from `http://mathoverflow.net/questions/71869/` but it actually is a well-known argument.

Now, U is an (L, L)-bimodule, thus (in particular) a left L-module. Hence, there exists a well-defined $\mathbb{Z}$-linear map $\rho : L \otimes_K U \to U$ which satisfies

$$(\rho\left(b \otimes_K v\right) = bv \qquad \text{for every } b \in L \text{ and } v \in U).$$

Applying this map ρ to the equation (23), we obtain

$$\sum_{i=1}^{n} d\left(at_i\right) s_i = \sum_{i=1}^{n} d\left(t_i\right) s_i a.$$

Thus,

$$\begin{aligned}
0 &= \sum_{i=1}^{n} d\left(at_i\right) s_i - \sum_{i=1}^{n} d\left(t_i\right) s_i a = \sum_{i=1}^{n} \left(\underbrace{d\left(at_i\right)}_{=d(a)t_i + ad(t_i) \text{ (since } d \text{ is a derivation)}} s_i - d\left(t_i\right) s_i a \right) \\
&= \sum_{i=1}^{n} \left(\left(d\left(a\right) t_i + ad\left(t_i\right)\right) s_i - d\left(t_i\right) s_i a \right) = \sum_{i=1}^{n} \left(d\left(a\right) t_i s_i + ad\left(t_i\right) s_i - d\left(t_i\right) s_i a \right) \\
&= d\left(a\right) \underbrace{\sum_{i=1}^{n} t_i s_i}_{=1} + a \sum_{i=1}^{n} d\left(t_i\right) s_i - \sum_{i=1}^{n} d\left(t_i\right) s_i a = d\left(a\right) + a \sum_{i=1}^{n} d\left(t_i\right) s_i - \sum_{i=1}^{n} d\left(t_i\right) s_i a.
\end{aligned}$$

Thus,

$$d\left(a\right) = \sum_{i=1}^{n} d\left(t_i\right) s_i a - a \sum_{i=1}^{n} d\left(t_i\right) s_i = a \underbrace{\left(- \sum_{i=1}^{n} d\left(t_i\right) s_i \right)}_{=u} - \underbrace{\left(- \sum_{i=1}^{n} d\left(t_i\right) s_i \right)}_{=u} a = au - ua.$$

We have thus proven that $d\left(a\right) = au - ua$ for every $u \in L$. This establishes Theorem 25.

Theorem 25 is way too general for us to use in its full glory; all we need is the following particular case:

Theorem 26. Let K be a commutative ring. Let L be a **commutative** separable K-algebra. Let U be an L-module. Let $d : L \to U$ be a derivation from L to U satisfying $d\left(K\right) = 0$. Then, $d = 0$.

Proof of Theorem 26. As we know, U, being an L-module, canonically becomes an (L, L)-bimodule, and the map $d : L \to U$, being a derivation from L to the L-module U, canonically becomes a derivation from L to the (L, L)-bimodule U. Thus, we can apply Theorem 25, and conclude that there exists an $u \in U$ such that

$$(d\left(a\right) = au - ua \qquad \text{for every } a \in L).$$

Consider this u. Then, every $a \in L$ satisfies $d\left(a\right) = au - ua = 0$ (because our (L, L)-bimodule was canonically constructed from the L-module U, and thus every $a \in L$ and $v \in U$ satisfy $av = va$, so that (in particular) $au = ua$ and thus $au - ua = 0$). In other words, $d = 0$. Theorem 26 is proven.

This gives us an alternative proof of Proposition 12:

Second proof of Proposition 12. Let $x \in L$. By Proposition 24, we see that $K[x]$ is a separable K-algebra.

On the other hand, since D is a derivation from L to U, it is clear that $D\mid_{K[x]}$ is a derivation from $K[x]$ to the $K[x]$-module U. Moreover, $\left(D\mid_{K[x]}\right)(K) = D(K) = 0$. Thus, Theorem 26 (applied to $K[x]$ and $D\mid_{K[x]}$ instead of L and d) yields $D\mid_{K[x]} = 0$. Since $x \in K[x]$, we have $D(x) = \underbrace{\left(D\mid_{K[x]}\right)}_{=0}(x) = 0(x) = 0$.

