The trace Cayley-Hamilton theorem

Darij Grinberg

October 17, 2022

Contents

1. Introduction

Let $\mathbb{K}$ be a commutative ring. The famous Cayley-Hamilton theorem says that if $\chi_A = \det(tI_n - A) \in \mathbb{K}[t]$ is the characteristic polynomial of an $n \times n$-matrix $A \in \mathbb{K}^{n\times n}$, then $\chi_A(A) = 0$. Speaking more explicitly, it means that if we write this polynomial χ_A in the form $\chi_A = \sum_{i=0}^{n} c_{n-i} t^i$ (with $c_{n-i} \in \mathbb{K}$), then $\sum_{i=0}^{n} c_{n-i} A^i = 0$. Various proofs of this theorem are well-known (we will present one in this note, but it could not be any farther from being new). A less standard fact, which I call the *trace Cayley-Hamilton theorem*, states that

$$kc_k + \sum_{i=1}^{k} \operatorname{Tr}\left(A^i\right) c_{k-i} = 0 \qquad \text{for every } k \in \mathbb{N} \tag{1}$$

(where $\sum_{i=0}^{n} c_{n-i} t^i$ is χ_A as before, and where we set $c_{n-i} = 0$ for every $i < 0$). In the case of $k \geq n$, this can easily be obtained from the Cayley-Hamilton theorem $\sum_{i=0}^{n} c_{n-i} A^i = 0$ by multiplying by A^{k-n} and taking traces[1]; no such simple proof exists in the general case, however. The result itself is not new (the $k \leq n$ case, for example, is [LomQui16, Chapter III, Exercise 14]), and is well-known e.g. to algebraic combinatorialists; however, it is hard to find an expository treatment.

When the ground ring $\mathbb{K}$ is a field, it is possible to prove the trace Cayley-Hamilton theorem by expressing both $\operatorname{Tr}\left(A^i\right)$ and the c_j through the eigenvalues of A (indeed, $\operatorname{Tr}\left(A^i\right)$ is the sum of the i-th powers of these eigenvalues, whereas c_j is $(-1)^j$ times their j-th elementary symmetric function); the identity (1) then boils down to the Newton identities for said eigenvalues. However, of course, the use of eigenvalues in this proof requires $\mathbb{K}$ to be a field. There are ways to adapt this proof to the case when $\mathbb{K}$ is a commutative ring. One is to apply the "method of universal identities" (see, e.g., [LomQui16, Chapter III, Exercise 14]; the method is also explained in [Conrad09]) to reduce the general case to the case when $\mathbb{K}$ is a field[2]. Another is to build up the theory of eigenvalues for square matrices over an arbitrary commutative ring $\mathbb{K}$; this is not as simple as for fields, but doable (see [Laksov13]).

In this note, I shall give a proof of both the Cayley-Hamilton and the trace Cayley-Hamilton theorems via a trick whose use in proving the former is well-known (see, e.g., [Heffer14, Chapter Five, Section IV, Lemma 1.9]). The trick is to observe that

[1] The details are left to the interested reader. The kc_k term on the left hand side appears off, but it actually is harmless: In the $k = n$ case, it can be rewritten as $\operatorname{Tr}\left(A^0\right) c_n$ and incorporated into the sum, whereas in the $k > n$ case, it simply vanishes.

[2] This relies on the observation that (1) (for a given k) is a polynomial identity in the entries of A.

the adjugate matrix $\operatorname{adj}(tI_n - A)$ can be written as $D_0 t^0 + D_1 t^1 + \cdots + D_{n-1} t^{n-1}$ for some n matrices $D_0, D_1, \ldots, D_{n-1} \in \mathbb{K}^{n \times n}$; then, a telescoping sum establishes the Cayley-Hamilton theorem. The same trick can be used for the trace Cayley-Hamilton theorem, although it requires more work; in particular, an intermediate step is necessary, establishing that the derivative of the characteristic polynomial $\chi_A = \det(tI_n - A)$ is $\operatorname{Tr}(\operatorname{adj}(tI_n - A))$. I hope that this writeup will have two uses: making the trace Cayley-Hamilton theorem more accessible, and demonstrating that the trick just mentioned can serve more than one purpose. Next, I shall show an application of the trace Cayley-Hamilton theorem, answering a question from [m.se1798703]. Finally, I shall discuss several other properties of the adjugate matrix as well as further applications of polynomial matrices in proving determinant identities.

2. Notations and theorems

2.1. Notations

Before we state the theorems that we will be occupying ourselves with, let us agree on the notations.

Definition 2.1. Throughout this note, the word "ring" will mean "associative ring with unity". We will always let $\mathbb{K}$ denote a commutative ring with unity. The word "matrix" shall always mean "matrix over $\mathbb{K}$", unless explicitly stated otherwise.

As usual, we let $\mathbb{K}[t]$ denote the polynomial ring in the indeterminate t over $\mathbb{K}$.

If $f \in \mathbb{K}[t]$ is a polynomial and n is an integer, then $[t^n] f$ will denote the coefficient of t^n in f. (If n is negative or greater than the degree of f, then this coefficient is understood to be 0.)

Let $\mathbb{N}$ denote the set $\{0, 1, 2, \ldots\}$.

If $n \in \mathbb{N}$ and $m \in \mathbb{N}$, and if we are given an element $a_{i,j} \in \mathbb{K}$ for every $(i,j) \in \{1, 2, \ldots, n\} \times \{1, 2, \ldots, m\}$, then we use the notation $(a_{i,j})_{1 \le i \le n,\ 1 \le j \le m}$ for the $n \times m$-matrix whose (i,j)-th entry is $a_{i,j}$ for all $(i,j) \in \{1, 2, \ldots, n\} \times \{1, 2, \ldots, m\}$.

For every $n \in \mathbb{N}$, we denote the $n \times n$ identity matrix by I_n.

For every $n \in \mathbb{N}$ and $m \in \mathbb{N}$, we denote the $n \times m$ zero matrix by $0_{n \times m}$.

If A is any $n \times n$-matrix, then we let $\det A$ denote the determinant of A, and we let $\operatorname{Tr} A$ denote the trace of A. (Recall that the trace of A is defined to be the sum of the diagonal entries of A.)

We consider $\mathbb{K}$ as a subring of $\mathbb{K}[t]$. Thus, for every $n \in \mathbb{N}$, every $n \times n$-matrix in $\mathbb{K}^{n \times n}$ can be considered as a matrix in $(\mathbb{K}[t])^{n \times n}$.

2.2. The main claims

We shall now state the results that we will prove further below. We begin with a basic fact:

Proposition 2.2. Let $n \in \mathbb{N}$. Let $A \in \mathbb{K}^{n\times n}$ and $B \in \mathbb{K}^{n\times n}$ be two $n \times n$-matrices. Consider the matrix $tA + B \in (\mathbb{K}[t])^{n\times n}$.

(a) Then, $\det(tA + B) \in \mathbb{K}[t]$ is a polynomial of degree $\leq n$ in t.

(b) We have $[t^0](\det(tA + B)) = \det B$.

(c) We have $[t^n](\det(tA + B)) = \det A$.

Definition 2.3. Let $n \in \mathbb{N}$. Let $A \in \mathbb{K}^{n\times n}$ be an $n \times n$-matrix. Then, we consider A as a matrix in $(\mathbb{K}[t])^{n\times n}$ as well (as explained above); thus, a matrix $tI_n - A \in (\mathbb{K}[t])^{n\times n}$ is defined. We let χ_A denote the polynomial $\det(tI_n - A) \in \mathbb{K}[t]$; we call χ_A the *characteristic polynomial* of A.

We notice that the notion of the characteristic polynomial is not standardized across the literature. Our definition of χ_A is identical with the definition in [Knapp2016, §V.3] (except that we use t instead of X as the indeterminate), but the definition in [Heffer14, Chapter Five, Section II, Definition 3.9] is different (it defines χ_A to be $\det(A - tI_n)$ instead). The two definitions differ merely in a sign (namely, one version of the characteristic polynomial is $(-1)^n$ times the other), whence any statement about one of them can easily be translated into a statement about the other; nevertheless this discrepancy creates some occasions for confusion. I shall, of course, use Definition 2.3 throughout this note.

Corollary 2.4. Let $n \in \mathbb{N}$. Let $A \in \mathbb{K}^{n\times n}$.

(a) Then, $\chi_A \in \mathbb{K}[t]$ is a polynomial of degree $\leq n$ in t.

(b) We have $[t^0]\chi_A = (-1)^n \det A$.

(c) We have $[t^n]\chi_A = 1$.

Of course, combining parts **(a)** and **(c)** of Corollary 2.4 shows that, for every $n \in \mathbb{N}$ and $A \in \mathbb{K}^{n\times n}$, the characteristic polynomial χ_A is a monic polynomial of degree n.

Let me now state the main two theorems in this note:

Theorem 2.5 (Cayley-Hamilton theorem)**.** Let $n \in \mathbb{N}$. Let $A \in \mathbb{K}^{n\times n}$. Then, $\chi_A(A) = 0_{n\times n}$. (Here, $\chi_A(A)$ denotes the result of substituting A for t in the polynomial χ_A. It does **not** denote the result of substituting A for t in the expression $\det(tI_n - A)$; in particular, $\chi_A(A)$ is an $n \times n$-matrix, not a determinant!)

Theorem 2.6 (trace Cayley-Hamilton theorem)**.** Let $n \in \mathbb{N}$. Let $A \in \mathbb{K}^{n\times n}$. For every $j \in \mathbb{Z}$, define an element $c_j \in \mathbb{K}$ by $c_j = [t^{n-j}]\chi_A$. Then,

$$kc_k + \sum_{i=1}^{k} \operatorname{Tr}\left(A^i\right) c_{k-i} = 0 \qquad \text{for every } k \in \mathbb{N}.$$

Theorem 2.5 is (as has already been said) well-known and a cornerstone of linear algebra. It appears (with proofs) in [Bernha11], [Brown93, Theorem 7.23], [Camero08, Theorem 2.16], [Climen13, Theorem 23.1], [Ford22, Theorem 4.6.12], [Garrett09, §28.10], [Heffer14, Chapter Five, Section IV, Lemma 1.9], [Knapp2016, Theorem 5.9], [Loehr14, §5.15], [Mate16, §4, Theorem 1], [McDona84, Theorem I.8], [Moore68, §6.1, Theorem 1.1], [Sage08, Seconde méthode (§3)], [Shurma15], [Stoll17, Theorem 3.1], [Straub83], [BroWil89, Theorem 7.10], [Zeilbe85, §3] and in many other sources[3]. The proof we will give below will essentially repeat the proof in [Heffer14, Chapter Five, Section IV, Lemma 1.9].

Theorem 2.6 is a less known result. It appears in [LomQui16, Chapter III, Exercise 14] (with a sketch of a proof), in [Zeilbe93, $(C-H)$] (with a beautiful short proof using exterior algebra) and in [Zeilbe85, Exercise 5] (without proof); its particular case when $\mathbb{K}$ is a field also tends to appear in representation-theoretical literature (mostly left as an exercise to the reader). We will prove it similarly to Theorem 2.5; this proof, to my knowledge, is new.

3. The proofs

3.1. Proposition 2.2 and Corollary 2.4

Let us now begin proving the results stated above. As a warmup, we will prove the (rather trivial) Proposition 2.2.

We first recall how the determinant of a matrix is defined: For any $n \in \mathbb{N}$, let S_n denote the n-th symmetric group (i.e., the group of all permutations of $\{1, 2, \ldots, n\}$). If $n \in \mathbb{N}$ and $\sigma \in S_n$, then $(-1)^\sigma$ denotes the sign of the permutation σ. If $n \in \mathbb{N}$, and if $A = \left(a_{i,j}\right)_{1 \leq i \leq n,\ 1 \leq j \leq n}$ is an $n \times n$-matrix, then

$$\det A = \sum_{\sigma \in S_n} (-1)^\sigma \prod_{i=1}^{n} a_{i,\sigma(i)}. \tag{2}$$

We prepare for the proof of Proposition 2.2 by stating a simple lemma:

Lemma 3.1. Let $n \in \mathbb{N}$. Let $x_1, x_2, \ldots, x_n$ be n elements of $\mathbb{K}$. Let $y_1, y_2, \ldots, y_n$ be n elements of $\mathbb{K}$. Define a polynomial $f \in \mathbb{K}[t]$ by $f = \prod_{i=1}^{n} (t x_i + y_i)$.

[3] All the sources we are citing (with the possible exception of [Garrett09, §28.10]) prove Theorem 2.5 in full generality, although some of them do not **state** Theorem 2.5 in full generality (indeed, they often state it under the additional requirement that $\mathbb{K}$ be a field). There are other sources which only prove Theorem 2.5 in the case when $\mathbb{K}$ is a field. The note [Sage08] gives four proofs of Theorem 2.5 for the case when $\mathbb{K} = \mathbb{C}$; the first of these proofs works for every field $\mathbb{K}$, whereas the second works for any commutative ring $\mathbb{K}$, and the third and the fourth actually require $\mathbb{K} = \mathbb{C}$.

Note that some authors decline to call Theorem 2.5 the Cayley-Hamilton theorem; they instead use this name for some related result. For instance, Hefferon, in [Heffer14], uses the name "Cayley-Hamilton theorem" for a corollary.

(a) Then, f is a polynomial of degree $\leq n$.

(b) We have $[t^n] f = \prod_{i=1}^{n} x_i$.

(c) We have $[t^0] f = \prod_{i=1}^{n} y_i$.

Proof of Lemma 3.1. Obvious by multiplying out the product $\prod_{i=1}^{n} (tx_i + y_i)$ (or, if one desires a formal proof, by a straightforward induction over n). $\square$

Proof of Proposition 2.2. Write the $n \times n$-matrix A in the form $A = (a_{i,j})_{1 \leq i \leq n,\ 1 \leq j \leq n}$. Thus, $a_{i,j} \in \mathbb{K}$ for every $(i,j) \in \{1,2,\ldots,n\}^2$ (since $A \in \mathbb{K}^{n \times n}$).

Write the $n \times n$-matrix B in the form $B = (b_{i,j})_{1 \leq i \leq n,\ 1 \leq j \leq n}$. Thus, $b_{i,j} \in \mathbb{K}$ for every $(i,j) \in \{1,2,\ldots,n\}^2$ (since $B \in \mathbb{K}^{n \times n}$).

For every $\sigma \in S_n$, define a polynomial $f_\sigma \in \mathbb{K}[t]$ by

$$f_\sigma = \prod_{i=1}^{n} \left(t a_{i,\sigma(i)} + b_{i,\sigma(i)} \right). \tag{3}$$

The following holds:

> *Fact 1:* For every $\sigma \in S_n$, the polynomial f_σ is a polynomial of degree $\leq n$.

[*Proof of Fact 1:* Let $\sigma \in S_n$. Then, Lemma 3.1 **(a)** (applied to $a_{i,\sigma(i)}$, $b_{i,\sigma(i)}$ and f_σ instead of x_i, y_i and f) shows that f_σ is a polynomial of degree $\leq n$. This proves Fact 1.]

From $A = (a_{i,j})_{1 \leq i \leq n,\ 1 \leq j \leq n}$ and $B = (b_{i,j})_{1 \leq i \leq n,\ 1 \leq j \leq n}$, we obtain $tA + B = (ta_{i,j} + b_{i,j})_{1 \leq i \leq n,\ 1 \leq j \leq n}$. Hence,

$$\begin{aligned}
\det(tA + B) &= \sum_{\sigma \in S_n} (-1)^\sigma \underbrace{\prod_{i=1}^{n} \left(t a_{i,\sigma(i)} + b_{i,\sigma(i)} \right)}_{\substack{= f_\sigma \\ \text{(by (3))}}} \\
&\qquad \left(\begin{array}{c} \text{by (2), applied to } \mathbb{K}[t],\ tA + B \text{ and } ta_{i,j} + b_{i,j} \\ \text{instead of } \mathbb{K},\ A \text{ and } a_{i,j} \end{array} \right) \\
&= \sum_{\sigma \in S_n} (-1)^\sigma f_\sigma.
\end{aligned}$$

Hence, $\det(tA + B)$ is a $\mathbb{K}$-linear combination of the polynomials f_σ for $\sigma \in S_n$. Since all of these polynomials are polynomials of degree $\leq n$ (by Fact 1), we thus conclude that $\det(tA + B)$ is a $\mathbb{K}$-linear combination of polynomials of degree $\leq n$.

Thus, $\det(tA+B)$ is itself a polynomial of degree $\leq n$. This proves Proposition 2.2 **(a)**.

(b) We have

$$\left[t^0\right] \underbrace{\left(\det(tA+B)\right)}_{=\sum_{\sigma\in S_n}(-1)^\sigma f_\sigma} = \left[t^0\right] \left(\sum_{\sigma\in S_n}(-1)^\sigma f_\sigma\right) = \sum_{\sigma\in S_n}(-1)^\sigma \underbrace{\left[t^0\right] f_\sigma}_{\substack{=\prod_{i=1}^n b_{i,\sigma(i)} \\ \text{(by Lemma 3.1 (c) (applied to} \\ a_{i,\sigma(i)},\ b_{i,\sigma(i)} \text{ and } f_\sigma \\ \text{instead of } x_i,\ y_i \text{ and } f))}}$$
$$= \sum_{\sigma\in S_n}(-1)^\sigma \prod_{i=1}^n b_{i,\sigma(i)}.$$

Comparing this with

$$\det B = \sum_{\sigma\in S_n}(-1)^\sigma \prod_{i=1}^n b_{i,\sigma(i)} \qquad \left(\begin{array}{c}\text{by (2), applied to } B \text{ and } b_{i,j} \\ \text{instead of } A \text{ and } a_{i,j}\end{array}\right),$$

we obtain $\left[t^0\right]\left(\det(tA+B)\right) = \det B$. This proves Proposition 2.2 **(b)**.

(c) We have

$$\left[t^n\right] \underbrace{\left(\det(tA+B)\right)}_{=\sum_{\sigma\in S_n}(-1)^\sigma f_\sigma} = \left[t^n\right] \left(\sum_{\sigma\in S_n}(-1)^\sigma f_\sigma\right) = \sum_{\sigma\in S_n}(-1)^\sigma \underbrace{\left[t^n\right] f_\sigma}_{\substack{=\prod_{i=1}^n a_{i,\sigma(i)} \\ \text{(by Lemma 3.1 (b) (applied to} \\ a_{i,\sigma(i)},\ b_{i,\sigma(i)} \text{ and } f_\sigma \\ \text{instead of } x_i,\ y_i \text{ and } f))}}$$
$$= \sum_{\sigma\in S_n}(-1)^\sigma \prod_{i=1}^n a_{i,\sigma(i)}.$$

Comparing this with (2), we obtain $\left[t^n\right]\left(\det(tA+B)\right) = \det A$. This proves Proposition 2.2 **(c)**. $\square$

Proof of Corollary 2.4. The definition of χ_A yields $\chi_A = \det\left(\underbrace{tI_n - A}_{=tI_n+(-A)}\right) = \det(tI_n + (-A))$. Hence, Corollary 2.4 follows from Proposition 2.2 (applied to I_n and $-A$ instead of A and B). (For part **(b)**, we need the additional observation that $\det(-A) = (-1)^n \det A$.) $\square$

Let me state one more trivial observation as a corollary:

Corollary 3.2. Let $n \in \mathbb{N}$. Let $A \in \mathbb{K}^{n \times n}$. For every $j \in \mathbb{Z}$, define an element $c_j \in \mathbb{K}$ by $c_j = \left[t^{n-j}\right] \chi_A$. Then, $\chi_A = \sum\limits_{k=0}^{n} c_{n-k} t^k$.

Proof of Corollary 3.2. For every $k \in \mathbb{Z}$, the definition of c_{n-k} yields

$$c_{n-k} = \left[t^{n-(n-k)}\right] \chi_A = \left[t^k\right] \chi_A. \tag{4}$$

We know that $\chi_A \in \mathbb{K}[t]$ is a polynomial of degree $\leq n$ in t (by Corollary 2.4 **(a)**). Hence,

$$\chi_A = \sum_{k=0}^{n} \underbrace{\left(\left[t^k\right] \chi_A\right)}_{\substack{=c_{n-k} \\ \text{(by (4))}}} t^k = \sum_{k=0}^{n} c_{n-k} t^k.$$

This proves Corollary 3.2. $\square$

3.2. Reminders on the adjugate

Let us now briefly introduce the adjugate of a matrix and state some of its properties.

We first recall the definitions (mostly quoting them from [Grinbe15, Chapter 6]):

Definition 3.3. Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. Let $A = \left(a_{i,j}\right)_{1 \leq i \leq n,\ 1 \leq j \leq m}$ be an $n \times m$-matrix. Let $i_1, i_2, \ldots, i_u$ be some elements of $\{1, 2, \ldots, n\}$; let $j_1, j_2, \ldots, j_v$ be some elements of $\{1, 2, \ldots, m\}$. Then, we define $\operatorname{sub}_{i_1, i_2, \ldots, i_u}^{j_1, j_2, \ldots, j_v} A$ to be the $u \times v$-matrix $\left(a_{i_x, j_y}\right)_{1 \leq x \leq u,\ 1 \leq y \leq v}$.

Definition 3.4. Let $n \in \mathbb{N}$. Let $a_1, a_2, \ldots, a_n$ be n objects. Let $i \in \{1, 2, \ldots, n\}$. Then, $(a_1, a_2, \ldots, \widehat{a_i}, \ldots, a_n)$ shall mean the list $(a_1, a_2, \ldots, a_{i-1}, a_{i+1}, a_{i+2}, \ldots, a_n)$ (that is, the list $(a_1, a_2, \ldots, a_n)$ with its i-th entry removed). (Thus, the "hat" over the a_i means that this a_i is being omitted from the list.)

For example, $\left(1^2, 2^2, \ldots, \widehat{5^2}, \ldots, 8^2\right) = \left(1^2, 2^2, 3^2, 4^2, 6^2, 7^2, 8^2\right)$.

Definition 3.5. Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. Let A be an $n \times m$-matrix. For every $i \in \{1, 2, \ldots, n\}$ and $j \in \{1, 2, \ldots, m\}$, we let $A_{\sim i, \sim j}$ be the $(n-1) \times (m-1)$-matrix $\operatorname{sub}_{1,2,\ldots,\widehat{i},\ldots,n}^{1,2,\ldots,\widehat{j},\ldots,m} A$. (Thus, $A_{\sim i, \sim j}$ is the matrix obtained from A by crossing out the i-th row and the j-th column.)

Definition 3.6. Let $n \in \mathbb{N}$. Let A be an $n \times n$-matrix. We define a new $n \times n$-matrix $\operatorname{adj} A$ by

$$\operatorname{adj} A = \left((-1)^{i+j} \det\left(A_{\sim j, \sim i}\right)\right)_{1 \leq i \leq n,\ 1 \leq j \leq n}.$$

This matrix $\operatorname{adj} A$ is called the *adjugate* of the matrix A.

The main property of the adjugate is the following fact:

Theorem 3.7. Let $n \in \mathbb{N}$. Let A be an $n \times n$-matrix. Then,

$$A \cdot \operatorname{adj} A = \operatorname{adj} A \cdot A = \det A \cdot I_n.$$

(Recall that I_n denotes the $n \times n$ identity matrix. Expressions such as $\operatorname{adj} A \cdot A$ and $\det A \cdot I_n$ have to be understood as $(\operatorname{adj} A) \cdot A$ and $(\det A) \cdot I_n$, respectively.)

Theorem 3.7 appears in almost any text on linear algebra that considers the adjugate; for example, it appears in [Heffer14, Chapter Four, Section III, Theorem 1.9], in [Knapp2016, Proposition 2.38], in [BroWil89, Theorem 4.11] and in [Grinbe15, Theorem 6.100]. (Again, most of these sources only state it in the case when $\mathbb{K}$ is a field, but the proofs given apply in all generality. Different texts use different notations. The source that is closest to my notations here is [Grinbe15], since Theorem 3.7 above is a verbatim copy of [Grinbe15, Theorem 6.100].)

Let us state a simple fact:

Lemma 3.8. Let $n \in \mathbb{N}$. Let u and v be two elements of $\{1, 2, \ldots, n\}$. Let λ and μ be two elements of $\mathbb{K}$. Let A and B be two $n \times n$-matrices. Then,

$$(\lambda A + \mu B)_{\sim u, \sim v} = \lambda A_{\sim u, \sim v} + \mu B_{\sim u, \sim v}.$$

Proof of Lemma 3.8. Obvious. □

Next, we prove a crucial, if simple, result:

Proposition 3.9. Let $n \in \mathbb{N}$. Let $A \in \mathbb{K}^{n \times n}$ be an $n \times n$-matrix. Then, there exist n matrices $D_0, D_1, \ldots, D_{n-1}$ in $\mathbb{K}^{n \times n}$ such that

$$\operatorname{adj}(tI_n - A) = \sum_{k=0}^{n-1} t^k D_k \qquad \text{in } (\mathbb{K}[t])^{n \times n}.$$

(Here, of course, the matrix D_k on the right hand side is understood as an element of $(\mathbb{K}[t])^{n \times n}$.)

Proof of Proposition 3.9. Fix $(u, v) \in \{1, 2, \ldots, n\}^2$. Then, Proposition 2.2 **(a)** (applied to $n-1$, $(I_n)_{\sim u, \sim v}$ and $(-A)_{\sim u, \sim v}$ instead of n, A and B) shows that $\det\left(t (I_n)_{\sim u, \sim v} + (-A)_{\sim u, \sim v}\right) \in \mathbb{K}[t]$ is a polynomial of degree $\leq n-1$ in t. In other words, there exists an n-tuple $(d_{u,v,0}, d_{u,v,1}, \ldots, d_{u,v,n-1}) \in \mathbb{K}^n$ such that

$$\det\left(t (I_n)_{\sim u, \sim v} + (-A)_{\sim u, \sim v}\right) = \sum_{k=0}^{n-1} d_{u,v,k} t^k.$$

Consider this $(d_{u,v,0}, d_{u,v,1}, \ldots, d_{u,v,n-1})$. But Lemma 3.8 (applied to $\mathbb{K}[t]$, t, 1, I_n and $-A$ instead of $\mathbb{K}$, λ, μ, A and B) yields $(tI_n - A)_{\sim u,\sim v} = t(I_n)_{\sim u,\sim v} + (-A)_{\sim u,\sim v}$ (after some simplifications). Thus,

$$\det\left((tI_n - A)_{\sim u,\sim v}\right) = \det\left(t(I_n)_{\sim u,\sim v} + (-A)_{\sim u,\sim v}\right) = \sum_{k=0}^{n-1} d_{u,v,k} t^k. \tag{5}$$

Now, forget that we fixed (u,v). Thus, for every $(u,v) \in \{1,2,\ldots,n\}^2$, we have constructed an n-tuple $(d_{u,v,0}, d_{u,v,1}, \ldots, d_{u,v,n-1}) \in \mathbb{K}^n$ satisfying (5).

