Integrality of matrices, finiteness of matrix semigroups, and dynamics of linear and additive cellular automata

Alberto Dennunzio* Enrico Formenti†

Darij Grinberg‡ Luciano Margara§

June 16, 2026

Abstract. Let $\mathbb{K}$ be a finite commutative ring, and let $\mathbb{L}$ be a commutative $\mathbb{K}$-algebra. Let A and B be two $n \times n$-matrices over $\mathbb{L}$ that have the same characteristic polynomial. The main result of this paper (Thm. 1.2) states that the set $\left\{A^0, A^1, A^2, \ldots\right\}$ is finite if and only if the set $\left\{B^0, B^1, B^2, \ldots\right\}$ is finite. We apply this result to the theory of discrete time dynamical systems. Indeed, it gives a complete and easy-to-check characterization of sensitivity to initial conditions and equicontinuity for linear cellular automata over the alphabet $\mathbb{K}^n$ for $\mathbb{K} = \mathbb{Z}/m\mathbb{Z}$ (Theorem 2.6), i.e., cellular automata in which the local rule is defined by $n \times n$-matrices with elements in $\mathbb{Z}/m\mathbb{Z}$.

To prove our main result, we rederive a classical integrality criterion for matrices (Thm. 1.18 and Prop. 1.17). Namely, let $\mathbb{K}$ be any commutative ring (not necessarily finite), and let $\mathbb{L}$ be a commutative $\mathbb{K}$-algebra. Consider any $n \times n$-matrix A over $\mathbb{L}$. Then, $A \in \mathbb{L}^{n \times n}$ is integral over $\mathbb{K}$ (that is, there exists a monic polynomial $f \in \mathbb{K}[t]$ satisfying $f(A) = 0$)

*Dipartimento di Informatica, Sistemistica e Comunicazione, Università degli Studi di Milano-Bicocca, Viale Sarca 336/14, 20126 Milano, Italy; alberto.dennunzio@unimib.it

†Université Côte d'Azur, CNRS, I3S, France; enrico.formenti@univ-cotedazur.fr

‡Drexel University, 15 S 33rd Street, Philadelphia PA, 19104, USA; Mathematisches Forschungsinstitut Oberwolfach, Schwarzwaldstr. 9-11, 77709 Oberwolfach-Walke, Germany; Institut Mittag-Leffler, Auravägen 17, SE-182 60 Djursholm, Sweden; darijgrinberg@gmail.com

§Department of Computer Science and Engineering, University of Bologna, Campus of Cesena, Via dell'università 50, Cesena, Italy; luciano.margara@unibo.it

if and only if all coefficients of the characteristic polynomial of A are integral over $\mathbb{K}$.

Furthermore, we extend the decidability result concerning sensitivity and equicontinuity to the wider class of additive cellular automata over a finite abelian group. For such cellular automata, we also prove the decidability of injectivity, surjectivity, topological transitivity and all the properties (as, for instance, ergodicity) that are equivalent to the latter. The proof of each of these decidabilty results exploits the corresponding decidability result for linear cellular automata.

Keywords. integrality, linear algebra over rings, commutative algebra, Cayley–Hamilton theorem, finiteness, semigroups, cellular automata, additive cellular automata, linear cellular automata, decidability, discrete dynamical systems, equicontinuity, sensitivity to the initial conditions, injectivity, surjectivity, topological transitivity, ergodicity

Contents

1. On matrices with finitely many distinct powers **3**
1.1. The main theorem . . . 3
1.2. Ingredient 1: Finite semigroups . . . 5
1.3. Ingredient 2: Integrality basics . . . 6
1.4. Characterizing integral matrices . . . 9
1.5. Gert Almkvist's exterior-power trick . . . 10
1.6. Characterizing integral matrices: the proof . . . 14
1.7. Two finiteness lemmas . . . 17
1.8. Proof of Theorem 1.2 . . . 18
1.9. Digression: Traces of nilpotent matrices . . . 20
1.10. Appendix: Second proof of Proposition 1.19 . . . 22
1.11. Appendix: Proof of Proposition 1.17 . . . 34

2. Dynamical Behaviour of Linear and Additive Cellular Automata **35**
2.1. Background on DTDS and Cellular Automata . . . 36
2.2. Additive and Linear Cellular Automata . . . 38
2.3. Decidability Results about Linear CA . . . 39
2.4. From Linear to Additive CA . . . 41
2.4.1. Sensitivity and Equicontinuity for Additive Cellular Automata 44
2.4.2. Surjectivity and Injectivity for Additive Cellular Automata . 45
2.4.3. Topological transitivity and ergodicity . . . 47

Acknowledgments

DG thanks the Mathematisches Forschungsinstitut Oberwolfach for its hospitality during part of the writing process. He also thanks the math.stackexchange user named "punctured dusk" for informing him of the appearance of Theorem 1.18 and Proposition 1.17 in [Bourba72].

1. On matrices with finitely many distinct powers

1.1. The main theorem

We recall a standard definition from linear algebra:[1]

Definition 1.1. Let $\mathbb{K}$ be a commutative ring. Let $n \in \mathbb{N}$. Let A be an $n \times n$-matrix over $\mathbb{K}$. Then, the *characteristic polynomial* χ_A of A is defined to be the polynomial $\det(tI_n - A) \in \mathbb{K}[t]$. Here, I_n stands for the $n \times n$ identity matrix, and $tI_n - A$ is considered as an $n \times n$-matrix over the polynomial ring $\mathbb{K}[t]$.

Our goal in this section is to prove the following theorem:

Theorem 1.2. Let $\mathbb{K}$ be a finite commutative ring. Let $\mathbb{L}$ be a commutative $\mathbb{K}$-algebra. Let $n \in \mathbb{N}$. Let A and B be two $n \times n$-matrices over $\mathbb{L}$ such that $\chi_A = \chi_B$. Then, the set $\{A^0, A^1, A^2, \ldots\}$ is finite if and only if the set $\{B^0, B^1, B^2, \ldots\}$ is finite.

Example 1.3. **(a)** Let $\mathbb{K} = \mathbb{Z}/4$ and $\mathbb{L} = (\mathbb{Z}/4)[x]$ (a polynomial ring) and $n = 2$ and $A = \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. Then, $\chi_A = (t-1)^2 = \chi_B$. Hence, Theorem 1.2 yields that the set $\{A^0, A^1, A^2, \ldots\}$ is finite if and only if the set $\{B^0, B^1, B^2, \ldots\}$ is finite. And indeed, both of these sets are finite: The former has 4 elements, while the latter has 1.

(b) Now, let $\mathbb{K} = \mathbb{Q}$ and $\mathbb{L} = \mathbb{Q}$ and $n = 2$ and $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. Then, $\chi_A = (t-1)^2 = \chi_B$. The ring $\mathbb{K}$ is not finite, so Theorem 1.2 does not apply here. And we see why: The set $\{B^0, B^1, B^2, \ldots\}$ is finite, but the set $\{A^0, A^1, A^2, \ldots\}$ is not.

We shall eventually prove Theorem 1.2, but first let us briefly discuss what rings $\mathbb{L}$ it applies to:

[1] Here and in the following, $\mathbb{N}$ denotes the set $\{0, 1, 2, \ldots\}$.

Proposition 1.4. Let $\mathbb{L}$ be a commutative ring. Then, the following two statements are equivalent:

- *Statement $\mathcal{K}$:* There exist a finite commutative ring $\mathbb{K}$ and a $\mathbb{K}$-algebra structure on $\mathbb{L}$.
- *Statement $\mathcal{M}$:* There exists a positive integer m such that $m \cdot 1_{\mathbb{L}} = 0$. (Here, we denote the unity of any ring $\mathbb{A}$ by $1_{\mathbb{A}}$.)

Proof of Proposition 1.4. We shall prove the two implications $\mathcal{K} \Longrightarrow \mathcal{M}$ and $\mathcal{M} \Longrightarrow \mathcal{K}$:

Proof of the implication $\mathcal{K} \Longrightarrow \mathcal{M}$: Assume that Statement $\mathcal{K}$ holds. In other words, there exist a finite commutative ring $\mathbb{K}$ and a $\mathbb{K}$-algebra structure on $\mathbb{L}$. Consider this ring $\mathbb{K}$ and this structure.

The ring $\mathbb{K}$ is finite. Hence, Lagrange's theorem (applied to the finite group $(\mathbb{K}, +)$) yields $|\mathbb{K}| \cdot a = 0$ for each $a \in \mathbb{K}$. Applying this to $a = 1_{\mathbb{K}}$, we obtain $|\mathbb{K}| \cdot 1_{\mathbb{K}} = 0$. Now, $|\mathbb{K}| \cdot \underbrace{1_{\mathbb{L}}}_{=1_{\mathbb{K}} \cdot 1_{\mathbb{L}}} = \underbrace{|\mathbb{K}| \cdot 1_{\mathbb{K}}}_{=0} \cdot 1_{\mathbb{L}} = 0$. Thus, there exists a positive integer m such that $m \cdot 1_{\mathbb{L}} = 0$ (namely, $m = |\mathbb{K}|$). In other words, Statement $\mathcal{M}$ holds. This proves the implication $\mathcal{K} \Longrightarrow \mathcal{M}$.

Proof of the implication $\mathcal{M} \Longrightarrow \mathcal{K}$: Assume that Statement $\mathcal{M}$ holds. In other words, there exists a positive integer m such that $m \cdot 1_{\mathbb{L}} = 0$. Consider this m. Then, $\mathbb{Z}/m\mathbb{Z}$ is a finite commutative ring. Now, the canonical ring homomorphism $\mathbb{Z} \to \mathbb{L}$, $a \mapsto a \cdot 1_{\mathbb{L}}$ factors through the quotient ring $\mathbb{Z}/m\mathbb{Z}$ (since it sends m to $m \cdot 1_{\mathbb{L}} = 0$, and thus its kernel contains m and therefore the whole ideal $m\mathbb{Z}$). Hence, we have found a ring homomorphism $\mathbb{Z}/m\mathbb{Z} \to \mathbb{L}$. This homomorphism makes $\mathbb{L}$ into a $\mathbb{Z}/m\mathbb{Z}$-algebra (since $\mathbb{L}$ and $\mathbb{Z}/m\mathbb{Z}$ are commutative). Thus, there exist a finite commutative ring $\mathbb{K}$ (namely, $\mathbb{Z}/m\mathbb{Z}$) and a $\mathbb{K}$-algebra structure on $\mathbb{L}$ (namely, the $\mathbb{Z}/m\mathbb{Z}$-algebra we have just found). In other words, Statement $\mathcal{K}$ holds. This proves the implication $\mathcal{M} \Longrightarrow \mathcal{K}$.

We have now proven both implications $\mathcal{K} \Longrightarrow \mathcal{M}$ and $\mathcal{M} \Longrightarrow \mathcal{K}$. Thus, Proposition 1.4 is proven. □

Using Proposition 1.4, we can restate Theorem 1.2 as follows:

Corollary 1.5. Let $\mathbb{L}$ be a commutative ring. Assume that there exists a positive integer m such that $m \cdot 1_{\mathbb{L}} = 0$. Let $n \in \mathbb{N}$. Let A and B be two $n \times n$-matrices over $\mathbb{L}$ such that $\chi_A = \chi_B$. Then, the set $\left\{A^0, A^1, A^2, \ldots\right\}$ is finite if and only if the set $\left\{B^0, B^1, B^2, \ldots\right\}$ is finite.

Remark 1.6. A converse of this corollary holds as well: Let $\mathbb{L}$ be a commutative ring for which there is **no** positive integer m such that $m \cdot 1_{\mathbb{L}} = 0$. Let $n \geq 2$ be an integer. Then, there exist two $n \times n$-matrices A and B over $\mathbb{L}$ such that $\chi_A = \chi_B$

and the set $\{A^0, A^1, A^2, \ldots\}$ is infinite but the set $\{B^0, B^1, B^2, \ldots\}$ is finite. Such matrices can easily be constructed by imitation of Example 1.3 **(b)**.

1.2. Ingredient 1: Finite semigroups

We now start preparing the ground for the proof of Theorem 1.2. The first ingredient of our proof are two basic facts about semigroups.

In the following, semigroups will always be written multiplicatively: That is, if M is a semigroup, then the operation of M will be written as multiplication (i.e., we will write ab for the image of $(a, b) \in M \times M$ under this operation).

Theorem 1.7. Let M be a finite semigroup. Let $a \in M$. Then, there exists a positive integer m such that $a^m = a^{2m}$.

Proof of Theorem 1.7. This is simply saying that the sub-semigroup $\{a^1, a^2, a^3, \ldots\}$ of M contains an idempotent. But this is well-known. See, e.g., [Steinbe16, Corollary 1.2] (where M and a are called S and u, respectively) or [Pin19, Proposition 6.31]. □

Proposition 1.8. Let M be a semigroup. Let $a \in M$. Let p and q be two positive integers such that $p > q$ and $a^p = a^q$. Then, $\{a^1, a^2, a^3, \ldots\} = \{a^1, a^2, \ldots, a^{p-1}\}$.

This proposition is also well-known, but proving it is easier than finding a reference:

Proof of Proposition 1.8. We claim that

$$a^i \in \left\{a^1, a^2, \ldots, a^{p-1}\right\} \qquad \text{for each positive integer } i. \tag{1}$$

[*Proof of (1):* We proceed by strong induction on i. Thus, we fix a positive integer j, and we assume that (1) holds for all $i < j$. We must then prove that (1) holds for $i = j$. In other words, we must prove that $a^j \in \{a^1, a^2, \ldots, a^{p-1}\}$. If $j \leq p - 1$, then this is obvious; thus, for the rest of this proof, we WLOG assume that $j > p - 1$. Hence, $j \geq p$, so that $j - p \geq 0$ and thus $q + \underbrace{(j-p)}_{\geq 0} \geq q$. Hence, $q + (j - p)$ is a positive integer (since q is a positive integer). Furthermore, $q + (j - p) = j + q - \underbrace{p}_{>q} < j + q - q = j$. Thus, (1) holds for $i = q + (j - p)$ (since we have assumed that (1) holds for all $i < j$). In other words, we have $a^{q+(j-p)} \in \{a^1, a^2, \ldots, a^{p-1}\}$.

But we have $a^p = a^q$. Thus,

$$a^{p+k} = a^{q+k} \qquad \text{for each } k \in \mathbb{N}. \tag{2}$$

(Indeed, if $k = 0$, then this follows from $a^p = a^q$; but in the other case it follows from $a^{p+k} = \underbrace{a^p}_{=a^q} a^k = a^q a^k = a^{q+k}$.)

We have $j - p \geq 0$, thus $j - p \in \mathbb{N}$. Hence, applying (2) to $k = j - p$, we find $a^{p+(j-p)} = a^{q+(j-p)}$. But $j = p + (j - p)$, so that $a^j = a^{p+(j-p)} = a^{q+(j-p)} \in \{a^1, a^2, \ldots, a^{p-1}\}$. In other words, (1) holds for $i = j$. This completes the induction step. Thus, (1) is proven.]