Now forget that we fixed x. We thus have shown that every $x \in L$ satisfies $D(x) = 0$. In other words, $D = 0$. Proposition 12 is proven.

Now, something really trivial before the main theorem:

Lemma 27. Let $n \in \mathbb{N}$. Let c_1, c_2, ..., c_n be n elements of a commutative ring. Then,

$$\left(\sum_{i=1}^{n} c_i\right)^2 = \sum_{i=1}^{n} c_i^2 + 2 \sum_{\substack{(i,j)\in\{1,2,\ldots,n\}^2; \\ i<j}} c_i c_j.$$

The main result of this third solution is now the following:

Theorem 28. Let K be a commutative ring. Let L be a **commutative** separable K-algebra. Let V, W and U be L-modules. Then, V, W and U canonically become K-modules

Let $h : V \times W \to U$ be a K-bilinear map (not necessarily an L-bilinear map). Assume that

$$h(xa, xb) = x^2 h(a, b) \qquad \text{for every } x \in L,\ a \in V \text{ and } b \in W. \tag{24}$$

Then, h is L-bilinear.

Proof of Theorem 28. The K-algebra L is separable. By Definition 22, this means that there exists an element $e \in L \otimes_K L$ satisfying $\varepsilon_{K,L}(e) = 1$ and ($ae = ea$ for all $a \in L$). Consider this e.

Since $e \in L \otimes_K L$ is a tensor, we can write it in the form $e = \sum_{i=1}^{n} t_i \otimes_K s_i$ for some $n \in \mathbb{N}$, some elements t_1, t_2, ..., t_n of L and some elements s_1, s_2, ..., s_n of L. Consider this n, these t_1, t_2, ..., t_n and these s_1, s_2, ..., s_n.

Just as in the proof of Theorem 25, we can show that $\sum_{i=1}^{n} t_i s_i = 1$. Hence, $\left(\sum_{i=1}^{n} t_i s_i\right)^2 = 1^2 = 1$. Since

$$\begin{aligned}\left(\sum_{i=1}^{n} t_i s_i\right)^2 &= \sum_{i=1}^{n} \underbrace{(t_i s_i)^2}_{=t_i^2 s_i^2} + 2 \sum_{\substack{(i,j)\in\{1,2,\ldots,n\}^2; \\ i<j}} t_i s_i t_j s_j \qquad \text{(by Lemma 27, applied to } c_i = t_i s_i\text{)} \\ &= \sum_{i=1}^{n} t_i^2 s_i^2 + 2 \sum_{\substack{(i,j)\in\{1,2,\ldots,n\}^2; \\ i<j}} t_i s_i t_j s_j,\end{aligned}$$

this rewrites as

$$\sum_{i=1}^{n} t_i^2 s_i^2 + 2 \sum_{\substack{(i,j)\in\{1,2,\ldots,n\}^2; \\ i<j}} t_i s_i t_j s_j = 1. \tag{25}$$

By the conditions of Theorem 28, we have

$$h(xa, xb) = x^2 h(a, b) \qquad \text{for every } x \in L,\ a \in V \text{ and } b \in W.$$

Hence, Lemma 13 yields that

$$h(xa, yb) + h(ya, xb) = 2xyh(a, b) \qquad \text{for every } x \in L,\ y \in L,\ a \in V \text{ and } b \in W. \tag{26}$$

Also,

$$\text{every derivation } D : L \to U \text{ satisfying } D(K) = 0 \text{ satisfies } D = 0 \tag{27}$$

(by Theorem 26, applied to $d = D$). Since (26) and (27) hold, we can apply Lemma 14.

Now, we can easily see that

$$\text{any } x \in L,\ a \in V \text{ and } b \in W \text{ satisfy } h\left(x^2 a, b\right) = x^2 h(a, b). \tag{28}$$

[12]

Fix any $a \in V$ and $b \in W$. Define a map $f : L \to L$ by

$$\left(f(x) = h(xa, b) \qquad \text{for all } x \in L\right).$$

Then, f is K-linear (since h is K-bilinear). As a consequence, $f \otimes_K \mathrm{id}_L$ is a well-defined map $L \otimes_K L \to L \otimes_K L$.