Now, the definition of $\operatorname{adj}(tI_n - A)$ yields

$$\begin{aligned}
\operatorname{adj}(tI_n - A) &= \left((-1)^{i+j} \underbrace{\det\left((tI_n - A)_{\sim j,\sim i}\right)}_{\substack{=\sum_{k=0}^{n-1} d_{j,i,k} t^k \\ \text{(by (5), applied to } (u,v)=(j,i))}} \right)_{1\leq i\leq n,\ 1\leq j\leq n} \\
&= \left(\underbrace{(-1)^{i+j} \sum_{k=0}^{n-1} d_{j,i,k} t^k}_{=\sum_{k=0}^{n-1} t^k (-1)^{i+j} d_{j,i,k}} \right)_{1\leq i\leq n,\ 1\leq j\leq n} = \left(\sum_{k=0}^{n-1} t^k (-1)^{i+j} d_{j,i,k} \right)_{1\leq i\leq n,\ 1\leq j\leq n}.
\end{aligned}$$

Comparing this with

$$\sum_{k=0}^{n-1} t^k \left((-1)^{i+j} d_{j,i,k} \right)_{1\leq i\leq n,\ 1\leq j\leq n} = \left(\sum_{k=0}^{n-1} t^k (-1)^{i+j} d_{j,i,k} \right)_{1\leq i\leq n,\ 1\leq j\leq n},$$

we obtain $\operatorname{adj}(tI_n - A) = \sum_{k=0}^{n-1} t^k \left((-1)^{i+j} d_{j,i,k} \right)_{1\leq i\leq n,\ 1\leq j\leq n}$. Hence, there exist n matrices $D_0, D_1, \ldots, D_{n-1}$ in $\mathbb{K}^{n\times n}$ such that

$$\operatorname{adj}(tI_n - A) = \sum_{k=0}^{n-1} t^k D_k \qquad \text{in } (\mathbb{K}[t])^{n\times n}$$

(namely, $D_k = \left((-1)^{i+j} d_{j,i,k} \right)_{1\leq i\leq n,\ 1\leq j\leq n}$ for every $k \in \{0,1,\ldots,n-1\}$). This proves Proposition 3.9. □

3.3. Polynomials with matrix entries: a trivial lemma

Lemma 3.10. Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. Let $(B_0, B_1, \ldots, B_m) \in (\mathbb{K}^{n \times n})^{m+1}$ and $(C_0, C_1, \ldots, C_m) \in (\mathbb{K}^{n \times n})^{m+1}$ be two $(m+1)$-tuples of matrices in $\mathbb{K}^{n \times n}$. Assume that

$$\sum_{k=0}^{m} t^k B_k = \sum_{k=0}^{m} t^k C_k \qquad \text{in } (\mathbb{K}[t])^{n \times n}.$$

Then, $B_k = C_k$ for every $k \in \{0, 1, \ldots, m\}$.

Proof of Lemma 3.10. For every $k \in \{0, 1, \ldots, m\}$, write the matrix $B_k \in \mathbb{K}^{n \times n}$ in the form $B_k = (b_{k,i,j})_{1 \leq i \leq n,\ 1 \leq j \leq n}$, and write the matrix $C_k \in \mathbb{K}^{n \times n}$ in the form $C_k = (c_{k,i,j})_{1 \leq i \leq n,\ 1 \leq j \leq n}$.

Now, $\sum\limits_{k=0}^{m} t^k B_k = \left(\sum\limits_{k=0}^{m} t^k b_{k,i,j} \right)_{1 \leq i \leq n,\ 1 \leq j \leq n}$ (since $B_k = (b_{k,i,j})_{1 \leq i \leq n,\ 1 \leq j \leq n}$ for every $k \in \{0, 1, \ldots, m\}$). Similarly, $\sum\limits_{k=0}^{m} t^k C_k = \left(\sum\limits_{k=0}^{m} t^k c_{k,i,j} \right)_{1 \leq i \leq n,\ 1 \leq j \leq n}$. Thus,

$$\left(\sum_{k=0}^{m} t^k b_{k,i,j} \right)_{1 \leq i \leq n,\ 1 \leq j \leq n} = \sum_{k=0}^{m} t^k B_k = \sum_{k=0}^{m} t^k C_k = \left(\sum_{k=0}^{m} t^k c_{k,i,j} \right)_{1 \leq i \leq n,\ 1 \leq j \leq n}.$$

In other words,

$$\sum_{k=0}^{m} t^k b_{k,i,j} = \sum_{k=0}^{m} t^k c_{k,i,j}$$

for every $(i, j) \in \{1, 2, \ldots, n\}^2$. Comparing coefficients on both sides of this equality, we obtain

$$b_{k,i,j} = c_{k,i,j}$$

for every $k \in \{0, 1, \ldots, m\}$ for every $(i, j) \in \{1, 2, \ldots, n\}^2$. Now, every $k \in \{0, 1, \ldots, m\}$ satisfies

$$B_k = \left(\underbrace{b_{k,i,j}}_{= c_{k,i,j}} \right)_{1 \leq i \leq n,\ 1 \leq j \leq n} = (c_{k,i,j})_{1 \leq i \leq n,\ 1 \leq j \leq n} = C_k.$$

This proves Lemma 3.10. $\square$

3.4. Proof of the Cayley-Hamilton theorem

We are now fully prepared for the proof of the Cayley-Hamilton theorem. However, we are going to organize the crucial part of this proof as a lemma, so that we can use it later in our proof of the trace Cayley-Hamilton theorem.

Lemma 3.11. Let $n \in \mathbb{N}$. Let $A \in \mathbb{K}^{n\times n}$. For every $j \in \mathbb{Z}$, define an element $c_j \in \mathbb{K}$ by $c_j = \left[t^{n-j}\right] \chi_A$.

Let $D_0, D_1, \ldots, D_{n-1}$ be n matrices in $\mathbb{K}^{n\times n}$ such that

$$\operatorname{adj}\left(tI_n - A\right) = \sum_{k=0}^{n-1} t^k D_k \qquad \text{in } \left(\mathbb{K}\left[t\right]\right)^{n\times n}. \tag{6}$$

Thus, an n-tuple $\left(D_0, D_1, \ldots, D_{n-1}\right)$ of matrices in $\mathbb{K}^{n\times n}$ is defined. Extend this n-tuple to a family $\left(D_k\right)_{k\in\mathbb{Z}}$ of matrices in $\mathbb{K}^{n\times n}$ by setting

$$\left(D_k = 0_{n\times n} \qquad \text{for every } k \in \mathbb{Z} \setminus \{0, 1, \ldots, n-1\}\right). \tag{7}$$

Then:

(a) We have $\chi_A = \sum\limits_{k=0}^{n} c_{n-k} t^k$.

(b) For every integer k, we have $c_{n-k} I_n = D_{k-1} - AD_k$.

(c) Every $k \in \mathbb{N}$ satisfies

$$\sum_{i=0}^{k} c_{k-i} A^i = D_{n-1-k}.$$

Proof of Lemma 3.11. **(a)** Lemma 3.11 **(a)** is just Corollary 3.2.

(b) We have

$$\begin{aligned}
\sum_{k=0}^{n} t^k D_{k-1} &= t^0 \underbrace{D_{0-1}}_{\substack{=D_{-1}=0_{n\times n} \\ \text{(by (7))}}} + \sum_{k=1}^{n} t^k D_{k-1} = \sum_{k=1}^{n} t^k D_{k-1} = \sum_{k=0}^{n-1} \underbrace{t^{k+1}}_{=tt^k} \underbrace{D_{(k+1)-1}}_{=D_k} \\
&\qquad \text{(here, we have substituted } k+1 \text{ for } k \text{ in the sum)} \\
&= \sum_{k=0}^{n-1} tt^k D_k = t \underbrace{\sum_{k=0}^{n-1} t^k D_k}_{\substack{=\operatorname{adj}(tI_n - A) \\ \text{(by (6))}}} = t \operatorname{adj}\left(tI_n - A\right)
\end{aligned} \tag{8}$$

and

$$\begin{aligned}
\sum_{k=0}^{n} t^k D_k &= t^n \underbrace{D_n}_{\substack{=0_{n\times n} \\ \text{(by (7))}}} + \sum_{k=0}^{n-1} t^k D_k = \sum_{k=0}^{n-1} t^k D_k \\
&= \operatorname{adj}\left(tI_n - A\right) \qquad \text{(by (6))}.
\end{aligned} \tag{9}$$

But Theorem 3.7 (applied to $\mathbb{K}\left[t\right]$ and $tI_n - A$ instead of $\mathbb{K}$ and A) shows that

$$\left(tI_n - A\right) \cdot \operatorname{adj}\left(tI_n - A\right) = \operatorname{adj}\left(tI_n - A\right) \cdot \left(tI_n - A\right) = \det\left(tI_n - A\right) \cdot I_n.$$

Thus, in particular,

$$(tI_n - A) \cdot \operatorname{adj}(tI_n - A) = \underbrace{\det(tI_n - A)}_{\substack{=\chi_A \\ \text{(by the definition of } \chi_A)}} \cdot I_n = \chi_A \cdot I_n,$$

so that

$$\begin{aligned}
\chi_A \cdot I_n &= (tI_n - A) \cdot \operatorname{adj}(tI_n - A) \\
&= \underbrace{t \operatorname{adj}(tI_n - A)}_{\substack{= \sum_{k=0}^{n} t^k D_{k-1} \\ \text{(by (8))}}} - A \underbrace{\operatorname{adj}(tI_n - A)}_{\substack{= \sum_{k=0}^{n} t^k D_k \\ \text{(by (9))}}} = \sum_{k=0}^{n} t^k D_{k-1} - \underbrace{A \sum_{k=0}^{n} t^k D_k}_{= \sum_{k=0}^{n} t^k A D_k} \\
&= \sum_{k=0}^{n} t^k D_{k-1} - \sum_{k=0}^{n} t^k A D_k = \sum_{k=0}^{n} t^k (D_{k-1} - AD_k).
\end{aligned}$$

Thus,

$$\begin{aligned}
\sum_{k=0}^{n} t^k (D_{k-1} - AD_k) &= \underbrace{\chi_A}_{\substack{= \sum_{k=0}^{n} c_{n-k} t^k \\ \text{(by Lemma 3.11 } \textbf{(a)})}} \cdot I_n = \left(\sum_{k=0}^{n} c_{n-k} t^k \right) \cdot I_n \\
&= \sum_{k=0}^{n} t^k c_{n-k} I_n.
\end{aligned}$$

Lemma 3.10 (applied to $m = n$, $B_k = D_{k-1} - AD_k$ and $C_k = c_{n-k} I_n$) thus shows that

$$D_{k-1} - AD_k = c_{n-k} I_n \qquad \text{for every } k \in \{0, 1, \ldots, n\}. \tag{10}$$

Now, let k be an integer. We must prove that $c_{n-k} I_n = D_{k-1} - AD_k$.

If $k \in \{0, 1, \ldots, n\}$, then this follows from (10). Thus, we WLOG assume that $k \notin \{0, 1, \ldots, n\}$. Hence, $k - 1 \in \mathbb{Z} \setminus \{0, 1, \ldots, n-1\}$, so that (7) (applied to $k-1$ instead of k) yields $D_{k-1} = 0_{n \times n}$. Also, $k \notin \{0, 1, \ldots, n\}$ leads to $k \in \mathbb{Z} \setminus \{0, 1, \ldots, n-1\}$; therefore, (7) yields $D_k = 0_{n \times n}$. Now, $\underbrace{D_{k-1}}_{=0_{n \times n}} - A \underbrace{D_k}_{=0_{n \times n}} = 0_{n \times n} - 0_{n \times n} = 0_{n \times n}$.

On the other hand, $c_{n-k} = 0$ [4]. Hence, $\underbrace{c_{n-k}}_{=0} I_n = 0_{n \times n}$. Compared with $D_{k-1} - AD_k = 0_{n \times n}$, this yields $c_{n-k} I_n = D_{k-1} - AD_k$.

Hence, $c_{n-k} I_n = D_{k-1} - AD_k$ is proven. In other words, Lemma 3.11 **(b)** is proven.

[4] *Proof.* Recall that χ_A is a polynomial of degree $\leq n$ (by Corollary 2.4 **(a)**). Hence, $\left[t^k\right] \chi_A = 0$ (since $k \notin \{0, 1, \ldots, n\}$). Now, (4) yields $c_{n-k} = \left[t^k\right] \chi_A = 0$.

(c) Let $k \in \mathbb{N}$. Then,

$$\begin{aligned}
\sum_{i=0}^{k} c_{k-i} A^i &= \sum_{i=n-k}^{n} \underbrace{c_{k-(k-n+i)}}_{=c_{n-i}} A^{k-n+i} \\
&\qquad \text{(here, we have substituted } k-n+i \text{ for } i \text{ in the sum)} \\
&= \sum_{i=n-k}^{n} \underbrace{c_{n-i} A^{k-n+i}}_{=A^{k-n+i} c_{n-i} I_n} = \sum_{i=n-k}^{n} A^{k-n+i} \underbrace{c_{n-i} I_n}_{\substack{=D_{i-1}-AD_i \\ \text{(by Lemma 3.11 (b),} \\ \text{applied to } i \text{ instead of } k)}} \\
&= \sum_{i=n-k}^{n} \underbrace{A^{k-n+i}\left(D_{i-1} - AD_i\right)}_{=A^{k-n+i}D_{i-1} - A^{k-n+i}AD_i} \\
&= \sum_{i=n-k}^{n} \left(A^{k-n+i} D_{i-1} - \underbrace{A^{k-n+i} A}_{=A^{k-n+i+1}=A^{k-n+(i+1)}} \underbrace{D_i}_{=D_{(i+1)-1}} \right) \\
&= \sum_{i=n-k}^{n} \left(A^{k-n+i} D_{i-1} - A^{k-n+(i+1)} D_{(i+1)-1} \right) \\
&= \underbrace{A^{k-n+(n-k)}}_{=A^0=I_n} D_{n-k-1} - A^{k-n+(n+1)} \underbrace{D_{(n+1)-1}}_{\substack{=D_n=0_{n\times n} \\ \text{(by (7))}}} \\
&\qquad \text{(by the telescope principle)} \\
&= D_{n-k-1} = D_{n-1-k}.
\end{aligned}$$

This proves Lemma 3.11 **(c)**. ☐

Proof of Theorem 2.5. For every $j \in \mathbb{Z}$, define an element $c_j \in \mathbb{K}$ by $c_j = \left[t^{n-j}\right] \chi_A$.

Proposition 3.9 shows that there exist n matrices $D_0, D_1, \ldots, D_{n-1}$ in $\mathbb{K}^{n\times n}$ such that

$$\operatorname{adj}(tI_n - A) = \sum_{k=0}^{n-1} t^k D_k \qquad \text{in } (\mathbb{K}[t])^{n\times n}.$$

Consider these $D_0, D_1, \ldots, D_{n-1}$. Thus, an n-tuple $(D_0, D_1, \ldots, D_{n-1})$ of matrices in $\mathbb{K}^{n\times n}$ is defined. Extend this n-tuple to a family $(D_k)_{k\in\mathbb{Z}}$ of matrices in $\mathbb{K}^{n\times n}$ by setting

$$D_k = 0_{n\times n} \qquad \text{for every } k \in \mathbb{Z} \setminus \{0, 1, \ldots, n-1\}.$$

Thus, in particular, $D_{-1} = 0_{n\times n}$.

Lemma 3.11 **(a)** shows that $\chi_A = \sum\limits_{k=0}^{n} c_{n-k} t^k = \sum\limits_{i=0}^{n} c_{n-i} t^i$. Substituting A for t in this equality, we obtain

$$\chi_A(A) = \sum_{i=0}^{n} c_{n-i} A^i = D_{n-1-n} \qquad \text{(by Lemma 3.11 **(c)**, applied to } k = n\text{)}$$
$$= D_{-1} = 0_{n\times n}.$$

This proves Theorem 2.5. □

3.5. Derivations and determinants

Now, let us make what seems to be a detour, and define $\mathbb{K}$-derivations of a $\mathbb{K}$-algebra[5]:

Definition 3.12. Let $\mathbb{L}$ be a $\mathbb{K}$-algebra. A $\mathbb{K}$-linear map $f : \mathbb{L} \to \mathbb{L}$ is said to be a $\mathbb{K}$-*derivation* if it satisfies

$$(f(ab) = af(b) + f(a)b \qquad \text{for every } a \in \mathbb{L} \text{ and } b \in \mathbb{L}). \tag{11}$$

The notion of a "$\mathbb{K}$-derivation" is a particular case of the notion of a "**k**-derivation" defined in [Grinbe16a, Definition 1.5]; specifically, it is obtained from the latter when setting $\mathbf{k} = \mathbb{K}$, $A = \mathbb{L}$ and $M = \mathbb{L}$. This particular case will suffice for us. Examples of $\mathbb{K}$-derivations abound (there are several in [Grinbe16a]), but the only one we will need is the following:

Proposition 3.13. Let $\partial : \mathbb{K}[t] \to \mathbb{K}[t]$ be the differentiation operator (i.e., the map that sends every polynomial $f \in \mathbb{K}[t]$ to the derivative of f). Then, $\partial : \mathbb{K}[t] \to \mathbb{K}[t]$ is a $\mathbb{K}$-derivation.

Proof of Proposition 3.13. This follows from the fact that $\partial(ab) = a\partial(b) + \partial(a)b$ for any two polynomials a and b (the well-known Leibniz law). □

A fundamental fact about $\mathbb{K}$-derivations is the following:

Proposition 3.14. Let $\mathbb{L}$ be a $\mathbb{K}$-algebra. Let $f : \mathbb{L} \to \mathbb{L}$ be a $\mathbb{K}$-derivation. Let $n \in \mathbb{N}$, and let $a_1, a_2, \ldots, a_n \in \mathbb{L}$. Then,

$$f(a_1 a_2 \cdots a_n) = \sum_{i=1}^{n} a_1 a_2 \cdots a_{i-1} f(a_i) a_{i+1} a_{i+2} \cdots a_n.$$

[5] See [Grinbe16a, Convention 1.1] for what we mean by a "$\mathbb{K}$-algebra". In a nutshell, we require $\mathbb{K}$-algebras to be associative and unital, and we require the multiplication map on a $\mathbb{K}$-algebra to be $\mathbb{K}$-bilinear.

This proposition is a particular case of [Grinbe16a, Theorem 1.14] (obtained by setting $\mathbf{k} = \mathbb{K}$, $A = \mathbb{L}$ and $M = \mathbb{L}$); it is also easy to prove[6].

What we are going to need is a formula for how a derivation acts on the determinant of a matrix. We first introduce a notation:

Definition 3.15. Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. Let $\mathbb{L}$ and $\mathbb{M}$ be rings. Let $f : \mathbb{L} \to \mathbb{M}$ be any map. Then, $f^{n\times m}$ will denote the map from $\mathbb{L}^{n\times m}$ to $\mathbb{M}^{n\times m}$ which sends every matrix $\left(a_{i,j}\right)_{1\leq i\leq n,\ 1\leq j\leq m} \in \mathbb{L}^{n\times m}$ to the matrix $\left(f\left(a_{i,j}\right)\right)_{1\leq i\leq n,\ 1\leq j\leq m} \in \mathbb{M}^{n\times m}$. (In other words, $f^{n\times m}$ is the map which takes an $n\times m$-matrix in $\mathbb{L}^{n\times m}$, and applies f to each entry of this matrix.)

Theorem 3.16. Let $\mathbb{L}$ be a commutative $\mathbb{K}$-algebra. Let $f : \mathbb{L} \to \mathbb{L}$ be a $\mathbb{K}$-derivation. Let $n \in \mathbb{N}$. Let $A \in \mathbb{L}^{n\times n}$. Then,

$$f\left(\det A\right) = \operatorname{Tr}\left(f^{n\times n}\left(A\right) \cdot \operatorname{adj} A\right).$$

Proving Theorem 3.16 will take us a while. Let us begin by stating three lemmas:

Lemma 3.17. Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. Let $A = \left(a_{i,j}\right)_{1\leq i\leq n,\ 1\leq j\leq m} \in \mathbb{K}^{n\times m}$ and $B = \left(b_{i,j}\right)_{1\leq i\leq m,\ 1\leq j\leq n} \in \mathbb{K}^{m\times n}$. Then,

$$\operatorname{Tr}\left(AB\right) = \sum_{i=1}^{n}\sum_{j=1}^{m} a_{i,j}b_{j,i}.$$

Proof of Lemma 3.17. The definition of AB yields $AB = \left(\sum_{k=1}^{m} a_{i,k}b_{k,j}\right)_{1\leq i\leq n,\ 1\leq j\leq n}$ (since $A = \left(a_{i,j}\right)_{1\leq i\leq n,\ 1\leq j\leq m}$ and $B = \left(b_{i,j}\right)_{1\leq i\leq m,\ 1\leq j\leq n}$). Hence,

$$\operatorname{Tr}\left(AB\right) = \sum_{i=1}^{n}\sum_{k=1}^{m} a_{i,k}b_{k,i} = \sum_{i=1}^{n}\sum_{j=1}^{m} a_{i,j}b_{j,i}$$

(here, we have renamed the summation index k as j in the second sum). This proves Lemma 3.17. $\square$

Lemma 3.18. Let $\mathbb{L}$ be a commutative $\mathbb{K}$-algebra. Let $f : \mathbb{L} \to \mathbb{L}$ be a $\mathbb{K}$-derivation. Let $n \in \mathbb{N}$, and let $a_1, a_2, \ldots, a_n \in \mathbb{L}$. Then,

$$f\left(a_1a_2\cdots a_n\right) = \sum_{k=1}^{n} f\left(a_k\right) \prod_{\substack{i\in\{1,2,\ldots,n\};\\ i\neq k}} a_i.$$

[6] First one should show that $f\left(1\right) = 0$ (by applying (11) to $a = 1$ and $b = 1$). Then, one can prove Proposition 3.14 by straightforward induction on n.

Proof of Lemma 3.18. Proposition 3.14 yields

$$\begin{aligned}
f(a_1a_2\cdots a_n) &= \sum_{i=1}^{n} a_1a_2\cdots a_{i-1}f(a_i)\,a_{i+1}a_{i+2}\cdots a_n \\
&= \sum_{k=1}^{n} \underbrace{a_1a_2\cdots a_{k-1}f(a_k)\,a_{k+1}a_{k+2}\cdots a_n}_{\substack{=f(a_k)(a_1a_2\cdots a_{k-1})(a_{k+1}a_{k+2}\cdots a_n) \\ \text{(since } \mathbb{L} \text{ is commutative)}}} \\
&\qquad \left(\text{here, we have renamed the summation index } i \text{ as } k\right) \\
&= \sum_{k=1}^{n} f(a_k)\underbrace{(a_1a_2\cdots a_{k-1})(a_{k+1}a_{k+2}\cdots a_n)}_{=\prod_{\substack{i\in\{1,2,\ldots,n\};\\ i\neq k}} a_i} = \sum_{k=1}^{n} f(a_k)\prod_{\substack{i\in\{1,2,\ldots,n\};\\ i\neq k}} a_i.
\end{aligned}$$

This proves Lemma 3.18. $\square$

Lemma 3.19. Let $n\in\mathbb{N}$. Let $A=(a_{i,j})_{1\leq i\leq n,\ 1\leq j\leq n}$ be an $n\times n$-matrix. Let $p\in\{1,2,\ldots,n\}$ and $q\in\{1,2,\ldots,n\}$. Then,

$$\sum_{\substack{\sigma\in S_n;\\ \sigma(p)=q}} (-1)^{\sigma}\prod_{\substack{i\in\{1,2,\ldots,n\};\\ i\neq p}} a_{i,\sigma(i)} = (-1)^{p+q}\det\left(A_{\sim p,\sim q}\right).$$

Lemma 3.19 is [Grinbe15, Lemma 6.84]; it is also easy to prove (it is the main step in the proof of the Laplace expansion formula for the determinant).