Now, (1) immediately yields that $\{a^1, a^2, a^3, \ldots\} \subseteq \{a^1, a^2, \ldots, a^{p-1}\}$. Combining this with the obvious fact that $\{a^1, a^2, \ldots, a^{p-1}\} \subseteq \{a^1, a^2, a^3, \ldots\}$, we obtain $\{a^1, a^2, a^3, \ldots\} = \{a^1, a^2, \ldots, a^{p-1}\}$. This proves Proposition 1.8. □

1.3. Ingredient 2: Integrality basics

Our proof will rely on some basic properties of integrality over a commutative ring. This concept is defined as follows:

Definition 1.9. Let $\mathbb{K}$ be a commutative ring. Let $\mathbb{L}$ be a $\mathbb{K}$-algebra (not necessarily commutative). An element $u \in \mathbb{L}$ is said to be *integral over* $\mathbb{K}$ if and only if there exists a monic polynomial $f \in \mathbb{K}[t]$ such that $f(u) = 0$.

Recall that a polynomial is said to be *monic* if its leading coefficient is 1. Definition 1.9 generalizes [SwaHun06, Definition 2.1.1] from commutative ring extensions to arbitrary algebras, and generalizes [AllKle14, Definition (10.21)] from commutative $\mathbb{K}$-algebras $\mathbb{L}$ to arbitrary $\mathbb{K}$-algebras $\mathbb{L}$.

Philosophically, there is a similarity between integral elements of a $\mathbb{K}$-algebra, and "finite-order" elements of a semigroup (i.e., elements a such that the set $\{a^1, a^2, a^3, \ldots\}$ is finite). In Proposition 1.25, we shall see a direct connection between these two concepts, but even before that, the similarity is helpful as a guide.

Definition 1.10. Let $\mathbb{K}$ be a commutative ring. Let M be a $\mathbb{K}$-module, and let $n \in \mathbb{N}$.

(a) If $m_1, m_2, \ldots, m_n$ are n elements of M, then we let $\langle m_1, m_2, \ldots, m_n \rangle_{\mathbb{K}}$ denote the $\mathbb{K}$-submodule of M spanned by $m_1, m_2, \ldots, m_n$. This $\mathbb{K}$-submodule is called the $\mathbb{K}$-*linear span* of $m_1, m_2, \ldots, m_n$. A similar notation will be used for spans of infinitely many elements.

(b) We say that the $\mathbb{K}$-module M is *n-generated* if and only if there exist n elements $m_1, m_2, \ldots, m_n \in M$ such that $M = \langle m_1, m_2, \ldots, m_n \rangle_{\mathbb{K}}$.

We notice that a $\mathbb{K}$-module M is finitely generated if and only if there exists some $n \in \mathbb{N}$ such that M is n-generated.

We recall one basic fact about finitely generated $\mathbb{K}$-modules:

Lemma 1.11. Let $\mathbb{K}$ be a commutative ring. Let M and N be two $\mathbb{K}$-modules such that M is finitely generated. Let $f : M \to N$ be a surjective $\mathbb{K}$-module homomorphism. Then, the $\mathbb{K}$-module N is finitely generated.

Proof of Lemma 1.11. The $\mathbb{K}$-module M is finitely generated. In other words, there exists a finite list $(m_1, m_2, \ldots, m_n)$ of elements of M such that $M = \langle m_1, m_2, \ldots, m_n \rangle_{\mathbb{K}}$. Consider this list. From $M = \langle m_1, m_2, \ldots, m_n \rangle_{\mathbb{K}}$, we obtain

$$f(M) = f(\langle m_1, m_2, \ldots, m_n \rangle_{\mathbb{K}}) = \langle f(m_1), f(m_2), \ldots, f(m_n) \rangle_{\mathbb{K}}$$

(since f is a $\mathbb{K}$-module homomorphism). But $f(M) = N$ (since f is surjective). Hence, $N = f(M) = \langle f(m_1), f(m_2), \ldots, f(m_n) \rangle_{\mathbb{K}}$. Thus, the $\mathbb{K}$-module N is finitely generated. This proves Lemma 1.11. □

The following fact provides several criteria for when an element of a commutative $\mathbb{K}$-algebra is integral over $\mathbb{K}$:

Theorem 1.12. Let $\mathbb{K}$ be a commutative ring. Let $\mathbb{L}$ be a commutative $\mathbb{K}$-algebra. Let $n \in \mathbb{N}$. Let $u \in \mathbb{L}$. Then, the following assertions $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ and $\mathcal{D}$ are equivalent:

- *Assertion $\mathcal{A}$:* There exists a monic polynomial $f \in \mathbb{K}[t]$ of degree n such that $f(u) = 0$.
- *Assertion $\mathcal{B}$:* There exist an $\mathbb{L}$-module C and an n-generated $\mathbb{K}$-submodule U of C such that $uU \subseteq U$ and such that every $v \in \mathbb{L}$ satisfying $vU = 0$ satisfies $v = 0$. (Here, we are making use of the fact that each $\mathbb{L}$-module canonically becomes a $\mathbb{K}$-module, since $\mathbb{L}$ is a $\mathbb{K}$-algebra.)
- *Assertion $\mathcal{C}$:* There exists an n-generated $\mathbb{K}$-submodule U of $\mathbb{L}$ such that $1 \in U$ and $uU \subseteq U$.
- *Assertion $\mathcal{D}$:* We have $\mathbb{K}[u] = \left\langle u^0, u^1, \ldots, u^{n-1} \right\rangle_{\mathbb{K}}$.

Proof of Theorem 1.12. Theorem 1.12 is precisely [Grinbe19, Theorem 1.1] (with A, B, X and P renamed as $\mathbb{K}$, $\mathbb{L}$, t and f). □

Note that Theorem 1.12 is just one of several "determinantal tricks" used in studying integrality over rings. See [Bourba72, Chapter V, Section 1.1, Theorem 1] or [ChaLoi14, Theorem 8.1.6] for another. We shall only use the implications $\mathcal{B} \Longrightarrow \mathcal{A}$ and $\mathcal{A} \Longrightarrow \mathcal{D}$ of Theorem 1.12.

We shall draw the following conclusion from Theorem 1.12:

Corollary 1.13. Let $\mathbb{K}$ be a commutative ring. Let $\mathbb{L}$ be a commutative $\mathbb{K}$-algebra. Let $u \in \mathbb{L}$. Let C be an $\mathbb{L}$-module. Let U be a finitely generated $\mathbb{K}$-submodule of C such that $uU \subseteq U$. Assume that every $v \in \mathbb{L}$ satisfying $vU = 0$ satisfies $v = 0$. (Here, we are making use of the fact that each $\mathbb{L}$-module canonically becomes a $\mathbb{K}$-module, since $\mathbb{L}$ is a $\mathbb{K}$-algebra.)

Then, $u \in \mathbb{L}$ is integral over $\mathbb{K}$.

First proof of Corollary 1.13. The $\mathbb{K}$-module U is finitely generated. In other words, it is n-generated for some $n \in \mathbb{N}$. Consider this n. Thus, Assertion $\mathcal{B}$ of Theorem 1.12 is satisfied. Hence, Assertion $\mathcal{A}$ of Theorem 1.12 is satisfied as well (since Theorem 1.12 shows that these two assertions are equivalent). In other words, there exists a monic polynomial $f \in \mathbb{K}[t]$ of degree n such that $f(u) = 0$. Hence, u is integral over $\mathbb{K}$. This proves Corollary 1.13. □

Second proof of Corollary 1.13 (sketched). The $\mathbb{K}$-module U is finitely generated. In other words, it is n-generated for some $n \in \mathbb{N}$. Consider this n.

Consider the commutative $\mathbb{K}$-subalgebra $\mathbb{K}[u]$ of $\mathbb{L}$. Then, the $\mathbb{L}$-module C is a $\mathbb{K}[u]$-module (by restriction). We have $uU \subseteq U$. Using this fact, it is easy to prove (by induction on k) that $u^k U \subseteq U$ for each $k \in \mathbb{N}$. Hence, $fU \subseteq U$ for each $f \in \mathbb{K}[u]$ (since each $f \in \mathbb{K}[u]$ is a $\mathbb{K}$-linear combination of the elements $u^0, u^1, u^2, \ldots$, and since U is a $\mathbb{K}$-module). Thus, U is a $\mathbb{K}[u]$-submodule of C.

Moreover, we assumed that every $v \in \mathbb{L}$ satisfying $vU = 0$ satisfies $v = 0$. Hence, every $v \in \mathbb{K}[u]$ satisfying $vU = 0$ satisfies $v = 0$ (since $v \in \mathbb{K}[u] \subseteq \mathbb{L}$). In the parlance of commutative algebra, this is saying that the $\mathbb{K}[u]$-module U is faithful. Hence, there is a faithful $\mathbb{K}[u]$-module which is n-generated when considered as a $\mathbb{K}$-module (namely, U). Thus, [AllKle14, Proposition (10.23), implication (4) $\Longrightarrow$ (1)] (applied to $R = \mathbb{K}$, $R' = \mathbb{K}[u]$ and $x = u$) shows that there exists a monic polynomial $f \in \mathbb{K}[t]$ of degree n such that $f(u) = 0$. Hence, u is integral over $\mathbb{K}$. This proves Corollary 1.13 again. □

The following proposition is a linear analogue of Proposition 1.8:

Proposition 1.14. Let $\mathbb{K}$ be a commutative ring. Let $\mathbb{L}$ be a $\mathbb{K}$-algebra. Let $u \in \mathbb{L}$ be integral over $\mathbb{K}$. Then, there exists a $g \in \mathbb{N}$ such that $\mathbb{K}[u] = \left\langle u^0, u^1, \ldots, u^{g-1} \right\rangle_{\mathbb{K}}$.

Proposition 1.14 appears, e.g., in [AllKle14, Proposition (10.23), implication (1) $\Longrightarrow$ (2)]. For the sake of self-containedness, let us prove it as well:

Proof of Proposition 1.14. We first observe that $\mathbb{K}[u]$ is a $\mathbb{K}$-subalgebra of $\mathbb{L}$, and that $u \in \mathbb{K}[u]$. The meaning of our assumption "u is integral over $\mathbb{K}$", and also of our claim "there exists a $g \in \mathbb{N}$ such that $\mathbb{K}[u] = \left\langle u^0, u^1, \ldots, u^{g-1} \right\rangle_{\mathbb{K}}$", does not depend on whether we consider u as an element of $\mathbb{L}$ or as an element of $\mathbb{K}[u]$. Thus, for the rest of this proof, we can WLOG assume that $\mathbb{L} = \mathbb{K}[u]$ (since otherwise, we can simply replace $\mathbb{L}$ by $\mathbb{K}[u]$). Assume this. Then, $\mathbb{L}$ is commutative (since $\mathbb{K}[u]$ is clearly commutative).

We have assumed that u is integral over $\mathbb{K}$. In other words, there exists a monic polynomial $f \in \mathbb{K}[t]$ such that $f(u) = 0$. Consider this f. Set $n = \deg f$. Then, Assertion $\mathcal{A}$ of Theorem 1.12 is satisfied (since f has degree n). Hence, Assertion $\mathcal{D}$ of Theorem 1.12 is satisfied as well (since Theorem 1.12 shows that these two assertions are equivalent). In other words, we have $\mathbb{K}[u] = \left\langle u^0, u^1, \ldots, u^{n-1} \right\rangle_{\mathbb{K}}$. Thus, there exists a $g \in \mathbb{N}$ such that $\mathbb{K}[u] = \left\langle u^0, u^1, \ldots, u^{g-1} \right\rangle_{\mathbb{K}}$ (namely, $g = n$). This proves Proposition 1.14. □

Theorem 1.15. Let $\mathbb{K}$ be a commutative ring. Let $\mathbb{L}$ be a commutative $\mathbb{K}$-algebra. Let $u_1, u_2, \ldots, u_m$ be a finite list of elements of $\mathbb{L}$. Assume that these m elements $u_1, u_2, \ldots, u_m$ are all integral over $\mathbb{K}$, and generate $\mathbb{L}$ as a $\mathbb{K}$-algebra. Then, the $\mathbb{K}$-module $\mathbb{L}$ is finitely generated.

Theorem 1.15 appears, e.g., in [AllKle14, Theorem (10.28), implication (2) $\Longrightarrow$ (3)]. For the sake of self-containedness, let us prove it as well:

Proof of Theorem 1.15. Fix $i \in \{1, 2, \ldots, m\}$. Then, $u_i \in \mathbb{L}$ is integral over $\mathbb{K}$ (by assumption). Hence, Proposition 1.14 (applied to $u = u_i$) shows that there exists a $g \in \mathbb{N}$ such that $\mathbb{K}[u_i] = \left\langle u_i^0, u_i^1, \ldots, u_i^{g-1} \right\rangle_{\mathbb{K}}$. Thus, the $\mathbb{K}$-module $\mathbb{K}[u_i]$ is finitely generated.

Now, forget that we fixed i. We thus have shown that for each $i \in \{1, 2, \ldots, m\}$, the $\mathbb{K}$-module $\mathbb{K}[u_i]$ is finitely generated. Hence, the $\mathbb{K}$-module[2] $\mathbb{K}[u_1] \otimes \mathbb{K}[u_2] \otimes \cdots \otimes \mathbb{K}[u_m]$ is also finitely generated[3]. The $\mathbb{K}$-module homomorphism

$$\begin{aligned} \pi : \mathbb{K}[u_1] \otimes \mathbb{K}[u_2] \otimes \cdots \otimes \mathbb{K}[u_m] &\to \mathbb{L}, \\ a_1 \otimes a_2 \otimes \cdots \otimes a_m &\mapsto a_1 a_2 \cdots a_m \end{aligned}$$

is surjective (since $u_1, u_2, \ldots, u_m$ generate $\mathbb{L}$ as a $\mathbb{K}$-algebra[4]). Hence, Lemma 1.11 (applied to $M = \mathbb{K}[u_1] \otimes \mathbb{K}[u_2] \otimes \cdots \otimes \mathbb{K}[u_m]$, $N = \mathbb{L}$ and $f = \pi$) shows that the $\mathbb{K}$-module $\mathbb{L}$ is finitely generated. This proves Theorem 1.15. $\square$

1.4. Characterizing integral matrices

The following is a simple consequence of the Cayley–Hamilton theorem:

[2] In this proof, the "$\otimes$" symbol always means a tensor product over $\mathbb{K}$.

[3] This is because of the following general fact: If $A_1, A_2, \ldots, A_m$ are m finitely generated $\mathbb{K}$-modules, then the $\mathbb{K}$-module $A_1 \otimes A_2 \otimes \cdots \otimes A_m$ is also finitely generated. (This fact can be proven as follows: For each $i \in \{1, 2, \ldots, m\}$, we fix a finite family $(a_{i,s})_{s \in S_i}$ of vectors in A_i that generates A_i. Then, the family

$$\left(a_{1,s_1} \otimes a_{2,s_2} \otimes \cdots \otimes a_{m,s_m}\right)_{(s_1, s_2, \ldots, s_m) \in S_1 \times S_2 \times \cdots \times S_m}$$

of vectors in $A_1 \otimes A_2 \otimes \cdots \otimes A_m$ is finite and generates the $\mathbb{K}$-module $A_1 \otimes A_2 \otimes \cdots \otimes A_m$. Hence, $A_1 \otimes A_2 \otimes \cdots \otimes A_m$ is finitely generated.)