The definition of f yields $f(1) = h\left(\underbrace{1a}_{=a}, b\right) = h(a, b)$.

It is easy (using Lemma 14 **(c)**) to show that any $x \in L$ satisfies

$$2f(x) = 2xf(1). \tag{29}$$

[13] Moreover, it is easy (using (28)) to show that any $y \in L$ and $z \in L$ satisfy

$$f\left(y^2 z\right) = y^2 f(z). \tag{30}$$

[12] *Proof of (16):* Let $x \in L$, $a \in V$ and $b \in W$. Applying Lemma 14 **(a)** to 1, xa and x instead of x, a and y, we obtain $h(1 \cdot xa, xb) = h(xxa, 1 \cdot b)$. Thus,

$$h\left(\underbrace{x^2}_{=xx} a, \underbrace{b}_{=1\cdot b}\right) = h(xxa, 1 \cdot b) = h\left(\underbrace{1 \cdot xa}_{=xa}, xb\right) = h(xa, xb) = x^2 h(a, b)$$

(by (24)). This proves (28).

[13] *Proof of (29):* Lemma 14 **(c)** yields $2h(xa, b) = 2xh(a, b)$ for every $x \in L$. Thus, every $x \in L$ satisfies $2\underbrace{f(x)}_{=h(xa,b)} = 2h(xa, b) = 2x\underbrace{h(a, b)}_{=f(1)} = 2xf(1)$. This proves (29).

[14]

Let $x \in L$. Then, $xe = ex$ (since $ae = ea$ for all $a \in L$).

Since L is a commutative K-algebra, the tensor product $L \otimes_K L$ also is a commutative K-algebra. Thus, $(1 \otimes x) \cdot e = e \cdot (1 \otimes x)$.

By the definition of the (L, L)-bimodule structure on $L \otimes_K L$, we have $xe^2 = (x \otimes 1) \cdot e^2$, $e^2 x = e^2 \cdot (1 \otimes x)$, $xe = (x \otimes 1) \cdot e$ and $ex = e \cdot (1 \otimes x)$. Thus,

$$xe^2 = (x \otimes 1) \cdot e^2 = \underbrace{(x \otimes 1) \cdot e}_{=xe=ex=e\cdot(1\otimes x)} \cdot e = e \cdot \underbrace{(1 \otimes x) \cdot e}_{=e\cdot(1\otimes x)} = e \cdot e \cdot (1 \otimes x) = e^2 \cdot (1 \otimes x) = e^2 x.$$

But since $e = \sum_{i=1}^{n} t_i \otimes_K s_i$, we have

$$e^2 = \left(\sum_{i=1}^{n} t_i \otimes_K s_i \right)^2 = \sum_{i=1}^{n} \underbrace{\left(t_i \otimes_K s_i\right)^2}_{=t_i^2 \otimes_K s_i^2} + 2 \sum_{\substack{(i,j) \in \{1,2,\ldots,n\}^2; \\ i<j}} \underbrace{\left(t_i \otimes_K s_i\right)\left(t_j \otimes_K s_j\right)}_{=t_i t_j \otimes_K s_i s_j}$$

(by Lemma 27, applied to $c_i = t_i \otimes_K s_i$)

$$= \sum_{i=1}^{n} t_i^2 \otimes_K s_i^2 + 2 \sum_{\substack{(i,j) \in \{1,2,\ldots,n\}^2; \\ i<j}} t_i t_j \otimes_K s_i s_j. \tag{31}$$

From (31), we obtain

$$xe^2 = x \left(\sum_{i=1}^{n} t_i^2 \otimes_K s_i^2 + 2 \sum_{\substack{(i,j) \in \{1,2,\ldots,n\}^2; \\ i<j}} t_i t_j \otimes_K s_i s_j \right) = \sum_{i=1}^{n} x t_i^2 \otimes_K s_i^2 + 2 \sum_{\substack{(i,j) \in \{1,2,\ldots,n\}^2; \\ i<j}} x t_i t_j \otimes_K s_i s_j.$$