Proof of Theorem 3.16. Write the matrix $A\in\mathbb{L}^{n\times n}$ in the form $A=(a_{i,j})_{1\leq i\leq n,\ 1\leq j\leq n}$. Hence, $f^{n\times n}(A)=\left(f(a_{i,j})\right)_{1\leq i\leq n,\ 1\leq j\leq n}$ (by the definition of $f^{n\times n}$). The definition of $\operatorname{adj}A$ shows that $\operatorname{adj}A=\left((-1)^{i+j}\det\left(A_{\sim j,\sim i}\right)\right)_{1\leq i\leq n,\ 1\leq j\leq n}$. Hence, Lemma 3.17 (applied to $\mathbb{L}$, n, $f^{n\times n}(A)$, $f(a_{i,j})$, $\operatorname{adj}A$ and $(-1)^{i+j}\det\left(A_{\sim j,\sim i}\right)$ instead of $\mathbb{K}$, m, A, $a_{i,j}$, B and $b_{i,j}$) yields

$$\begin{aligned}
\operatorname{Tr}\left(f^{n\times n}(A)\cdot\operatorname{adj}A\right) &= \sum_{i=1}^{n}\sum_{j=1}^{n} f(a_{i,j})\,(-1)^{j+i}\det\left(A_{\sim i,\sim j}\right) \\
&= \sum_{k=1}^{n}\sum_{j=1}^{n} f(a_{k,j})\underbrace{(-1)^{j+k}}_{=(-1)^{k+j}}\det\left(A_{\sim k,\sim j}\right) \\
&\qquad \left(\begin{array}{c}\text{here, we have renamed the summation index } i \\ \text{as } k \text{ in the outer sum}\end{array}\right) \\
&= \sum_{k=1}^{n}\sum_{j=1}^{n} f(a_{k,j})\,(-1)^{k+j}\det\left(A_{\sim k,\sim j}\right). \qquad (12)
\end{aligned}$$

But the map f is a $\mathbb{K}$-derivation, and thus is $\mathbb{K}$-linear. Now, (2) (applied to $\mathbb{L}$ instead of $\mathbb{K}$) yields $\det A = \sum\limits_{\sigma \in S_n} (-1)^\sigma \prod\limits_{i=1}^n a_{i,\sigma(i)}$. Applying f to both sides of this equality, we find

$$\begin{aligned}
& f(\det A) \\
&= f\left(\sum_{\sigma \in S_n} (-1)^\sigma \prod_{i=1}^n a_{i,\sigma(i)}\right) \\
&= \sum_{\sigma \in S_n} (-1)^\sigma f\left(\underbrace{\prod_{i=1}^n a_{i,\sigma(i)}}_{=a_{1,\sigma(1)}a_{2,\sigma(2)}\cdots a_{n,\sigma(n)}}\right) \qquad \text{(since the map } f \text{ is } \mathbb{K}\text{-linear)} \\
&= \sum_{\sigma \in S_n} (-1)^\sigma \underbrace{f\left(a_{1,\sigma(1)}a_{2,\sigma(2)} \cdots a_{n,\sigma(n)}\right)}_{\substack{= \sum_{k=1}^n f\left(a_{k,\sigma(k)}\right) \prod_{i \in \{1,2,\ldots,n\};\ i \neq k} a_{i,\sigma(i)} \\ \text{(by Lemma 3.18, applied to} \\ a_{i,\sigma(i)} \text{ instead of } a_i)}} \\
&= \sum_{\sigma \in S_n} (-1)^\sigma \sum_{k=1}^n f\left(a_{k,\sigma(k)}\right) \prod_{\substack{i \in \{1,2,\ldots,n\}; \\ i \neq k}} a_{i,\sigma(i)} \\
&= \sum_{k=1}^n \sum_{\sigma \in S_n} (-1)^\sigma f\left(a_{k,\sigma(k)}\right) \prod_{\substack{i \in \{1,2,\ldots,n\}; \\ i \neq k}} a_{i,\sigma(i)}. \qquad (13)
\end{aligned}$$

But every $k \in \{1, 2, \ldots, n\}$ satisfies

$$\begin{aligned}
&\underbrace{\sum_{\sigma \in S_n}}_{\substack{= \sum\limits_{j \in \{1,2,\ldots,n\}} \sum\limits_{\substack{\sigma \in S_n; \\ \sigma(k)=j}} \\ \text{(since } \sigma(k) \in \{1,2,\ldots,n\} \text{ for each } \sigma \in S_n\text{)}}} (-1)^{\sigma} f\left(a_{k,\sigma(k)}\right) \prod_{\substack{i \in \{1,2,\ldots,n\}; \\ i \neq k}} a_{i,\sigma(i)} \\
&= \sum_{j \in \{1,2,\ldots,n\}} \sum_{\substack{\sigma \in S_n; \\ \sigma(k)=j}} (-1)^{\sigma} f\left(\underbrace{a_{k,\sigma(k)}}_{\substack{= a_{k,j} \\ \text{(since } \sigma(k)=j\text{)}}}\right) \prod_{\substack{i \in \{1,2,\ldots,n\}; \\ i \neq k}} a_{i,\sigma(i)} \\
&= \sum_{j \in \{1,2,\ldots,n\}} \sum_{\substack{\sigma \in S_n; \\ \sigma(k)=j}} (-1)^{\sigma} f\left(a_{k,j}\right) \prod_{\substack{i \in \{1,2,\ldots,n\}; \\ i \neq k}} a_{i,\sigma(i)} \\
&= \underbrace{\sum_{j \in \{1,2,\ldots,n\}}}_{= \sum\limits_{j=1}^{n}} f\left(a_{k,j}\right) \underbrace{\sum_{\substack{\sigma \in S_n; \\ \sigma(k)=j}} (-1)^{\sigma} \prod_{\substack{i \in \{1,2,\ldots,n\}; \\ i \neq k}} a_{i,\sigma(i)}}_{\substack{= (-1)^{k+j} \det\left(A_{\sim k, \sim j}\right) \\ \text{(by Lemma 3.19, applied to} \\ \mathbb{L},\ k \text{ and } j \text{ instead of } \mathbb{K},\ p \text{ and } q\text{)}}} \\
&= \sum_{j=1}^{n} f\left(a_{k,j}\right) (-1)^{k+j} \det\left(A_{\sim k, \sim j}\right).
\end{aligned}$$

Hence, (13) becomes

$$\begin{aligned}
f(\det A) &= \sum_{k=1}^{n} \underbrace{\sum_{\sigma \in S_n} (-1)^{\sigma} f\left(a_{k,\sigma(k)}\right) \prod_{\substack{i \in \{1,2,\ldots,n\}; \\ i \neq k}} a_{i,\sigma(i)}}_{= \sum\limits_{j=1}^{n} f\left(a_{k,j}\right) (-1)^{k+j} \det\left(A_{\sim k, \sim j}\right)} \\
&= \sum_{k=1}^{n} \sum_{j=1}^{n} f\left(a_{k,j}\right) (-1)^{k+j} \det\left(A_{\sim k, \sim j}\right) = \operatorname{Tr}\left(f^{n \times n}(A) \cdot \operatorname{adj} A\right)
\end{aligned}$$

(by (12)). This proves Theorem 3.16. □

3.6. The derivative of the characteristic polynomial

The characteristic polynomial χ_A of a square matrix A is, first of all, a polynomial; and a polynomial has a derivative. We shall have need for a formula for this derivative:

Theorem 3.20. Let $n \in \mathbb{N}$. Let $A \in \mathbb{K}^{n\times n}$. Let $\partial : \mathbb{K}[t] \to \mathbb{K}[t]$ be the differentiation operator (i.e., the map that sends every polynomial $f \in \mathbb{K}[t]$ to the derivative of f). Then,

$$\partial \chi_A = \operatorname{Tr}(\operatorname{adj}(tI_n - A)).$$

Proof of Theorem 3.20. Proposition 3.13 shows that $\partial : \mathbb{K}[t] \to \mathbb{K}[t]$ is a $\mathbb{K}$-derivation. Now, consider the map $\partial^{n\times n} : (\mathbb{K}[t])^{n\times n} \to (\mathbb{K}[t])^{n\times n}$ (defined according to Definition 3.15). It is easy to see that

$$\partial^{n\times n}(tB - A) = B \tag{14}$$

for any $n \times n$-matrix $B \in \mathbb{K}^{n\times n}$ [7]. Applying this to $B = I_n$, we obtain $\partial^{n\times n}(tI_n - A) = I_n$.

The definition of χ_A yields $\chi_A = \det(tI_n - A)$. Applying the map ∂ to both sides of this equality, we obtain

$$\begin{aligned} \partial \chi_A &= \partial(\det(tI_n - A)) = \operatorname{Tr}\left(\underbrace{\partial^{n\times n}(tI_n - A)}_{=I_n} \cdot \operatorname{adj}(tI_n - A)\right) \\ &\qquad \left(\begin{array}{c}\text{by Theorem 3.16 (applied to } \mathbb{K}[t], \partial \text{ and } tI_n - A \\ \text{instead of } \mathbb{L}, f \text{ and } A)\end{array}\right) \\ &= \operatorname{Tr}\left(\underbrace{I_n \cdot \operatorname{adj}(tI_n - A)}_{=\operatorname{adj}(tI_n - A)}\right) = \operatorname{Tr}(\operatorname{adj}(tI_n - A)). \end{aligned}$$

This proves Theorem 3.20. □

[7] *Proof.* Let $B \in \mathbb{K}^{n\times n}$ be an $n \times n$-matrix. Write the matrix B in the form $B = (b_{i,j})_{1\le i\le n,\ 1\le j\le n}$. Write the matrix A in the form $A = (a_{i,j})_{1\le i\le n,\ 1\le j\le n}$. Both matrices A and B belong to $\mathbb{K}^{n\times n}$; thus, every $(i,j) \in \{1,2,\ldots,n\}^2$ satisfies $a_{i,j} \in \mathbb{K}$ and $b_{i,j} \in \mathbb{K}$ and therefore $\partial(tb_{i,j} - a_{i,j}) = b_{i,j}$ (since ∂ is the differentiation operator).

Now,

$$\begin{aligned} &t \underbrace{B}_{=(b_{i,j})_{1\le i\le n,\ 1\le j\le n}} - \underbrace{A}_{=(a_{i,j})_{1\le i\le n,\ 1\le j\le n}} \\ &= t(b_{i,j})_{1\le i\le n,\ 1\le j\le n} - (a_{i,j})_{1\le i\le n,\ 1\le j\le n} = (tb_{i,j} - a_{i,j})_{1\le i\le n,\ 1\le j\le n}. \end{aligned}$$

Hence, the definition of the map $\partial^{n\times n}$ yields

$$\partial^{n\times n}(tB - A) = \left(\underbrace{\partial(tb_{i,j} - a_{i,j})}_{=b_{i,j}}\right)_{1\le i\le n,\ 1\le j\le n} = (b_{i,j})_{1\le i\le n,\ 1\le j\le n} = B,$$

qed.

We can use Theorem 3.20 to obtain the following result:

Proposition 3.21. Let $n \in \mathbb{N}$. Let $A \in \mathbb{K}^{n \times n}$. For every $j \in \mathbb{Z}$, define an element $c_j \in \mathbb{K}$ by $c_j = \left[t^{n-j}\right] \chi_A$.

Let $D_0, D_1, \ldots, D_{n-1}$ be n matrices in $\mathbb{K}^{n \times n}$ satisfying (6). Thus, an n-tuple $(D_0, D_1, \ldots, D_{n-1})$ of matrices in $\mathbb{K}^{n \times n}$ is defined. Extend this n-tuple to a family $(D_k)_{k \in \mathbb{Z}}$ of matrices in $\mathbb{K}^{n \times n}$ by setting (7). Then, every $k \in \mathbb{Z}$ satisfies

$$\operatorname{Tr}(D_k) = (k+1) c_{n-(k+1)}. \tag{15}$$

Proof of Proposition 3.21. Let $\partial : \mathbb{K}[t] \to \mathbb{K}[t]$ be the differentiation operator (i.e., the map that sends every polynomial $f \in \mathbb{K}[t]$ to the derivative of f).

Lemma 3.11 **(a)** yields $\chi_A = \sum_{k=0}^{n} c_{n-k} t^k$. Applying the map ∂ to both sides of this equality, we obtain

$$\begin{aligned}
\partial \chi_A &= \partial \left(\sum_{k=0}^{n} c_{n-k} t^k \right) = \sum_{k=1}^{n} c_{n-k} k t^{k-1} \qquad \left(\begin{array}{c} \text{since } \partial \text{ is the differentiation} \\ \text{operator} \end{array} \right) \\
&= \sum_{k=1}^{n} k c_{n-k} t^{k-1} = \sum_{k=0}^{n-1} (k+1) c_{n-(k+1)} t^k
\end{aligned}$$

(here, we have substituted $k+1$ for k in the sum). Comparing this with

$$\begin{aligned}
\partial \chi_A &= \operatorname{Tr} \left(\underbrace{\operatorname{adj}(tI_n - A)}_{\substack{= \sum_{k=0}^{n-1} t^k D_k \\ \text{(by (6))}}} \right) \qquad \text{(by Theorem 3.20)} \\
&= \operatorname{Tr} \left(\sum_{k=0}^{n-1} t^k D_k \right) = \sum_{k=0}^{n-1} t^k \operatorname{Tr}(D_k) = \sum_{k=0}^{n-1} \operatorname{Tr}(D_k) t^k,
\end{aligned}$$

we obtain

$$\sum_{k=0}^{n-1} \operatorname{Tr}(D_k) t^k = \sum_{k=0}^{n-1} (k+1) c_{n-(k+1)} t^k.$$

This is an identity between two polynomials in $\mathbb{K}[t]$. Comparing coefficients on both sides of this identity, we conclude that

$$\operatorname{Tr}(D_k) = (k+1) c_{n-(k+1)} \qquad \text{for every } k \in \{0, 1, \ldots, n-1\}. \tag{16}$$

Now, let $k \in \mathbb{Z}$. We must prove (15).

If $k \in \{0, 1, \ldots, n-1\}$, then (15) follows immediately from (16). Hence, for the rest of this proof, we WLOG assume that we don't have $k \in \{0, 1, \ldots, n-1\}$.

We don't have $k \in \{0, 1, \ldots, n-1\}$. Thus, $k \in \mathbb{Z} \setminus \{0, 1, \ldots, n-1\}$. Hence, (7) yields $D_k = 0_{n \times n}$, so that $\operatorname{Tr}(D_k) = \operatorname{Tr}(0_{n \times n}) = 0$.

Recall again that $k \in \mathbb{Z} \setminus \{0, 1, \ldots, n-1\}$. In other words, we have either $k < 0$ or $k \geq n$. Thus, we are in one of the following two cases:

Case 1: We have $k < 0$.

Case 2: We have $k \geq n$.

Let us first consider Case 1. In this case, we have $k < 0$. If $k = -1$, then (15) holds[8]. Hence, for the rest of this proof, we WLOG assume that $k \neq -1$. Combining $k < 0$ with $k \neq -1$, we obtain $k < -1$. Hence, $k + 1 < 0$.

The definition of $c_{n-(k+1)}$ yields $c_{n-(k+1)} = \left[\underbrace{t^{n-(n-(k+1))}}_{=t^{k+1}}\right] \chi_A = \left[t^{k+1}\right] \chi_A = 0$ (since $k + 1 < 0$, but χ_A is a polynomial). Hence, $(k+1) \underbrace{c_{n-(k+1)}}_{=0} = 0$. Comparing this with $\operatorname{Tr}(D_k) = 0$, we obtain $\operatorname{Tr}(D_k) = (k+1) c_{n-(k+1)}$. Hence, (15) is proven in Case 1.

Let us now consider Case 2. In this case, we have $k \geq n$. Thus, $k + 1 \geq n + 1 > n$. But χ_A is a polynomial of degree $\leq n$. Hence, $[t^m] \chi_A = 0$ for every integer $m > n$. Applying this to $m = k + 1$, we obtain $\left[t^{k+1}\right] \chi_A = 0$ (since $k + 1 > n$).

The definition of $c_{n-(k+1)}$ yields $c_{n-(k+1)} = \left[\underbrace{t^{n-(n-(k+1))}}_{=t^{k+1}}\right] \chi_A = \left[t^{k+1}\right] \chi_A = 0$. Hence, $(k+1) \underbrace{c_{n-(k+1)}}_{=0} = 0$. Comparing this with $\operatorname{Tr}(D_k) = 0$, we obtain $\operatorname{Tr}(D_k) = (k+1) c_{n-(k+1)}$. Hence, (15) is proven in Case 2.

We have now proven (15) in each of the two Cases 1 and 2. Thus, (15) always holds. Thus, Proposition 3.21 is proven. $\square$

3.7. Proof of the trace Cayley-Hamilton theorem

Now, we can finally prove the trace Cayley-Hamilton theorem itself:

Proof of Theorem 2.6. Proposition 3.9 shows that there exist n matrices $D_0, D_1, \ldots, D_{n-1}$ in $\mathbb{K}^{n \times n}$ such that

$$\operatorname{adj}(tI_n - A) = \sum_{k=0}^{n-1} t^k D_k \qquad \text{in } (\mathbb{K}[t])^{n \times n}.$$

[8] *Proof.* Assume that $k = -1$. Then, $k + 1 = 0$, so that $\underbrace{(k+1)}_{=0} c_{n-(k+1)} = 0$. Comparing this with $\operatorname{Tr}(D_k) = 0$, we obtain $\operatorname{Tr}(D_k) = (k+1) c_{n-(k+1)}$; hence, (15) holds, qed.

Consider these $D_0, D_1, \ldots, D_{n-1}$. Thus, an n-tuple $(D_0, D_1, \ldots, D_{n-1})$ of matrices in $\mathbb{K}^{n\times n}$ is defined. Extend this n-tuple to a family $(D_k)_{k\in\mathbb{Z}}$ of matrices in $\mathbb{K}^{n\times n}$ by setting

$$(D_k = 0_{n\times n} \qquad \text{for every } k \in \mathbb{Z} \setminus \{0, 1, \ldots, n-1\}).$$

Now, let $k \in \mathbb{N}$. Then, Proposition 3.21 (applied to $n-1-k$ instead of k) yields

$$\operatorname{Tr}(D_{n-1-k}) = \underbrace{((n-1-k)+1)}_{=n-k} \underbrace{c_{n-((n-1-k)+1)}}_{\substack{=c_k \\ (\text{since } n-((n-1-k)+1)=k)}} = (n-k)\, c_k.$$

Thus,

$$\begin{aligned}
(n-k)\, c_k &= \operatorname{Tr}\left(\underbrace{D_{n-1-k}}_{\substack{= \sum_{i=0}^{k} c_{k-i} A^i \\ (\text{by Lemma 3.11 } \textbf{(c)})}} \right) = \operatorname{Tr}\left(\sum_{i=0}^{k} c_{k-i} A^i \right) \\
&= \sum_{i=0}^{k} c_{k-i} \operatorname{Tr}\left(A^i\right) = \underbrace{c_{k-0}}_{=c_k} \operatorname{Tr}\left(\underbrace{A^0}_{=I_n} \right) + \sum_{i=1}^{k} c_{k-i} \operatorname{Tr}\left(A^i\right) \\
&\qquad (\text{here, we have split off the addend for } i=0 \text{ from the sum}) \\
&= c_k \underbrace{\operatorname{Tr}(I_n)}_{=n} + \sum_{i=1}^{k} \underbrace{c_{k-i} \operatorname{Tr}\left(A^i\right)}_{=\operatorname{Tr}\left(A^i\right) c_{k-i}} = c_k n + \sum_{i=1}^{k} \operatorname{Tr}\left(A^i\right) c_{k-i}.
\end{aligned}$$

Solving this equation for $\sum_{i=1}^{k} \operatorname{Tr}\left(A^i\right) c_{k-i}$, we obtain

$$\sum_{i=1}^{k} \operatorname{Tr}\left(A^i\right) c_{k-i} = \underbrace{(n-k)\, c_k}_{=nc_k - kc_k} - \underbrace{c_k n}_{=nc_k} = nc_k - kc_k - nc_k = -kc_k.$$

Adding kc_k to both sides of this equation, we obtain $kc_k + \sum_{i=1}^{k} \operatorname{Tr}\left(A^i\right) c_{k-i} = 0$. This proves Theorem 2.6. □

3.8. A corollary

The following fact (which can also be easily proven by other means) follows readily from Theorem 2.6:

Corollary 3.22. Let $n \in \mathbb{N}$. Let $A \in \mathbb{K}^{n \times n}$. Then, $\left[t^{n-1}\right] \chi_A = -\operatorname{Tr} A$.

Proof of Corollary 3.22. For every $j \in \mathbb{Z}$, define an element $c_j \in \mathbb{K}$ by $c_j = \left[t^{n-j}\right] \chi_A$. The definition of c_1 yields $c_1 = \left[t^{n-1}\right] \chi_A$. The definition of c_0 yields $c_0 = \underbrace{\left[t^{n-0}\right]}_{=[t^n]} \chi_A = [t^n] \chi_A = 1$ (by Corollary 2.4 **(c)**).

Theorem 2.6 (applied to $k = 1$) yields $1c_1 + \sum\limits_{i=1}^{1} \operatorname{Tr}\left(A^i\right) c_{1-i} = 0$. Thus,

$$1c_1 = -\underbrace{\sum_{i=1}^{1} \operatorname{Tr}\left(A^i\right) c_{1-i}}_{=\operatorname{Tr}\left(A^1\right)c_{1-1}} = -\operatorname{Tr}\left(\underbrace{A^1}_{=A}\right) \underbrace{c_{1-1}}_{=c_0=1} = -\operatorname{Tr} A.$$

Comparing this with $1c_1 = c_1 = \left[t^{n-1}\right] \chi_A$, we obtain $\left[t^{n-1}\right] \chi_A = -\operatorname{Tr} A$. This proves Corollary 3.22. □

4. Application: Nilpotency and traces

4.1. A nilpotency criterion

As an application of Theorem 2.6, let us now prove the following fact (generalizing [m.se1798703]):

Corollary 4.1. Let $n \in \mathbb{N}$. Let $A \in \mathbb{K}^{n \times n}$. Assume that

$$\operatorname{Tr}\left(A^i\right) = 0 \qquad \text{for every } i \in \{1, 2, \ldots, n\}. \tag{17}$$

(a) Then, $n! A^n = 0_{n \times n}$.
(b) If $\mathbb{K}$ is a commutative $\mathbb{Q}$-algebra, then $A^n = 0_{n \times n}$.
(c) We have $n! \chi_A = n! t^n$.
(d) If $\mathbb{K}$ is a commutative $\mathbb{Q}$-algebra, then $\chi_A = t^n$.

Proof of Corollary 4.1. For every $j \in \mathbb{Z}$, define an element $c_j \in \mathbb{K}$ by $c_j = \left[t^{n-j}\right] \chi_A$. The definition of c_0 yields $c_0 = \underbrace{\left[t^{n-0}\right]}_{=[t^n]} \chi_A = [t^n] \chi_A = 1$ (by Corollary 2.4 **(c)**).

We now claim that

$$kc_k = 0 \qquad \text{for every } k \in \{1, 2, \ldots, n\}. \tag{18}$$

[*Proof of (18):* Let $k \in \{1, 2, \ldots, n\}$. Then, every $i \in \{1, 2, \ldots, k\}$ satisfies $i \in \{1, 2, \ldots, n\}$ and therefore also

$$\operatorname{Tr}\left(A^i\right) = 0 \tag{19}$$

(by (17)). Now, Theorem 2.6 yields

$$kc_k + \sum_{i=1}^{k} \operatorname{Tr}\left(A^i\right) c_{k-i} = 0.$$

Solving this equation for kc_k, we obtain

$$kc_k = -\sum_{i=1}^{k} \underbrace{\operatorname{Tr}\left(A^i\right)}_{\substack{=0 \\ \text{(by (19))}}} c_{k-i} = \underbrace{-\sum_{i=1}^{k} 0c_{k-i}}_{=0} = -0 = 0.$$

This proves (18).]

Now, we claim that

$$n!c_k = 0 \qquad \text{for every } k \in \{1,2,\ldots,n\}. \tag{20}$$

[*Proof of (20):* Let $k \in \{1,2,\ldots,n\}$. The product $1 \cdot 2 \cdot \cdots \cdot n$ contains k as a factor, and thus is a multiple of k; in other words, $n!$ is a multiple of k (since $n! = 1 \cdot 2 \cdot \cdots \cdot n$). Hence, $n!c_k$ is a multiple of kc_k. Thus, (20) follows from (18).]

Finally, we observe that

$$n!c_{n-k} = 0 \qquad \text{for every } k \in \{0,1,\ldots,n-1\}. \tag{21}$$

[*Proof of (21):* Let $k \in \{0,1,\ldots,n-1\}$. Then, $n-k \in \{1,2,\ldots,n\}$. Hence, (20) (applied to $n-k$ instead of k) yields $n!c_{n-k} = 0$. This proves (21).]

Now, Corollary 3.2 yields $\chi_A = \sum_{k=0}^{n} c_{n-k} t^k$. Substituting A for t in this equality, we obtain $\chi_A(A) = \sum_{k=0}^{n} c_{n-k} A^k$. Multiplying both sides of the latter equality by $n!$, we obtain

$$\begin{aligned} n!\chi_A(A) &= n! \sum_{k=0}^{n} c_{n-k} A^k = \sum_{k=0}^{n} n!c_{n-k} A^k = \sum_{k=0}^{n-1} \underbrace{n!c_{n-k}}_{\substack{=0 \\ \text{(by (21))}}} A^k + n! \underbrace{c_{n-n}}_{=c_0=1} A^n \\ &\qquad \text{(here, we have split off the addend for } k=n \text{ from the sum)} \\ &= \underbrace{\sum_{k=0}^{n-1} 0A^k}_{=0} + n!A^n = n!A^n. \end{aligned}$$

Hence,

$$n!A^n = n! \underbrace{\chi_A(A)}_{\substack{=0_{n\times n} \\ \text{(by Theorem 2.5)}}} = 0_{n\times n}.$$

This proves Corollary 4.1 **(a)**.

(b) Assume that $\mathbb{K}$ is a commutative $\mathbb{Q}$-algebra. Corollary 4.1 **(a)** yields $n!A^n = 0_{n\times n}$. Now, $\dfrac{1}{n!} \in \mathbb{Q}$, so that we can multiply an $n \times n$-matrix in $\mathbb{K}^{n\times n}$ by $\dfrac{1}{n!}$ (since $\mathbb{K}$ is a $\mathbb{Q}$-algebra). We have $\underbrace{\dfrac{1}{n!}n!}_{=1} A^n = A^n$. Hence, $A^n = \dfrac{1}{n!} \underbrace{n!A^n}_{=0_{n\times n}} = \dfrac{1}{n!}0_{n\times n} = 0_{n\times n}$. This proves Corollary 4.1 **(b)**.

(c) Multiplying the equality $\chi_A = \sum\limits_{k=0}^{n} c_{n-k}t^k$ by $n!$, we obtain

$$n!\chi_A = n! \sum_{k=0}^{n} c_{n-k}t^k = \sum_{k=0}^{n} n!c_{n-k}t^k = \sum_{k=0}^{n-1} \underbrace{n!c_{n-k}}_{\substack{=0 \\ \text{(by (21))}}} t^k + n! \underbrace{c_{n-n}}_{=c_0=1} t^n$$

$$\left(\text{here, we have split off the addend for } k = n \text{ from the sum}\right)$$

$$= \underbrace{\sum_{k=0}^{n-1} 0t^k}_{=0} + n!t^n = n!t^n.$$

This proves Corollary 4.1 **(c)**.

(d) Assume that $\mathbb{K}$ is a commutative $\mathbb{Q}$-algebra. Corollary 4.1 **(c)** yields $n!\chi_A = n!t^n$.

Now, $\dfrac{1}{n!} \in \mathbb{Q}$, so that we can multiply any polynomial in $\mathbb{K}[t]$ by $\dfrac{1}{n!}$ (since $\mathbb{K}$ is a $\mathbb{Q}$-algebra). We have $\underbrace{\dfrac{1}{n!}n!}_{=1} \chi_A = \chi_A$. Hence, $\chi_A = \dfrac{1}{n!} \underbrace{n!\chi_A}_{=n!t^n} = \dfrac{1}{n!}n!t^n = t^n$. This proves Corollary 4.1 **(d)**. □

4.2. A converse direction

The following result – in a sense, a converse of Corollary 4.1 **(d)** – also follows from Theorem 2.6:

Corollary 4.2. Let $n \in \mathbb{N}$. Let $A \in \mathbb{K}^{n\times n}$. Assume that $\chi_A = t^n$. Then, $\operatorname{Tr}\left(A^i\right) = 0$ for every positive integer i.