[4] To be more precise: The image of π is a $\mathbb{K}$-submodule of $\mathbb{L}$ (since π is a $\mathbb{K}$-module homomorphism). But we have assumed that $u_1, u_2, \ldots, u_m$ generate $\mathbb{L}$ as a $\mathbb{K}$-algebra. Thus, each element of $\mathbb{L}$ can be written as a polynomial in the $u_1, u_2, \ldots, u_m$ with coefficients in $\mathbb{K}$ (since $\mathbb{L}$ is commutative). In other words, each element of $\mathbb{L}$ can be written as a $\mathbb{K}$-linear combination of products of the form $u_1^{n_1} u_2^{n_2} \cdots u_m^{n_m}$ with $n_1, n_2, \ldots, n_m \in \mathbb{N}$. But each of the latter products belongs to the image of π (because $u_1^{n_1} u_2^{n_2} \cdots u_m^{n_m} = \pi\left(u_1^{n_1} \otimes u_2^{n_2} \otimes \cdots \otimes u_m^{n_m}\right)$ for all $n_1, n_2, \ldots, n_m \in \mathbb{N}$). Hence, each element of $\mathbb{L}$ can be written as a $\mathbb{K}$-linear combination of elements of the image of π, and thus itself belongs to the image of π (since the image of π is a $\mathbb{K}$-submodule of $\mathbb{L}$). In other words, π is surjective.

Proposition 1.16. Let $\mathbb{K}$ be a commutative ring. Let $n \in \mathbb{N}$. Let A be an $n \times n$-matrix over $\mathbb{K}$. Then, A is integral over $\mathbb{K}$ (as an element of the $\mathbb{K}$-algebra $\mathbb{K}^{n \times n}$).

Proof of Proposition 1.16. The characteristic polynomial χ_A of A is a monic polynomial in $\mathbb{K}[t]$. The Cayley–Hamilton theorem yields $\chi_A(A) = 0$. Thus, there exists a monic polynomial $f \in \mathbb{K}[t]$ such that $f(A) = 0$ (namely, $f = \chi_A$). In other words, A is integral over $\mathbb{K}$. This proves Proposition 1.16. $\square$

It is not hard to prove a generalization of Proposition 1.16:

Proposition 1.17. Let $\mathbb{K}$ be a commutative ring. Let $n \in \mathbb{N}$. Let $\mathbb{L}$ be a commutative $\mathbb{K}$-algebra. Let A be an $n \times n$-matrix over $\mathbb{L}$. Assume that each coefficient of the characteristic polynomial $\chi_A \in \mathbb{L}[t]$ is integral over $\mathbb{K}$. Then, A is integral over $\mathbb{K}$ (as an element of the $\mathbb{K}$-algebra $\mathbb{L}^{n \times n}$).

We will not need this proposition, so we banish its proof into Section 1.11. However, we will use its converse:

Theorem 1.18. Let $\mathbb{K}$ be a commutative ring. Let $n \in \mathbb{N}$. Let $\mathbb{L}$ be a commutative $\mathbb{K}$-algebra. Let A be an $n \times n$-matrix over $\mathbb{L}$. Assume that A is integral over $\mathbb{K}$ (as an element of the $\mathbb{K}$-algebra $\mathbb{L}^{n \times n}$). Then, each coefficient of the characteristic polynomial $\chi_A \in \mathbb{L}[t]$ is integral over $\mathbb{K}$.

Proposition 1.17 and Theorem 1.18 are parts of [Bourba72, Chapter V, Section 1.6, Proposition 17, (a) $\Longleftrightarrow$ (c)]. For convenience and expository value, we shall nevertheless reprove them here.

1.5. Gert Almkvist's exterior-power trick

Our following proof of Theorem 1.18 will rely on the notion of exterior powers of an $\mathbb{L}$-module (where $\mathbb{L}$ is a commutative ring). See [Bourba74, Chapter III, §7] or [Conrad13a] for the relevant background. Our method is inspired by Gert Almkvist's exterior-power trick ([Almkvi73, proof of Theorem 1.7], [Zeilbe93]). We shall need the following proposition (which is essentially the equality $(*')$ in [Zeilbe93], or the equality $(**)$ in [Almkvi73, proof of Theorem 1.7]):

Proposition 1.19. Let $\mathbb{K}$ be a commutative ring. Let $n \in \mathbb{N}$. Let $A \in \mathbb{K}^{n \times n}$ be an $n \times n$-matrix. Let V be the free $\mathbb{K}$-module $\mathbb{K}^n$ (consisting of column vectors of size n). Consider A as an endomorphism of the free $\mathbb{K}$-module $V = \mathbb{K}^n$ (in the usual way: i.e., we let $A(v) = Av$ for each column vector $v \in \mathbb{K}^n$). Consider the n-th exterior power $\Lambda^n V$ of the $\mathbb{K}$-module V.

Fix $k \in \mathbb{N}$. Let $a_k \in \mathbb{K}$ be the coefficient of t^k in the characteristic polynomial $\chi_A \in \mathbb{K}[t]$. Then, for each $w_1, w_2, \ldots, w_n \in V$, we have

$$a_k \cdot w_1 \wedge w_2 \wedge \cdots \wedge w_n = (-1)^{n-k} \sum_{\substack{i_1, i_2, \ldots, i_n \in \{0,1\}; \\ i_1 + i_2 + \cdots + i_n = n-k}} A^{i_1} w_1 \wedge A^{i_2} w_2 \wedge \cdots \wedge A^{i_n} w_n.$$

Before we prove this proposition, we need a well-known lemma that connects exterior powers with determinants:

Lemma 1.20. Let $\mathbb{L}$ be a commutative ring. Let $n \in \mathbb{N}$. If M is an $\mathbb{L}$-module, then $\Lambda^n_{\mathbb{L}} M$ shall denote the n-th exterior power of the $\mathbb{L}$-module M.

(a) If M is an $\mathbb{L}$-module, then each endomorphism u of the $\mathbb{L}$-module M induces an endomorphism $\Lambda^n_{\mathbb{L}} u$ of the $\mathbb{L}$-module $\Lambda^n_{\mathbb{L}} M$, defined by

$$(\Lambda^n_{\mathbb{L}} u)(w_1 \wedge w_2 \wedge \cdots \wedge w_n) = uw_1 \wedge uw_2 \wedge \cdots \wedge uw_n \quad \text{for all } w_1, w_2, \ldots, w_n \in M.$$

(b) For each endomorphism u of the free $\mathbb{L}$-module $\mathbb{L}^n$ and each $p \in \Lambda^n_{\mathbb{L}}(\mathbb{L}^n)$, we have

$$(\Lambda^n_{\mathbb{L}} u)\, p = \det u \cdot p. \tag{3}$$

(Here, $\det u$ stands for the determinant of u, which is defined as the determinant of the $n \times n$-matrix representing u.)

Proof of Lemma 1.20. Lemma 1.20 **(a)** appears (e.g.) in [Bourba74, Chapter III, §7.2, equation (4)] and in [Conrad13a, Theorem 5.1]. Lemma 1.20 **(b)** appears (e.g.) in [Conrad13a, Theorem 6.1] and (implicitly) in [Bourba74] as well (indeed, the formula (3) is how $\det u$ is defined in [Bourba74, Chapter III, §8, Section 1, Definition 1] (applied to $A = \mathbb{L}$ and $M = \mathbb{L}^n$)). □

We shall also need a lemma about n-th exterior powers of free modules of rank n:

Lemma 1.21. Let $\mathbb{L}$ be a commutative ring. Let $n \in \mathbb{N}$. Let M be an $\mathbb{L}$-module with a basis $(b_1, b_2, \ldots, b_n)$. Then, the 1-tuple $(b_1 \wedge b_2 \wedge \cdots \wedge b_n)$ is a basis of the $\mathbb{L}$-module $\Lambda^n_{\mathbb{L}} M$. (Here, $\Lambda^n_{\mathbb{L}} M$ denotes the n-th exterior power of the $\mathbb{L}$-module M.)

Proof of Lemma 1.21. See [Conrad13a, Theorem 4.2] or [Bourba74, Chapter III, §7.9, Corollary 1]. □

We shall now give a proof of Proposition 1.19; a second proof (somewhat more elementary, but more laborious) will be provided in Section 1.10.

First proof of Proposition 1.19. Note that V is a free $\mathbb{K}$-module of rank n; thus, $\Lambda^n V$ is a free $\mathbb{K}$-module of rank $\binom{n}{n} = 1$. In other words, $\Lambda^n V \cong \mathbb{K}$ as a $\mathbb{K}$-module.

Let $\mathbb{L}$ be the polynomial ring $\mathbb{K}[t]$. Then, $\mathbb{L}$ is a commutative $\mathbb{K}$-algebra. We consider the commutative ring $\mathbb{K}$ as a subring of the polynomial ring $\mathbb{L} = \mathbb{K}[t]$ (which is also commutative). Thus, the free $\mathbb{K}$-module $V = \mathbb{K}^n$ canonically embeds into the free $\mathbb{L}$-module $\mathbb{L}^n$, and its n-th exterior power[5] $\Lambda^n V$ canonically embeds into

[5] In the following, the symbol "Λ^n" without a subscript will always mean an n-th exterior power over the base ring $\mathbb{K}$.

the corresponding n-th exterior power $\Lambda_{\mathbb{L}}^{n}\left(\mathbb{L}^{n}\right)$, where the subscript "$_{\mathbb{L}}$" signals that this is an exterior power over the base ring $\mathbb{L}$. (This is indeed an embedding, since both modules $\Lambda^{n}V$ and $\Lambda_{\mathbb{L}}^{n}\left(\mathbb{L}^{n}\right)$ have bases consisting of 1 element only (by Lemma 1.21), and the canonical map $\Lambda^{n}V \to \Lambda_{\mathbb{L}}^{n}\left(\mathbb{L}^{n}\right)$ sends the basis of one to the basis of the other.)

But $\mathbb{L} \otimes_{\mathbb{K}} \underbrace{V}_{=\mathbb{K}^{n}} = \mathbb{L} \otimes_{\mathbb{K}} \mathbb{K}^{n} \cong \mathbb{L}^{n}$ as $\mathbb{L}$-modules. Hence, the canonical $\mathbb{L}$-module homomorphism

$$\begin{gathered}\mathbb{L} \otimes_{\mathbb{K}} \left(\Lambda^{n}V\right) \to \Lambda_{\mathbb{L}}^{n}\left(\mathbb{L}^{n}\right), \\ \ell \otimes \left(w_{1} \wedge w_{2} \wedge \cdots \wedge w_{n}\right) \mapsto \ell \cdot \left(w_{1} \wedge w_{2} \wedge \cdots \wedge w_{n}\right)\end{gathered}$$

is an $\mathbb{L}$-module isomorphism (see, e.g., [Conrad13b, Theorem 1]; see also [Bourba74, Chapter III, §7.5, Proposition 8] for the inverse of this isomorphism). We use this isomorphism to identify the $\mathbb{L}$-module $\Lambda_{\mathbb{L}}^{n}\left(\mathbb{L}^{n}\right)$ with $\mathbb{L} \otimes_{\mathbb{K}} \left(\Lambda^{n}V\right)$. Thus,

$$\Lambda_{\mathbb{L}}^{n}\left(\mathbb{L}^{n}\right) = \underbrace{\mathbb{L}}_{=\mathbb{K}[t]} \otimes_{\mathbb{K}} \left(\Lambda^{n}V\right) = \mathbb{K}\left[t\right] \otimes_{\mathbb{K}} \left(\Lambda^{n}V\right) \cong \left(\Lambda^{n}V\right)\left[t\right]. \tag{4}$$

Concretely, this means that every element of $\Lambda_{\mathbb{L}}^{n}\left(\mathbb{L}^{n}\right)$ can be written as a polynomial in t with coefficients in $\Lambda^{n}V$.

Note that our canonical embedding $\Lambda^{n}V \hookrightarrow \Lambda_{\mathbb{L}}^{n}\left(\mathbb{L}^{n}\right)$ sends each $p \in \Lambda^{n}V$ to $1 \otimes p \in \mathbb{L} \otimes_{\mathbb{K}} \left(\Lambda^{n}V\right) = \Lambda_{\mathbb{L}}^{n}\left(\mathbb{L}^{n}\right)$.

Consider the matrix $tI_{n} - A \in \left(\mathbb{K}\left[t\right]\right)^{n\times n} = \mathbb{L}^{n\times n}$ (since $\mathbb{K}\left[t\right] = \mathbb{L}$). It satisfies

$$tI_{n} - A = \sum_{i\in\{0,1\}} t^{1-i}\left(-1\right)^{i} A^{i}$$

(since $\sum_{i\in\{0,1\}} t^{1-i}\left(-1\right)^{i} A^{i} = \underbrace{t^{1-0}}_{=t}\underbrace{\left(-1\right)^{0}}_{=1}\underbrace{A^{0}}_{=I_{n}} + \underbrace{t^{1-1}}_{=t^{0}=1}\underbrace{\left(-1\right)^{1}}_{=-1}\underbrace{A^{1}}_{=A} = tI_{n} + \left(-1\right)A = tI_{n} - A$).

The $n \times n$-matrix $tI_{n} - A \in \mathbb{L}^{n\times n}$ can be viewed as an endomorphism of the free $\mathbb{L}$-module $\mathbb{L}^{n}$ (since any $n \times n$-matrix over $\mathbb{L}$ can be viewed as such an endomorphism). Applying (3) to $u = tI_{n} - A$ (or, more precisely, to the $\mathbb{L}$-module endomorphism we just mentioned), we obtain

$$\left(\Lambda_{\mathbb{L}}^{n}\left(tI_{n} - A\right)\right)p = \det\left(tI_{n} - A\right) \cdot p \tag{5}$$

for each $p \in \Lambda_{\mathbb{L}}^{n}\left(\mathbb{L}^{n}\right)$. (Here, of course, the meaning of $\Lambda_{\mathbb{L}}^{n}\left(tI_{n} - A\right)$ is as in Lemma 1.20.)

Now, fix $w_{1}, w_{2}, \ldots, w_{n} \in V$, and set $p = w_{1} \wedge w_{2} \wedge \cdots \wedge w_{n} \in \Lambda^{n}V$. Note that $p \in \Lambda^{n}V \subseteq \Lambda_{\mathbb{L}}^{n}\left(\mathbb{L}^{n}\right)$.