[14] *Proof of (30):* Let $y \in L$ and $z \in L$. By the definition of f, we have $f\left(y^2 z\right) = h\left(y^2 za, b\right)$ and $f(z) = h(za, b)$. Applying (28) to y and za instead of x and a, we obtain $h\left(y^2 za, b\right) = y^2 h(za, b)$. Thus, $f\left(y^2 z\right) = h\left(y^2 za, b\right) = y^2 \underbrace{h(za, b)}_{=f(z)} = y^2 f(z)$, so that (30) is proven.

Hence,

$$
\begin{aligned}
&\left(f \otimes_K \mathrm{id}_L\right)\left(xe^2\right) \\
&= \left(f \otimes_K \mathrm{id}_L\right)\left(\sum_{i=1}^{n} xt_i^2 \otimes_K s_i^2 + 2 \sum_{\substack{(i,j)\in\{1,2,\ldots,n\}^2; \\ i<j}} xt_it_j \otimes_K s_is_j\right) \\
&= \sum_{i=1}^{n} f\left(\underbrace{xt_i^2}_{=t_i^2x}\right) \otimes_K s_i^2 + 2 \sum_{\substack{(i,j)\in\{1,2,\ldots,n\}^2; \\ i<j}} f\left(xt_it_j\right) \otimes_K s_is_j \\
&= \sum_{i=1}^{n} \underbrace{f\left(t_i^2x\right)}_{\substack{=t_i^2f(x) \\ \text{(by (30), applied to} \\ y=t_i \text{ and } z=x\text{)}}} \otimes_K s_i^2 + 2 \sum_{\substack{(i,j)\in\{1,2,\ldots,n\}^2; \\ i<j}} f\left(xt_it_j\right) \otimes_K s_is_j \\
&= \sum_{i=1}^{n} t_i^2 f\left(x\right) \otimes_K s_i^2 + \sum_{\substack{(i,j)\in\{1,2,\ldots,n\}^2; \\ i<j}} \underbrace{2f\left(xt_it_j\right)}_{\substack{=2xt_it_jf(1) \\ \text{(by (29), applied to } xt_it_j \\ \text{instead of } x\text{)}}} \otimes_K s_is_j \\
&= \sum_{i=1}^{n} t_i^2 f\left(x\right) \otimes_K s_i^2 + \sum_{\substack{(i,j)\in\{1,2,\ldots,n\}^2; \\ i<j}} 2xt_it_jf\left(1\right) \otimes_K s_is_j,
\end{aligned}
$$

so that

$$
\begin{aligned}
&\varepsilon_{K,L}\left(\left(f\otimes_K \mathrm{id}_L\right)\left(xe^2\right)\right)\\
&=\varepsilon_{K,L}\left(\sum_{i=1}^{n} t_i^2 f\left(x\right)\otimes_K s_i^2+\sum_{\substack{(i,j)\in\{1,2,\ldots,n\}^2;\\ i<j}} 2xt_it_jf\left(1\right)\otimes_K s_is_j\right)\\
&=\underbrace{\sum_{i=1}^{n} t_i^2 f\left(x\right)s_i^2}_{=f(x)\sum\limits_{i=1}^{n} t_i^2s_i^2}+\underbrace{\sum_{\substack{(i,j)\in\{1,2,\ldots,n\}^2;\\ i<j}} 2xt_it_jf\left(1\right)s_is_j}_{=2f(1)x\sum\limits_{\substack{(i,j)\in\{1,2,\ldots,n\}^2;\\ i<j}} t_is_it_js_j}\\
&\qquad\left(\text{by the definition of }\varepsilon_{K,L}\right)\\
&=f\left(x\right)\sum_{i=1}^{n} t_i^2s_i^2+\underbrace{2xf\left(1\right)}_{\substack{=2f(x)\\ \text{(by (29))}}}\sum_{\substack{(i,j)\in\{1,2,\ldots,n\}^2;\\ i<j}} t_is_it_js_j=f\left(x\right)\sum_{i=1}^{n} t_i^2s_i^2+2f\left(x\right)\sum_{\substack{(i,j)\in\{1,2,\ldots,n\}^2;\\ i<j}} t_is_it_js_j\\
&=f\left(x\right)\cdot\underbrace{\left(\sum_{i=1}^{n} t_i^2s_i^2+2\sum_{\substack{(i,j)\in\{1,2,\ldots,n\}^2;\\ i<j}} t_is_it_js_j\right)}_{\substack{=1\\ \text{(by (25))}}}=f\left(x\right). \qquad (32)
\end{aligned}
$$