Proof of Corollary 4.2. For every $j \in \mathbb{Z}$, define an element $c_j \in \mathbb{K}$ by $c_j = \left[t^{n-j}\right] \chi_A$. Then, each positive integer j satisfies

$$c_j = 0. \tag{22}$$

[*Proof of (22):* Let j be a positive integer. Thus, $j \neq 0$, so that $n - j \neq n$ and thus $\left[t^{n-j}\right]\left(t^n\right) = 0$. In view of $\chi_A = t^n$, this rewrites as $\left[t^{n-j}\right] \chi_A = 0$. But the definition of c_j yields $c_j = \left[t^{n-j}\right] \chi_A = 0$. This proves (22).]

The definition of c_0 yields $c_0 = \underbrace{\left[t^{n-0}\right]}_{=[t^n]} \chi_A = [t^n] \chi_A = 1$ (by Corollary 2.4 **(c)**).

Now, we claim that

$$\operatorname{Tr}(A^p) = 0 \qquad \text{for every positive integer } p. \tag{23}$$

[*Proof of (23):* We shall prove (23) by strong induction on p:

Induction step: Fix a positive integer k. Assume (as the induction hypothesis) that (23) holds whenever $p < k$. We must now prove that (23) holds for $p = k$.

From (22) (applied to $j = k$), we obtain $c_k = 0$.

We have assumed that (23) holds whenever $p < k$. In other words,

$$\operatorname{Tr}(A^p) = 0 \qquad \text{for every positive integer } p < k. \tag{24}$$

Now, Theorem 2.6 yields

$$kc_k + \sum_{i=1}^{k} \operatorname{Tr}\left(A^i\right) c_{k-i} = 0.$$

Hence,

$$0 = k \underbrace{c_k}_{=0} + \sum_{i=1}^{k} \operatorname{Tr}\left(A^i\right) c_{k-i} = \sum_{i=1}^{k} \operatorname{Tr}\left(A^i\right) c_{k-i} = \sum_{i=1}^{k-1} \underbrace{\operatorname{Tr}\left(A^i\right)}_{\substack{=0 \\ \text{(by (24),} \\ \text{applied} \\ \text{to } p=i)}} c_{k-i} + \operatorname{Tr}\left(A^k\right) \underbrace{c_{k-k}}_{=c_0=1}$$

$$\text{(here, we have split off the addend for } i = k \text{ from the sum)}$$

$$= \underbrace{\sum_{i=1}^{k-1} 0 c_{k-i}}_{=0} + \operatorname{Tr}\left(A^k\right) = \operatorname{Tr}\left(A^k\right).$$

Thus, $\operatorname{Tr}\left(A^k\right) = 0$. In other words, (23) holds for $p = k$. This completes the induction step. Thus, (23) is proven by strong induction.]

We have thus proven that $\operatorname{Tr}(A^p) = 0$ for every positive integer p. Renaming the variable p as i in this statement, we conclude that $\operatorname{Tr}\left(A^i\right) = 0$ for every positive integer i. This proves Corollary 4.2. □

5. More on the adjugate

I shall now discuss various other properties of the adjugate $\operatorname{adj} A$ of a square matrix A.

5.1. Functoriality

For any $n \in \mathbb{N}$ and $m \in \mathbb{N}$, a homomorphism $f : \mathbb{L} \to \mathbb{M}$ between two rings $\mathbb{L}$ and $\mathbb{M}$ gives rise to a map $f^{n\times m} : \mathbb{L}^{n\times m} \to \mathbb{M}^{n\times m}$ (as defined in Definition 3.15). We recall some classical properties of these maps $f^{n\times m}$:

Proposition 5.1. Let $\mathbb{L}$ and $\mathbb{M}$ be two commutative rings. Let $f : \mathbb{L} \to \mathbb{M}$ be a ring homomorphism.

(a) For every $n \in \mathbb{N}$ and $m \in \mathbb{N}$, the map $f^{n\times m} : \mathbb{L}^{n\times m} \to \mathbb{M}^{n\times m}$ is a homomorphism of additive groups.

(b) Every $n \in \mathbb{N}$ satisfies $f^{n\times n}(I_n) = I_n$.

(c) For every $n \in \mathbb{N}$, $m \in \mathbb{N}$, $p \in \mathbb{N}$, $A \in \mathbb{L}^{n\times m}$ and $B \in \mathbb{L}^{m\times p}$, we have $f^{n\times p}(AB) = f^{n\times m}(A) \cdot f^{m\times p}(B)$.

(d) For every $n \in \mathbb{N}$ and $m \in \mathbb{N}$ and every $A \in \mathbb{L}^{n\times m}$ and $\lambda \in \mathbb{L}$, we have $f^{n\times m}(\lambda A) = f(\lambda) f^{n\times m}(A)$.

Now, let me state the classical (and simple) fact which is often (somewhat incompletely) subsumed under the slogan "ring homomorphisms preserve determinants and adjugates":

Proposition 5.2. Let $\mathbb{L}$ and $\mathbb{M}$ be two commutative rings. Let $f : \mathbb{L} \to \mathbb{M}$ be a ring homomorphism. Let $n \in \mathbb{N}$. Let $A \in \mathbb{L}^{n\times n}$.

(a) We have $f(\det A) = \det(f^{n\times n}(A))$.

(b) Any two elements u and v of $\{1, 2, \ldots, n\}$ satisfy $f^{(n-1)\times(n-1)}(A_{\sim u,\sim v}) = (f^{n\times n}(A))_{\sim u,\sim v}$.

(c) We have $f^{n\times n}(\operatorname{adj} A) = \operatorname{adj}(f^{n\times n}(A))$.

Proof of Proposition 5.2. Proving Proposition 5.2 is completely straightforward, and left to the reader. □

5.2. The evaluation homomorphism

We shall apply the above to relate the determinant and the adjugate of a matrix A with those of the matrix $tI_n + A$:

Proposition 5.3. Let $\varepsilon : \mathbb{K}[t] \to \mathbb{K}$ be the map which sends every polynomial $p \in \mathbb{K}[t]$ to its value $p(0)$. It is well-known that ε is a $\mathbb{K}$-algebra homomorphism.

Let $n \in \mathbb{N}$. Let $A \in \mathbb{K}^{n\times n}$. Consider the matrix $tI_n + A \in (\mathbb{K}[t])^{n\times n}$. Then:

(a) We have $\varepsilon(\det(tI_n + A)) = \det A$.

(b) We have $\varepsilon^{n\times n}(\operatorname{adj}(tI_n + A)) = \operatorname{adj} A$.

(c) We have $\varepsilon^{n\times n}(tI_n + A) = A$.

Proof of Proposition 5.3. We have

$$\varepsilon^{n\times n}(tB + A) = A$$

for every $B \in \mathbb{K}^{n \times n}$ [9]. Applying this to $B = I_n$, we obtain $\varepsilon^{n \times n} (tI_n + A) = A$. This proves Proposition 5.3 **(c)**.

(a) Proposition 5.2 **(a)** (applied to $\mathbb{K}[t]$, $\mathbb{K}$, ε and $tI_n + A$ instead of $\mathbb{L}$, $\mathbb{M}$, f and A) yields

$$\varepsilon (\det (tI_n + A)) = \det \left(\underbrace{\varepsilon^{n \times n} (tI_n + A)}_{=A} \right) = \det A.$$

This proves Proposition 5.3 **(a)**.

(b) Proposition 5.2 **(c)** (applied to $\mathbb{K}[t]$, $\mathbb{K}$, ε and $tI_n + A$ instead of $\mathbb{L}$, $\mathbb{M}$, f and A) yields

$$\varepsilon^{n \times n} (\operatorname{adj} (tI_n + A)) = \operatorname{adj} \left(\underbrace{\varepsilon^{n \times n} (tI_n + A)}_{=A} \right) = \operatorname{adj} A.$$

This proves Proposition 5.3 **(b)**. □

If $A \in \mathbb{K}^{n \times n}$ is a square matrix, then the matrix $tI_n + A \in (\mathbb{K}[t])^{n \times n}$ has a property which the matrix A might not have: namely, its determinant is regular. Let us first define what this means:

Definition 5.4. Let $\mathbb{A}$ be a commutative ring. Let $a \in \mathbb{A}$. The element a of $\mathbb{A}$ is said to be *regular* if and only if every $x \in \mathbb{A}$ satisfying $ax = 0$ satisfies $x = 0$.

Instead of saying that a is regular, one can also say that "a is *cancellable*", or that "a is a *non-zero-divisor*".

A basic property of regular elements is the following:

Lemma 5.5. Let $\mathbb{A}$ be a commutative ring. Let a be a regular element of $\mathbb{A}$. Let b and c be two elements of $\mathbb{A}$ such that $ab = ac$. Then, $b = c$.

Proof of Lemma 5.5. We have $a(b - c) = \underbrace{ab}_{=ac} - ac = ac - ac = 0$.

Now, recall that the element a of $\mathbb{A}$ is regular if and only if every $x \in \mathbb{A}$ satisfying $ax = 0$ satisfies $x = 0$ (by the definition of "regular"). Hence, every $x \in \mathbb{A}$ satisfying $ax = 0$ satisfies $x = 0$ (because the element a of $\mathbb{A}$ is regular). Applying this to $x = b - c$, we obtain $b - c = 0$ (since $a(b - c) = 0$). Thus, $b = c$. This proves Lemma 5.5. □

Regular elements, of course, can also be cancelled from matrix equations:

Lemma 5.6. Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. Let a be a regular element of $\mathbb{K}$. Let $B \in \mathbb{K}^{n \times m}$ and $C \in \mathbb{K}^{n \times m}$ be such that $aB = aC$. Then, $B = C$.

[9] *Proof.* This equality is similar to (14), and is proven analogously.

Proof of Lemma 5.6. Write the $n \times m$-matrices B and C in the forms $B = \left(b_{i,j}\right)_{1 \leq i \leq n,\ 1 \leq j \leq m}$ and $C = \left(c_{i,j}\right)_{1 \leq i \leq n,\ 1 \leq j \leq m}$. Then, $aB = \left(ab_{i,j}\right)_{1 \leq i \leq n,\ 1 \leq j \leq m}$ and $aC = \left(ac_{i,j}\right)_{1 \leq i \leq n,\ 1 \leq j \leq m}$. Hence,

$$\left(ab_{i,j}\right)_{1 \leq i \leq n,\ 1 \leq j \leq m} = aB = aC = \left(ac_{i,j}\right)_{1 \leq i \leq n,\ 1 \leq j \leq m}.$$

In other words,

$$ab_{i,j} = ac_{i,j} \qquad \text{for every } (i,j) \in \{1,2,\ldots,n\} \times \{1,2,\ldots,m\}.$$

Thus,

$$b_{i,j} = c_{i,j} \qquad \text{for every } (i,j) \in \{1,2,\ldots,n\} \times \{1,2,\ldots,m\}$$

(by Lemma 5.5, applied to $b = b_{i,j}$ and $c = c_{i,j}$). Hence, $\left(b_{i,j}\right)_{1 \leq i \leq n,\ 1 \leq j \leq m} = \left(c_{i,j}\right)_{1 \leq i \leq n,\ 1 \leq j \leq m}$. Thus, $B = \left(b_{i,j}\right)_{1 \leq i \leq n,\ 1 \leq j \leq m} = \left(c_{i,j}\right)_{1 \leq i \leq n,\ 1 \leq j \leq m} = C$. Lemma 5.6 is proven. □

One important way to construct regular elements is the following fact:

Proposition 5.7. Let $n \in \mathbb{N}$. Let $p \in \mathbb{K}[t]$ be a monic polynomial of degree n. Then, the element p of $\mathbb{K}[t]$ is regular.

Proof of Proposition 5.7. Proposition 5.7 is precisely [Grinbe16b, Corollary 3.15]. □

Corollary 5.8. Let $n \in \mathbb{N}$. Let $A \in \mathbb{K}^{n \times n}$. Consider the matrix $tI_n + A \in (\mathbb{K}[t])^{n \times n}$.

Then, the element $\det(tI_n + A)$ of $\mathbb{K}[t]$ is regular.

Proof of Corollary 5.8. Proposition 2.2 **(a)** (applied to I_n and A instead of A and B) yields that $\det(tI_n + A) \in \mathbb{K}[t]$ is a polynomial of degree $\leq n$ in t. Proposition 2.2 **(c)** (applied to I_n and A instead of A and B) yields that $[t^n](\det(tI_n + A)) = \det(I_n) = 1$.

So we know that the polynomial $\det(tI_n + A) \in \mathbb{K}[t]$ is a polynomial of degree $\leq n$, and that the coefficient of t^n in this polynomial is $[t^n](\det(tI_n + A)) = 1$. In other words, the polynomial $\det(tI_n + A) \in \mathbb{K}[t]$ is monic of degree n. Thus, Proposition 5.7 (applied to $p = \det(tI_n + A)$) shows that the element $\det(tI_n + A)$ of $\mathbb{K}[t]$ is regular. This proves Corollary 5.8. □

A square matrix whose determinant is regular can be cancelled from equations, as the following lemma shows:

Lemma 5.9. Let $n \in \mathbb{N}$. Let $A \in \mathbb{K}^{n \times n}$. Assume that the element $\det A$ of $\mathbb{K}$ is regular. Let $m \in \mathbb{N}$.

(a) If $B \in \mathbb{K}^{n \times m}$ and $C \in \mathbb{K}^{n \times m}$ are such that $AB = AC$, then $B = C$.

(b) If $B \in \mathbb{K}^{m \times n}$ and $C \in \mathbb{K}^{m \times n}$ are such that $BA = CA$, then $B = C$.

Proof of Lemma 5.9. Define an element a of $\mathbb{K}$ by $a = \det A$. Recall that the element $\det A$ of $\mathbb{K}$ is regular. In other words, the element a of $\mathbb{K}$ is regular (since $a = \det A$). Theorem 3.7 yields $A \cdot \operatorname{adj} A = \operatorname{adj} A \cdot A = \det A \cdot I_n$.

(a) Let $B \in \mathbb{K}^{n \times m}$ and $C \in \mathbb{K}^{n \times m}$ be such that $AB = AC$. We must prove that $B = C$.

We have

$$\underbrace{\operatorname{adj} A \cdot A}_{= \det A \cdot I_n} B = \underbrace{\det A}_{= a} \cdot \underbrace{I_n B}_{= B} = aB.$$

Thus,

$$aB = \operatorname{adj} A \cdot \underbrace{AB}_{= AC} = \underbrace{\operatorname{adj} A \cdot A}_{= \det A \cdot I_n} C = \underbrace{\det A}_{= a} \cdot \underbrace{I_n C}_{= C} = aC.$$

Lemma 5.6 thus yields $B = C$. This proves Lemma 5.9 **(a)**.

(b) The proof of Lemma 5.9 **(b)** is similar to the proof of Lemma 5.9 **(a)** (but now we need to work with $BA \cdot \operatorname{adj} A$ and $CA \cdot \operatorname{adj} A$ instead of $\operatorname{adj} A \cdot AB$ and $\operatorname{adj} A \cdot AC$). The details are left to the reader. □

5.3. The adjugate of a product

Corollary 5.8 can be put to use in several circumstances. Here is a simple example:

Theorem 5.10. Let $n \in \mathbb{N}$. Let A and B be two $n \times n$-matrices. Then,

$$\operatorname{adj}(AB) = \operatorname{adj} B \cdot \operatorname{adj} A.$$

Theorem 5.10 is the statement of [Grinbe15, Exercise 6.33]; see [Grinbe15, solution of Exercise 6.33] for a proof of this theorem. We shall show a different proof of it now.

We begin by showing a particular case of Theorem 5.10:

Lemma 5.11. Let $n \in \mathbb{N}$. Let A and B be two $n \times n$-matrices. Assume that the elements $\det A$ and $\det B$ of $\mathbb{K}$ are regular. Then, $\operatorname{adj}(AB) = \operatorname{adj} B \cdot \operatorname{adj} A$.

Proof of Lemma 5.11. Theorem 3.7 yields

$$A \cdot \operatorname{adj} A = \operatorname{adj} A \cdot A = \det A \cdot I_n.$$

Theorem 3.7 (applied to B instead of A) yields

$$B \cdot \operatorname{adj} B = \operatorname{adj} B \cdot B = \det B \cdot I_n.$$

Theorem 3.7 (applied to AB instead of A) yields

$$AB \cdot \operatorname{adj}(AB) = \operatorname{adj}(AB) \cdot AB = \det(AB) \cdot I_n.$$

Now,

$$\begin{aligned} A\underbrace{B\cdot\operatorname{adj}B}_{=\det B\cdot I_n}\cdot\operatorname{adj}A &= \underbrace{A\cdot\det B\cdot I_n}_{=\det B\cdot A}\cdot\operatorname{adj}A \\ &= \det B\cdot\underbrace{A\cdot\operatorname{adj}A}_{=\det A\cdot I_n} = \det B\cdot\det A\cdot I_n \\ &= \det A\cdot\det B\cdot I_n. \end{aligned}$$

Comparing this with

$$AB\cdot\operatorname{adj}(AB) = \underbrace{\det(AB)}_{=\det A\cdot\det B}\cdot I_n = \det A\cdot\det B\cdot I_n,$$

we obtain $AB\cdot\operatorname{adj}B\cdot\operatorname{adj}A = AB\cdot\operatorname{adj}(AB)$. Lemma 5.9 **(a)** (applied to n, $B\cdot\operatorname{adj}B\cdot\operatorname{adj}A$ and $B\cdot\operatorname{adj}(AB)$ instead of m, B and C) therefore yields $B\cdot\operatorname{adj}B\cdot\operatorname{adj}A = B\cdot\operatorname{adj}(AB)$ (since the element $\det A$ of $\mathbb{K}$ is regular). Thus, Lemma 5.9 **(a)** (applied to n, B, $\operatorname{adj}B\cdot\operatorname{adj}A$ and $\operatorname{adj}(AB)$ instead of m, A, B and C) yields $\operatorname{adj}B\cdot\operatorname{adj}A = \operatorname{adj}(AB)$ (since the element $\det B$ of $\mathbb{K}$ is regular). This proves Lemma 5.11. □

We now derive Theorem 5.10 from this lemma:

Proof of Theorem 5.10. Define the $\mathbb{K}$-algebra homomorphism $\varepsilon:\mathbb{K}[t]\to\mathbb{K}$ as in Proposition 5.3.

Define two matrices $\widetilde{A}$ and $\widetilde{B}$ in $(\mathbb{K}[t])^{n\times n}$ by $\widetilde{A} = tI_n + A$ and $\widetilde{B} = tI_n + B$.

From $\widetilde{A} = tI_n + A$, we obtain $\varepsilon^{n\times n}\left(\operatorname{adj}\widetilde{A}\right) = \varepsilon^{n\times n}(\operatorname{adj}(tI_n + A)) = \operatorname{adj}A$ (by Proposition 5.3 **(b)**). Similarly, $\varepsilon^{n\times n}\left(\operatorname{adj}\widetilde{B}\right) = \operatorname{adj}B$.

From $\widetilde{A} = tI_n + A$, we obtain $\varepsilon^{n\times n}\left(\widetilde{A}\right) = \varepsilon^{n\times n}(tI_n + A) = A$ (by Proposition 5.3 **(c)**). Similarly, $\varepsilon^{n\times n}\left(\widetilde{B}\right) = B$.

Corollary 5.8 shows that the element $\det(tI_n + A)$ of $\mathbb{K}[t]$ is regular. In other words, the element $\det\widetilde{A}$ of $\mathbb{K}[t]$ is regular (since $\widetilde{A} = tI_n + A$). Similarly, the element $\det\widetilde{B}$ of $\mathbb{K}[t]$ is regular. Lemma 5.11 (applied to $\mathbb{K}[t]$, $\widetilde{A}$ and $\widetilde{B}$ instead of $\mathbb{K}$, A and B) thus yields

$$\operatorname{adj}\left(\widetilde{A}\widetilde{B}\right) = \operatorname{adj}\widetilde{B}\cdot\operatorname{adj}\widetilde{A}.$$

Applying the map $\varepsilon^{n\times n}$ to both sides of this equality, we obtain

$$\begin{aligned} \varepsilon^{n\times n}\left(\operatorname{adj}\left(\widetilde{A}\widetilde{B}\right)\right) &= \varepsilon^{n\times n}\left(\operatorname{adj}\widetilde{B}\cdot\operatorname{adj}\widetilde{A}\right) = \underbrace{\varepsilon^{n\times n}\left(\operatorname{adj}\widetilde{B}\right)}_{=\operatorname{adj}B}\cdot\underbrace{\varepsilon^{n\times n}\left(\operatorname{adj}\widetilde{A}\right)}_{=\operatorname{adj}A} \\ &\qquad \left(\begin{array}{c}\text{by Proposition 5.1 } \textbf{(c)}\text{, applied to} \\ \mathbb{K}[t],\ \mathbb{K},\ \varepsilon,\ n,\ n,\ \operatorname{adj}\widetilde{B} \text{ and } \operatorname{adj}\widetilde{A} \\ \text{instead of } \mathbb{L},\ \mathbb{M},\ f,\ m,\ p,\ A \text{ and } B\end{array}\right) \\ &= \operatorname{adj}B\cdot\operatorname{adj}A. \end{aligned}$$

Hence,

$$\operatorname{adj} B \cdot \operatorname{adj} A = \varepsilon^{n\times n}\left(\operatorname{adj}\left(\widetilde{A}\widetilde{B}\right)\right) = \operatorname{adj}\left(\varepsilon^{n\times n}\left(\widetilde{A}\widetilde{B}\right)\right) \tag{25}$$

(by Proposition 5.2 **(c)**, applied to $\mathbb{K}[t]$, $\mathbb{K}$, ε and $\widetilde{A}\widetilde{B}$ instead of $\mathbb{L}$, $\mathbb{M}$, f and A).

But Proposition 5.1 **(c)** (applied to $\mathbb{K}[t]$, $\mathbb{K}$, ε, n, n, $\widetilde{A}$ and $\widetilde{B}$ instead of $\mathbb{L}$, $\mathbb{M}$, f, m, p, A and B) shows that

$$\varepsilon^{n\times n}\left(\widetilde{A}\widetilde{B}\right) = \underbrace{\varepsilon^{n\times n}\left(\widetilde{A}\right)}_{=A} \cdot \underbrace{\varepsilon^{n\times n}\left(\widetilde{B}\right)}_{=B} = AB.$$

Hence, (25) becomes

$$\operatorname{adj} B \cdot \operatorname{adj} A = \operatorname{adj}\left(\underbrace{\varepsilon^{n\times n}\left(\widetilde{A}\widetilde{B}\right)}_{=AB}\right) = \operatorname{adj}(AB).$$

This proves Theorem 5.10. □

5.4. Determinant and adjugate of an adjugate

Our next target is the following result:

Theorem 5.12. Let $n \in \mathbb{N}$. Let A be an $n \times n$-matrix.

(a) If $n \geq 1$, then $\det(\operatorname{adj} A) = (\det A)^{n-1}$.

(b) If $n \geq 2$, then $\operatorname{adj}(\operatorname{adj} A) = (\det A)^{n-2} A$.

Again, we shall first prove it in a particular case:

Lemma 5.13. Let $n \in \mathbb{N}$. Let A be an $n \times n$-matrix. Assume that the element $\det A$ of $\mathbb{K}$ is regular.

(a) If $n \geq 1$, then $\det(\operatorname{adj} A) = (\det A)^{n-1}$.

(b) If $n \geq 2$, then $\operatorname{adj}(\operatorname{adj} A) = (\det A)^{n-2} A$.

Before we start proving Lemma 5.13, let us first recall the following fact: If $n \in \mathbb{N}$, $\lambda \in \mathbb{K}$ and $C \in \mathbb{K}^{n\times n}$, then

$$\det(\lambda C) = \lambda^n \det C. \tag{26}$$

(In fact, this is precisely [Grinbe15, Proposition 6.12] (applied to C instead of A).)

Proof of Lemma 5.13. Theorem 3.7 yields

$$A \cdot \operatorname{adj} A = \operatorname{adj} A \cdot A = \det A \cdot I_n.$$

(a) Assume that $n \geq 1$. Now,

$$\det\left(\underbrace{A \cdot \operatorname{adj} A}_{=\det A \cdot I_n}\right) = \det\left(\det A \cdot I_n\right) = \left(\det A\right)^n \underbrace{\det\left(I_n\right)}_{=1}$$
$$\text{(by (26) (applied to } \det A \text{ and } I_n \text{ instead of } \lambda \text{ and } C))$$
$$= \left(\det A\right)^n = \det A \cdot \left(\det A\right)^{n-1}.$$

Thus,

$$\det A \cdot \left(\det A\right)^{n-1} = \det\left(A \cdot \operatorname{adj} A\right) = \det A \cdot \det\left(\operatorname{adj} A\right).$$

Hence, Lemma 5.5 (applied to $\mathbb{A} = \mathbb{K}$, $a = \det A$, $b = \left(\det A\right)^{n-1}$ and $c = \det\left(\operatorname{adj} A\right)$) yields $\left(\det A\right)^{n-1} = \det\left(\operatorname{adj} A\right)$ (since $\det A$ is a regular element of $\mathbb{K}$). This proves Lemma 5.13 **(a)**.

(b) Assume that $n \geq 2$. Thus, $n - 1 \geq 1$ and $n \geq 2 \geq 1$. Now, Lemma 5.13 **(a)** yields

$$\det\left(\operatorname{adj} A\right) = \left(\det A\right)^{n-1} = \det A \cdot \underbrace{\left(\det A\right)^{(n-1)-1}}_{=\left(\det A\right)^{n-2}} \qquad \text{(since } n - 1 \geq 1\text{)}$$
$$= \det A \cdot \left(\det A\right)^{n-2}.$$

But Theorem 3.7 (applied to $\operatorname{adj} A$ instead of A) yields

$$\operatorname{adj} A \cdot \operatorname{adj}\left(\operatorname{adj} A\right) = \operatorname{adj}\left(\operatorname{adj} A\right) \cdot \operatorname{adj} A = \det\left(\operatorname{adj} A\right) \cdot I_n.$$

Now,

$$A \cdot \underbrace{\operatorname{adj} A \cdot \operatorname{adj}\left(\operatorname{adj} A\right)}_{=\det(\operatorname{adj} A) \cdot I_n} = A \cdot \det\left(\operatorname{adj} A\right) \cdot I_n = \underbrace{\det\left(\operatorname{adj} A\right)}_{=\det A \cdot (\det A)^{n-2}} A$$
$$= \det A \cdot \left(\det A\right)^{n-2} A.$$

Hence,

$$\det A \cdot \left(\det A\right)^{n-2} A = \underbrace{A \cdot \operatorname{adj} A}_{=\det A \cdot I_n} \cdot \operatorname{adj}\left(\operatorname{adj} A\right) = \det A \cdot I_n \cdot \operatorname{adj}\left(\operatorname{adj} A\right)$$
$$= \det A \cdot \operatorname{adj}\left(\operatorname{adj} A\right).$$

Hence, Lemma 5.6 (applied to n, $\det A$, $\left(\det A\right)^{n-2} A$ and $\operatorname{adj}\left(\operatorname{adj} A\right)$ instead of m, a, B and C) yields $\left(\det A\right)^{n-2} A = \operatorname{adj}\left(\operatorname{adj} A\right)$ (since $\det A$ is a regular element of $\mathbb{K}$). This proves Lemma 5.13 **(b)**. $\square$

Let us now derive Theorem 5.12 from this lemma:

Proof of Theorem 5.12. Define the $\mathbb{K}$-algebra homomorphism $\varepsilon : \mathbb{K}[t] \to \mathbb{K}$ as in Proposition 5.3.