Definition 1.1 yields $\chi_A = \det(tI_n - A)$. Hence,

$$\begin{aligned}
&\chi_A \cdot p \\
&= \det(tI_n - A) \cdot p \\
&= \left(\Lambda^n_{\mathbb{L}}(tI_n - A)\right) p \qquad \text{(by (5))} \\
&= \left(\Lambda^n_{\mathbb{L}}(tI_n - A)\right)(w_1 \wedge w_2 \wedge \cdots \wedge w_n) \qquad (\text{since } p = w_1 \wedge w_2 \wedge \cdots \wedge w_n) \\
&= (tI_n - A)\, w_1 \wedge (tI_n - A)\, w_2 \wedge \cdots \wedge (tI_n - A)\, w_n \\
&\qquad \text{(by the definition of } \Lambda^n_{\mathbb{L}}(tI_n - A)) \\
&= \left(\sum_{i \in \{0,1\}} t^{1-i}(-1)^i A^i\right) w_1 \wedge \left(\sum_{i \in \{0,1\}} t^{1-i}(-1)^i A^i\right) w_2 \wedge \cdots \\
&\qquad \wedge \left(\sum_{i \in \{0,1\}} t^{1-i}(-1)^i A^i\right) w_n \qquad \left(\text{since } tI_n - A = \sum_{i \in \{0,1\}} t^{1-i}(-1)^i A^i\right) \\
&= \sum_{i_1, i_2, \ldots, i_n \in \{0,1\}} \underbrace{t^{1-i_1}(-1)^{i_1} A^{i_1} w_1 \wedge t^{1-i_2}(-1)^{i_2} A^{i_2} w_2 \wedge \cdots \wedge t^{1-i_n}(-1)^{i_n} A^{i_n} w_n}_{\substack{= t^{(1-i_1)+(1-i_2)+\cdots+(1-i_n)}(-1)^{i_1+i_2+\cdots+i_n} \cdot A^{i_1} w_1 \wedge A^{i_2} w_2 \wedge \cdots \wedge A^{i_n} w_n \\ \text{(by the multilinearity of the exterior product)}}} \\
&\qquad \text{(by the multilinearity of the exterior product)} \\
&= \sum_{i_1, i_2, \ldots, i_n \in \{0,1\}} \underbrace{t^{(1-i_1)+(1-i_2)+\cdots+(1-i_n)}}_{= t^{n-(i_1+i_2+\cdots+i_n)}} (-1)^{i_1+i_2+\cdots+i_n} \cdot A^{i_1} w_1 \wedge A^{i_2} w_2 \wedge \cdots \wedge A^{i_n} w_n \\
&= \sum_{i_1, i_2, \ldots, i_n \in \{0,1\}} t^{n-(i_1+i_2+\cdots+i_n)} (-1)^{i_1+i_2+\cdots+i_n} \cdot A^{i_1} w_1 \wedge A^{i_2} w_2 \wedge \cdots \wedge A^{i_n} w_n.
\end{aligned}$$

This is an equality in $\Lambda^n_{\mathbb{L}}(\mathbb{L}^n)$. In view of (4), this becomes an equality in $(\Lambda^n V)[t]$. Hence, by comparing the coefficients of t^k on both sides of this equality, we obtain

$$a_k \cdot p = \sum_{\substack{i_1, i_2, \ldots, i_n \in \{0,1\}; \\ n-(i_1+i_2+\cdots+i_n)=k}} (-1)^{i_1+i_2+\cdots+i_n} \cdot A^{i_1} w_1 \wedge A^{i_2} w_2 \wedge \cdots \wedge A^{i_n} w_n$$

(since the coefficient of t^k in χ_A is a_k, and thus the coefficient of t^k in $\chi_A \cdot p$ is $a_k \cdot p$). Since the condition "$n - (i_1 + i_2 + \cdots + i_n) = k$" under the summation sign is equivalent to "$i_1 + i_2 + \cdots + i_n = n - k$", we can rewrite this as follows:

$$\begin{aligned}
a_k \cdot p &= \sum_{\substack{i_1, i_2, \ldots, i_n \in \{0,1\}; \\ i_1+i_2+\cdots+i_n=n-k}} \underbrace{(-1)^{i_1+i_2+\cdots+i_n}}_{\substack{= (-1)^{n-k} \\ (\text{since } i_1+i_2+\cdots+i_n=n-k)}} \cdot A^{i_1} w_1 \wedge A^{i_2} w_2 \wedge \cdots \wedge A^{i_n} w_n \\
&= (-1)^{n-k} \sum_{\substack{i_1, i_2, \ldots, i_n \in \{0,1\}; \\ i_1+i_2+\cdots+i_n=n-k}} A^{i_1} w_1 \wedge A^{i_2} w_2 \wedge \cdots \wedge A^{i_n} w_n.
\end{aligned}$$

In view of $p = w_1 \wedge w_2 \wedge \cdots \wedge w_n$, this rewrites as

$$a_k \cdot w_1 \wedge w_2 \wedge \cdots \wedge w_n = (-1)^{n-k} \sum_{\substack{i_1, i_2, \ldots, i_n \in \{0,1\}; \\ i_1 + i_2 + \cdots + i_n = n-k}} A^{i_1} w_1 \wedge A^{i_2} w_2 \wedge \cdots \wedge A^{i_n} w_n.$$

This proves Proposition 1.19. □

1.6. Characterizing integral matrices: the proof

Now, everything needed for proving Theorem 1.18 is in place, so we can start the proof:

Proof of Theorem 1.18. Let V be the free $\mathbb{L}$-module $\mathbb{L}^n$. In the following, $\Lambda^n V$ shall always mean the n-th exterior power of the $\mathbb{L}$-module V.

We have assumed that A is integral over $\mathbb{K}$ (as an element of the $\mathbb{K}$-algebra $\mathbb{L}^{n \times n}$). Hence, Proposition 1.14 (applied to $\mathbb{L}^{n \times n}$ and A instead of $\mathbb{L}$ and u) shows that there exists a $g \in \mathbb{N}$ such that $\mathbb{K}[A] = \left\langle A^0, A^1, \ldots, A^{g-1} \right\rangle_{\mathbb{K}}$. Consider this g.

Let C be the $\mathbb{L}$-module $\Lambda^n V$. It is easy (using Lemma 1.21) to show that $C \cong \mathbb{L}$ as an $\mathbb{L}$-module, but we shall not use this.

Let U be the $\mathbb{K}$-submodule of C spanned by all elements of the form

$$B_1 v_1 \wedge B_2 v_2 \wedge \cdots \wedge B_n v_n \qquad \text{with } B_j \in \mathbb{K}[A] \text{ and } v_j \in \mathbb{K}^n. \tag{6}$$

Here, we are using the canonical $\mathbb{K}$-linear map $\mathbb{K}^n \to \mathbb{L}^n$ (induced by the $\mathbb{K}$-algebra morphism $\mathbb{K} \to \mathbb{L}$ given by the $\mathbb{K}$-algebra structure on $\mathbb{L}$) to turn the vectors $v_j \in \mathbb{K}^n$ into vectors in $\mathbb{L}^n$ (so that the matrices $B_j \in \mathbb{K}[A] \subseteq \mathbb{L}^{n \times n}$ can be multiplied onto them, yielding new vectors $B_j v_j \in \mathbb{L}^n$). Now, we claim the following:

Claim 1: The $\mathbb{K}$-module U is finitely generated.

[*First proof of Claim 1:* Let $(e_1, e_2, \ldots, e_n)$ be the standard basis of the $\mathbb{K}$-module $\mathbb{K}^n$ (so that e_i is the vector with a 1 in its i-th entry and 0 everywhere else).

We have $\mathbb{K}[A] = \left\langle A^0, A^1, \ldots, A^{g-1} \right\rangle_{\mathbb{K}}$. Thus, each $B \in \mathbb{K}[A]$ is a $\mathbb{K}$-linear combination of $A^0, A^1, \ldots, A^{g-1}$. We can thus easily see that each element of the form (6) is a $\mathbb{K}$-linear combination of elements of the form

$$A^{i_1} v_1 \wedge A^{i_2} v_2 \wedge \cdots \wedge A^{i_n} v_n \qquad \text{with } i_j \in \{0, 1, \ldots, g-1\} \text{ and } v_j \in \mathbb{K}^n. \tag{7}$$

Thus, the $\mathbb{K}$-module U is spanned by all elements of the form (7) (since it is spanned by all elements of the form (6)).

But $(e_1, e_2, \ldots, e_n)$ is the standard basis of the $\mathbb{K}$-module $\mathbb{K}^n$. Thus, the elements $e_1, e_2, \ldots, e_n$ span the $\mathbb{K}$-module $\mathbb{K}^n$. Hence, each $v \in \mathbb{K}^n$ is a $\mathbb{K}$-linear combination

of $e_1, e_2, \ldots, e_n$. Hence, each element of the form (7) is a $\mathbb{K}$-linear combination of elements of the form

$$A^{i_1} v_1 \wedge A^{i_2} v_2 \wedge \cdots \wedge A^{i_n} v_n \qquad \text{with } i_j \in \{0, 1, \ldots, g-1\} \text{ and } v_j \in \{e_1, e_2, \ldots, e_n\}. \tag{8}$$

Thus, the $\mathbb{K}$-module U is spanned by all elements of the form (8) (since it is spanned by all elements of the form (7)). Hence, the $\mathbb{K}$-module U is finitely generated (since there are only finitely many elements of the form (8)). This proves Claim 1.]

[*Second proof of Claim 1:* Here is a more formal way of stating the same proof. The $\mathbb{K}$-module $\mathbb{K}[A]$ is finitely generated (since $\mathbb{K}[A] = \left\langle A^0, A^1, \ldots, A^{g-1} \right\rangle_{\mathbb{K}}$), and so is the $\mathbb{K}$-module $\mathbb{K}^n$. Hence, the $\mathbb{K}$-module[6]

$$M := \underbrace{\mathbb{K}[A] \otimes \mathbb{K}[A] \otimes \cdots \otimes \mathbb{K}[A]}_{n \text{ times}} \otimes \underbrace{\mathbb{K}^n \otimes \mathbb{K}^n \otimes \cdots \otimes \mathbb{K}^n}_{n \text{ times}}$$

is finitely generated as well[7].

Now, for any $B_1, B_2, \ldots, B_n \in \mathbb{K}[A]$ and any $v_1, v_2, \ldots, v_n \in \mathbb{K}^n$, the element $B_1 v_1 \wedge B_2 v_2 \wedge \cdots \wedge B_n v_n \in C$ belongs to U (since it is an element of the form (6)). Thus, the $\mathbb{K}$-linear map

$$\pi : M \to U,$$
$$B_1 \otimes B_2 \otimes \cdots \otimes B_n \otimes v_1 \otimes v_2 \otimes \cdots \otimes v_n \mapsto B_1 v_1 \wedge B_2 v_2 \wedge \cdots \wedge B_n v_n$$

is well-defined. This $\mathbb{K}$-linear map π is surjective (since each element of the form (6) belongs to its image[8]). Thus, Lemma 1.11 (applied to $N = U$ and $f = \pi$) shows that the $\mathbb{K}$-module U is finitely generated (since the $\mathbb{K}$-module M is finitely generated). This proves Claim 1 again.]

Now, fix $k \in \mathbb{N}$. Let $a_k \in \mathbb{L}$ be the coefficient of t^k in the characteristic polynomial $\chi_A \in \mathbb{L}[t]$. We are going to show the following:

Claim 2: We have $a_k U \subseteq U$.

[*Proof of Claim 2:* It suffices to show that $a_k u \in U$ for each $u \in U$. So let us fix $u \in U$; thus we must prove that $a_k u \in U$.

We can WLOG assume that u is an element of the form (6) (since U is spanned by elements of this form). So assume this. Then,

$$u = B_1 v_1 \wedge B_2 v_2 \wedge \cdots \wedge B_n v_n \qquad \text{for some } B_j \in \mathbb{K}[A] \text{ and } v_j \in \mathbb{K}^n.$$

[6] In this proof, the "$\otimes$" symbol always means a tensor product over $\mathbb{K}$.

[7] This is because of the following general fact: If $A_1, A_2, \ldots, A_m$ are m finitely generated $\mathbb{K}$-modules, then the $\mathbb{K}$-module $A_1 \otimes A_2 \otimes \cdots \otimes A_m$ is also finitely generated. (We already proved this fact while proving Theorem 1.15.)

[8] Namely:

$$B_1 v_1 \wedge B_2 v_2 \wedge \cdots \wedge B_n v_n = \pi\left(B_1 \otimes B_2 \otimes \cdots \otimes B_n \otimes v_1 \otimes v_2 \otimes \cdots \otimes v_n\right).$$

Consider these B_j and these v_j. Note that for each n-tuple $(i_1, i_2, \ldots, i_n) \in \{0,1\}^n$, we have

$$A^{i_1} B_1 v_1 \wedge A^{i_2} B_2 v_2 \wedge \cdots \wedge A^{i_n} B_n v_n \in U, \tag{9}$$

since $A^{i_1} B_1 v_1 \wedge A^{i_2} B_2 v_2 \wedge \cdots \wedge A^{i_n} B_n v_n$ is an element of the form (6) (because each $j \in \{1, 2, \ldots, n\}$ satisfies $\underbrace{A^{i_j}}_{\in \mathbb{K}[A]} \underbrace{B_j}_{\in \mathbb{K}[A]} \in \mathbb{K}[A] \cdot \mathbb{K}[A] \subseteq \mathbb{K}[A]$).

Multiplying both sides of the equality $u = B_1 v_1 \wedge B_2 v_2 \wedge \cdots \wedge B_n v_n$ with $a_k \in \mathbb{L}$, we obtain

$$\begin{aligned} a_k u &= a_k \cdot B_1 v_1 \wedge B_2 v_2 \wedge \cdots \wedge B_n v_n \\ &= (-1)^{n-k} \sum_{\substack{i_1, i_2, \ldots, i_n \in \{0,1\}; \\ i_1 + i_2 + \cdots + i_n = n-k}} \underbrace{A^{i_1} B_1 v_1 \wedge A^{i_2} B_2 v_2 \wedge \cdots \wedge A^{i_n} B_n v_n}_{\substack{\in U \\ \text{(by (9))}}} \\ &\qquad \left(\text{by Proposition 1.19, applied to } \mathbb{L} \text{ and } B_j v_j \text{ instead of } \mathbb{K} \text{ and } w_j\right) \\ &\in (-1)^{n-k} \sum_{\substack{i_1, i_2, \ldots, i_n \in \{0,1\}; \\ i_1 + i_2 + \cdots + i_n = n-k}} U \subseteq U \end{aligned}$$

(since U is a $\mathbb{K}$-module). This completes the proof of Claim 2.]

Claim 3: Every $v \in \mathbb{L}$ satisfying $vU = 0$ satisfies $v = 0$.

[*Proof of Claim 3:* Let $v \in \mathbb{L}$ satisfy $vU = 0$. We must prove that $v = 0$.

Let $(e_1, e_2, \ldots, e_n)$ be the standard basis of the $\mathbb{L}$-module $\mathbb{L}^n$ (so that e_i is the vector with a 1 in its i-th entry and 0 everywhere else). Also, let $(e'_1, e'_2, \ldots, e'_n)$ be the standard basis of the $\mathbb{K}$-module $\mathbb{K}^n$. The canonical mapping from $\mathbb{K}^n$ to $\mathbb{L}^n$ sends each e'_j to the corresponding e_j. Hence, the element $e_1 \wedge e_2 \wedge \cdots \wedge e_n$ of C has the form (6) (namely, for $B_j = I_n$ and $v_j = e'_j$). Hence, this element belongs to U (by the definition of U). In other words, $e_1 \wedge e_2 \wedge \cdots \wedge e_n \in U$. Thus, $v \underbrace{(e_1 \wedge e_2 \wedge \cdots \wedge e_n)}_{\in U} \in vU = 0$, so that $v(e_1 \wedge e_2 \wedge \cdots \wedge e_n) = 0$.