On the other hand, from (31), we get

$$
e^2x=\left(\sum_{i=1}^{n} t_i^2\otimes_K s_i^2+2\sum_{\substack{(i,j)\in\{1,2,\ldots,n\}^2;\\ i<j}} t_it_j\otimes_K s_is_j\right)x=\sum_{i=1}^{n} t_i^2\otimes_K s_i^2x+2\sum_{\substack{(i,j)\in\{1,2,\ldots,n\}^2;\\ i<j}} t_it_j\otimes_K s_is_jx,
$$

so that

$$
\begin{aligned}
(f \otimes_K \mathrm{id}_L)\left(e^2 x\right) &= (f \otimes_K \mathrm{id}_L)\left(\sum_{i=1}^{n} t_i^2 \otimes_K s_i^2 x + 2 \sum_{\substack{(i,j)\in\{1,2,\ldots,n\}^2; \\ i<j}} t_i t_j \otimes_K s_i s_j x\right) \\
&= \sum_{i=1}^{n} f\left(\underbrace{t_i^2}_{=t_i^2\cdot 1}\right) \otimes_K s_i^2 x + 2 \sum_{\substack{(i,j)\in\{1,2,\ldots,n\}^2; \\ i<j}} f\left(t_i t_j\right) \otimes_K s_i s_j x \\
&= \sum_{i=1}^{n} \underbrace{f\left(t_i^2 \cdot 1\right)}_{\substack{=t_i^2 f(1) \\ \text{(by (30), applied to } y=t_i \\ \text{and } z=1)}} \otimes_K s_i^2 x + 2 \sum_{\substack{(i,j)\in\{1,2,\ldots,n\}^2; \\ i<j}} f\left(t_i t_j\right) \otimes_K s_i s_j x \\
&= \sum_{i=1}^{n} t_i^2 f(1) \otimes_K s_i^2 x + 2 \sum_{\substack{(i,j)\in\{1,2,\ldots,n\}^2; \\ i<j}} f\left(t_i t_j\right) \otimes_K s_i s_j x \\
&= \sum_{i=1}^{n} t_i^2 f(1) \otimes_K s_i^2 x + \sum_{\substack{(i,j)\in\{1,2,\ldots,n\}^2; \\ i<j}} \underbrace{2 f\left(t_i t_j\right)}_{\substack{=2 t_i t_j f(1) \\ \text{(by (29), applied to } t_i t_j \\ \text{instead of } x)}} \otimes_K s_i s_j x \\
&= \sum_{i=1}^{n} t_i^2 f(1) \otimes_K s_i^2 x + \sum_{\substack{(i,j)\in\{1,2,\ldots,n\}^2; \\ i<j}} 2 t_i t_j f(1) \otimes_K s_i s_j x,
\end{aligned}
$$