Define a matrix $\widetilde{A} \in (\mathbb{K}[t])^{n\times n}$ by $\widetilde{A} = tI_n + A$. Corollary 5.8 shows that the element $\det(tI_n + A)$ of $\mathbb{K}[t]$ is regular. In other words, the element $\det \widetilde{A}$ of $\mathbb{K}[t]$ is regular (since $\widetilde{A} = tI_n + A$).

We have $\varepsilon^{n\times n}\left(\operatorname{adj} \underbrace{\widetilde{A}}_{=tI_n+A}\right) = \varepsilon^{n\times n}(\operatorname{adj}(tI_n + A)) = \operatorname{adj} A$ (by Proposition 5.3 **(b)**). Also, $\varepsilon\left(\det \underbrace{\widetilde{A}}_{=tI_n+A}\right) = \varepsilon(\det(tI_n + A)) = \det A$ (by Proposition 5.3 **(a)**).

(a) Assume that $n \geq 1$. Lemma 5.13 **(a)** (applied to $\mathbb{K}[t]$ and $\widetilde{A}$ instead of $\mathbb{K}$ and A) yields $\det\left(\operatorname{adj} \widetilde{A}\right) = \left(\det \widetilde{A}\right)^{n-1}$.

Now, Proposition 5.2 **(a)** (applied to $\mathbb{K}[t]$, $\mathbb{K}$, ε and $\operatorname{adj} \widetilde{A}$ instead of $\mathbb{L}$, $\mathbb{M}$, f and A) yields

$$\varepsilon\left(\det\left(\operatorname{adj} \widetilde{A}\right)\right) = \det\left(\underbrace{\varepsilon^{n\times n}\left(\operatorname{adj} \widetilde{A}\right)}_{=\operatorname{adj} A}\right) = \det(\operatorname{adj} A).$$

Hence,

$$\begin{aligned}
\det(\operatorname{adj} A) &= \varepsilon\left(\underbrace{\det\left(\operatorname{adj} \widetilde{A}\right)}_{=\left(\det \widetilde{A}\right)^{n-1}}\right) = \varepsilon\left(\left(\det \widetilde{A}\right)^{n-1}\right) \\
&= \left(\underbrace{\varepsilon\left(\det \widetilde{A}\right)}_{=\det A}\right)^{n-1} \qquad \text{(since } \varepsilon \text{ is a } \mathbb{K}\text{-algebra homomorphism)} \\
&= (\det A)^{n-1}.
\end{aligned}$$

This proves Theorem 5.12 **(a)**.

(b) Assume that $n \geq 2$. Lemma 5.13 **(b)** (applied to $\mathbb{K}[t]$ and $\widetilde{A}$ instead of $\mathbb{K}$ and A) yields $\operatorname{adj}\left(\operatorname{adj} \widetilde{A}\right) = \left(\det \widetilde{A}\right)^{n-2} \widetilde{A}$. We have $\varepsilon^{n\times n}\left(\underbrace{\widetilde{A}}_{=tI_n+A}\right) = \varepsilon^{n\times n}(tI_n + A) = A$ (by Proposition 5.3 **(c)**). Proposition 5.2 **(c)** (applied to $\mathbb{K}[t]$, $\mathbb{K}$, ε and $\operatorname{adj} \widetilde{A}$

instead of $\mathbb{L}$, $\mathbb{M}$, f and A) yields

$$\varepsilon^{n\times n}\left(\operatorname{adj}\left(\operatorname{adj}\widetilde{A}\right)\right)=\operatorname{adj}\left(\underbrace{\varepsilon^{n\times n}\left(\operatorname{adj}\widetilde{A}\right)}_{=\operatorname{adj}A}\right)=\operatorname{adj}\left(\operatorname{adj}A\right).$$

Thus,

$$\begin{aligned}
\operatorname{adj}\left(\operatorname{adj}A\right)&=\varepsilon^{n\times n}\left(\underbrace{\operatorname{adj}\left(\operatorname{adj}\widetilde{A}\right)}_{=\left(\det\widetilde{A}\right)^{n-2}\widetilde{A}}\right)=\varepsilon^{n\times n}\left(\left(\det\widetilde{A}\right)^{n-2}\widetilde{A}\right)\\
&=\underbrace{\varepsilon\left(\left(\det\widetilde{A}\right)^{n-2}\right)}_{\substack{=\left(\varepsilon\left(\det\widetilde{A}\right)\right)^{n-2}\\ \text{(since }\varepsilon\text{ is a }\mathbb{K}\text{-algebra}\\ \text{homomorphism)}}}\underbrace{\varepsilon^{n\times n}\left(\widetilde{A}\right)}_{=A}\\
&\qquad\left(\begin{array}{c}\text{by Proposition 5.1 (applied to }\mathbb{K}\left[t\right],\mathbb{K},\\ \varepsilon, n, \widetilde{A}\text{ and }\left(\det\widetilde{A}\right)^{n-2}\text{ instead of }\mathbb{L}, \mathbb{M}, f, m, A\text{ and }\lambda)\end{array}\right)\\
&=\left(\underbrace{\varepsilon\left(\det\widetilde{A}\right)}_{=\det A}\right)^{n-2}A=\left(\det A\right)^{n-2}A.
\end{aligned}$$

This proves Theorem 5.12 **(b)**. □

5.5. The adjugate of A as a polynomial in A

Next, let us show that the adjugate of a square matrix A is a polynomial in A (with coefficients that depend on A, but are scalars – not matrices):

Theorem 5.14. Let $n\in\mathbb{N}$. Let $A\in\mathbb{K}^{n\times n}$. For every $j\in\mathbb{Z}$, define an element $c_j\in\mathbb{K}$ by $c_j=\left[t^{n-j}\right]\chi_A$. Then,

$$\operatorname{adj}A=\left(-1\right)^{n-1}\sum_{i=0}^{n-1}c_{n-1-i}A^i.$$

One consequence of Theorem 5.14 is that every $n\times n$-matrix which commutes with a given $n\times n$-matrix A must also commute with $\operatorname{adj}A$.

We prepare for the proof of Theorem 5.14 with two really simple facts:

Lemma 5.15. Let $n \in \mathbb{N}$. Let u and v be two elements of $\{1, 2, \ldots, n\}$. Let $\lambda \in \mathbb{K}$. Let A be an $n \times n$-matrix. Then,

$$(\lambda A)_{\sim u, \sim v} = \lambda A_{\sim u, \sim v}.$$

Proof of Lemma 5.15. This follows from Lemma 3.8 (applied to $\mu = 0$ and $B = A$). □

Proposition 5.16. Let n be a positive integer. Let $A \in \mathbb{K}^{n \times n}$ and $\lambda \in \mathbb{K}$. Then, $\operatorname{adj}(\lambda A) = \lambda^{n-1} \operatorname{adj} A$.

Proof of Proposition 5.16. Recalling the definitions of $\operatorname{adj}(\lambda A)$ and $\operatorname{adj} A$ (and using Lemma 5.15), the reader can easily reduce Proposition 5.16 to (26) (applied to $n-1$ and $A_{\sim j, \sim i}$ instead of n and C). □

Now, let me show a slightly simpler variant of Theorem 5.14:

Lemma 5.17. Let n be a positive integer. Let $A \in \mathbb{K}^{n \times n}$. For every $j \in \mathbb{Z}$, define an element $c_j \in \mathbb{K}$ by $c_j = \left[t^{n-j}\right] \chi_A$. Then,

$$\operatorname{adj}(-A) = \sum_{i=0}^{n-1} c_{n-1-i} A^i.$$

Proof of Lemma 5.17. For every $j \in \mathbb{Z}$, define an element $c_j \in \mathbb{K}$ by $c_j = \left[t^{n-j}\right] \chi_A$.

Proposition 3.9 shows that there exist n matrices $D_0, D_1, \ldots, D_{n-1}$ in $\mathbb{K}^{n \times n}$ such that

$$\operatorname{adj}(tI_n - A) = \sum_{k=0}^{n-1} t^k D_k \qquad \text{in } (\mathbb{K}[t])^{n \times n}. \tag{27}$$

Consider these $D_0, D_1, \ldots, D_{n-1}$. Thus, an n-tuple $(D_0, D_1, \ldots, D_{n-1})$ of matrices in $\mathbb{K}^{n \times n}$ is defined. Extend this n-tuple to a family $(D_k)_{k \in \mathbb{Z}}$ of matrices in $\mathbb{K}^{n \times n}$ by setting (7). Lemma 3.11 **(c)** (applied to $k = n-1$) yields

$$\sum_{i=0}^{n-1} c_{n-1-i} A^i = D_{n-1-(n-1)} = D_0. \tag{28}$$

On the other hand, define the $\mathbb{K}$-algebra homomorphism $\varepsilon : \mathbb{K}[t] \to \mathbb{K}$ as in Proposition 5.3. This homomorphism ε satisfies $\varepsilon(t) = 0$. Also, it satisfies $\varepsilon(u) = u$ for every $u \in \mathbb{K}$. Hence, the map $\varepsilon^{n \times n} : (\mathbb{K}[t])^{n \times n} \to \mathbb{K}^{n \times n}$ (defined as in Definition 3.15) satisfies

$$\varepsilon^{n \times n}(F) = F \qquad \text{for every } F \in \mathbb{K}^{n \times n}. \tag{29}$$

But Proposition 5.1 **(a)** (applied to $\mathbb{L} = \mathbb{K}[t]$, $\mathbb{M} = \mathbb{K}$, $f = \varepsilon$ and $m = n$) yields that the map $\varepsilon^{n\times n} : (\mathbb{K}[t])^{n\times n} \to \mathbb{K}^{n\times n}$ is a homomorphism of additive groups. Hence,

$$\begin{aligned}
\varepsilon^{n\times n}\left(\sum_{k=0}^{n-1} t^k D_k\right) &= \sum_{k=0}^{n-1} \underbrace{\varepsilon^{n\times n}\left(t^k D_k\right)}_{\substack{=\varepsilon\left(t^k\right)\varepsilon^{n\times n}(D_k) \\ \text{(by Proposition 5.1 (d) (applied to} \\ \mathbb{K}[t],\ \mathbb{K},\ \varepsilon,\ n,\ D_k \text{ and } t^k \\ \text{instead of } \mathbb{L},\ \mathbb{M},\ f,\ m,\ A \text{ and } \lambda))}} \\
&= \sum_{k=0}^{n-1} \underbrace{\varepsilon\left(t^k\right)}_{\substack{=(\varepsilon(t))^k \\ \text{(since } \varepsilon \text{ is a ring} \\ \text{homomorphism)}}} \underbrace{\varepsilon^{n\times n}\left(D_k\right)}_{\substack{=D_k \\ \text{(by (29)} \\ \text{(applied to } F=D_k))}} \\
&= \sum_{k=0}^{n-1} \Bigl(\underbrace{\varepsilon(t)}_{=0}\Bigr)^k D_k = \sum_{k=0}^{n-1} 0^k D_k = \underbrace{0^0}_{=1} D_0 + \sum_{k=1}^{n-1} \underbrace{0^k}_{\substack{=0 \\ \text{(since } k\geq 1)}} D_k \\
&\qquad \left(\begin{array}{c}\text{here, we have split off the addend for } k=0 \text{ from the sum} \\ \text{(since } 0 \in \{0,1,\ldots,n-1\}\text{)}\end{array}\right) \\
&= D_0 + \underbrace{\sum_{k=1}^{n-1} 0 D_k}_{=0_{n\times n}} = D_0. \qquad (30)
\end{aligned}$$

But applying the map $\varepsilon^{n\times n}$ to both sides of the equality (27), we obtain

$$\varepsilon^{n\times n}\left(\operatorname{adj}\left(tI_n - A\right)\right) = \varepsilon^{n\times n}\left(\sum_{k=0}^{n-1} t^k D_k\right) = D_0$$

(by (30)). Thus,

$$D_0 = \varepsilon^{n\times n}\left(\operatorname{adj}\left(\underbrace{tI_n - A}_{=tI_n+(-A)}\right)\right) = \varepsilon^{n\times n}\left(\operatorname{adj}\left(tI_n + (-A)\right)\right) = \operatorname{adj}(-A)$$

(by Proposition 5.3 **(b)**, applied to $-A$ instead of A). Hence, (28) becomes

$$\sum_{i=0}^{n-1} c_{n-1-i} A^i = D_0 = \operatorname{adj}(-A).$$

This proves Lemma 5.17. □

Finally, we are ready to prove Theorem 5.14:

Proof of Theorem 5.14. We must prove the equality $\operatorname{adj} A = (-1)^{n-1} \sum_{i=0}^{n-1} c_{n-1-i} A^i$. This is an equality between two $n \times n$-matrices, and thus obviously holds if $n = 0$. Hence, we WLOG assume that $n \neq 0$. Thus, n is a positive integer. Hence, Proposition 5.16 (applied to $\lambda = -1$) yields

$$\operatorname{adj}(-A) = (-1)^{n-1} \operatorname{adj} A.$$

Therefore,

$$\operatorname{adj} A = (-1)^{n-1} \underbrace{\operatorname{adj}(-A)}_{\substack{= \sum_{i=0}^{n-1} c_{n-1-i} A^i \\ \text{(by Lemma 5.17)}}} = (-1)^{n-1} \sum_{i=0}^{n-1} c_{n-1-i} A^i.$$

This proves Theorem 5.14. □

5.6. Minors of the adjugate: Jacobi's theorem

A *minor* of a matrix A is defined to be a determinant of a square submatrix of A. A theorem due to Jacobi connects the minors of $\operatorname{adj} A$ (for a square matrix A) with the minors of A. Before we can state this theorem, let us introduce some notations:

Definition 5.18. Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. Let $A = (a_{i,j})_{1 \leq i \leq n,\ 1 \leq j \leq m}$ be an $n \times m$-matrix. Let $i_1, i_2, \ldots, i_u$ be some elements of $\{1, 2, \ldots, n\}$; let $j_1, j_2, \ldots, j_v$ be some elements of $\{1, 2, \ldots, m\}$. Then, we shall use $\operatorname{sub}_{(i_1, i_2, \ldots, i_u)}^{(j_1, j_2, \ldots, j_v)} A$ as a synonym for the $u \times v$-matrix $\operatorname{sub}_{i_1, i_2, \ldots, i_u}^{j_1, j_2, \ldots, j_v} A$. Thus, for every $\mathbf{i} \in \{1, 2, \ldots, n\}^u$ and $\mathbf{j} \in \{1, 2, \ldots, m\}^v$, a $u \times v$-matrix $\operatorname{sub}_{\mathbf{i}}^{\mathbf{j}} A$ is defined.

Definition 5.19. If I is a finite set of integers, then $\sum I$ shall denote the sum of all elements of I. (Thus, $\sum I = \sum_{i \in I} i$.)

Definition 5.20. If I is a finite set of integers, then $w(I)$ shall denote the list of all elements of I in increasing order (with no repetitions). (For example, $w(\{3, 4, 8\}) = (3, 4, 8)$.)

The following fact is obvious:

Remark 5.21. Let $n \in \mathbb{N}$. Let I be a subset of $\{1, 2, \ldots, n\}$. Then, $w(I) \in \{1, 2, \ldots, n\}^{|I|}$.

Now, we can state Jacobi's theorem[10]:

[10] This is [Grinbe15, Corollary 7.256]. It also appears in [Prasol94, Theorem 2.5.2] (in a different form).

Theorem 5.22. Let $n \in \mathbb{N}$. For any subset I of $\{1, 2, \ldots, n\}$, we let $\widetilde{I}$ denote the complement $\{1, 2, \ldots, n\} \setminus I$ of I.

Let A be an $n \times n$-matrix.

Let P and Q be two subsets of $\{1, 2, \ldots, n\}$ such that $|P| = |Q| \geq 1$. Then,

$$\det \left(\operatorname{sub}_{w(P)}^{w(Q)} (\operatorname{adj} A) \right) = (-1)^{\sum P + \sum Q} (\det A)^{|Q|-1} \det \left(\operatorname{sub}_{w(\widetilde{Q})}^{w(\widetilde{P})} A \right).$$

We shall not give a standalone proof of this theorem; instead, we will merely derive it from results proven in [Grinbe15]. Namely, in [Grinbe15, Corollary 7.255], the following was proven:

Lemma 5.23. Let $n \in \mathbb{N}$. For any subset I of $\{1, 2, \ldots, n\}$, we let $\widetilde{I}$ denote the complement $\{1, 2, \ldots, n\} \setminus I$ of I.

Let A be an $n \times n$-matrix.

Let P and Q be two subsets of $\{1, 2, \ldots, n\}$ such that $|P| = |Q|$. Then,

$$\det A \cdot \det \left(\operatorname{sub}_{w(P)}^{w(Q)} (\operatorname{adj} A) \right) = (-1)^{\sum P + \sum Q} (\det A)^{|Q|} \det \left(\operatorname{sub}_{w(\widetilde{Q})}^{w(\widetilde{P})} A \right).$$

We shall also use the following obvious lemma:

Lemma 5.24. Let $\mathbb{L}$ and $\mathbb{M}$ be two commutative rings. Let $f : \mathbb{L} \to \mathbb{M}$ be any map. Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. Let $A \in \mathbb{L}^{n \times m}$.

Let $u \in \mathbb{N}$ and $v \in \mathbb{N}$. Let $\mathbf{i} \in \{1, 2, \ldots, n\}^u$ and $\mathbf{j} \in \{1, 2, \ldots, m\}^v$. Then,

$$f^{u \times v} \left(\operatorname{sub}_{\mathbf{i}}^{\mathbf{j}} A \right) = \operatorname{sub}_{\mathbf{i}}^{\mathbf{j}} \left(f^{n \times m} (A) \right).$$

Proof of Theorem 5.22. Define the $\mathbb{K}$-algebra homomorphism $\varepsilon : \mathbb{K}[t] \to \mathbb{K}$ as in Proposition 5.3.

Define a matrix $\widetilde{A} \in (\mathbb{K}[t])^{n \times n}$ by $\widetilde{A} = tI_n + A$. Corollary 5.8 shows that the element $\det(tI_n + A)$ of $\mathbb{K}[t]$ is regular. In other words, the element $\det \widetilde{A}$ of $\mathbb{K}[t]$ is regular (since $\widetilde{A} = tI_n + A$).

We have $|Q| - 1 \in \mathbb{N}$ (since $|Q| \geq 1$). Lemma 5.23 (applied to $\mathbb{K}[t]$ and $\widetilde{A}$ instead of $\mathbb{K}$ and A) yields

$$\begin{aligned}
\det \widetilde{A} \cdot \det\left(\operatorname{sub}_{w(P)}^{w(Q)}\left(\operatorname{adj}\widetilde{A}\right)\right) &= (-1)^{\sum P + \sum Q} \underbrace{\left(\det \widetilde{A}\right)^{|Q|}}_{\substack{=\left(\det \widetilde{A}\right)\left(\det \widetilde{A}\right)^{|Q|-1} \\ \text{(since } |Q| \geq 1\text{)}}} \det\left(\operatorname{sub}_{w(\widetilde{Q})}^{w(\widetilde{P})} \widetilde{A}\right) \\
&= (-1)^{\sum P + \sum Q} \left(\det \widetilde{A}\right) \left(\det \widetilde{A}\right)^{|Q|-1} \det\left(\operatorname{sub}_{w(\widetilde{Q})}^{w(\widetilde{P})} \widetilde{A}\right) \\
&= \det \widetilde{A} \cdot (-1)^{\sum P + \sum Q} \left(\det \widetilde{A}\right)^{|Q|-1} \det\left(\operatorname{sub}_{w(\widetilde{Q})}^{w(\widetilde{P})} \widetilde{A}\right).
\end{aligned}$$

Hence, Lemma 5.5 (applied to $\mathbb{A} = \mathbb{K}[t]$, $a = \det \widetilde{A}$, $b = \det\left(\operatorname{sub}_{w(P)}^{w(Q)}\left(\operatorname{adj}\widetilde{A}\right)\right)$ and $c = (-1)^{\sum P + \sum Q} \left(\det \widetilde{A}\right)^{|Q|-1} \det\left(\operatorname{sub}_{w(\widetilde{Q})}^{w(\widetilde{P})} \widetilde{A}\right)$) yields

$$\det\left(\operatorname{sub}_{w(P)}^{w(Q)}\left(\operatorname{adj}\widetilde{A}\right)\right) = (-1)^{\sum P + \sum Q} \left(\det \widetilde{A}\right)^{|Q|-1} \det\left(\operatorname{sub}_{w(\widetilde{Q})}^{w(\widetilde{P})} \widetilde{A}\right)$$

(since the element $\det \widetilde{A}$ of $\mathbb{K}[t]$ is regular). Applying the map ε to both sides of this equality, we obtain

$$\begin{aligned}
&\varepsilon\left(\det\left(\operatorname{sub}_{w(P)}^{w(Q)}\left(\operatorname{adj}\widetilde{A}\right)\right)\right) \\
&= \varepsilon\left((-1)^{\sum P + \sum Q} \left(\det \widetilde{A}\right)^{|Q|-1} \det\left(\operatorname{sub}_{w(\widetilde{Q})}^{w(\widetilde{P})} \widetilde{A}\right)\right) \\
&= (-1)^{\sum P + \sum Q} \left(\varepsilon\left(\det \widetilde{A}\right)\right)^{|Q|-1} \varepsilon\left(\det\left(\operatorname{sub}_{w(\widetilde{Q})}^{w(\widetilde{P})} \widetilde{A}\right)\right) \qquad (31)
\end{aligned}$$

(since ε is a $\mathbb{K}$-algebra homomorphism).

The definition of $\widetilde{P}$ yields $\widetilde{P} = \{1, 2, \ldots, n\} \setminus P$. Hence,

$$\begin{aligned}
\left|\widetilde{P}\right| &= \underbrace{|\{1, 2, \ldots, n\}|}_{=n} - |P| \qquad \text{(since } P \subseteq \{1, 2, \ldots, n\}\text{)} \\
&= n - |P|.
\end{aligned}$$

Similarly, $\left|\widetilde{Q}\right| = n - |Q|$. Notice that $\left|\widetilde{P}\right| = n - \underbrace{|P|}_{=|Q|} = n - |Q|$.

But Remark 5.21 (applied to $I = P$) yields $w(P) \in \{1, 2, \ldots, n\}^{|P|} = \{1, 2, \ldots, n\}^{|Q|}$ (since $|P| = |Q|$). Also, Remark 5.21 (applied to $I = Q$) yields $w(Q) \in \{1, 2, \ldots, n\}^{|Q|}$. Furthermore, Remark 5.21 (applied to $I = \widetilde{P}$) yields $w\left(\widetilde{P}\right) \in \{1, 2, \ldots, n\}^{\left|\widetilde{P}\right|} = \{1, 2, \ldots, n\}^{n-|Q|}$ (since $\left|\widetilde{P}\right| = n - |Q|$). Finally, Remark 5.21 (applied to $I = \widetilde{Q}$) yields $w\left(\widetilde{Q}\right) \in \{1, 2, \ldots, n\}^{\left|\widetilde{Q}\right|} = \{1, 2, \ldots, n\}^{n-|Q|}$ (since $\left|\widetilde{Q}\right| = n - |Q|$).

Recall that $\operatorname{adj}\widetilde{A} \in (\mathbb{K}[t])^{n\times n}$. Furthermore,

$$\varepsilon^{n\times n}\left(\operatorname{adj}\underbrace{\widetilde{A}}_{=tI_n+A}\right) = \varepsilon^{n\times n}\left(\operatorname{adj}\left(tI_n + A\right)\right) = \operatorname{adj} A$$

(by Proposition 5.3 **(b)**).

We have $w(P) \in \{1,2,\ldots,n\}^{|Q|}$ and $w(Q) \in \{1,2,\ldots,n\}^{|Q|}$. Hence,

$$\operatorname{sub}_{w(P)}^{w(Q)}\left(\operatorname{adj}\widetilde{A}\right) \in (\mathbb{K}[t])^{|Q|\times|Q|}.$$

Thus, Proposition 5.2 **(a)** (applied to $\mathbb{K}[t]$, $\mathbb{K}$, ε, $|Q|$ and $\operatorname{sub}_{w(P)}^{w(Q)}\left(\operatorname{adj}\widetilde{A}\right)$ instead of $\mathbb{L}$, $\mathbb{M}$, f, n and A) yields

$$\begin{aligned}
\varepsilon\left(\det\left(\operatorname{sub}_{w(P)}^{w(Q)}\left(\operatorname{adj}\widetilde{A}\right)\right)\right) &= \det\left(\underbrace{\varepsilon^{|Q|\times|Q|}\left(\operatorname{sub}_{w(P)}^{w(Q)}\left(\operatorname{adj}\widetilde{A}\right)\right)}_{\substack{=\operatorname{sub}_{w(P)}^{w(Q)}\left(\varepsilon^{n\times n}\left(\operatorname{adj}\widetilde{A}\right)\right) \\ \text{(by Lemma 5.24 (applied to } \mathbb{K}[t],\ \mathbb{K},\ \varepsilon,\ n,\ \operatorname{adj}\widetilde{A},\ |Q|,\ |Q|, \\ w(P) \text{ and } w(Q) \text{ instead of } \mathbb{L},\ \mathbb{M},\ f,\ m,\ A,\ u,\ v,\ \mathbf{i} \text{ and } \mathbf{j}))}}\right) \\
&= \det\left(\operatorname{sub}_{w(P)}^{w(Q)}\left(\underbrace{\varepsilon^{n\times n}\left(\operatorname{adj}\widetilde{A}\right)}_{=\operatorname{adj} A}\right)\right) \\
&= \det\left(\operatorname{sub}_{w(P)}^{w(Q)}\left(\operatorname{adj} A\right)\right). \qquad (32)
\end{aligned}$$

Comparing this with (31), we obtain

$$\begin{aligned}
&\det\left(\operatorname{sub}_{w(P)}^{w(Q)}\left(\operatorname{adj} A\right)\right) \\
&= (-1)^{\sum P+\sum Q}\left(\varepsilon\left(\det\widetilde{A}\right)\right)^{|Q|-1}\varepsilon\left(\det\left(\operatorname{sub}_{w(\widetilde{Q})}^{w(\widetilde{P})}\widetilde{A}\right)\right). \qquad (33)
\end{aligned}$$

Recall that $\varepsilon^{n\times n}\left(\underbrace{\widetilde{A}}_{=tI_n+A}\right) = \varepsilon^{n\times n}\left(tI_n + A\right) = A$ (by Proposition 5.3 **(c)**).