But recall that the n-tuple $(e_1, e_2, \ldots, e_n)$ is a basis of the $\mathbb{L}$-module $\mathbb{L}^n$. Hence, Lemma 1.21 (applied to $M = \mathbb{L}^n$ and $b_j = e_j$) shows that the 1-tuple $(e_1 \wedge e_2 \wedge \cdots \wedge e_n)$ is a basis of the $\mathbb{L}$-module $\Lambda^n \underbrace{(\mathbb{L}^n)}_{=V} = \Lambda^n V$. Hence, this 1-tuple is $\mathbb{L}$-linearly independent. In other words, every $w \in \mathbb{L}$ satisfying $w(e_1 \wedge e_2 \wedge \cdots \wedge e_n) = 0$ satisfies $w = 0$. Applying this to $w = v$, we obtain $v = 0$ (since $v(e_1 \wedge e_2 \wedge \cdots \wedge e_n) = 0$). This proves Claim 3.]

Now, Corollary 1.13 can be applied to $u = a_k$ (since Claim 1, Claim 2 and Claim 3 ensure that the assumptions of Corollary 1.13 are satisfied). Thus, we conclude that $a_k \in \mathbb{L}$ is integral over $\mathbb{K}$. In other words, the coefficient of t^k in the characteristic polynomial $\chi_A \in \mathbb{L}[t]$ is integral over $\mathbb{K}$ (since a_k was defined to be this coefficient).

Now, forget that we fixed k. We thus have shown that for each $k \in \mathbb{N}$, the coefficient of t^k in the characteristic polynomial $\chi_A \in \mathbb{L}[t]$ is integral over $\mathbb{K}$. This proves Theorem 1.18. $\square$

Corollary 1.22. Let $\mathbb{K}$ be a commutative ring. Let $n \in \mathbb{N}$. Let $\mathbb{L}$ be a commutative $\mathbb{K}$-algebra. Let A be an $n \times n$-matrix over $\mathbb{L}$. Assume that A is integral over $\mathbb{K}$ (as an element of the $\mathbb{K}$-algebra $\mathbb{L}^{n \times n}$). Let $\mathbb{M}$ be the $\mathbb{K}$-subalgebra of $\mathbb{L}$ generated by the coefficients of the characteristic polynomial $\chi_A \in \mathbb{L}[t]$. Then, $\mathbb{M}$ is a finitely generated $\mathbb{K}$-module.

Proof of Corollary 1.22. Let $u_1, u_2, \ldots, u_m$ be the coefficients of the polynomial χ_A. These coefficients $u_1, u_2, \ldots, u_m$ are integral over $\mathbb{K}$ (by Theorem 1.18), and generate $\mathbb{M}$ as a $\mathbb{K}$-algebra (by the definition of $\mathbb{M}$); thus, in particular, they are elements of $\mathbb{M}$. Hence, Theorem 1.15 (applied to $\mathbb{M}$ instead of $\mathbb{L}$) yields that the $\mathbb{K}$-module $\mathbb{M}$ is finitely generated. This proves Corollary 1.22. □

1.7. Two finiteness lemmas

We need two more lemmas about finite generation of certain modules:

Lemma 1.23. Let $\mathbb{K}$ be a finite commutative ring. Let M be a finitely generated $\mathbb{K}$-module. Then, M is finite (as a set).

Proof of Lemma 1.23. The $\mathbb{K}$-module M is finitely generated. In other words, there exist finitely many vectors $a_1, a_2, \ldots, a_m \in M$ that generate M as a $\mathbb{K}$-module. Consider these $a_1, a_2, \ldots, a_m$. Thus, each element of M is a $\mathbb{K}$-linear combination of $a_1, a_2, \ldots, a_m$ (since $a_1, a_2, \ldots, a_m$ generate M as a $\mathbb{K}$-module).

There exist only finitely many $\mathbb{K}$-linear combinations of $a_1, a_2, \ldots, a_m$ (because a $\mathbb{K}$-linear combination $\lambda_1 a_1 + \lambda_2 a_2 + \cdots + \lambda_m a_m$ of $a_1, a_2, \ldots, a_m$ is uniquely determined by choosing its m coefficients $\lambda_1, \lambda_2, \ldots, \lambda_m \in \mathbb{K}$, but each of these m coefficients can be chosen in only finitely many ways[9]). Hence, there are only finitely many elements of M (since each element of M is a $\mathbb{K}$-linear combination of $a_1, a_2, \ldots, a_m$). In other words, M is finite. This proves Lemma 1.23. □

Lemma 1.24. Let $\mathbb{K}$ be a commutative ring. Let $f \in \mathbb{K}[t]$ be a monic polynomial. Then, the $\mathbb{K}$-module $\mathbb{K}[t] / (f)$ is finitely generated.

First proof of Lemma 1.24. Much more can be said: For each $u \in \mathbb{K}[t]$, we let $\overline{u}$ denote the projection of u onto $\mathbb{K}[t] / (f)$. Then, the $\mathbb{K}$-module $\mathbb{K}[t] / (f)$ is free with basis $\left(\overline{t^0}, \overline{t^1}, \ldots, \overline{t^{n-1}}\right)$, where $n = \deg f$. This is a well-known fact[10] and follows easily from "Euclidean division of polynomials". Of course, this entails that the $\mathbb{K}$-module $\mathbb{K}[t] / (f)$ is finitely generated. This proves Lemma 1.24. □

Second proof of Lemma 1.24. For each $u \in \mathbb{K}[t]$, we let $\overline{u}$ denote the projection of u onto $\mathbb{K}[t] / (f)$. The element $\overline{t}$ of $\mathbb{K}[t] / (f)$ satisfies $f\left(\overline{t}\right) = \overline{f(t)} = 0$ (since

[9] since $\mathbb{K}$ is finite

[10] See, e.g., [Aluffi09, Chapter III, Proposition 4.6] for an equivalent version of this fact (restated in terms of an isomorphism $\mathbb{K}[t] / (f) \to \mathbb{K}^{\oplus n}$).

$f(t) = f \in (f)$). Hence, $\bar{t} \in \mathbb{K}[t] / (f)$ is integral over $\mathbb{K}$ (by the definition of "integral"), since the polynomial $f \in \mathbb{K}[t]$ is monic.

Moreover, the $\mathbb{K}$-algebra $\mathbb{K}[t]$ is generated by t; thus, its quotient $\mathbb{K}$-algebra $\mathbb{K}[t] / (f)$ is generated by $\bar{t}$. Hence, Theorem 1.15 (applied to $\mathbb{L} = \mathbb{K}[t] / (f)$, $m = 1$ and $u_1 = \bar{t}$) yields that the $\mathbb{K}$-module $\mathbb{K}[t] / (f)$ is finitely generated. This proves Lemma 1.24. □

1.8. Proof of Theorem 1.2

The following fact will bring us very close to Theorem 1.2:

Proposition 1.25. Let $\mathbb{K}$ be a finite commutative ring. Let $n \in \mathbb{N}$. Let $\mathbb{L}$ be a commutative $\mathbb{K}$-algebra. Let A be an $n \times n$-matrix over $\mathbb{L}$. Then, the following three assertions are equivalent:

- *Assertion $\mathcal{U}$:* The set $\{A^0, A^1, A^2, \ldots\}$ is finite.
- *Assertion $\mathcal{V}$:* The matrix A is integral over $\mathbb{K}$ (as an element of the $\mathbb{K}$-algebra $\mathbb{L}^{n \times n}$).
- *Assertion $\mathcal{W}$:* There exists a positive integer m such that the polynomial $t^{2m} - t^m$ is a multiple of χ_A in $\mathbb{L}[t]$.

Proof of Proposition 1.25. We shall prove the implications $\mathcal{U} \Longrightarrow \mathcal{V}$ and $\mathcal{V} \Longrightarrow \mathcal{W}$ and $\mathcal{W} \Longrightarrow \mathcal{U}$:

Proof of the implication $\mathcal{U} \Longrightarrow \mathcal{V}$: Assume that Assertion $\mathcal{U}$ holds. We must prove that Assertion $\mathcal{V}$ holds.

The set $\{A^0, A^1, A^2, \ldots\}$ is closed under multiplication. Thus, this set (equipped with multiplication) is a semigroup. Furthermore, this set is finite (since Assertion $\mathcal{U}$ holds), and thus is a finite semigroup. Hence, Theorem 1.7 (applied to $M = \{A^0, A^1, A^2, \ldots\}$ and $a = A$) shows that there exists a positive integer m such that $A^m = A^{2m}$ (since $A = A^1 \in \{A^0, A^1, A^2, \ldots\}$). Consider this m. Let $g \in \mathbb{K}[t]$ be the polynomial $t^{2m} - t^m$. Then, g is monic (since $m > 0$) and satisfies $g(A) = A^{2m} - A^m = 0$ (since $A^m = A^{2m}$). Hence, there exists a monic polynomial $f \in \mathbb{K}[t]$ such that $f(A) = 0$ (namely, $f = g$). In other words, A is integral over $\mathbb{K}$. In other words, Assertion $\mathcal{V}$ holds. This proves the implication $\mathcal{U} \Longrightarrow \mathcal{V}$.

Proof of the implication $\mathcal{V} \Longrightarrow \mathcal{W}$: Assume that Assertion $\mathcal{V}$ holds. We must prove that Assertion $\mathcal{W}$ holds.

We have assumed that Assertion $\mathcal{V}$ holds. In other words, A is integral over $\mathbb{K}$. Let $\mathbb{M}$ be the $\mathbb{K}$-subalgebra of $\mathbb{L}$ generated by the coefficients of the characteristic polynomial $\chi_A \in \mathbb{L}[t]$. Then, the coefficients of χ_A belong to this $\mathbb{K}$-subalgebra $\mathbb{M}$; thus, $\chi_A \in \mathbb{M}[t]$. Furthermore, Corollary 1.22 shows that $\mathbb{M}$ is a finitely generated $\mathbb{K}$-module. Thus, Lemma 1.23 (applied to $M = \mathbb{M}$) shows that $\mathbb{M}$ is finite (as a set).

The polynomial $\chi_A \in \mathbb{M}[t]$ is monic. Thus, the $\mathbb{M}$-module $\mathbb{M}[t] / (\chi_A)$ is finitely generated (by Lemma 1.24, applied to $\mathbb{M}$ and χ_A instead of $\mathbb{K}$ and f). Thus, Lemma 1.23 (applied to $\mathbb{M}$ and $\mathbb{M}[t] / (\chi_A)$ instead of $\mathbb{K}$ and M) shows that $\mathbb{M}[t] / (\chi_A)$ is finite (as a set). This ring $\mathbb{M}[t] / (\chi_A)$ becomes a semigroup when equipped with its multiplication. This semigroup $\mathbb{M}[t] / (\chi_A)$ is finite (since we have just shown that $\mathbb{M}[t] / (\chi_A)$ is finite).

For each $u \in \mathbb{M}[t]$, we let $\overline{u}$ denote the projection of u onto $\mathbb{M}[t] / (\chi_A)$. Then, Theorem 1.7 (applied to $M = \mathbb{M}[t] / (\chi_A)$ and $a = \overline{t}$) yields that there exists a positive integer m such that $\overline{t}^m = \overline{t}^{2m}$. Consider this m. Then, $\overline{t^m} = \overline{t}^m = \overline{t}^{2m} = \overline{t^{2m}}$; in other words, we have the congruence $t^m \equiv t^{2m} \bmod \chi_A$ in the ring $\mathbb{M}[t]$. In other words, the polynomial $t^{2m} - t^m$ is a multiple of χ_A in $\mathbb{M}[t]$. Hence, the polynomial $t^{2m} - t^m$ is a multiple of χ_A in $\mathbb{L}[t]$ (since $\mathbb{M}[t]$ is a subring of $\mathbb{L}[t]$). Thus, Assertion $\mathcal{W}$ holds. This proves the implication $\mathcal{V} \Longrightarrow \mathcal{W}$.

Proof of the implication $\mathcal{W} \Longrightarrow \mathcal{U}$: Assume that Assertion $\mathcal{W}$ holds. We must prove that Assertion $\mathcal{U}$ holds.

We have assumed that Assertion $\mathcal{W}$ holds. In other words, there exists a positive integer m such that the polynomial $t^{2m} - t^m$ is a multiple of χ_A in $\mathbb{L}[t]$. Consider this m. Note that $2m$ and m are positive integers satisfying $2m > m$. Consider the ring $\mathbb{L}^{n \times n}$ as a semigroup (equipped with its multiplication).

Now, there exists a polynomial $g \in \mathbb{L}[t]$ such that $t^{2m} - t^m = \chi_A \cdot g$ (since the polynomial $t^{2m} - t^m$ is a multiple of χ_A in $\mathbb{L}[t]$). Consider this g. Evaluating both sides of the polynomial identity $t^{2m} - t^m = \chi_A \cdot g$ at A, we obtain

$$A^{2m} - A^m = \underbrace{\chi_A(A)}_{\substack{=0 \\ \text{(by the Cayley–Hamilton} \\ \text{theorem)}}} \cdot g(A) = 0.$$

In other words, $A^{2m} = A^m$. Hence, Proposition 1.8 (applied to $M = \mathbb{L}^{n \times n}$, $a = A$, $p = 2m$ and $q = m$) yields $\{A^1, A^2, A^3, \ldots\} = \{A^1, A^2, \ldots, A^{2m-1}\}$. Thus, the set $\{A^1, A^2, A^3, \ldots\}$ is finite (since the set $\{A^1, A^2, \ldots, A^{2m-1}\}$ is clearly finite). Hence, the set $\{A^0, A^1, A^2, \ldots\}$ is also finite (since this set is $\{A^1, A^2, A^3, \ldots\} \cup \{A^0\}$). In other words, Assertion $\mathcal{U}$ holds. This proves the implication $\mathcal{W} \Longrightarrow \mathcal{U}$.

We have now proven all three implications $\mathcal{U} \Longrightarrow \mathcal{V}$ and $\mathcal{V} \Longrightarrow \mathcal{W}$ and $\mathcal{W} \Longrightarrow \mathcal{U}$. Hence, $\mathcal{U} \Longleftrightarrow \mathcal{V} \Longleftrightarrow \mathcal{W}$. This proves Proposition 1.25. □

We can now easily prove Theorem 1.2:

Proof of Theorem 1.2. Proposition 1.25 (or, more precisely, the equivalence of the Assertions $\mathcal{U}$ and $\mathcal{W}$ in this proposition) shows that the set $\{A^0, A^1, A^2, \ldots\}$ is finite if and only if there exists a positive integer m such that the polynomial $t^{2m} - t^m$ is

a multiple of χ_A in $\mathbb{L}[t]$. In other words, we have the logical equivalence

$$\begin{aligned}&\left(\text{the set } \left\{A^0, A^1, A^2, \ldots\right\} \text{ is finite}\right)\\ \Longleftrightarrow\ &\left(\text{there exists a positive integer } m \text{ such that the}\right.\\ &\left.\text{polynomial } t^{2m} - t^m \text{ is a multiple of } \chi_A \text{ in } \mathbb{L}[t]\right).\end{aligned}$$

The same argument (applied to B instead of A) yields the logical equivalence

$$\begin{aligned}&\left(\text{the set } \left\{B^0, B^1, B^2, \ldots\right\} \text{ is finite}\right)\\ \Longleftrightarrow\ &\left(\text{there exists a positive integer } m \text{ such that the}\right.\\ &\left.\text{polynomial } t^{2m} - t^m \text{ is a multiple of } \chi_B \text{ in } \mathbb{L}[t]\right).\end{aligned}$$

But the right hand sides of these two equivalences are equivalent (since $\chi_A = \chi_B$). Hence, their left hand sides are equivalent as well. In other words, we have the equivalence

$$\left(\text{the set } \left\{A^0, A^1, A^2, \ldots\right\} \text{ is finite}\right) \Longleftrightarrow \left(\text{the set } \left\{B^0, B^1, B^2, \ldots\right\} \text{ is finite}\right).$$

This proves Theorem 1.2. □

1.9. Digression: Traces of nilpotent matrices

While this is unrelated to Theorem 1.2, let us illustrate the usefulness of Theorem 1.18 on a different application:

Corollary 1.26. Let $\mathbb{K}$ be a commutative ring. Let $n \in \mathbb{N}$. Let $\mathbb{L}$ be a commutative $\mathbb{K}$-algebra. Let A be an $n \times n$-matrix over $\mathbb{L}$. Assume that A is integral over $\mathbb{K}$ (as an element of the $\mathbb{K}$-algebra $\mathbb{L}^{n\times n}$). Then, the trace $\operatorname{Tr} A \in \mathbb{L}$ is integral over $\mathbb{K}$.