so that

$$\begin{aligned}
&\varepsilon_{K,L}\left(\left(f\otimes_K \mathrm{id}_L\right)\left(e^2x\right)\right)\\
&=\varepsilon_{K,L}\left(\sum_{i=1}^{n} t_i^2 f\left(1\right)\otimes_K s_i^2 x+\sum_{\substack{(i,j)\in\{1,2,\ldots,n\}^2;\\ i<j}} 2t_it_j f\left(1\right)\otimes_K s_is_jx\right)\\
&=\underbrace{\sum_{i=1}^{n} t_i^2 f\left(1\right) s_i^2 x}_{=xf(1)\sum\limits_{i=1}^{n} t_i^2s_i^2}+\underbrace{\sum_{\substack{(i,j)\in\{1,2,\ldots,n\}^2;\\ i<j}} 2t_it_j f\left(1\right) s_is_jx}_{=2xf(1)\sum\limits_{\substack{(i,j)\in\{1,2,\ldots,n\}^2;\\ i<j}} t_is_it_js_j}\\
&\qquad\left(\text{by the definition of }\varepsilon_{K,L}\right)\\
&=xf\left(1\right)\sum_{i=1}^{n} t_i^2s_i^2+2xf\left(1\right)\sum_{\substack{(i,j)\in\{1,2,\ldots,n\}^2;\\ i<j}} t_is_it_js_j\\
&=xf\left(1\right)\cdot\underbrace{\left(\sum_{i=1}^{n} t_i^2s_i^2+2\sum_{\substack{(i,j)\in\{1,2,\ldots,n\}^2;\\ i<j}} t_is_it_js_j\right)}_{\substack{=1\\ \text{(by (25))}}}=xf\left(1\right). \qquad (33)
\end{aligned}$$

Now, by (32), we have

$$f\left(x\right)=\varepsilon_{K,L}\left(\left(f\otimes_K \mathrm{id}_L\right)\underbrace{\left(xe^2\right)}_{=e^2x}\right)=\varepsilon_{K,L}\left(\left(f\otimes_K \mathrm{id}_L\right)\left(e^2x\right)\right)=xf\left(1\right)$$

(by (33)). Since $f\left(x\right)=h\left(xa,b\right)$ and $f\left(1\right)=h\left(a,b\right)$, this rewrites as $h\left(xa,b\right)=xh\left(a,b\right)$.

Forget that we fixed $x\in L$, $a\in V$ and $b\in W$ now. We thus have proven that $h\left(xa,b\right)=xh\left(a,b\right)$ for every $x\in L$, $a\in V$ and $b\in W$. Combined with the fact that $h\left(a+a',b\right)=h\left(a,b\right)+h\left(a',b\right)$ for all $a\in V$, $a'\in V$ and $b\in W$ (since h is K-bilinear), this yields the map h is L-linear in its first variable. Similarly, the map h is L-linear in its second variable. Hence, the map h is L-linear in each of its two variables, i. e., an L-bilinear map. This proves Theorem 28.

Theorem 28 does not **immediately** apply to our problem since separable algebraic field extensions of a field K need not be separable as K-algebras (unless they are finite). But what applies is the following slight generalization:

> **Corollary 29.** Let K be a commutative ring. Let L be a **commutative** K-algebra such that for every $y\in L$, there exists a separable K-subalgebra of L containing y. Let V, W and U be L-modules. Then, V, W and U canonically become K-modules

Let $h : V \times W \to U$ be a K-bilinear map (not necessarily an L-bilinear map). Assume that

$$h(xa, xb) = x^2 h(a, b) \qquad \text{for every } x \in L,\ a \in V \text{ and } b \in W. \tag{34}$$

Then, h is L-bilinear.

Proof of Corollary 29. Let $y \in L$. According to the conditions of Corollary 29, there exists a separable K-subalgebra of L containing y. Let S be such a K-subalgebra.

Clearly,

$$h(xa, xb) = x^2 h(a, b) \qquad \text{for every } x \in S,\ a \in V \text{ and } b \in W$$

(by (34), since $S \subseteq L$). Hence, Theorem 28 (applied to S instead of L) yields that h is S-bilinear. Thus, $h(ya, b) = yh(a, b)$ (because $y \in S$).

Now, forget that we fixed y. We have thus proven that $h(ya, b) = yh(a, b)$ for every $y \in L$, $a \in V$ and $b \in W$. Combined with the fact that $h(a + a', b) = h(a, b) + h(a', b)$ for all $a \in V$, $a' \in V$ and $b \in W$ (since h is K-bilinear), this yields the map h is L-linear in its first variable. Similarly, the map h is L-linear in its second variable. Hence, the map h is L-linear in each of its two variables, i. e., an L-bilinear map. This proves Corollary 29.

We can now finally start solving the problem. Let K, L, V, W, U and h be as defined in the problem.