On the other hand, $w\left(\widetilde{P}\right) \in \{1,2,\ldots,n\}^{n-|Q|}$ and $w\left(\widetilde{Q}\right) \in \{1,2,\ldots,n\}^{n-|Q|}$. Hence,

$$\operatorname{sub}_{w(\widetilde{Q})}^{w(\widetilde{P})}\widetilde{A} \in (\mathbb{K}[t])^{(n-|Q|)\times(n-|Q|)}.$$

Hence, Proposition 5.2 **(a)** (applied to $\mathbb{K}[t]$, $\mathbb{K}$, ε, $n-|Q|$ and $\operatorname{sub}_{w(\widetilde{Q})}^{w(\widetilde{P})} \widetilde{A}$ instead of $\mathbb{L}$, $\mathbb{M}$, f, n and A) yields

$$\begin{aligned}
\varepsilon\left(\det\left(\operatorname{sub}_{w(\widetilde{Q})}^{w(\widetilde{P})} \widetilde{A}\right)\right) &= \det\left(\underbrace{\varepsilon^{(n-|Q|)\times(n-|Q|)}\left(\operatorname{sub}_{w(\widetilde{Q})}^{w(\widetilde{P})} \widetilde{A}\right)}_{\substack{=\operatorname{sub}_{w(\widetilde{Q})}^{w(\widetilde{P})}\left(\varepsilon^{n\times n}\left(\widetilde{A}\right)\right) \\ \text{(by Lemma 5.24 (applied to } \mathbb{K}[t],\ \mathbb{K},\ \varepsilon,\ n,\ \widetilde{A},\ n-|Q|,\ n-|Q|, \\ w(\widetilde{Q}) \text{ and } w(\widetilde{P}) \text{ instead of } \mathbb{L},\ \mathbb{M},\ f,\ m,\ A,\ u,\ v,\ \mathbf{i} \text{ and } \mathbf{j}))}}\right) \\
&= \det\left(\operatorname{sub}_{w(\widetilde{Q})}^{w(\widetilde{P})}\left(\underbrace{\varepsilon^{n\times n}\left(\widetilde{A}\right)}_{=A}\right)\right) \\
&= \det\left(\operatorname{sub}_{w(\widetilde{Q})}^{w(\widetilde{P})} A\right). \qquad (34)
\end{aligned}$$

Also, $\varepsilon\left(\det \underbrace{\widetilde{A}}_{=tI_n+A}\right) = \varepsilon\left(\det\left(tI_n + A\right)\right) = \det A$ (by Proposition 5.3 **(a)**).

Now, (33) becomes

$$\begin{aligned}
&\det\left(\operatorname{sub}_{w(P)}^{w(Q)}\left(\operatorname{adj} A\right)\right) \\
&= (-1)^{\sum P+\sum Q}\left(\underbrace{\varepsilon\left(\det \widetilde{A}\right)}_{=\det A}\right)^{|Q|-1} \underbrace{\varepsilon\left(\det\left(\operatorname{sub}_{w(\widetilde{Q})}^{w(\widetilde{P})} \widetilde{A}\right)\right)}_{\substack{=\det\left(\operatorname{sub}_{w(\widetilde{Q})}^{w(\widetilde{P})} A\right) \\ \text{(by (34))}}} \\
&= (-1)^{\sum P+\sum Q}\left(\det A\right)^{|Q|-1} \det\left(\operatorname{sub}_{w(\widetilde{Q})}^{w(\widetilde{P})} A\right).
\end{aligned}$$

This proves Theorem 5.22. □

5.7. Another application of the $tI_n + A$ strategy

The strategy that we have used to prove Theorem 5.10, Theorem 5.12 and Theorem 5.22 (namely, replacing a matrix $A \in \mathbb{K}^{n\times n}$ by the matrix $tI_n + A \in \left(\mathbb{K}[t]\right)^{n\times n}$, whose determinant is a regular element of $\mathbb{K}[t]$; and then applying the homomorphism ε to get back to A) has many applications; not all of them concern the adjugate of a matrix. As an example of such an application, let us prove a neat property of commuting matrices:

Theorem 5.25. Let $n \in \mathbb{N}$. Let A, B and S be three $n \times n$-matrices such that $AB = BA$. Then,

$$\det(AS + B) = \det(SA + B).$$

Again, we start by showing a particular case of this theorem:

Lemma 5.26. Let $n \in \mathbb{N}$. Let A, B and S be three $n \times n$-matrices such that $AB = BA$. Assume that the element $\det A$ of $\mathbb{K}$ is regular. Then,

$$\det(AS + B) = \det(SA + B).$$

Proof of Lemma 5.26. Define two $n \times n$-matrices X and Y by $X = AS + B$ and $Y = SA + B$. Comparing

$$\underbrace{X}_{=AS+B} A = (AS + B) A = ASA + BA$$

with

$$A \underbrace{Y}_{=SA+B} = A(SA + B) = ASA + \underbrace{AB}_{=BA} = ASA + BA,$$

we obtain $XA = AY$. Now, comparing

$$\det\left(\underbrace{XA}_{=AY}\right) = \det(AY) = \det A \cdot \det Y$$

with

$$\det(XA) = \det X \cdot \det A = \det A \cdot \det X,$$

we obtain $\det A \cdot \det X = \det A \cdot \det Y$. Lemma 5.5 (applied to $\mathbb{K}$, $\det A$, $\det X$ and $\det Y$ instead of $\mathbb{A}$, a, b and c) thus yields $\det X = \det Y$ (since the element $\det A$ of $\mathbb{K}$ is regular). In view of $X = AS + B$ and $Y = SA + B$, this rewrites as $\det(AS + B) = \det(SA + B)$. This proves Lemma 5.26. $\square$

Proof of Theorem 5.25. Define the $\mathbb{K}$-algebra homomorphism $\varepsilon : \mathbb{K}[t] \to \mathbb{K}$ as in Proposition 5.3. Thus, ε is a ring homomorphism. Hence, Proposition 5.1 **(a)** (applied to $\mathbb{L} = \mathbb{K}[t]$, $\mathbb{M} = \mathbb{K}$ and $m = n$) shows that the map $\varepsilon^{n\times n} : (\mathbb{K}[t])^{n\times n} \to \mathbb{K}^{n\times n}$ is a homomorphism of additive groups.

Recall that every $n \times n$-matrix in $\mathbb{K}^{n\times n}$ can be considered as a matrix in $(\mathbb{K}[t])^{n\times n}$. In other words, for each $F \in \mathbb{K}^{n\times n}$, we can consider F as a matrix in $(\mathbb{K}[t])^{n\times n}$; therefore, $\varepsilon^{n\times n}(F)$ is well-defined. We have

$$\varepsilon^{n\times n}(F) = F \qquad \text{for every } F \in \mathbb{K}^{n\times n}. \tag{35}$$

(In fact, the proof of (35) is identical with the proof of (29) we gave above.)

Let $\widetilde{A}$ be the matrix $tI_n + A \in (\mathbb{K}[t])^{n\times n}$. Thus, $\widetilde{A} = tI_n + A$. Applying the map $\varepsilon^{n\times n}$ to both sides of this equality, we find $\varepsilon^{n\times n}\left(\widetilde{A}\right) = \varepsilon^{n\times n}(tI_n + A) = A$ (by Proposition 5.3 **(c)**).

Corollary 5.8 shows that the element $\det(tI_n + A)$ of $\mathbb{K}[t]$ is regular. In other words, the element $\det \widetilde{A}$ of $\mathbb{K}[t]$ is regular (since $\widetilde{A} = tI_n + A$).

Let us consider the matrix $S \in \mathbb{K}^{n\times n}$ as a matrix in $(\mathbb{K}[t])^{n\times n}$ (since every $n \times n$-matrix in $\mathbb{K}^{n\times n}$ can be considered as a matrix in $(\mathbb{K}[t])^{n\times n}$).

Similarly, let us consider the matrix $B \in \mathbb{K}^{n\times n}$ as a matrix in $(\mathbb{K}[t])^{n\times n}$. Then,

$$\begin{aligned}
\underbrace{\widetilde{A}}_{=tI_n+A} B &= (tI_n + A) B = t \underbrace{I_n B}_{=B=BI_n} + \underbrace{AB}_{=BA} = \underbrace{tBI_n}_{=B\cdot tI_n} + BA \\
&= B \cdot tI_n + BA = B \underbrace{(tI_n + A)}_{=\widetilde{A}} = B\widetilde{A}.
\end{aligned}$$

Hence, Lemma 5.26 (applied to $\mathbb{K}[t]$ and $\widetilde{A}$ instead of $\mathbb{K}$ and A) yields

$$\det\left(\widetilde{A}S + B\right) = \det\left(S\widetilde{A} + B\right). \tag{36}$$

Proposition 5.2 **(a)** (applied to $\mathbb{K}[t]$, $\mathbb{K}$, ε and $\widetilde{A}S + B$ instead of $\mathbb{L}$, $\mathbb{M}$, f and A) yields

$$\varepsilon\left(\det\left(\widetilde{A}S + B\right)\right) = \det\left(\varepsilon^{n\times n}\left(\widetilde{A}S + B\right)\right).$$

In view of

$$\begin{aligned}
\varepsilon^{n\times n}\left(\widetilde{A}S + B\right) &= \underbrace{\varepsilon^{n\times n}\left(\widetilde{A}S\right)}_{\substack{=\varepsilon^{n\times n}\left(\widetilde{A}\right)\cdot\varepsilon^{n\times n}(S) \\ \text{(by Proposition 5.1 \textbf{(b)}} \\ \text{(applied to } \mathbb{K}[t],\ \mathbb{K},\ \varepsilon,\ n,\ n,\ \widetilde{A} \text{ and } S \\ \text{instead of } \mathbb{L},\ \mathbb{M},\ f,\ m,\ p,\ A \text{ and } B))}} + \varepsilon^{n\times n}(B) \\
&\qquad \left(\begin{array}{c}\text{since the map } \varepsilon^{n\times n} \text{ is a homomorphism} \\ \text{of additive groups}\end{array}\right) \\
&= \underbrace{\varepsilon^{n\times n}\left(\widetilde{A}\right)}_{=A} \cdot \underbrace{\varepsilon^{n\times n}(S)}_{\substack{=S \\ \text{(by (35)} \\ \text{(applied to } F=S))}} + \underbrace{\varepsilon^{n\times n}(B)}_{\substack{=B \\ \text{(by (35)} \\ \text{(applied to } F=B))}} \\
&= AS + B,
\end{aligned}$$

this becomes

$$\varepsilon\left(\det\left(\widetilde{A}S + B\right)\right) = \det\left(\underbrace{\varepsilon^{n\times n}\left(\widetilde{A}S + B\right)}_{=AS+B}\right) = \det(AS + B). \tag{37}$$

Similarly,

$$\varepsilon\left(\det\left(S\widetilde{A}+B\right)\right)=\det\left(SA+B\right).$$

Comparing this with

$$\varepsilon\left(\underbrace{\det\left(S\widetilde{A}+B\right)}_{\substack{=\det\left(\widetilde{A}S+B\right)\\ \text{(by (36))}}}\right)=\varepsilon\left(\det\left(\widetilde{A}S+B\right)\right)=\det\left(AS+B\right) \qquad \text{(by (37))},$$

we obtain $\det(AS+B)=\det(SA+B)$. This proves Theorem 5.25. □

5.8. Another application of the strategy: block matrices

The same strategy (replacing $A\in\mathbb{K}^{n\times n}$ by $tI_n+A\in(\mathbb{K}[t])^{n\times n}$) turns out to be useful in proving a formula for determinants of block matrices with a certain property.

We will use [Grinbe15, Definition 6.89] in this section. Roughly speaking, this definition says that if n, n', m and m' are four nonnegative integers, and if $A\in\mathbb{K}^{n\times m}$, $B\in\mathbb{K}^{n\times m'}$, $C\in\mathbb{K}^{n'\times m}$ and $D\in\mathbb{K}^{n'\times m'}$ are four matrices, then $\begin{pmatrix} A & B\\ C & D\end{pmatrix}$ shall denote the $(n+n')\times(m+m')$-matrix obtained by "gluing the matrices A, B, C and D together" in the way the notation suggests (i.e., the matrix B is glued to the right edge of A, and then the matrices C and D are glued to the bottom edges of A and B, respectively). For example, if $n=2$, $n'=2$, $m=2$ and $m'=2$, and if

$$A=\begin{pmatrix} a_{1,1} & a_{1,2}\\ a_{2,1} & a_{2,2}\end{pmatrix}, \qquad B=\begin{pmatrix} b_{1,1} & b_{1,2}\\ b_{2,1} & b_{2,2}\end{pmatrix}, \tag{38}$$

$$C=\begin{pmatrix} c_{1,1} & c_{1,2}\\ c_{2,1} & c_{2,2}\end{pmatrix}, \qquad \text{and } D=\begin{pmatrix} d_{1,1} & d_{1,2}\\ d_{2,1} & d_{2,2}\end{pmatrix}, \tag{39}$$

then

$$\begin{pmatrix} A & B\\ C & D\end{pmatrix}=\begin{pmatrix} a_{1,1} & a_{1,2} & b_{1,1} & b_{1,2}\\ a_{2,1} & a_{2,2} & b_{2,1} & b_{2,2}\\ c_{1,1} & c_{1,2} & d_{1,1} & d_{1,2}\\ c_{2,1} & c_{2,2} & d_{2,1} & d_{2,2}\end{pmatrix}.$$

There are more general versions of this "gluing operation" that allow for more than four matrices; but we will only concern ourselves with the case of four matrices.

We are aiming to prove the following theorem:

Theorem 5.27. Let $n \in \mathbb{N}$. Let A, B, C and D be four $n \times n$-matrices such that $AC = CA$. Then, the $(2n) \times (2n)$-matrix $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ satisfies

$$\det \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det (AD - CB).$$

Theorem 5.27 appears, e.g., in [Silves00, (14)]. Our proof of this theorem will closely follow [Silves00, proof of Lemma 2]. We will use the following obvious lemma:

Lemma 5.28. Let $\mathbb{L}$ and $\mathbb{M}$ be two commutative rings. Let $f : \mathbb{L} \to \mathbb{M}$ be any map. Let n, n', m and m' be four nonnegative integers. Let $A \in \mathbb{L}^{n \times m}$, $B \in \mathbb{L}^{n \times m'}$, $C \in \mathbb{L}^{n' \times m}$ and $D \in \mathbb{L}^{n' \times m'}$ be four matrices. Then,

$$f^{(n+n') \times (m+m')} \left(\begin{pmatrix} A & B \\ C & D \end{pmatrix} \right) = \begin{pmatrix} f^{n \times m}(A) & f^{n \times m'}(B) \\ f^{n' \times m}(C) & f^{n' \times m'}(D) \end{pmatrix}.$$

Example 5.29. For this example, set $n = 2$ and $n' = 2$ and $m = 2$ and $m' = 2$, and let the 2×2-matrices A, B, C and D be given by (38) and (39). Then, Lemma 5.28 says that

$$f^{4 \times 4} \left(\begin{pmatrix} A & B \\ C & D \end{pmatrix} \right) = \begin{pmatrix} f^{2 \times 2}(A) & f^{2 \times 2}(B) \\ f^{2 \times 2}(C) & f^{2 \times 2}(D) \end{pmatrix}.$$

Both the left and the right hand side of this equality are easily seen to equal

$$\begin{pmatrix} f(a_{1,1}) & f(a_{1,2}) & f(b_{1,1}) & f(b_{1,2}) \\ f(a_{2,1}) & f(a_{2,2}) & f(b_{2,1}) & f(b_{2,2}) \\ f(c_{1,1}) & f(c_{1,2}) & f(d_{1,1}) & f(d_{1,2}) \\ f(c_{2,1}) & f(c_{2,2}) & f(d_{2,1}) & f(d_{2,2}) \end{pmatrix}.$$

Next, let us recall a result from [Grinbe15] (a version of the Schur complement theorem):

Proposition 5.30. Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. Let $A \in \mathbb{K}^{n \times n}$, $B \in \mathbb{K}^{n \times m}$, $C \in \mathbb{K}^{m \times n}$ and $D \in \mathbb{K}^{m \times m}$. Furthermore, let $W \in \mathbb{K}^{m \times m}$ and $V \in \mathbb{K}^{m \times n}$ be such that $VA = -WC$. Then,

$$\det W \cdot \det \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det A \cdot \det (VB + WD).$$

Proposition 5.30 appears (with proof) in [Grinbe15, Exercise 6.35], so we will not prove it here.

Let us next prove the particular case of Theorem 5.27 in which we assume $\det A$ to be regular:

Lemma 5.31. Let $n \in \mathbb{N}$. Let A, B, C and D be four $n \times n$-matrices such that $AC = CA$. Assume that the element $\det A$ of $\mathbb{K}$ is regular. Then, the $(2n) \times (2n)$-matrix $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ satisfies

$$\det \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det (AD - CB).$$

Proof of Lemma 5.31. The matrix $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ is an $(n+n) \times (n+n)$-matrix (by its definition), i.e., a $(2n) \times (2n)$-matrix (since $n + n = 2n$).

We have $CA = -(-A)C$ (since $-(-A)C = AC = CA$). Thus, Proposition 5.30 (applied to $V = C$ and $W = -A$) yields

$$\det(-A) \cdot \det \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det A \cdot \det \Bigg(\underbrace{CB + (-A) D}_{=CB-AD} \Bigg) = \det A \cdot \det (CB - AD). \tag{40}$$

But (26) (applied to -1 and $CB - AD$ instead of λ and C) yields

$$\det((-1)(CB - AD)) = (-1)^n \det (CB - AD). \tag{41}$$

Also, (26) (applied to -1 and $-A$ instead of λ and C) yields $\det((-1)(-A)) = (-1)^n \det(-A)$. In view of $(-1)(-A) = A$, this rewrites as $\det A = (-1)^n \det(-A)$. Hence,

$$\begin{aligned}
&\underbrace{\det A}_{=(-1)^n \det(-A)} \cdot \det \begin{pmatrix} A & B \\ C & D \end{pmatrix} \\
&= (-1)^n \underbrace{\det(-A) \cdot \det \begin{pmatrix} A & B \\ C & D \end{pmatrix}}_{\substack{=\det A \cdot \det(CB-AD) \\ \text{(by (40))}}} \\
&= (-1)^n \det A \cdot \det(CB - AD) = \det A \cdot \underbrace{(-1)^n \det(CB - AD)}_{\substack{=\det((-1)(CB-AD)) \\ \text{(by (41))}}} \\
&= \det A \cdot \det \Bigg(\underbrace{(-1)(CB - AD)}_{=AD-CB} \Bigg) = \det A \cdot \det (AD - CB).
\end{aligned}$$

Lemma 5.5 (applied to $\mathbb{K}$, $\det A$, $\det \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ and $\det (AD - CB)$ instead of $\mathbb{A}$, a, b and c) thus yields $\det \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det (AD - CB)$ (since the element $\det A$ of $\mathbb{K}$ is regular). This proves Lemma 5.31. $\square$

We are now ready to prove Theorem 5.27:

Proof of Theorem 5.27. The matrix $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ is a $(2n) \times (2n)$-matrix. (This is proven in the same way as in our proof of Lemma 5.31.)

Define the $\mathbb{K}$-algebra homomorphism $\varepsilon : \mathbb{K}[t] \to \mathbb{K}$ as in Proposition 5.3. Thus, ε is a ring homomorphism. Hence, Proposition 5.1 **(a)** (applied to $\mathbb{L} = \mathbb{K}[t]$, $\mathbb{M} = \mathbb{K}$ and $m = n$) shows that the map $\varepsilon^{n\times n} : (\mathbb{K}[t])^{n\times n} \to \mathbb{K}^{n\times n}$ is a homomorphism of additive groups.

Recall that every $n \times n$-matrix in $\mathbb{K}^{n\times n}$ can be considered as a matrix in $(\mathbb{K}[t])^{n\times n}$. In other words, for each $F \in \mathbb{K}^{n\times n}$, we can consider F as a matrix in $(\mathbb{K}[t])^{n\times n}$; therefore, $\varepsilon^{n\times n}(F)$ is well-defined. We have

$$\varepsilon^{n\times n}(F) = F \qquad \text{for every } F \in \mathbb{K}^{n\times n}. \tag{42}$$

(In fact, the proof of (42) is identical with the proof of (29) we gave above.)

Let $\widetilde{A}$ be the matrix $tI_n + A \in (\mathbb{K}[t])^{n\times n}$. Thus, $\widetilde{A} = tI_n + A$. Applying the map $\varepsilon^{n\times n}$ to both sides of this equality, we find $\varepsilon^{n\times n}\left(\widetilde{A}\right) = \varepsilon^{n\times n}(tI_n + A) = A$ (by Proposition 5.3 **(c)**).

Corollary 5.8 shows that the element $\det(tI_n + A)$ of $\mathbb{K}[t]$ is regular. In other words, the element $\det \widetilde{A}$ of $\mathbb{K}[t]$ is regular (since $\widetilde{A} = tI_n + A$).

Let us consider the matrix $B \in \mathbb{K}^{n\times n}$ as a matrix in $(\mathbb{K}[t])^{n\times n}$ (since every $n \times n$-matrix in $\mathbb{K}^{n\times n}$ can be considered as a matrix in $(\mathbb{K}[t])^{n\times n}$). Similarly, let us consider the matrices C and D as matrices in $(\mathbb{K}[t])^{n\times n}$.

Notice that (42) (applied to $F = B$) yields $\varepsilon^{n\times n}(B) = B$. Similarly, $\varepsilon^{n\times n}(C) = C$ and $\varepsilon^{n\times n}(D) = D$.

Now,

$$\begin{aligned} \underbrace{\widetilde{A}}_{=tI_n+A} C &= (tI_n + A)\, C = t \underbrace{I_n C}_{=C=CI_n} + \underbrace{AC}_{=CA} = \underbrace{tCI_n}_{=C\cdot tI_n} + CA \\ &= C \cdot tI_n + CA = C \underbrace{(tI_n + A)}_{=\widetilde{A}} = C\widetilde{A}. \end{aligned}$$

Thus, Lemma 5.31 (applied to $\mathbb{K}[t]$ and $\widetilde{A}$ instead of $\mathbb{K}$ and A) yields

$$\det \begin{pmatrix} \widetilde{A} & B \\ C & D \end{pmatrix} = \det \left(\widetilde{A}D - CB\right)$$

(since the element $\det \widetilde{A}$ of $\mathbb{K}[t]$ is regular). Applying the map ε to both sides of this equality, we find

$$\varepsilon\left(\det\begin{pmatrix}\widetilde{A} & B\\ C & D\end{pmatrix}\right)=\varepsilon\left(\det\left(\widetilde{A}D-CB\right)\right). \tag{43}$$

But Proposition 5.2 **(a)** (applied to $\mathbb{K}[t]$, $\mathbb{K}$, ε and $\widetilde{A}D-CB$ instead of $\mathbb{L}$, $\mathbb{M}$, f and A) yields

$$\varepsilon\left(\det\left(\widetilde{A}D-CB\right)\right)=\det\left(\varepsilon^{n\times n}\left(\widetilde{A}D-CB\right)\right).$$

In view of

$$\begin{aligned}\varepsilon^{n\times n}\left(\widetilde{A}D-CB\right)&=\underbrace{\varepsilon^{n\times n}\left(\widetilde{A}D\right)}_{\substack{=\varepsilon^{n\times n}\left(\widetilde{A}\right)\cdot\varepsilon^{n\times n}(D)\\ \text{(by Proposition 5.1 (b)}\\ \text{(applied to }\mathbb{K}[t],\ \mathbb{K},\ \varepsilon,\ n,\ n,\ \widetilde{A}\text{ and }D\\ \text{instead of }\mathbb{L},\ \mathbb{M},\ f,\ m,\ p,\ A\text{ and }B))}}-\underbrace{\varepsilon^{n\times n}(CB)}_{\substack{=\varepsilon^{n\times n}(C)\cdot\varepsilon^{n\times n}(B)\\ \text{(by Proposition 5.1 (b)}\\ \text{(applied to }\mathbb{K}[t],\ \mathbb{K},\ \varepsilon,\ n,\ n,\ C\text{ and }B\\ \text{instead of }\mathbb{L},\ \mathbb{M},\ f,\ m,\ p,\ A\text{ and }B))}}\\ &\qquad\left(\begin{array}{c}\text{since the map }\varepsilon^{n\times n}\text{ is a homomorphism}\\ \text{of additive groups}\end{array}\right)\\ &=\underbrace{\varepsilon^{n\times n}\left(\widetilde{A}\right)}_{=A}\cdot\underbrace{\varepsilon^{n\times n}(D)}_{=D}-\underbrace{\varepsilon^{n\times n}(C)}_{=C}\cdot\underbrace{\varepsilon^{n\times n}(B)}_{=B}\\ &=AD-CB,\end{aligned}$$

this becomes

$$\begin{aligned}\varepsilon\left(\det\left(\widetilde{A}D-CB\right)\right)&=\det\left(\underbrace{\varepsilon^{n\times n}\left(\widetilde{A}D-CB\right)}_{=AD-CB}\right)\\ &=\det\left(AD-CB\right).\end{aligned} \tag{44}$$

But Proposition 5.2 **(a)** (applied to $\mathbb{K}[t]$, $\mathbb{K}$, ε, $n+n$ and $\begin{pmatrix}\widetilde{A} & B\\ C & D\end{pmatrix}$ instead of $\mathbb{L}$, $\mathbb{M}$, f, n and A) yields

$$\varepsilon\left(\det\begin{pmatrix}\widetilde{A} & B\\ C & D\end{pmatrix}\right)=\det\left(\varepsilon^{(n+n)\times(n+n)}\left(\begin{pmatrix}\widetilde{A} & B\\ C & D\end{pmatrix}\right)\right). \tag{45}$$

On the other hand, Lemma 5.28 (applied to $\mathbb{K}[t]$, $\mathbb{K}$, ε, n, n, n and $\widetilde{A}$ instead of $\mathbb{L}$, $\mathbb{M}$, f, n', m, m' and A) yields

$$\varepsilon^{(n+n)\times(n+n)}\left(\begin{pmatrix}\widetilde{A} & B\\ C & D\end{pmatrix}\right)=\begin{pmatrix}\varepsilon^{n\times n}\left(\widetilde{A}\right) & \varepsilon^{n\times n}(B)\\ \varepsilon^{n\times n}(C) & \varepsilon^{n\times n}(D)\end{pmatrix}=\begin{pmatrix}A & B\\ C & D\end{pmatrix}$$

(since $\varepsilon^{n\times n}\left(\widetilde{A}\right) = A$ and $\varepsilon^{n\times n}(B) = B$ and $\varepsilon^{n\times n}(C) = C$ and $\varepsilon^{n\times n}(D) = D$). Taking determinants on both sides of this equality, we find

$$\det\left(\varepsilon^{(n+n)\times(n+n)}\left(\left(\begin{array}{cc}\widetilde{A} & B\\ C & D\end{array}\right)\right)\right) = \det\left(\begin{array}{cc}A & B\\ C & D\end{array}\right).$$

Hence,

$$\begin{aligned}\det\left(\begin{array}{cc}A & B\\ C & D\end{array}\right) &= \det\left(\varepsilon^{(n+n)\times(n+n)}\left(\left(\begin{array}{cc}\widetilde{A} & B\\ C & D\end{array}\right)\right)\right)\\ &= \varepsilon\left(\det\left(\begin{array}{cc}\widetilde{A} & B\\ C & D\end{array}\right)\right) \qquad (\text{by } (45))\\ &= \varepsilon\left(\det\left(\widetilde{A}D - CB\right)\right) \qquad (\text{by } (43))\\ &= \det(AD - CB) \qquad (\text{by } (44)).\end{aligned}$$

This completes the proof of Theorem 5.27. □

5.9. The trace of the adjugate

The following neat result follows so easily from Theorem 5.14 and Theorem 2.6 that it would be strange not to mention it:

Theorem 5.32. Let $n \in \mathbb{N}$. Let $A \in \mathbb{K}^{n\times n}$. For every $j \in \mathbb{Z}$, define an element $c_j \in \mathbb{K}$ by $c_j = \left[t^{n-j}\right]\chi_A$. Then,

$$\operatorname{Tr}(\operatorname{adj} A) = (-1)^{n-1} c_{n-1} = (-1)^{n-1}\left[t^1\right]\chi_A.$$

In other words, the trace of the adjugate $\operatorname{adj} A$ of an $n\times n$-matrix is the coefficient of t in the characteristic polynomial χ_A.