Proof of Corollary 1.26. If $n = 0$, then this is obvious (since $\operatorname{Tr} A = 0$ in this case). Thus, we assume that $n \neq 0$ from now on. Hence, the coefficient of t^{n-1} in the characteristic polynomial $\chi_A \in \mathbb{L}[t]$ is known to be $-\operatorname{Tr} A$. But on the other hand, the coefficient of t^{n-1} in the characteristic polynomial $\chi_A \in \mathbb{L}[t]$ is integral over $\mathbb{K}$ (by Theorem 1.18). In other words, $-\operatorname{Tr} A$ is integral over $\mathbb{K}$ (since this coefficient is $-\operatorname{Tr} A$). Thus, $\operatorname{Tr} A$ is integral over $\mathbb{K}$ (because it is easy to see that if $u \in \mathbb{L}$ is an element such that $-u$ is integral over $\mathbb{K}$, then u is integral over $\mathbb{K}$). This proves Corollary 1.26. □

Corollary 1.27. Let $\mathbb{K}$ be a commutative ring. Let $n \in \mathbb{N}$. Let $A \in \mathbb{K}^{n\times n}$ be a nilpotent matrix. Then, its trace $\operatorname{Tr} A$ is nilpotent.

This is a generalization of the classical result that a nilpotent square matrix over a field must have trace 0.

There is actually a stronger version of Corollary 1.27, which says that if $A^{m+1} = 0$ for some $m \in \mathbb{N}$, then $(\operatorname{Tr} A)^{mn+1} = 0$ (see [Zeilbe93], and [Almkvi73, Theorem 1.7 (i)] for an even more general result). We shall only prove Corollary 1.27. The proof relies on the following neat fact, which reveals nilpotence to be an instance of integrality:

Lemma 1.28. Let $\mathbb{K}$ be a commutative ring. Let $\mathbb{L}$ be a $\mathbb{K}$-algebra. Let $a \in \mathbb{L}$. Consider the polynomial ring $\mathbb{L}[t]$. Then, a is nilpotent if and only if $at \in \mathbb{L}[t]$ is integral over $\mathbb{K}$.

Proof of Lemma 1.28. $\Longrightarrow$: Assume that a is nilpotent. Thus, there exists some $m \in \mathbb{N}$ such that $a^m = 0$. Consider this m. Now, let f be the polynomial $t^m \in \mathbb{K}[t]$. Then, this polynomial f is monic and satisfies $f(at) = (at)^m = \underbrace{a^m}_{=0} t^m = 0$. Hence, $at \in \mathbb{L}[t]$ is integral over $\mathbb{K}$ (by the definition of "integral"). This proves the "$\Longrightarrow$" direction of Lemma 1.28.

$\Longleftarrow$: Assume that $at \in \mathbb{L}[t]$ is integral over $\mathbb{K}$. Thus, there exists a monic polynomial $f \in \mathbb{K}[t]$ such that $f(at) = 0$. Consider this f.

Write the polynomial f in the form $f = k_0 t^0 + k_1 t^1 + \cdots + k_n t^n$, where $n = \deg f$ and $k_0, k_1, \ldots, k_n \in \mathbb{K}$. Then, $k_n = 1$ (since f is monic). Furthermore, from $f = k_0 t^0 + k_1 t^1 + \cdots + k_n t^n$, we obtain

$$f(at) = k_0 (at)^0 + k_1 (at)^1 + \cdots + k_n (at)^n = k_0 a^0 t^0 + k_1 a^1 t^1 + \cdots + k_n a^n t^n.$$

Comparing this with $f(at) = 0$, we obtain

$$k_0 a^0 t^0 + k_1 a^1 t^1 + \cdots + k_n a^n t^n = 0.$$

This is an equality between two polynomials in $\mathbb{L}[t]$. Comparing the coefficients of t^n on both sides of this equality, we conclude that $k_n a^n = 0$. Since $k_n = 1$, this rewrites as $a^n = 0$. Hence, a is nilpotent. This proves the "$\Longleftarrow$" direction of Lemma 1.28. $\square$

Proof of Corollary 1.27. Consider the polynomial ring $\mathbb{K}^{n \times n}[t]$. This is a $\mathbb{K}$-algebra which may be noncommutative, but t belongs to its center. The element $A \in \mathbb{K}^{n \times n}$ is nilpotent. Hence, the "$\Longrightarrow$" direction of Lemma 1.28 (applied to $\mathbb{K}^{n \times n}$ and A instead of $\mathbb{L}$ and a) yields that $At \in \mathbb{K}^{n \times n}[t]$ is integral over $\mathbb{K}$. Now, identify the $\mathbb{K}$-algebra $\mathbb{K}^{n \times n}[t]$ with $(\mathbb{K}[t])^{n \times n}$ in the usual way (i.e., in the same way as one identifies polynomial matrices with polynomials over matrix rings in linear algebra). Thus, $At \in \mathbb{K}^{n \times n}[t] = (\mathbb{K}[t])^{n \times n}$. The trace of this matrix At is $\operatorname{Tr}(At) = (\operatorname{Tr} A) \cdot t$ (since the trace is linear).

This matrix At is integral over $\mathbb{K}$ (as we know). Hence, Corollary 1.26 (applied to $\mathbb{K}[t]$ and At instead of $\mathbb{L}$ and A) yields that the trace $\operatorname{Tr}(At) \in \mathbb{K}[t]$ is integral over $\mathbb{K}$. In other words, $(\operatorname{Tr} A) \cdot t \in \mathbb{K}[t]$ is integral over $\mathbb{K}$ (since $\operatorname{Tr}(At) = (\operatorname{Tr} A) \cdot t$). Hence, the "$\Longleftarrow$" direction of Lemma 1.28 (applied to $\mathbb{L} = \mathbb{K}$ and $a = \operatorname{Tr} A$) yields that $\operatorname{Tr} A$ is nilpotent. This proves Corollary 1.27. $\square$

1.10. Appendix: Second proof of Proposition 1.19

Let us also sketch a second proof of Proposition 1.19, which avoids exterior powers over $\mathbb{K}[t]$ but instead uses determinantal identities.

In this section, we shall use the following notations: Fix a commutative ring $\mathbb{K}$ and an $n \in \mathbb{N}$. We let $[n]$ denote the set $\{1, 2, \ldots, n\}$.

Furthermore, if $A \in \mathbb{K}^{n \times n}$ is an $n \times n$-matrix, and if U and V are two subsets of $[n]$, then $\operatorname{sub}_U^V A$ shall denote the $|U| \times |V|$-matrix obtained from A by removing all the rows whose indices[11] don't belong to U and all the columns whose indices don't belong to V. (Formally speaking, this $\operatorname{sub}_U^V A$ is defined by

$$\operatorname{sub}_U^V A = \left(a_{u_i, v_j}\right)_{1 \leq i \leq p,\ 1 \leq j \leq q},$$

where we have written the matrix A in the form $A = \left(a_{i,j}\right)_{1 \leq i \leq n,\ 1 \leq j \leq n}$ and where we have written the two subsets U and V as $U = \left\{u_1 < u_2 < \cdots < u_p\right\}$ and $V = \left\{v_1 < v_2 < \cdots < v_q\right\}$.)

For example,

$$\operatorname{sub}_{\{2,4\}}^{\{1,4\}} \begin{pmatrix} a & b & c & d \\ a' & b' & c' & d' \\ a'' & b'' & c'' & d'' \\ a''' & b''' & c''' & d''' \end{pmatrix} = \begin{pmatrix} a' & d' \\ a''' & d''' \end{pmatrix}.$$

We shall now prove (or cite proofs of) a sequence of basic properties of submatrices and their determinants.

Lemma 1.29. Let $A \in \mathbb{K}^{n \times n}$ be a matrix. For each $x \in \mathbb{K}$, we have

$$\det\left(A + xI_n\right) = \sum_{P \subseteq [n]} \det\left(\operatorname{sub}_P^P A\right) x^{n - |P|}.$$

Lemma 1.29 is precisely the first equality sign of [Grinbe15, Corollary 6.164] (up to notation[12]). Thus, we don't need to prove it here.

Corollary 1.30. Let $A \in \mathbb{K}^{n \times n}$ be a matrix. Fix $k \in \mathbb{N}$. Let $a_k \in \mathbb{K}$ be the coefficient of t^k in the characteristic polynomial $\chi_A \in \mathbb{K}[t]$. Then,

$$a_k = (-1)^{n-k} \sum_{\substack{P \subseteq [n]; \\ |P| = n-k}} \det\left(\operatorname{sub}_P^P A\right).$$

[11] The *index* of a row in a matrix means the number saying which row it is. In other words, the index of the i-th row in a matrix means the number i. Similar terminology is used for columns.

[12] Specifically, our notations differ from those in [Grinbe15, Corollary 6.164] in two ways: Firstly, we use the shorthand $[n]$ for $\{1, 2, \ldots, n\}$; secondly, what we call $\operatorname{sub}_U^V A$ is called $\operatorname{sub}_{w(U)}^{w(V)} A$ in the notation of [Grinbe15, Definition 6.78 and Definition 6.153].

Proof of Corollary 1.30. Let us consider $A \in \mathbb{K}^{n\times n}$ as an $n \times n$-matrix over the polynomial ring $\mathbb{K}[t]$. Then,

$$\chi_A = \det \underbrace{(tI_n - A)}_{=-(A+(-t)I_n)} = \det\left(-\left(A + (-t)\, I_n\right)\right) = (-1)^n \underbrace{\det\left(A + (-t)\, I_n\right)}_{\substack{= \sum\limits_{P\subseteq[n]} \det\left(\operatorname{sub}_P^P A\right)(-t)^{n-|P|} \\ \text{(by Lemma 1.29,} \\ \text{applied to } \mathbb{K}[t] \text{ and } -t \\ \text{instead of } \mathbb{K} \text{ and } x)}}$$

$$= (-1)^n \sum_{P\subseteq[n]} \det\left(\operatorname{sub}_P^P A\right) \underbrace{(-t)^{n-|P|}}_{=(-1)^{n-|P|}t^{n-|P|}}$$

$$= (-1)^n \sum_{P\subseteq[n]} \det\left(\operatorname{sub}_P^P A\right) (-1)^{n-|P|}\, t^{n-|P|}$$

$$= \sum_{P\subseteq[n]} \det\left(\operatorname{sub}_P^P A\right) \underbrace{(-1)^n\, (-1)^{n-|P|}}_{=(-1)^{|P|}} t^{n-|P|} = \sum_{P\subseteq[n]} \det\left(\operatorname{sub}_P^P A\right) (-1)^{|P|}\, t^{n-|P|}.$$

Hence,

$$\left(\text{the coefficient of } t^k \text{ in } \chi_A\right)$$

$$= \left(\text{the coefficient of } t^k \text{ in } \sum_{P\subseteq[n]} \det\left(\operatorname{sub}_P^P A\right) (-1)^{|P|}\, t^{n-|P|}\right)$$

$$= \sum_{\substack{P\subseteq[n]; \\ n-|P|=k}} \det\left(\operatorname{sub}_P^P A\right) (-1)^{|P|} = \sum_{\substack{P\subseteq[n]; \\ |P|=n-k}} \det\left(\operatorname{sub}_P^P A\right) \underbrace{(-1)^{|P|}}_{\substack{=(-1)^{n-k} \\ \text{(since } |P|=n-k)}}$$

$$\left(\begin{array}{c}\text{since the condition "}n - |P| = k\text{" on a subset } P \text{ of } [n] \\ \text{is equivalent to the condition "}|P| = n - k\text{"}\end{array}\right)$$

$$= (-1)^{n-k} \sum_{\substack{P\subseteq[n]; \\ |P|=n-k}} \det\left(\operatorname{sub}_P^P A\right).$$

Now, by the definition of a_k, we have

$$a_k = \left(\text{the coefficient of } t^k \text{ in } \chi_A\right) = (-1)^{n-k} \sum_{\substack{P\subseteq[n]; \\ |P|=n-k}} \det\left(\operatorname{sub}_P^P A\right).$$

Corollary 1.30 is thus proven. □

We introduce two more notations:

- If S is any subset of $[n]$, then $\widetilde{S}$ shall denote the complement $[n] \setminus S$ of S.

- If S is any subset of $[n]$, then $\sum S$ shall denote the sum of the elements of S.

For example, if $n = 5$, then $\widetilde{\{1,3\}} = \{2,4,5\}$ and $\sum \{1,3\} = 1 + 3 = 4$.

Lemma 1.31. Let P and Q be two subsets of $[n]$. Let $A = \left(a_{i,j}\right)_{1 \leq i \leq n,\ 1 \leq j \leq n} \in \mathbb{K}^{n \times n}$ be an $n \times n$-matrix such that

$$\text{every } i \in P \text{ and } j \in Q \text{ satisfy } a_{i,j} = 0.$$

If $|P| + |Q| = n$, then

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

Lemma 1.31 is precisely [Grinbe15, Exercise 6.47 **(b)**] (up to notation[13]). Thus, we don't need to prove it here.

Our next proposition tells us what happens to the determinant of a matrix if we replace some columns of the matrix by the respective columns of the identity matrix I_n. To state this proposition, we need the following notation: If A is an $n \times m$-matrix, and if $j \in \{1, 2, \ldots, m\}$, then $\operatorname{Col}_j A$ shall denote the j-th column of A. For example, $\operatorname{Col}_2 \begin{pmatrix} a & b & c \\ a' & b' & c' \end{pmatrix} = \begin{pmatrix} b \\ b' \end{pmatrix}$.

Proposition 1.32. Let $A \in \mathbb{K}^{n \times n}$ be a matrix. Let P be a subset of $[n]$. Let $B \in \mathbb{K}^{n \times n}$ be the $n \times n$-matrix defined by setting

$$\operatorname{Col}_j B = \begin{cases} \operatorname{Col}_j A, & \text{if } j \in P; \\ \operatorname{Col}_j (I_n), & \text{if } j \notin P \end{cases} \qquad \text{for all } j \in [n]. \tag{10}$$

(That is, the columns of B whose indices lie in P equal the corresponding columns of A, while the other columns equal the corresponding columns of I_n.)