For every $y \in L$, the K-subalgebra $K[y]$ of L is separable (by Proposition 24, applied to $x = y$) and contains y. Thus, for every $y \in L$, there exists a separable K-subalgebra of L containing y.

Also,

$$h(xa, xb) = x^2 h(a, b) \qquad \text{for every } x \in L,\ a \in V \text{ and } b \in W$$

(by the condition of the problem). Thus, Corollary 29 yields that h is L-bilinear. The problem is solved.

Remarks

Remark 1

1) In the third solution, we found a generalization of the problem (Corollary 29). But the first solution also shows that the problem can be generalized. Namely, the problem will still be valid if we replace "Let L/K be a separable algebraic extension of fields" by "Let K and L be commutative rings with 1 such that L is a K-algebra" and add the assumption that "For every $x \in L$ and every $u \in U$, there exists a polynomial $P \in K[X]$ such that $P(x) = 0$ and such that (if $P'(x)\, u = 0$ then $u = 0$)". (This assumption is what replaces the assumption that L/K be separable. It is used in our proof of Lemma 2.)

This generalization of the problem seems to be neither more nor less strong than Corollary 29 (the generalization obtained in the third solution).

Remark 2

2) In the second solution of our problem, we used Proposition 12. But we could, conversely, derive Proposition 12 from our problem:

Third proof of Proposition 12. Define a map $h : L \times L \to U$ by

$$\left(h\left(a,b\right) = aD\left(b\right) - bD\left(a\right) \qquad \text{for all } a \in L \text{ and } b \in L\right).$$

Then, any $a \in L$, $b \in L$ and $b' \in L$ satisfy

$$\begin{aligned} h\left(a, b+b'\right) &= a \underbrace{D\left(b+b'\right)}_{\substack{=D(b)+D(b') \\ \text{(since } D \text{ is a derivation and thus} \\ \text{a homomorphism of additive groups)}}} - \underbrace{\left(b+b'\right) D\left(a\right)}_{=bD(a)+b'D(a)} \qquad \text{(by the definition of } h\left(a, b+b'\right)\text{)} \\ &= a\left(D\left(b\right) + D\left(b'\right)\right) - \left(bD\left(a\right) + b'D\left(a\right)\right) = aD\left(b\right) + aD\left(b'\right) - bD\left(a\right) - b'D\left(a\right) \\ &= \underbrace{\left(aD\left(b\right) - bD\left(a\right)\right)}_{\substack{=h(a,b) \\ \text{(since } h(a,b) \text{ was defined as } aD(b)-bD(a)\text{)}}} + \underbrace{\left(aD\left(b'\right) - b'D\left(a\right)\right)}_{\substack{=h(a,b') \\ \text{(since } h(a,b') \text{ was defined as } aD(b')-b'D(a)\text{)}}} \\ &= h\left(a,b\right) + h\left(a,b'\right). \end{aligned} \tag{35}$$

Also, any $a \in L$, $b \in L$ and $x \in K$ satisfy

$$\begin{aligned} h\left(a, xb\right) &= a \underbrace{D\left(xb\right)}_{\substack{=D(x)\cdot b+x\cdot D(b) \\ \text{(since } D \text{ is a derivation)}}} - xbD\left(a\right) \qquad \text{(by the definition of } h\left(a, xb\right)\text{)} \\ &= a\left(\underbrace{D\left(x\right)}_{\substack{=0 \\ \text{(since } x\in K \text{ and thus } D(x)\in D(K)=0\text{)}}} \cdot b + x \cdot D\left(b\right)\right) - xbD\left(a\right) \\ &= a\left(\underbrace{0 \cdot b}_{=0} + x \cdot D\left(b\right)\right) - xbD\left(a\right) = ax \cdot D\left(b\right) - xbD\left(a\right) \\ &= x\underbrace{\left(aD\left(b\right) - bD\left(a\right)\right)}_{=h(a,b)} = xh\left(a,b\right). \end{aligned} \tag{36}$$

The map h is K-linear in its second variable (since any $a \in L$, $b \in L$ and $b' \in L$ satisfy (35), and since any $a \in L$, $b \in L$ and $x \in K$ satisfy (36)), and K-linear in its first variable (for similar reasons). Hence, the map h is K-bilinear.