Proof of Theorem 5.32. The definition of c_{n-1} yields $c_{n-1} = \left[t^{n-(n-1)}\right]\chi_A = \left[t^1\right]\chi_A$ (since $n-(n-1) = 1$).

It is easy to see that Theorem 5.32 holds for $n = 0$ [11]. Thus, for the rest of this proof, we can WLOG assume that we don't have $n = 0$. Assume this. Hence, $n \neq 0$, so that $n \geq 1$ (since $n \in \mathbb{N}$). Therefore, $n - 1 \in \mathbb{N}$.

[11] *Proof.* Assume that $n = 0$. Thus, $1 > 0 = n$. But Corollary 2.4 **(a)** yields that $\chi_A \in \mathbb{K}[t]$ is a polynomial of degree $\leq n$ in t. Hence, $[t^m]\chi_A = 0$ for every integer $m > n$. Applying this to $m = 1$, we obtain $\left[t^1\right]\chi_A = 0$ (since $1 > n$). Also, $\operatorname{adj} A$ is an $n\times n$-matrix, and thus a 0×0-matrix (since $n = 0$). Hence, $\operatorname{Tr}(\operatorname{adj} A) = 0$ (since the trace of a 0×0-matrix is 0). Comparing this with $(-1)^{n-1}\underbrace{c_{n-1}}_{=[t^1]\chi_A=0} = 0$, we obtain $\operatorname{Tr}(\operatorname{adj} A) = (-1)^{n-1}\underbrace{c_{n-1}}_{=[t^1]\chi_A} = (-1)^{n-1}\left[t^1\right]\chi_A$. Hence, we have proven Theorem 5.32 under the assumption that $n = 0$.

Thus, Theorem 2.6 (applied to $k = n-1$) yields

$$(n-1)\, c_{n-1} + \sum_{i=1}^{n-1} \operatorname{Tr}\left(A^i\right) c_{n-1-i} = 0.$$

Subtracting $(n-1)\, c_{n-1}$ from both sides of this equation, we obtain

$$\sum_{i=1}^{n-1} \operatorname{Tr}\left(A^i\right) c_{n-1-i} = -\,(n-1)\, c_{n-1}. \tag{46}$$

But Theorem 5.14 yields

$$\operatorname{adj} A = (-1)^{n-1} \sum_{i=0}^{n-1} c_{n-1-i} A^i.$$

Applying the map $\operatorname{Tr} : \mathbb{K}^{n\times n} \to \mathbb{K}$ to both sides of this equality, we obtain

$$\begin{aligned} \operatorname{Tr}(\operatorname{adj} A) &= \operatorname{Tr}\left((-1)^{n-1} \sum_{i=0}^{n-1} c_{n-1-i} A^i \right) \\ &= (-1)^{n-1} \sum_{i=0}^{n-1} c_{n-1-i} \operatorname{Tr}\left(A^i\right) \end{aligned} \tag{47}$$

(since the map $\operatorname{Tr} : \mathbb{K}^{n\times n} \to \mathbb{K}$ is $\mathbb{K}$-linear). But $n-1 \geq 0$ (since $n \geq 1$); therefore, $0 \in \{0, 1, \ldots, n-1\}$. Hence, we can split off the addend for $i = 0$ from the sum $\sum\limits_{i=0}^{n-1} c_{n-1-i} \operatorname{Tr}\left(A^i\right)$. We thus obtain

$$\begin{aligned} \sum_{i=0}^{n-1} c_{n-1-i} \operatorname{Tr}\left(A^i\right) &= \underbrace{c_{n-1-0}}_{=c_{n-1}} \operatorname{Tr}\left(\underbrace{A^0}_{=I_n} \right) + \sum_{i=1}^{n-1} \underbrace{c_{n-1-i} \operatorname{Tr}\left(A^i\right)}_{=\operatorname{Tr}\left(A^i\right) c_{n-1-i}} \\ &= c_{n-1} \underbrace{\operatorname{Tr}(I_n)}_{=n} + \underbrace{\sum_{i=1}^{n-1} \operatorname{Tr}\left(A^i\right) c_{n-1-i}}_{\substack{=-(n-1)c_{n-1} \\ \text{(by (46))}}} = c_{n-1} n + \left(-\,(n-1)\, c_{n-1}\right) \\ &= \underbrace{(n-(n-1))}_{=1} c_{n-1} = c_{n-1}. \end{aligned}$$

Hence, (47) becomes

$$\operatorname{Tr}(\operatorname{adj} A) = (-1)^{n-1} \underbrace{\sum_{i=0}^{n-1} c_{n-1-i} \operatorname{Tr}\left(A^i\right)}_{=c_{n-1}} = (-1)^{n-1} \underbrace{c_{n-1}}_{=\left[t^1\right]\chi_A} = (-1)^{n-1} \left[t^1\right] \chi_A.$$

This proves Theorem 5.32. □

5.10. Yet another application to block matrices

Let us show one further formula for determinants of certain block matrices that can be proved using our "replace A by $tI_n + A$" strategy.

We will again use [Grinbe15, Definition 6.89] in this section. We shall furthermore use the following notation:

Definition 5.33. If B is any 1×1-matrix, then $\operatorname{ent} B$ will denote the $(1,1)$-th entry of B. (This entry is, of course, the only entry of B. Thus, the 1×1-matrix B satisfies $B = \left(\operatorname{ent} B \right)$.)

We now claim the following:

Theorem 5.34. Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. Let $A \in \mathbb{K}^{n \times n}$ and $D \in \mathbb{K}^{m \times m}$ be two square matrices. Let $p \in \mathbb{K}^{n \times 1}$ and $q \in \mathbb{K}^{m \times 1}$ be two column vectors. Let $v \in \mathbb{K}^{1 \times m}$ and $u \in \mathbb{K}^{1 \times n}$ be two row vectors. Then,

$$\det \begin{pmatrix} A & pv \\ qu & D \end{pmatrix} = \det A \cdot \det D - \operatorname{ent}\left(u \left(\operatorname{adj} A\right) p\right) \cdot \operatorname{ent}\left(v \left(\operatorname{adj} D\right) q\right).$$

Example 5.35. Let us see what Theorem 5.34 says in the case when $n = 2$ and $m = 2$. Indeed, let $n = 2$ and $m = 2$ and

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix} \qquad \text{and} \qquad D = \begin{pmatrix} d_{1,1} & d_{1,2} \\ d_{2,1} & d_{2,2} \end{pmatrix} \qquad \text{and}$$
$$p = \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} \qquad \text{and} \qquad q = \begin{pmatrix} q_1 \\ q_2 \end{pmatrix} \qquad \text{and}$$
$$v = \begin{pmatrix} v_1 & v_2 \end{pmatrix} \qquad \text{and} \qquad u = \begin{pmatrix} u_1 & u_2 \end{pmatrix}.$$

Then,

$$pv = \begin{pmatrix} p_1 v_1 & p_1 v_2 \\ p_2 v_1 & p_2 v_2 \end{pmatrix} \qquad \text{and} \qquad qu = \begin{pmatrix} q_1 u_1 & q_1 u_2 \\ q_2 u_1 & q_2 u_2 \end{pmatrix}.$$

Hence,

$$\begin{pmatrix} A & pv \\ qu & D \end{pmatrix} = \begin{pmatrix} a_{1,1} & a_{1,2} & p_1 v_1 & p_1 v_2 \\ a_{2,1} & a_{2,2} & p_2 v_1 & p_2 v_2 \\ q_1 u_1 & q_1 u_2 & d_{1,1} & d_{1,2} \\ q_2 u_1 & q_2 u_2 & d_{2,1} & d_{2,2} \end{pmatrix}.$$

Hence, the claim of Theorem 5.34 rewrites as follows in our case:

$$\det\begin{pmatrix} a_{1,1} & a_{1,2} & p_1v_1 & p_1v_2 \\ a_{2,1} & a_{2,2} & p_2v_1 & p_2v_2 \\ q_1u_1 & q_1u_2 & d_{1,1} & d_{1,2} \\ q_2u_1 & q_2u_2 & d_{2,1} & d_{2,2} \end{pmatrix}$$
$$= \det\begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix} \cdot \det\begin{pmatrix} d_{1,1} & d_{1,2} \\ d_{2,1} & d_{2,2} \end{pmatrix}$$
$$- \operatorname{ent}\left(\begin{pmatrix} u_1 & u_2 \end{pmatrix} \left(\operatorname{adj}\begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix} \right) \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} \right)$$
$$\cdot \operatorname{ent}\left(\begin{pmatrix} v_1 & v_2 \end{pmatrix} \left(\operatorname{adj}\begin{pmatrix} d_{1,1} & d_{1,2} \\ d_{2,1} & d_{2,2} \end{pmatrix} \right) \begin{pmatrix} q_1 \\ q_2 \end{pmatrix} \right).$$

In order to prove Theorem 5.34, we will need one standard result (known as the *matrix determinant lemma*):

Theorem 5.36. Let $n \in \mathbb{N}$. Let u be a column vector with n entries, and let v be a row vector with n entries. (Thus, uv is an $n \times n$-matrix, whereas vu is a 1×1-matrix.) Let A be an $n \times n$-matrix. Then,

$$\det(A + uv) = \det A + \operatorname{ent}(v(\operatorname{adj} A)u).$$

See [Grinbe15, Theorem 7.262] for a proof of Theorem 5.36.

We will furthermore need the following three trivial lemmas:

Lemma 5.37. Let $C \in \mathbb{K}^{1\times 1}$ be an 1×1-matrix. Let $\lambda \in \mathbb{K}$. Then, $\operatorname{ent}(\lambda C) = \lambda \operatorname{ent} C$.

Lemma 5.38. Let B and C be two 1×1-matrices. Then, $\operatorname{ent}(BC) = \operatorname{ent} B \cdot \operatorname{ent} C$.

Lemma 5.39. Let $\mathbb{L}$ and $\mathbb{M}$ be rings. Let $f : \mathbb{L} \to \mathbb{M}$ be any map. Let $B \in \mathbb{L}^{1\times 1}$. Then, $\operatorname{ent}\left(f^{1\times 1}(B)\right) = f(\operatorname{ent} B)$.

Finally, we will need a simple property of regular elements in a commutative ring:

Lemma 5.40. Let $\mathbb{A}$ be a commutative ring. Let a be a regular element of $\mathbb{A}$. Let $m \in \mathbb{N}$. Then, a^m is a regular element of $\mathbb{A}$.

Proof of Lemma 5.40. The element a is regular. In other words,

$$\text{every } x \in \mathbb{A} \text{ satisfying } ax = 0 \text{ satisfies } x = 0 \tag{48}$$

(by the definition of "regular").

Now, let $x \in \mathbb{A}$ satisfy $a^m x = 0$. We shall show that $x = 0$. Indeed, we shall first prove that

$$a^{m-i} x = 0 \qquad \text{for each } i \in \{0, 1, \ldots, m\}. \tag{49}$$

[*Proof of (49):* We proceed by induction on i:

Induction base: We have $a^{m-0}x = a^m x = 0$ (by assumption). Hence, (49) holds for $i = 0$.

Induction step: Let $j \in \{1, 2, \ldots, m\}$. Assume that (49) holds for $i = j - 1$. We must show that (49) holds for $i = j$ as well.

We have assumed that (49) holds for $i = j - 1$. In other words, $a^{m-(j-1)}x = 0$. In other words, $a^{m-j+1}x = 0$ (since $m - (j-1) = m - j + 1$). In other words, $aa^{m-j}x = 0$ (since $a^{m-j+1} = aa^{m-j}$). Hence, (48) (applied to $a^{m-j}x$ instead of x) yields $a^{m-j}x = 0$. In other words, (49) holds for $i = j$. This completes the induction step. Thus, (49) is proved by induction.]

Now, (49) (applied to $i = m$) yields $a^{m-m}x = 0$. Since $\underbrace{a^{m-m}}_{=a^0=1} x = x$, this rewrites as $x = 0$.

Forget that we fixed x. We thus have shown that every $x \in \mathbb{A}$ satisfying $a^m x = 0$ satisfies $x = 0$. In other words, the element a^m of $\mathbb{A}$ is regular (by the definition of "regular"). This proves Lemma 5.40. $\square$

Now, we can approach the proof of Theorem 5.34 using the same technique as various theorems proved above. We begin by proving it in the case when $\det A$ is regular:

Lemma 5.41. Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. Let $A \in \mathbb{K}^{n \times n}$ and $D \in \mathbb{K}^{m \times m}$ be two square matrices. Assume that the element $\det A$ of $\mathbb{K}$ is regular. Let $p \in \mathbb{K}^{n \times 1}$ and $q \in \mathbb{K}^{m \times 1}$ be two column vectors. Let $v \in \mathbb{K}^{1 \times m}$ and $u \in \mathbb{K}^{1 \times n}$ be two row vectors. Then,

$$\det \begin{pmatrix} A & pv \\ qu & D \end{pmatrix} = \det A \cdot \det D - \operatorname{ent}\left(u \left(\operatorname{adj} A\right) p\right) \cdot \operatorname{ent}\left(v \left(\operatorname{adj} D\right) q\right).$$

Proof of Lemma 5.41. Set $\lambda = \det A$. Thus, the element λ of $\mathbb{K}$ is regular (since the element $\det A$ of $\mathbb{K}$ is regular). Hence, Lemma 5.40 (applied to $\mathbb{A} = \mathbb{K}$ and $a = \lambda$) shows that λ^m is a regular element of $\mathbb{K}$.

It is furthermore easy to see that

$$\lambda \operatorname{adj}(\lambda D) = \lambda^m \operatorname{adj} D. \tag{50}$$

[*Proof of (50):* If $m = 0$, then (50) holds for trivial reasons[12]. Thus, for the rest of this proof, we WLOG assume that we don't have $m = 0$. Hence, m is a positive

[12] *Proof.* Assume that $m = 0$. The matrices $\lambda \operatorname{adj}(\lambda D)$ and $\lambda^m \operatorname{adj} D$ are $m \times m$-matrices (since D is an $m \times m$-matrix). In other words, the matrices $\lambda \operatorname{adj}(\lambda D)$ and $\lambda^m \operatorname{adj} D$ are 0×0-matrices (since

integer (since $m \in \mathbb{N}$). Therefore, Proposition 5.16 (applied to m and D instead of n and A) yields $\operatorname{adj}(\lambda D) = \lambda^{m-1} \operatorname{adj} D$. Hence, $\lambda \underbrace{\operatorname{adj}(\lambda D)}_{=\lambda^{m-1}\operatorname{adj} D} = \underbrace{\lambda \lambda^{m-1}}_{=\lambda^m} \operatorname{adj} D = \lambda^m \operatorname{adj} D$. This proves (50).]

Define two matrices $W \in \mathbb{K}^{m \times m}$ and $V \in \mathbb{K}^{m \times n}$ by

$$W = \lambda I_m \qquad \text{and} \qquad V = -qu \operatorname{adj} A.$$

Then,

$$\begin{aligned}
\underbrace{V}_{=-qu\operatorname{adj} A} A &= -qu \underbrace{\operatorname{adj} A \cdot A}_{\substack{=\det A \cdot I_n \\ \text{(by Theorem 3.7)}}} = -qu \underbrace{\det A}_{\substack{=\lambda \\ \text{(since } \lambda = \det A)}} \cdot I_n = -qu\lambda I_n \\
&= -\lambda qu I_n = -\lambda qu = -Wqu
\end{aligned}$$

(since $-\underbrace{W}_{=\lambda I_m} qu = -\lambda I_m qu = -\lambda qu$). Hence, Proposition 5.30 (applied to $B = pv$ and $C = qu$) yields

$$\det W \cdot \det \begin{pmatrix} A & pv \\ qu & D \end{pmatrix} = \det A \cdot \det (Vpv + WD).$$

In view of

$$\begin{aligned}
\det \underbrace{W}_{=\lambda I_m} &= \det (\lambda I_m) = \lambda^m \underbrace{\det (I_m)}_{=1} \qquad \left(\begin{array}{c} \text{by (26), applied to } m \text{ and } I_m \\ \text{instead of } n \text{ and } C \end{array} \right) \\
&= \lambda^m
\end{aligned}$$

and $\det A = \lambda$, we can rewrite this as

$$\lambda^m \cdot \det \begin{pmatrix} A & pv \\ qu & D \end{pmatrix} = \lambda \cdot \det (Vpv + WD). \tag{51}$$

Let us define two 1×1-matrices B and C by

$$B = u (\operatorname{adj} A) p \in \mathbb{K}^{1 \times 1} \qquad \text{and} \qquad C = v \cdot \operatorname{adj}(\lambda D) \cdot (-q) \in \mathbb{K}^{1 \times 1}.$$

Next, we observe that

$$\begin{aligned}
&\underbrace{V}_{=-qu\operatorname{adj} A} pv + \underbrace{W}_{=\lambda I_m} D \\
&= \underbrace{(-qu \operatorname{adj} A) pv}_{=(-q)u(\operatorname{adj} A)pv} + \lambda \underbrace{I_m D}_{=D} = (-q) \underbrace{u (\operatorname{adj} A) p}_{\substack{=B \\ \text{(since } B = u(\operatorname{adj} A)p)}} v + \lambda D = (-q) Bv + \lambda D \\
&= \lambda D + (-q) Bv = \lambda D + (-q)(Bv).
\end{aligned} \tag{52}$$

$m = 0$). Hence, these two matrices $\lambda \operatorname{adj}(\lambda D)$ and $\lambda^m \operatorname{adj} D$ are equal (since there exists only one 0×0-matrix, and therefore any two 0×0-matrices are equal). In other words, $\lambda \operatorname{adj}(\lambda D) = \lambda^m \operatorname{adj} D$. Thus, we have proved (50) under the assumption that $m = 0$.

Note that $\lambda D \in \mathbb{K}^{m\times m}$ and $-q \in \mathbb{K}^{m\times 1}$ and $Bv \in \mathbb{K}^{1\times m}$. Hence, Theorem 5.36 (applied to m, $-q$, Bv and D instead of n, u, v and A) yields

$$\begin{aligned}
&\det\left(\lambda D + (-q)(Bv)\right) \\
&= \underbrace{\det(\lambda D)}_{\substack{=\lambda^m \det D \\ \text{(by (26), applied to } m \text{ and } D \\ \text{instead of } n \text{ and } A\text{)}}} + \operatorname{ent}\left(\underbrace{(Bv)\cdot \operatorname{adj}(\lambda D)\cdot(-q)}_{=B(v\cdot \operatorname{adj}(\lambda D)\cdot(-q))}\right) \\
&= \lambda^m \det D + \operatorname{ent}\left(B\underbrace{\left(v\cdot \operatorname{adj}(\lambda D)\cdot(-q)\right)}_{\substack{=C \\ \text{(since } C = v\cdot\operatorname{adj}(\lambda D)\cdot(-q))}}\right) = \lambda^m \det D + \underbrace{\operatorname{ent}(BC)}_{\substack{=\operatorname{ent} B\cdot \operatorname{ent} C \\ \text{(by Lemma 5.38)}}} \\
&= \lambda^m \det D + \operatorname{ent} B \cdot \operatorname{ent} C.
\end{aligned}$$

In view of (52), we can rewrite this as

$$\det(Vpv + WD) = \lambda^m \det D + \operatorname{ent} B \cdot \operatorname{ent} C.$$

Thus, (51) becomes

$$\begin{aligned}
\lambda^m \cdot \det\begin{pmatrix} A & pv \\ qu & D \end{pmatrix} &= \lambda \cdot \underbrace{\det(Vpv + WD)}_{=\lambda^m \det D + \operatorname{ent} B\cdot \operatorname{ent} C} \\
&= \lambda \cdot \left(\lambda^m \det D + \operatorname{ent} B \cdot \operatorname{ent} C\right) \\
&= \lambda\lambda^m \det D + \underbrace{\lambda \operatorname{ent} B \cdot \operatorname{ent} C}_{=\lambda \operatorname{ent} C\cdot \operatorname{ent} B} \\
&= \lambda\lambda^m \det D + \lambda \operatorname{ent} C \cdot \operatorname{ent} B \\
&= \lambda\lambda^m \det D - (-\lambda) \operatorname{ent} C \cdot \operatorname{ent} B. \qquad (53)
\end{aligned}$$

However, Lemma 5.37 (applied to $-\lambda$ instead of λ) yields $\operatorname{ent}((-\lambda) C) = (-\lambda) \operatorname{ent} C$. Thus,

$$\begin{aligned}
(-\lambda)\operatorname{ent} C &= \operatorname{ent}\left((-\lambda)\underbrace{C}_{=v\cdot\operatorname{adj}(\lambda D)\cdot(-q)}\right) = \operatorname{ent}\left(\underbrace{(-\lambda)\, v\cdot \operatorname{adj}(\lambda D)\cdot(-q)}_{\substack{=\lambda v\cdot \operatorname{adj}(\lambda D)\cdot q \\ = v\cdot\lambda \operatorname{adj}(\lambda D)\cdot q}}\right) \\
&= \operatorname{ent}\left(v\cdot \underbrace{\lambda \operatorname{adj}(\lambda D)}_{\substack{=\lambda^m \operatorname{adj} D \\ \text{(by (50))}}}\cdot q\right) = \operatorname{ent}\left(\underbrace{v\cdot\left(\lambda^m \operatorname{adj} D\right)\cdot q}_{=\lambda^m\cdot v(\operatorname{adj} D)q}\right) \\
&= \operatorname{ent}\left(\lambda^m \cdot v\left(\operatorname{adj} D\right) q\right) = \lambda^m \operatorname{ent}\left(v\left(\operatorname{adj} D\right) q\right)
\end{aligned}$$

(by Lemma 5.37, applied to λ^m and $v(\operatorname{adj} D) q$ instead of λ and C). Thus, (53) becomes

$$\begin{aligned}
\lambda^m \cdot \det \begin{pmatrix} A & pv \\ qu & D \end{pmatrix} &= \lambda\lambda^m \det D - \underbrace{(-\lambda) \operatorname{ent} C}_{=\lambda^m \operatorname{ent}(v(\operatorname{adj} D)q)} \cdot \operatorname{ent} B \\
&= \lambda\lambda^m \det D - \lambda^m \operatorname{ent}(v(\operatorname{adj} D) q) \cdot \operatorname{ent} B \\
&= \lambda^m \cdot (\lambda \det D - \operatorname{ent}(v(\operatorname{adj} D) q) \cdot \operatorname{ent} B).
\end{aligned}$$

Since λ^m is a regular element of $\mathbb{K}$, we can thus conclude that

$$\det \begin{pmatrix} A & pv \\ qu & D \end{pmatrix} = \lambda \det D - \operatorname{ent}(v(\operatorname{adj} D) q) \cdot \operatorname{ent} B$$

(by Lemma 5.5, applied to $\mathbb{A} = \mathbb{K}$ and $a = \lambda^m$ and $b = \det \begin{pmatrix} A & pv \\ qu & D \end{pmatrix}$ and $c = \lambda \det D - \operatorname{ent}(v(\operatorname{adj} D) q) \cdot \operatorname{ent} B$). In view of $\lambda = \det A$ and $B = u(\operatorname{adj} A) p$, we can rewrite this as

$$\begin{aligned}
\det \begin{pmatrix} A & pv \\ qu & D \end{pmatrix} &= \det A \cdot \det D - \operatorname{ent}(v(\operatorname{adj} D) q) \cdot \operatorname{ent}(u(\operatorname{adj} A) p) \\
&= \det A \cdot \det D - \operatorname{ent}(u(\operatorname{adj} A) p) \cdot \operatorname{ent}(v(\operatorname{adj} D) q).
\end{aligned}$$

This proves Lemma 5.41. □

We can now derive Theorem 5.34 from Lemma 5.41 by the same recipe as before:

Proof of Theorem 5.34. Define the $\mathbb{K}$-algebra homomorphism $\varepsilon : \mathbb{K}[t] \to \mathbb{K}$ as in Proposition 5.3. Thus, ε is a ring homomorphism.