Then,

$$\det B = \det\left(\operatorname{sub}_P^P A\right).$$

[13] Specifically, our notations differ from those in [Grinbe15, Exercise 6.47] in two ways: Firstly, we use the shorthand $[n]$ for $\{1, 2, \ldots, n\}$; secondly, what we call $\operatorname{sub}_U^V A$ is called $\operatorname{sub}_{w(U)}^{w(V)} A$ in the notation of [Grinbe15, Definition 6.78 and Definition 6.153].

Example 1.33. Let $n = 4$ and $A = \begin{pmatrix} a & b & c & d \\ a' & b' & c' & d' \\ a'' & b'' & c'' & d'' \\ a''' & b''' & c''' & d''' \end{pmatrix}$ and $P = \{2, 4\}$. Then, the matrix B in Proposition 1.32 is given by

$$B = \begin{pmatrix} 1 & b & 0 & d \\ 0 & b' & 0 & d' \\ 0 & b'' & 1 & d'' \\ 0 & b''' & 0 & d''' \end{pmatrix}.$$

Proposition 1.32 states that this matrix satisfies

$$\det B = \det\left(\operatorname{sub}_P^P A\right) = \det\left(\operatorname{sub}_{\{2,4\}}^{\{2,4\}} A\right) = \det\begin{pmatrix} b' & d' \\ b''' & d''' \end{pmatrix}.$$

Proof of Proposition 1.32. For each $j \in P$, we have

$$\begin{aligned} \operatorname{Col}_j B &= \begin{cases} \operatorname{Col}_j A, & \text{if } j \in P; \\ \operatorname{Col}_j (I_n), & \text{if } j \notin P \end{cases} \qquad \text{(by (10))} \\ &= \operatorname{Col}_j A \qquad \text{(since } j \in P). \end{aligned} \tag{11}$$

In other words, the columns of B with indices $j \in P$ equal the corresponding columns of A. Hence, the submatrix $\operatorname{sub}_P^P B$ of B equals the corresponding submatrix $\operatorname{sub}_P^P A$ of A (because these two submatrices are contained entirely in the columns with indices $j \in P$). In other words,

$$\operatorname{sub}_P^P B = \operatorname{sub}_P^P A.$$

Define a subset Q of $[n]$ by $Q = \widetilde{P}$. Thus, Q is the complement of P in the n-element set $[n]$; hence, $|Q| = n - |P|$. In other words, $|P| + |Q| = n$. Moreover, from $Q = \widetilde{P}$, we obtain $P = \widetilde{Q}$, so that $\widetilde{Q} = P$. For each $j \in Q$, we have

$$\begin{aligned} \operatorname{Col}_j B &= \begin{cases} \operatorname{Col}_j A, & \text{if } j \in P; \\ \operatorname{Col}_j (I_n), & \text{if } j \notin P \end{cases} \qquad \text{(by (10))} \\ &= \operatorname{Col}_j (I_n) \qquad \left(\text{since } j \notin P \text{ (because } j \in Q = \widetilde{P} = [n] \setminus P)\right). \end{aligned}$$

In other words, the columns of B with indices $j \in Q$ equal the corresponding columns of I_n. Hence, the submatrix $\operatorname{sub}_Q^Q B$ of B equals the corresponding submatrix $\operatorname{sub}_Q^Q (I_n)$ of I_n (because these two submatrices are contained entirely in the columns with indices $j \in Q$). In other words,

$$\operatorname{sub}_Q^Q B = \operatorname{sub}_Q^Q (I_n) = I_{|Q|}$$

(since any principal submatrix of an identity matrix is itself an identity matrix).

Write the matrix $B \in \mathbb{K}^{n\times n}$ in the form $B = \left(b_{i,j}\right)_{1\leq i\leq n,\ 1\leq j\leq n}$. Thus,

$$b_{i,j} = (\text{the } (i,j)\text{-th entry of } B) \tag{12}$$

for all $i, j \in [n]$. Furthermore,

$$\text{every } i \in P \text{ and } j \in Q \text{ satisfy } b_{i,j} = 0$$

[14]. Hence, Lemma 1.31 (applied to B and $b_{i,j}$ instead of A and $a_{i,j}$) shows that

$$\begin{aligned}
\det B &= (-1)^{\sum P + \sum \widetilde{Q}} \det\left(\operatorname{sub}_P^{\widetilde{Q}} B\right) \det\left(\operatorname{sub}_{\widetilde{P}}^{Q} B\right) \\
&= \underbrace{(-1)^{\sum P + \sum P}}_{=(-1)^{2\sum P}=1} \det \underbrace{\left(\operatorname{sub}_P^P B\right)}_{=\operatorname{sub}_P^P A} \det \underbrace{\left(\operatorname{sub}_Q^Q B\right)}_{=I_{|Q|}} \qquad \left(\text{since } \widetilde{Q} = P \text{ and } \widetilde{P} = Q\right) \\
&= \det\left(\operatorname{sub}_P^P A\right) \underbrace{\det\left(I_{|Q|}\right)}_{=1} = \det\left(\operatorname{sub}_P^P A\right).
\end{aligned}$$

This proves Proposition 1.32. □

Proposition 1.32 has the following consequence for exterior powers:

---

[14] *Proof.* Let $i \in P$ and $j \in Q$. Then, $j \in Q = \widetilde{P} = [n] \setminus P$ (by the definition of $\widetilde{P}$), so that $j \notin P$. Hence, $i \neq j$ (since otherwise, we would have $i = j \notin P$, which would contradict $i \in P$). Now, the definition of B yields

$$\begin{aligned}
\operatorname{Col}_j B &= \begin{cases} \operatorname{Col}_j A, & \text{if } j \in P; \\ \operatorname{Col}_j (I_n), & \text{if } j \notin P \end{cases} = \operatorname{Col}_j (I_n) \qquad (\text{since } j \notin P) \\
&= \left(\underbrace{0, 0, \ldots, 0}_{j-1 \text{ zeroes}}, 1, \underbrace{0, 0, \ldots, 0}_{n-j \text{ zeroes}}\right)^T
\end{aligned}$$

(by the definition of I_n). Hence,

$$\left(\text{the } i\text{-th entry of the vector } \operatorname{Col}_j B\right) = \begin{cases} 1, & \text{if } i = j; \\ 0, & \text{if } i \neq j \end{cases} = 0 \qquad (\text{since } i \neq j).$$

But (12) yields

$$\begin{aligned}
b_{i,j} &= (\text{the } (i,j)\text{-th entry of } B) \\
&= \left(\text{the } i\text{-th entry of the vector } \operatorname{Col}_j B\right) \\
&\qquad \left(\text{since } \operatorname{Col}_j B \text{ is the } j\text{-th column of } B\right) \\
&= 0.
\end{aligned}$$

Qed.

Corollary 1.34. Let $A \in \mathbb{K}^{n \times n}$ be a matrix.

Let V be the free $\mathbb{K}$-module $\mathbb{K}^n$. Consider A as an endomorphism of the free $\mathbb{K}$-module $V = \mathbb{K}^n$. Consider the n-th exterior power $\Lambda^n V$ of the $\mathbb{K}$-module V.

Let $(e_1, e_2, \ldots, e_n)$ be the standard basis of the $\mathbb{K}$-module $\mathbb{K}^n$ (so that e_i is the vector with a 1 in its i-th entry and 0 everywhere else).

Let $(i_1, i_2, \ldots, i_n) \in \{0,1\}^n$. Define a subset P of $[n]$ by $P = \{p \in [n] \mid i_p = 1\}$. Then, in $\Lambda^n V$, we have

$$\det\left(\operatorname{sub}_P^P A\right) \cdot e_1 \wedge e_2 \wedge \cdots \wedge e_n = A^{i_1} e_1 \wedge A^{i_2} e_2 \wedge \cdots \wedge A^{i_n} e_n.$$

Proof of Corollary 1.34. Define an $n \times n$-matrix $B \in \mathbb{K}^{n \times n}$ as in Proposition 1.32. Consider B as an endomorphism of the free $\mathbb{K}$-module $V = \mathbb{K}^n$ as well.

We have $A^{i_p} e_p = B e_p$ for every $p \in [n]$ [15]. Combining these equalities, we obtain

$$A^{i_1} e_1 \wedge A^{i_2} e_2 \wedge \cdots \wedge A^{i_n} e_n = B e_1 \wedge B e_2 \wedge \cdots \wedge B e_n. \tag{13}$$

[15] *Proof.* Let $p \in [n]$. Recall that $(e_1, e_2, \ldots, e_n)$ is the standard basis of the $\mathbb{K}$-module $\mathbb{K}^n$. Thus, e_p is the column vector with a 1 in its p-th entry and 0's everywhere else. Hence, for every $n \times n$-matrix $C \in \mathbb{K}^{n \times n}$, we have

$$C e_p = (\text{the } p\text{-th column of } C) = \operatorname{Col}_p C$$

(by the definition of $\operatorname{Col}_p C$). Applying this to $C = B$, we obtain $B e_p = \operatorname{Col}_p B$. The same argument (using A instead of B) shows that $A e_p = \operatorname{Col}_p A$.

We are in one of the following two cases:

Case 1: We have $p \in P$.

Case 2: We have $p \notin P$.

Let us first consider Case 1. In this case, we have $p \in P$. In other words, $i_p = 1$ (by the definition of P). Thus, $A^{i_p} = A^1 = A$. Hence, $A^{i_p} e_p = A e_p = \operatorname{Col}_p A$. Comparing this with

$$\begin{aligned} B e_p = \operatorname{Col}_p B &= \begin{cases} \operatorname{Col}_p A, & \text{if } p \in P; \\ \operatorname{Col}_p (I_n), & \text{if } p \notin P \end{cases} \qquad \text{(by (10), applied to } j = p) \\ &= \operatorname{Col}_p A \qquad (\text{since } p \in P), \end{aligned}$$

we obtain $A^{i_p} e_p = B e_p$. Hence, $A^{i_p} e_p = B e_p$ is proven in Case 1.

Let us next consider Case 2. In this case, we have $p \notin P$. In other words, $i_p \neq 1$ (by the definition of P). But $i_p \in \{0,1\}$ (since $(i_1, i_2, \ldots, i_n) \in \{0,1\}^n$). Hence, $i_p = 0$ (since $i_p \neq 1$). Thus, $A^{i_p} = A^0 = I_n$. Hence, $A^{i_p} e_p = I_n e_p = e_p$. Comparing this with

$$\begin{aligned} B e_p = \operatorname{Col}_p B &= \begin{cases} \operatorname{Col}_p A, & \text{if } p \in P; \\ \operatorname{Col}_p (I_n), & \text{if } p \notin P \end{cases} \qquad \text{(by (10), applied to } j = p) \\ &= \operatorname{Col}_p (I_n) \qquad (\text{since } p \notin P) \\ &= e_p \qquad (\text{since the columns of } I_n \text{ are } e_1, e_2, \ldots, e_n), \end{aligned}$$

we obtain $A^{i_p} e_p = B e_p$. Hence, $A^{i_p} e_p = B e_p$ is proven in Case 2.

We have now proven $A^{i_p} e_p = B e_p$ in both Cases 1 and 2. Thus, the proof of $A^{i_p} e_p = B e_p$ is complete.

Recall that B is an endomorphism of the free $\mathbb{K}$-module $\mathbb{K}^n$. Hence, Lemma 1.20 **(b)** (applied to $\mathbb{L} = \mathbb{K}$, $u = B$ and $p = e_1 \wedge e_2 \wedge \cdots \wedge e_n$) yields

$$\begin{aligned}(\Lambda^n B)(e_1 \wedge e_2 \wedge \cdots \wedge e_n) &= \underbrace{\det B}_{\substack{=\det\left(\operatorname{sub}_P^P A\right) \\ \text{(by Proposition 1.32)}}} \cdot e_1 \wedge e_2 \wedge \cdots \wedge e_n \\ &= \det\left(\operatorname{sub}_P^P A\right) \cdot e_1 \wedge e_2 \wedge \cdots \wedge e_n.\end{aligned}$$

Hence,

$$\begin{aligned}&\det\left(\operatorname{sub}_P^P A\right) \cdot e_1 \wedge e_2 \wedge \cdots \wedge e_n \\ &= (\Lambda^n B)(e_1 \wedge e_2 \wedge \cdots \wedge e_n) \\ &= Be_1 \wedge Be_2 \wedge \cdots \wedge Be_n \qquad \text{(by the definition of the map } \Lambda^n B) \\ &= A^{i_1} e_1 \wedge A^{i_2} e_2 \wedge \cdots \wedge A^{i_n} e_n \qquad \text{(by (13))}.\end{aligned}$$

This proves Corollary 1.34. □

Lemma 1.35. Let V be any $\mathbb{K}$-module. Let A be an endomorphism of the $\mathbb{K}$-module V. Let $w_1, w_2, \ldots, w_n \in V$ be arbitrary vectors. Assume that two of these n vectors $w_1, w_2, \ldots, w_n$ are equal. Let $k \in \mathbb{Z}$. Then,

$$\sum_{\substack{i_1, i_2, \ldots, i_n \in \{0,1\}; \\ i_1 + i_2 + \cdots + i_n = n-k}} A^{i_1} w_1 \wedge A^{i_2} w_2 \wedge \cdots \wedge A^{i_n} w_n = 0$$

in the exterior power $\Lambda^n V$ of the $\mathbb{K}$-module V.

Proof of Lemma 1.35. We have assumed that two of the n vectors $w_1, w_2, \ldots, w_n$ are equal. In other words, there exist two elements u and v of $[n]$ such that $u < v$ and $w_u = w_v$. Consider these u and v.

Let Z be the set $\left\{(i_1, i_2, \ldots, i_n) \in \{0,1\}^n \mid i_1 + i_2 + \cdots + i_n = n - k\right\}$. Then, the summation sign "$\sum\limits_{\substack{i_1, i_2, \ldots, i_n \in \{0,1\}; \\ i_1 + i_2 + \cdots + i_n = n-k}}$" can be rewritten as "$\sum\limits_{(i_1, i_2, \ldots, i_n) \in Z}$".

Now, let σ be the permutation of the set $[n]$ that swaps u with v while leaving all other elements unchanged. (This permutation is known as a transposition.) We write σp for the image of an element $p \in [n]$ under this permutation σ. The definition of σ yields $\sigma u = v$ and $\sigma v = u$. Hence, it is easy to see that

$$w_{\sigma p} = w_p \qquad \text{for all } p \in [n] \tag{14}$$

[16].