For every $x \in L$, $a \in L$ and $b \in L$, we have

$$\begin{aligned} h\left(xa, xb\right) &= xa \underbrace{D\left(xb\right)}_{\substack{=D(x)\cdot b+x\cdot D(b) \\ \text{(since } D \text{ is a derivation)}}} - xb \underbrace{D\left(xa\right)}_{\substack{=D(x)\cdot a+x\cdot D(a) \\ \text{(since } D \text{ is a derivation)}}} \qquad \text{(by the definition of } h\left(xa, xb\right)\text{)} \\ &= xa\left(D\left(x\right)\cdot b + x\cdot D\left(b\right)\right) - xb\left(D\left(x\right)\cdot a + x\cdot D\left(a\right)\right) \\ &= \underbrace{xaD\left(x\right)\cdot b}_{=xD(x)\cdot ab} + \underbrace{xax\cdot D\left(b\right)}_{=x^2aD(b)} - \underbrace{xbD\left(x\right)\cdot a}_{=xD(x)\cdot ab} - \underbrace{xbx\cdot D\left(a\right)}_{=x^2bD(a)} \\ &= xD\left(x\right)\cdot ab + x^2aD\left(b\right) - xD\left(x\right)\cdot ab - x^2bD\left(a\right) \\ &= x^2aD\left(b\right) - x^2bD\left(a\right) = x^2\underbrace{\left(aD\left(b\right) - bD\left(a\right)\right)}_{=h(a,b)} = x^2h\left(a,b\right). \end{aligned}$$

Hence, our problem (applied to $V = L$ and $W = L$) yields that h is L-bilinear. Thus, every $x \in L$ satisfies $h(x \cdot 1, 1) = x \cdot h(1, 1)$. But since

$$h\left(\underbrace{x \cdot 1}_{=x}, 1\right) = h(x, 1) = x \underbrace{D(1)}_{\substack{=0 \\ \text{(since } 1 \in K \text{ and thus } D(1) \in D(K) = 0\text{)}}} - \underbrace{1D(x)}_{=D(x)} \qquad \text{(by the definition of } h(x, 1)\text{)}$$
$$= -D(x)$$

and

$$h(1, 1) = 1D(1) - 1D(1) \qquad \text{(by the definition of } h(1, 1)\text{)}$$
$$= 0,$$

this rewrites as $-D(x) = x \cdot 0$. Thus, every $x \in L$ satisfies $D(x) = -x \cdot 0 = 0$. In other words, $D = 0$. Proposition 12 is thus proven.

Remark 3

3) The condition that L/K be separable cannot be removed from the problem (without a proper replacement). In fact, if we let p be any prime, and consider the algebraic field extension $K = \mathbb{F}_p(T^p) \subseteq \mathbb{F}_p(T) = L$ (the classical example of a purely inseparable field extension) and let $V = L$, $U = L$ and $W = L$, then we can define an $\mathbb{F}_p$-bilinear map

$$h : V \times W \to U, \qquad \left(T^a, T^b\right) \mapsto (a - b) T^{a+b};$$

this map is K-bilinear but not L-bilinear, although it satisfies (1). [15]

Note that this counterexample is not as weird as it looks like; in fact, the form $h : V \times W \to U$ constructed in this counterexample can also be characterized as the map $L \times L \to L$, $(u, v) \mapsto -u\dfrac{d}{dT}v + v\dfrac{d}{dT}u$, so that it (up to sign) is an example of the same construction that we made in the second proof of Proposition 12.

Using this construction, we can show a partial converse of the problem: If L/K is a *finitely generated but nonseparable* field extension, then there exists a K-bilinear map $h : L \times L \to L$ which satisfies (1) (for $V = L$ and $W = L$) without being L-bilinear. I don't know what can be said about non-finitely generated field extensions.

[15] Note that this map h is the Lie bracket of the infinite-dimensional Witt algebra over $\mathbb{F}_p$.