Recall that every $n \times n$-matrix in $\mathbb{K}^{n\times n}$ can be considered as a matrix in $(\mathbb{K}[t])^{n\times n}$. More generally, every $k \times \ell$-matrix in $\mathbb{K}^{k\times \ell}$ (for any nonnegative integers k and ℓ) can be considered as a matrix in $(\mathbb{K}[t])^{k\times \ell}$. In other words, for each $F \in \mathbb{K}^{k\times \ell}$, we can consider F as a matrix in $(\mathbb{K}[t])^{k\times \ell}$; therefore, $\varepsilon^{k\times \ell}(F)$ is well-defined. We have

$$\varepsilon^{k\times \ell}(F) = F \qquad \text{for every } F \in \mathbb{K}^{k\times \ell}. \tag{54}$$

(In fact, the proof of (54) is analogous to the proof of (29) we gave above.)

Let $\widetilde{A}$ be the matrix $tI_n + A \in (\mathbb{K}[t])^{n\times n}$. Thus, $\widetilde{A} = tI_n + A$. Applying the map $\varepsilon^{n\times n}$ to both sides of this equality, we find $\varepsilon^{n\times n}\left(\widetilde{A}\right) = \varepsilon^{n\times n}(tI_n + A) = A$ (by Proposition 5.3 **(c)**).

Corollary 5.8 shows that the element $\det(tI_n + A)$ of $\mathbb{K}[t]$ is regular. In other words, the element $\det \widetilde{A}$ of $\mathbb{K}[t]$ is regular (since $\widetilde{A} = tI_n + A$).

Let us consider the matrix $A \in \mathbb{K}^{n\times n}$ as a matrix in $(\mathbb{K}[t])^{n\times n}$ (since every $n \times n$-matrix in $\mathbb{K}^{n\times n}$ can be considered as a matrix in $(\mathbb{K}[t])^{n\times n}$). Similarly, let us consider the matrices $D \in \mathbb{K}^{m\times m}$, $p \in \mathbb{K}^{n\times 1}$, $q \in \mathbb{K}^{m\times 1}$, $v \in \mathbb{K}^{1\times m}$ and $u \in$

$\mathbb{K}^{1\times n}$ as matrices in $(\mathbb{K}[t])^{m\times m}$, $(\mathbb{K}[t])^{n\times 1}$, $(\mathbb{K}[t])^{m\times 1}$, $(\mathbb{K}[t])^{1\times m}$ and $(\mathbb{K}[t])^{1\times n}$, respectively.

Notice that (54) (applied to $k = n$ and $\ell = n$ and $F = A$) yields $\varepsilon^{n\times n}(A) = A$. Similarly, $\varepsilon^{m\times m}(D) = D$ and $\varepsilon^{n\times 1}(p) = p$ and $\varepsilon^{m\times 1}(q) = q$ and $\varepsilon^{1\times m}(v) = v$ and $\varepsilon^{1\times n}(u) = u$ and $\varepsilon^{n\times m}(pv) = pv$ and $\varepsilon^{m\times n}(qu) = qu$.

In the proof of Theorem 5.12, we have already shown that

$$\varepsilon^{n\times n}\left(\operatorname{adj}\widetilde{A}\right) = \operatorname{adj}A \qquad \text{and} \qquad \varepsilon\left(\det\widetilde{A}\right) = \det A.$$

Next, we claim that

$$\varepsilon\left(\operatorname{ent}\left(u\left(\operatorname{adj}\widetilde{A}\right)p\right)\right) = \operatorname{ent}\left(u\left(\operatorname{adj}A\right)p\right). \tag{55}$$

[*Proof of (55):* Define the 1×1-matrix $B = u\left(\operatorname{adj}\widetilde{A}\right)p \in (\mathbb{K}[t])^{1\times 1}$. Then, Lemma 5.39 (applied to $\mathbb{L} = \mathbb{K}[t]$ and $\mathbb{M} = \mathbb{K}$ and $f = \varepsilon$) yields $\operatorname{ent}\left(\varepsilon^{1\times 1}(B)\right) = \varepsilon(\operatorname{ent}B)$.

However, from $B = u\left(\operatorname{adj}\widetilde{A}\right)p = u\left(\left(\operatorname{adj}\widetilde{A}\right)p\right)$, we obtain

$$\begin{aligned}
\varepsilon^{1\times 1}(B) &= \varepsilon^{1\times 1}\left(u\left(\left(\operatorname{adj}\widetilde{A}\right)p\right)\right) \\
&= \underbrace{\varepsilon^{1\times n}(u)}_{=u}\cdot\underbrace{\varepsilon^{n\times 1}\left(\left(\operatorname{adj}\widetilde{A}\right)p\right)}_{\substack{=\varepsilon^{n\times n}\left(\operatorname{adj}\widetilde{A}\right)\cdot\varepsilon^{n\times 1}(p) \\ \text{(by Theorem 5.1 }\textbf{(c)}\text{)}}} \qquad \text{(by Theorem 5.1 }\textbf{(c)}\text{)} \\
&= u\cdot\underbrace{\varepsilon^{n\times n}\left(\operatorname{adj}\widetilde{A}\right)}_{=\operatorname{adj}A}\cdot\underbrace{\varepsilon^{n\times 1}(p)}_{=p} = u\left(\operatorname{adj}A\right)p.
\end{aligned}$$

In view of this, we can rewrite the equality $\operatorname{ent}\left(\varepsilon^{1\times 1}(B)\right) = \varepsilon(\operatorname{ent}B)$ (which we have proved in the previous paragraph) as

$$\operatorname{ent}\left(u\left(\operatorname{adj}A\right)p\right) = \varepsilon(\operatorname{ent}B) = \varepsilon\left(\operatorname{ent}\left(u\left(\operatorname{adj}\widetilde{A}\right)p\right)\right)$$

(since $B = u\left(\operatorname{adj}\widetilde{A}\right)p$). This proves (55).]

Furthermore, we have $\varepsilon(\lambda) = \lambda$ for each $\lambda \in \mathbb{K}$ (by the definition of ε). Applying this to $\lambda = \det D$, we find

$$\varepsilon(\det D) = \det D.$$

Also, $v\left(\operatorname{adj}D\right)q \in \mathbb{K}^{1\times 1}$ and thus $\operatorname{ent}\left(v\left(\operatorname{adj}D\right)q\right) \in \mathbb{K}$. Recall again that we have $\varepsilon(\lambda) = \lambda$ for each $\lambda \in \mathbb{K}$. Applying this to $\lambda = \operatorname{ent}\left(v\left(\operatorname{adj}D\right)q\right)$, we obtain

$$\varepsilon\left(\operatorname{ent}\left(v\left(\operatorname{adj}D\right)q\right)\right) = \operatorname{ent}\left(v\left(\operatorname{adj}D\right)q\right) \tag{56}$$

(since $\operatorname{ent}\left(v\left(\operatorname{adj}D\right)q\right) \in \mathbb{K}$).

Now, Lemma 5.41 (applied to $\mathbb{K}[t]$ and $\widetilde{A}$ instead of $\mathbb{K}$ and A) yields

$$\det \begin{pmatrix} \widetilde{A} & pv \\ qu & D \end{pmatrix} = \det \widetilde{A} \cdot \det D - \operatorname{ent}\left(u\left(\operatorname{adj} \widetilde{A}\right) p\right) \cdot \operatorname{ent}\left(v\left(\operatorname{adj} D\right) q\right)$$

(since the element $\det \widetilde{A}$ of $\mathbb{K}[t]$ is regular). Applying the map ε to both sides of this equality, we find

$$\begin{aligned} &\varepsilon\left(\det \begin{pmatrix} \widetilde{A} & pv \\ qu & D \end{pmatrix}\right) \\ &= \varepsilon\left(\det \widetilde{A} \cdot \det D - \operatorname{ent}\left(u\left(\operatorname{adj} \widetilde{A}\right) p\right) \cdot \operatorname{ent}\left(v\left(\operatorname{adj} D\right) q\right)\right) \\ &= \underbrace{\varepsilon\left(\det \widetilde{A}\right)}_{=\det A} \cdot \underbrace{\varepsilon\left(\det D\right)}_{=\det D} - \underbrace{\varepsilon\left(\operatorname{ent}\left(u\left(\operatorname{adj} \widetilde{A}\right) p\right)\right)}_{\substack{=\operatorname{ent}(u(\operatorname{adj} A)p) \\ \text{(by (55))}}} \cdot \underbrace{\varepsilon\left(\operatorname{ent}\left(v\left(\operatorname{adj} D\right) q\right)\right)}_{\substack{=\operatorname{ent}(v(\operatorname{adj} D)q) \\ \text{(by (56))}}} \\ &\qquad \left(\text{since } \varepsilon \text{ is a ring homomorphism}\right) \\ &= \det A \cdot \det D - \operatorname{ent}\left(u\left(\operatorname{adj} A\right) p\right) \cdot \operatorname{ent}\left(v\left(\operatorname{adj} D\right) q\right). \end{aligned} \tag{57}$$

However, $\begin{pmatrix} \widetilde{A} & pv \\ qu & D \end{pmatrix}$ is an $(n+m) \times (n+m)$-matrix in $(\mathbb{K}[t])^{(n+m)\times(n+m)}$. Thus, Proposition 5.2 **(a)** (applied to $\mathbb{K}[t]$, $\mathbb{K}$, ε, $n+m$ and $\begin{pmatrix} \widetilde{A} & pv \\ qu & D \end{pmatrix}$ instead of $\mathbb{L}$, $\mathbb{M}$, f, n and A) yields

$$\varepsilon\left(\det \begin{pmatrix} \widetilde{A} & pv \\ qu & D \end{pmatrix}\right) = \det\left(\varepsilon^{(n+m)\times(n+m)} \begin{pmatrix} \widetilde{A} & pv \\ qu & D \end{pmatrix}\right). \tag{58}$$

On the other hand, Lemma 5.28 (applied to $\mathbb{K}[t]$, $\mathbb{K}$, ε, m, n, m, $\widetilde{A}$, pv and qu instead of $\mathbb{L}$, $\mathbb{M}$, f, n', m, m', A, B and C) yields

$$\varepsilon^{(n+m)\times(n+m)}\left(\begin{pmatrix} \widetilde{A} & pv \\ qu & D \end{pmatrix}\right) = \begin{pmatrix} \varepsilon^{n\times n}\left(\widetilde{A}\right) & \varepsilon^{n\times m}\left(pv\right) \\ \varepsilon^{m\times n}\left(qu\right) & \varepsilon^{m\times m}\left(D\right) \end{pmatrix} = \begin{pmatrix} A & pv \\ qu & D \end{pmatrix}$$

(since $\varepsilon^{n\times n}\left(\widetilde{A}\right) = A$ and $\varepsilon^{n\times m}(pv) = pv$ and $\varepsilon^{m\times n}(qu) = qu$ and $\varepsilon^{m\times m}(D) = D$). Taking determinants on both sides of this equality, we find

$$\det\left(\varepsilon^{(n+m)\times(n+m)}\left(\begin{pmatrix} \widetilde{A} & pv \\ qu & D \end{pmatrix}\right)\right) = \det \begin{pmatrix} A & pv \\ qu & D \end{pmatrix}.$$

Hence,

$$\det \begin{pmatrix} A & pv \\ qu & D \end{pmatrix} = \det \left(\varepsilon^{(n+m)\times(n+m)} \left(\begin{pmatrix} \widetilde{A} & pv \\ qu & D \end{pmatrix} \right) \right)$$
$$= \varepsilon \left(\det \begin{pmatrix} \widetilde{A} & pv \\ qu & D \end{pmatrix} \right) \qquad \text{(by (58))}$$
$$= \det A \cdot \det D - \operatorname{ent}\left(u \left(\operatorname{adj} A\right) p\right) \cdot \operatorname{ent}\left(v \left(\operatorname{adj} D\right) q\right) \qquad \text{(by (57))}.$$

This completes the proof of Theorem 5.34. ☐

Note that Theorem 5.34 generalizes the following known fact (e.g., [Grinbe15, Exercise 6.60 **(a)**]):

Corollary 5.42. Let $n \in \mathbb{N}$. Let $u \in \mathbb{K}^{n\times 1}$ be a column vector with n entries, and let $v \in \mathbb{K}^{1\times n}$ be a row vector with n entries. (Thus, uv is an $n \times n$-matrix, whereas vu is a 1×1-matrix.) Let $h \in \mathbb{K}$. Let H be the 1×1-matrix $\begin{pmatrix} h \end{pmatrix} \in \mathbb{K}^{1\times 1}$. Let $A \in \mathbb{K}^{n\times n}$ be an $n \times n$-matrix. Then,

$$\det \begin{pmatrix} A & u \\ v & H \end{pmatrix} = h \det A - \operatorname{ent}\left(v \left(\operatorname{adj} A\right) u\right).$$

Proof of Corollary 5.42 (sketched). Consider the 1×1 identity matrix $I_1 = \begin{pmatrix} 1 \end{pmatrix}$. Then, I_1 is both a row vector and a column vector, and we have $u = uI_1$ and $v = I_1 v$. Moreover, the adjugate of any 1×1-matrix is $\begin{pmatrix} 1 \end{pmatrix}$ (since the determinant of a 0×0-matrix is defined to be 1). Thus, in particular, $\operatorname{adj} H = \begin{pmatrix} 1 \end{pmatrix} = I_1$. Furthermore, from $H = \begin{pmatrix} h \end{pmatrix}$, we obtain $\det H = h$. Now, from $u = uI_1$ and $v = I_1 v$, we obtain

$$\det \begin{pmatrix} A & u \\ v & H \end{pmatrix} = \det \begin{pmatrix} A & uI_1 \\ I_1 v & H \end{pmatrix}$$
$$= \det A \cdot \underbrace{\det H}_{=h} - \operatorname{ent}\left(v \left(\operatorname{adj} A\right) u\right) \cdot \operatorname{ent}\left(I_1 \underbrace{\left(\operatorname{adj} H\right)}_{=I_1} I_1 \right)$$
$$\left(\begin{array}{c} \text{by Theorem 5.34, applied to } 1, H, u, I_1, I_1 \text{ and } v \\ \text{instead of } m, D, p, q, v \text{ and } u \end{array} \right)$$
$$= \underbrace{\det A \cdot h}_{=h \det A} - \operatorname{ent}\left(v \left(\operatorname{adj} A\right) u\right) \cdot \operatorname{ent} \underbrace{\left(I_1 \cdot I_1 \cdot I_1\right)}_{=I_1}$$
$$= h \det A - \operatorname{ent}\left(v \left(\operatorname{adj} A\right) u\right) \cdot \underbrace{\operatorname{ent}\left(I_1\right)}_{=1}$$
$$= h \det A - \operatorname{ent}\left(v \left(\operatorname{adj} A\right) u\right).$$

This proves Corollary 5.42. ☐

Another particular case of Theorem 5.34 is the following:

Corollary 5.43. Let n and m be two positive integers. Let $A \in \mathbb{K}^{n\times n}$ and $D \in \mathbb{K}^{m\times m}$ be two square matrices. Let B be the $n \times m$-matrix whose $(n,1)$-th entry is 1 and whose all other entries are 0. Let C be the $m \times n$-matrix whose $(1,n)$-th entry is 1 and whose all other entries are 0. Then,

$$\det \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det A \cdot \det D - \det\left(A_{\sim n,\sim n}\right) \cdot \det\left(D_{\sim 1,\sim 1}\right).$$

(Recall that we are using the notations from Definition 3.5.)

Example 5.44. Let us see what Corollary 5.43 says in the case when $n = 2$ and $m = 3$. Indeed, let $n = 2$ and $m = 3$ and

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix} \qquad \text{and} \qquad D = \begin{pmatrix} d_{1,1} & d_{1,2} & d_{1,3} \\ d_{2,1} & d_{2,2} & d_{2,3} \\ d_{3,1} & d_{3,2} & d_{3,3} \end{pmatrix}.$$

Then, the matrices B and C defined in Corollary 5.43 are

$$B = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \qquad \text{and} \qquad C = \begin{pmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Hence,

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} a_{1,1} & a_{1,2} & 0 & 0 & 0 \\ a_{2,1} & a_{2,2} & 1 & 0 & 0 \\ 0 & 1 & d_{1,1} & d_{1,2} & d_{1,3} \\ 0 & 0 & d_{2,1} & d_{2,2} & d_{2,3} \\ 0 & 0 & d_{3,1} & d_{3,2} & d_{3,3} \end{pmatrix}.$$

Therefore, the claim of Corollary 5.43 rewrites as follows in our case:

$$\det \begin{pmatrix} a_{1,1} & a_{1,2} & 0 & 0 & 0 \\ a_{2,1} & a_{2,2} & 1 & 0 & 0 \\ 0 & 1 & d_{1,1} & d_{1,2} & d_{1,3} \\ 0 & 0 & d_{2,1} & d_{2,2} & d_{2,3} \\ 0 & 0 & d_{3,1} & d_{3,2} & d_{3,3} \end{pmatrix}$$
$$= \det \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix} \cdot \det \begin{pmatrix} d_{1,1} & d_{1,2} & d_{1,3} \\ d_{2,1} & d_{2,2} & d_{2,3} \\ d_{3,1} & d_{3,2} & d_{3,3} \end{pmatrix} - \det \begin{pmatrix} a_{1,1} \end{pmatrix} \cdot \det \begin{pmatrix} d_{2,2} & d_{2,3} \\ d_{3,2} & d_{3,3} \end{pmatrix}.$$

Proof of Corollary 5.43 (sketched). What follows is by far not the easiest proof of Corollary 5.43, but it puts the corollary in the context of Theorem 5.34.

We let $p \in \mathbb{K}^{n\times 1}$ be the column vector whose n-th entry is 1 and whose all other entries are 0.

We let $q \in \mathbb{K}^{m\times 1}$ be the column vector whose 1-st entry is 1 and whose all other entries are 0.

We let $v \in \mathbb{K}^{1\times m}$ be the row vector whose 1-st entry is 1 and whose all other entries are 0.

We let $u \in \mathbb{K}^{1\times n}$ be the row vector whose n-th entry is 1 and whose all other entries are 0.

Thus,

$$p = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \in \mathbb{K}^{n\times 1}, \qquad q = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{K}^{m\times 1},$$
$$v = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \end{pmatrix} \in \mathbb{K}^{1\times m}, \qquad u = \begin{pmatrix} 0 & 0 & \cdots & 0 & 1 \end{pmatrix} \in \mathbb{K}^{1\times n}.$$

Now, it is easy to see (using just the definition of a product of two matrices) that the following four claims hold:

Claim 1: We have $qu = C$.

Claim 2: We have $pv = B$.

Claim 3: Every $m \times m$-matrix Y satisfies

$$\operatorname{ent}(vYq) = (\text{the } (1,1)\text{-th entry of } Y). \tag{59}$$

Claim 4: Every $n \times n$-matrix X satisfies

$$\operatorname{ent}(uXp) = (\text{the } (n,n)\text{-th entry of } X). \tag{60}$$

Now, Theorem 5.34 yields

$$\det \begin{pmatrix} A & pv \\ qu & D \end{pmatrix} = \det A \cdot \det D - \operatorname{ent}(u(\operatorname{adj} A)p) \cdot \operatorname{ent}(v(\operatorname{adj} D)q).$$

In view of $pv = B$ and $qu = C$ and

$$\begin{aligned} \operatorname{ent}(u(\operatorname{adj} A)p) &= (\text{the } (n,n)\text{-th entry of } \operatorname{adj} A) \\ &\qquad (\text{by (60), applied to } X = \operatorname{adj} A) \\ &= \underbrace{(-1)^{n+n}}_{\substack{=(-1)^{2n}=1 \\ (\text{since } 2n \text{ is even})}} \det(A_{\sim n,\sim n}) \qquad (\text{by the definition of an adjugate}) \\ &= \det(A_{\sim n,\sim n}) \end{aligned}$$

and

$$\begin{aligned}
\operatorname{ent}\left(v\left(\operatorname{adj} D\right)q\right) &= \left(\text{the } (1,1)\text{-th entry of } \operatorname{adj} D\right) \\
&\qquad\qquad \left(\text{by (59), applied to } Y = \operatorname{adj} D\right) \\
&= \underbrace{(-1)^{1+1}}_{=1} \det\left(D_{\sim 1,\sim 1}\right) \qquad \left(\text{by the definition of an adjugate}\right) \\
&= \det\left(D_{\sim 1,\sim 1}\right),
\end{aligned}$$

we can rewrite this as

$$\det\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det A \cdot \det D - \det\left(A_{\sim n,\sim n}\right) \cdot \det\left(D_{\sim 1,\sim 1}\right).$$

This proves Corollary 5.43. □

References

[Bernha11] Chris Bernhardt, *A proof of the Cayley Hamilton theorem*, 10 January 2011.
`http://faculty.fairfield.edu/cbernhardt/cayleyhamilton.pdf`
Published in: The American Mathematical Monthly **116**(5), pp. 456–457.

[BroWil89] Joel G. Broida and S. Gill Williamson, *Comprehensive Introduction to Linear Algebra*, Addison-Wesley 1989.
`http://cseweb.ucsd.edu/~gill/CILASite/`

[Brown93] William Clough Brown, *Matrices over commutative rings*, Monographs and textbooks in pure and applied mathematics #169, Marcel Dekker 1993.

[Camero08] Peter J. Cameron, *Notes on Linear Algebra*, 5 September 2008.
`https://webspace.maths.qmul.ac.uk/p.j.cameron/notes/linalg.pdf`

[Climen13] Vaughn Climenhaga, *Lecture notes Math 4377/6308 – Advanced Linear Algebra I*, 3 December 2013.
`https://www.math.uh.edu/~climenha/math4377/4377-course-notes.pdf`

[Conrad09] Keith Conrad, *Universal identities*, 13 February 2021.
`http://www.math.uconn.edu/~kconrad/blurbs/linmultialg/univid.pdf`

[Ford22] Timothy J. Ford, *Abstract Algebra*, draft of a text, 25 April 2022.
`http://math.fau.edu/ford/preprints/Algebra_Book/Algebra_Book.pdf`

[Garrett09] Paul Garrett, *Abstract Algebra*, lecture notes, 2009.
`http://www.math.umn.edu/~garrett/m/algebra/notes/Whole.pdf`

[Grinbe15] Darij Grinberg, *Notes on the combinatorial fundamentals of algebra*, 26 May 2022.
`http://www.cip.ifi.lmu.de/~grinberg/primes2015/sols.pdf`
The numbering of theorems and formulas in this link might shift when the project gets updated; for a "frozen" version whose numbering is guaranteed to match that in the citations above, see `https://github.com/darijgr/detnotes/releases/tag/2022-05-26` .

[Grinbe16a] Darij Grinberg, *Collected trivialities on algebra derivations*, 22 October 2020.
`http://www.cip.ifi.lmu.de/~grinberg/algebra/derivat.pdf`

[Grinbe16b] Darij Grinberg, *Regular elements of a ring, monic polynomials and "lcm-coprimality"*, 22 May 2021.
`http://www.cip.ifi.lmu.de/~grinberg/algebra/regpol.pdf`

[Heffer14] Jim Hefferon, *Linear Algebra*, version of 26 April 2020.
`http://joshua.smcvt.edu/linearalgebra/book.pdf`

[Knapp2016] Anthony W. Knapp, *Basic Algebra*, Digital Second Edition, 2016.
`http://www.math.stonybrook.edu/~aknapp/download.html`

[Laksov13] Dan Laksov, *Diagonalization of matrices over rings*, Journal of Algebra, Volume **376**, 15 February 2013, pp. 123–138.

[Loehr14] Nicholas Loehr, *Advanced Linear Algebra*, CRC Press 2014.

[LomQui16] Henri Lombardi, Claude Quitté, *Commutative algebra: Constructive methods. Finite projective modules*, Series Algebra and Applications, Vol. 20, Translated from the French (Calvage & Mounet, 2011, revised and extended by the authors) by Tania K. Roblot, Springer, 2015. Published (with corrections) on arXiv as arXiv:1605.04832v3.

[Mate16] Attila Máté, *The Cayley-Hamilton Theorem*, version 28 March 2016.
`http://www.sci.brooklyn.cuny.edu/~mate/misc/cayley_hamilton.pdf`

[McDona84] Bernard R. McDonald, *Linear algebra over commutative rings*, Pure and applied mathematics #**87**, Marcel Dekker 1987.

[Moore68] John T. Moore, *Elements of Linear Algebra and Matrix Theory*, McGraw-Hill 1968.

[m.se1798703] J.E.M.S, *Traces of powers of a matrix A over an algebra are zero implies A nilpotent*, Mathematics Stack Exchange question #1798703, `http://math.stackexchange.com/q/1798703`.

[Prasol94] Viktor V. Prasolov, *Problems and Theorems in Linear Algebra*, Translations of Mathematical Monographs, vol. #134, AMS 1994.

[Sage08] Marc Sage, *Théorème de Cayley-Hamilton : quatre démonstrations*, 15 March 2018.
https://www.normalesup.org/~sage/Enseignement/Cours/CaylHami.pdf

[Shurma15] Jerry Shurman, *The Cayley-Hamilton theorem via multilinear algebra*, http://people.reed.edu/~jerry/332/28ch.pdf . Part of the collection *Course Materials for Mathematics 332: Algebra*, available at http://people.reed.edu/~jerry/332/mat.html .

[Silves00] John R. Silvester, *Determinants of Block Matrices*, The Mathematical Gazette, Vol. **84**, No. 501 (Nov., 2000), pp. 460–467.

[Stoll17] Michael Stoll, *Linear Algebra II (Spring 2007)*, with some additions by Ronald van Luijk, 8 October 2017.
https://websites.math.leidenuniv.nl/algebra/linalg2.pdf
(See https://websites.math.leidenuniv.nl/algebra/linalg1.pdf for volume I.)

[Straub83] Howard Straubing, *A combinatorial proof of the Cayley-Hamilton theorem*, Discrete Mathematics, Volume **43**, Issues 2–3, 1983, pp. 273–279.
https://doi.org/10.1016/0012-365X(83)90164-4

[Zeilbe85] Doron Zeilberger, *A combinatorial approach to matrix algebra*, Discrete Mathematics **56** (1985), pp. 61–72.

[Zeilbe93] Doron Zeilberger, *Gert Almkvist's generalization of a mistake of Bourbaki*, Contemporary Mathematics **143** (1993) (special volume in memory of Emil Grosswald, edited by Marvin Knopp and Mark Sheingorn), pp. 609–612.
http://sites.math.rutgers.edu/~zeilberg/mamarim/mamarimhtml/gert.html