[16] *Proof of (14):* Let $p \in [n]$. We must prove that $w_{\sigma p} = w_p$. We have either $p = u$ or $p = v$ or $p \notin \{u, v\}$. In each of these three cases, let us check that $w_{\sigma p} = w_p$ holds:

Now, each $(i_1, i_2, \ldots, i_n) \in Z$ satisfies either $i_u < i_v$ or $i_u = i_v$ or $i_u > i_v$. Hence, the sum $\sum\limits_{(i_1,i_2,\ldots,i_n)\in Z} A^{i_1}w_1 \wedge A^{i_2}w_2 \wedge \cdots \wedge A^{i_n}w_n$ can be split up as follows:

$$\sum_{(i_1,i_2,\ldots,i_n)\in Z} A^{i_1}w_1 \wedge A^{i_2}w_2 \wedge \cdots \wedge A^{i_n}w_n = \alpha + \beta + \gamma, \tag{15}$$

where

$$\alpha = \sum_{\substack{(i_1,i_2,\ldots,i_n)\in Z; \\ i_u<i_v}} A^{i_1}w_1 \wedge A^{i_2}w_2 \wedge \cdots \wedge A^{i_n}w_n; \tag{16}$$

$$\beta = \sum_{\substack{(i_1,i_2,\ldots,i_n)\in Z; \\ i_u=i_v}} A^{i_1}w_1 \wedge A^{i_2}w_2 \wedge \cdots \wedge A^{i_n}w_n; \tag{17}$$

$$\gamma = \sum_{\substack{(i_1,i_2,\ldots,i_n)\in Z; \\ i_u>i_v}} A^{i_1}w_1 \wedge A^{i_2}w_2 \wedge \cdots \wedge A^{i_n}w_n. \tag{18}$$

Consider these α, β, γ.

If $(i_1, i_2, \ldots, i_n) \in Z$ satisfies $i_u = i_v$, then $\underbrace{A^{i_u}}_{\substack{=A^{i_v} \\ (\text{since } i_u=i_v)}} \underbrace{w_u}_{=w_v} = A^{i_v}w_v$ and therefore

$$A^{i_1}w_1 \wedge A^{i_2}w_2 \wedge \cdots \wedge A^{i_n}w_n = 0 \tag{19}$$

(because the exterior product is alternating). Hence,

$$\beta = \sum_{\substack{(i_1,i_2,\ldots,i_n)\in Z; \\ i_u=i_v}} \underbrace{A^{i_1}w_1 \wedge A^{i_2}w_2 \wedge \cdots \wedge A^{i_n}w_n}_{\substack{=0 \\ (\text{by } (19))}} = 0.$$

Furthermore, fix $(i_1, i_2, \ldots, i_n) \in Z$. Then, the exterior product $A^{i_{\sigma 1}}w_{\sigma 1} \wedge A^{i_{\sigma 2}}w_{\sigma 2} \wedge \cdots \wedge A^{i_{\sigma n}}w_{\sigma n}$ is obtained from $A^{i_1}w_1 \wedge A^{i_2}w_2 \wedge \cdots \wedge A^{i_n}w_n$ by swapping two factors (because σ is a transposition). Therefore,

$$A^{i_{\sigma 1}}w_{\sigma 1} \wedge A^{i_{\sigma 2}}w_{\sigma 2} \wedge \cdots \wedge A^{i_{\sigma n}}w_{\sigma n} = -A^{i_1}w_1 \wedge A^{i_2}w_2 \wedge \cdots \wedge A^{i_n}w_n$$

- If $p = u$, then

$$\begin{aligned} w_{\sigma p} &= w_v && \left(\text{since } \sigma \underbrace{p}_{=u} = \sigma u = v\right) \\ &= w_u && (\text{since } w_u = w_v) \\ &= w_p && (\text{since } u = p)\,. \end{aligned}$$

  Thus, we have proven $w_{\sigma p} = w_p$ if $p = u$.
- A similar computation proves $w_{\sigma p} = w_p$ if $p = v$.
- If $p \notin \{u, v\}$, then $\sigma p = p$ (by the definition of σ) and thus $w_{\sigma p} = w_p$.

Thus, $w_{\sigma p} = w_p$ always holds. This proves (14).

(since the exterior product is antisymmetric). Comparing this with

$$A^{i_{\sigma 1}} \underbrace{w_{\sigma 1}}_{\substack{=w_1 \\ \text{(by (14))}}} \wedge A^{i_{\sigma 2}} \underbrace{w_{\sigma 2}}_{\substack{=w_2 \\ \text{(by (14))}}} \wedge \cdots \wedge A^{i_{\sigma n}} \underbrace{w_{\sigma n}}_{\substack{=w_n \\ \text{(by (14))}}} = A^{i_{\sigma 1}} w_1 \wedge A^{i_{\sigma 2}} w_2 \wedge \cdots \wedge A^{i_{\sigma n}} w_n,$$

we obtain

$$A^{i_{\sigma 1}} w_1 \wedge A^{i_{\sigma 2}} w_2 \wedge \cdots \wedge A^{i_{\sigma n}} w_n = -A^{i_1} w_1 \wedge A^{i_2} w_2 \wedge \cdots \wedge A^{i_n} w_n. \tag{20}$$

Now, forget that we fixed $(i_1, i_2, \ldots, i_n)$. We thus have proven (20) for each $(i_1, i_2, \ldots, i_n) \in Z$.

But it is easy to see that the map

$$\begin{aligned} \{(i_1, i_2, \ldots, i_n) \in Z \mid i_u < i_v\} &\to \{(i_1, i_2, \ldots, i_n) \in Z \mid i_u > i_v\}, \\ (i_1, i_2, \ldots, i_n) &\mapsto (i_{\sigma 1}, i_{\sigma 2}, \ldots, i_{\sigma n}) \end{aligned} \tag{21}$$

is well-defined and is a bijection[17]. Hence, we can substitute $(i_{\sigma 1}, i_{\sigma 2}, \ldots, i_{\sigma n})$ for $(i_1, i_2, \ldots, i_n)$ in the sum on the right hand side of (18). We thus find

$$\begin{aligned} &\sum_{\substack{(i_1, i_2, \ldots, i_n) \in Z; \\ i_u > i_v}} A^{i_1} w_1 \wedge A^{i_2} w_2 \wedge \cdots \wedge A^{i_n} w_n \\ &= \sum_{\substack{(i_1, i_2, \ldots, i_n) \in Z; \\ i_u < i_v}} \underbrace{A^{i_{\sigma 1}} w_1 \wedge A^{i_{\sigma 2}} w_2 \wedge \cdots \wedge A^{i_{\sigma n}} w_n}_{\substack{= -A^{i_1} w_1 \wedge A^{i_2} w_2 \wedge \cdots \wedge A^{i_n} w_n \\ \text{(by (20))}}} \\ &= \underbrace{-\sum_{\substack{(i_1, i_2, \ldots, i_n) \in Z; \\ i_u < i_v}} A^{i_1} w_1 \wedge A^{i_2} w_2 \wedge \cdots \wedge A^{i_n} w_n}_{\substack{=\alpha \\ \text{(by (16))}}} = -\alpha. \end{aligned}$$

Hence, (18) becomes

$$\gamma = \sum_{\substack{(i_1, i_2, \ldots, i_n) \in Z; \\ i_u > i_v}} A^{i_1} w_1 \wedge A^{i_2} w_2 \wedge \cdots \wedge A^{i_n} w_n = -\alpha.$$

Now, (15) becomes

$$\sum_{(i_1, i_2, \ldots, i_n) \in Z} A^{i_1} w_1 \wedge A^{i_2} w_2 \wedge \cdots \wedge A^{i_n} w_n = \alpha + \underbrace{\beta}_{=0} + \underbrace{\gamma}_{=-\alpha} = \alpha + (-\alpha) = 0.$$

[17] Indeed, all this map does is swapping the u-th and the v-th entry of the n-tuple it is being applied to (because σ swaps u with v while leaving all other numbers unchanged). Hence, it preserves the sum of all entries of the n-tuple. Thus, it sends an n-tuple in Z to an n-tuple in Z. Furthermore, if we apply this map to an n-tuple $(i_1, i_2, \ldots, i_n)$ satisfying $i_u < i_v$, then its image under this map will be an n-tuple $(i_1, i_2, \ldots, i_n)$ satisfying $i_u > i_v$ (since it swaps the u-th and the v-th entry of the n-tuple). This shows that this map is well-defined. In order to prove that it is a bijection, we just need to construct its inverse; this is easily done (it is given by the same recipe (21) as our original map, but it goes in the opposite direction).

In other words,

$$\sum_{\substack{i_1,i_2,\ldots,i_n\in\{0,1\};\\ i_1+i_2+\cdots+i_n=n-k}} A^{i_1}w_1\wedge A^{i_2}w_2\wedge\cdots\wedge A^{i_n}w_n=0$$

(since the summation sign "$\sum\limits_{\substack{i_1,i_2,\ldots,i_n\in\{0,1\};\\ i_1+i_2+\cdots+i_n=n-k}}$" can be rewritten as "$\sum\limits_{(i_1,i_2,\ldots,i_n)\in Z}$"). This proves Lemma 1.35. □

We now finally can prove Proposition 1.19 again:

Second proof of Proposition 1.19. The map

$$\begin{aligned} &V\times V\times\cdots\times V\to\Lambda^n V,\\ &(w_1,w_2,\ldots,w_n)\mapsto\sum_{\substack{i_1,i_2,\ldots,i_n\in\{0,1\};\\ i_1+i_2+\cdots+i_n=n-k}} A^{i_1}w_1\wedge A^{i_2}w_2\wedge\cdots\wedge A^{i_n}w_n \end{aligned}$$

is $\mathbb{K}$-multilinear and alternating[18]. Hence, the universal property of $\Lambda^n V$ (see, e.g., [Conrad13a, Theorem 3.3]) shows that there is a unique $\mathbb{K}$-linear map $\Phi:\Lambda^n V\to\Lambda^n V$ that satisfies

$$\Phi\left(w_1\wedge w_2\wedge\cdots\wedge w_n\right)=\sum_{\substack{i_1,i_2,\ldots,i_n\in\{0,1\};\\ i_1+i_2+\cdots+i_n=n-k}} A^{i_1}w_1\wedge A^{i_2}w_2\wedge\cdots\wedge A^{i_n}w_n$$
$$\text{for all } w_1,w_2,\ldots,w_n\in V.$$

Consider this Φ.

Let $(e_1,e_2,\ldots,e_n)$ be the standard basis of the $\mathbb{K}$-module $\mathbb{K}^n$ (so that e_i is the vector with a 1 in its i-th entry and 0 everywhere else). Thus, $(e_1,e_2,\ldots,e_n)$ is a basis of the $\mathbb{K}$-module V (since $\mathbb{K}^n=V$). Hence, Lemma 1.21 (applied to $\mathbb{L}=\mathbb{K}$, $M=V$ and $b_i=e_i$) shows that the 1-tuple $(e_1\wedge e_2\wedge\cdots\wedge e_n)$ is a basis of the $\mathbb{K}$-module $\Lambda^n V$.

Let

$$G:\{0,1\}^n\to\{\text{subsets of }[n]\}$$

be the map that sends each n-tuple $(i_1,i_2,\ldots,i_n)\in\{0,1\}^n$ to the subset $\{p\in[n]\mid i_p=1\}$ of $[n]$. This map G is a bijection (and is, in fact, the famous correspondence between bitstrings and subsets of $[n]$). Furthermore, each n-tuple $(i_1,i_2,\ldots,i_n)\in\{0,1\}^n$ satisfies $G(i_1,i_2,\ldots,i_n)=\{p\in[n]\mid i_p=1\}$ (by the definition of G) and therefore

$$\det\left(\operatorname{sub}_{G(i_1,i_2,\ldots,i_n)}^{G(i_1,i_2,\ldots,i_n)}A\right)\cdot e_1\wedge e_2\wedge\cdots\wedge e_n=A^{i_1}e_1\wedge A^{i_2}e_2\wedge\cdots\wedge A^{i_n}e_n \tag{22}$$

[18]Indeed, it is easy to show that it is $\mathbb{K}$-multilinear. But then, Lemma 1.35 shows that it is alternating.

(by Corollary 1.34, applied to $P = G(i_1, i_2, \ldots, i_n)$). Moreover, each n-tuple $(i_1, i_2, \ldots, i_n) \in \{0,1\}^n$ satisfies

$$|G(i_1, i_2, \ldots, i_n)| = i_1 + i_2 + \cdots + i_n, \tag{23}$$

because

$$\begin{aligned} i_1 + i_2 + \cdots + i_n &= \sum_{p \in [n]} i_p = \sum_{\substack{p \in [n]; \\ i_p \neq 1}} \underbrace{i_p}_{\substack{=0 \\ \text{(since } i_p \in \{0,1\} \\ \text{(because } (i_1,i_2,\ldots,i_n) \in \{0,1\}^n) \\ \text{but } i_p \neq 1)}} + \sum_{\substack{p \in [n]; \\ i_p = 1}} \underbrace{i_p}_{=1} \\ &= \underbrace{\sum_{\substack{p \in [n]; \\ i_p \neq 1}} 0}_{=0} + \sum_{\substack{p \in [n]; \\ i_p = 1}} 1 = \sum_{\substack{p \in [n]; \\ i_p = 1}} 1 = \left| \left\{ p \in [n] \mid i_p = 1 \right\} \right| \cdot 1 \\ &= \left| \underbrace{\left\{ p \in [n] \mid i_p = 1 \right\}}_{\substack{= G(i_1,i_2,\ldots,i_n) \\ \text{(by the definition of } G)}} \right| = |G(i_1, i_2, \ldots, i_n)|. \end{aligned}$$

Now,

$$
\begin{aligned}
&= (-1)^{n-k} \underbrace{\sum_{\substack{P \subseteq [n]; \\ |P| = n-k}}}_{\substack{\text{(by Corollary 1.30)}}} \det\left(\operatorname{sub}_P^P A\right) \cdot e_1 \wedge e_2 \wedge \cdots \wedge e_n \\
&= (-1)^{n-k} \sum_{\substack{P \subseteq [n]; \\ |P| = n-k}} \det\left(\operatorname{sub}_P^P A\right) \cdot e_1 \wedge e_2 \wedge \cdots \wedge e_n \\
&= (-1)^{n-k} \underbrace{\sum_{\substack{(i_1, i_2, \ldots, i_n) \in \{0,1\}^n; \\ |G(i_1, i_2, \ldots, i_n)| = n-k}}}_{\substack{= \sum\limits_{\substack{(i_1, i_2, \ldots, i_n) \in \{0,1\}^n; \\ i_1 + i_2 + \cdots + i_n = n-k}} \\ \text{(by (23))}}} \underbrace{\det\left(\operatorname{sub}_{G(i_1, i_2, \ldots, i_n)}^{G(i_1, i_2, \ldots, i_n)} A\right) \cdot e_1 \wedge e_2 \wedge \cdots \wedge e_n}_{\substack{= A^{i_1} e_1 \wedge A^{i_2} e_2 \wedge \cdots \wedge A^{i_n} e_n \\ \text{(by (22))}}} \\
&\qquad \left(\begin{array}{c} \text{here, we have substituted } G(i_1, i_2, \ldots, i_n) \text{ for } P \text{ in the sum,} \\ \text{since the map } G : \{0,1\}^n \to \{\text{subsets of } [n]\} \text{ is a bijection} \end{array} \right) \\
&= (-1)^{n-k} \underbrace{\sum_{\substack{(i_1, i_2, \ldots, i_n) \in \{0,1\}^n; \\ i_1 + i_2 + \cdots + i_n = n-k}}}_{= \sum\limits_{\substack{i_1, i_2, \ldots, i_n \in \{0,1\}; \\ i_1 + i_2 + \cdots + i_n = n-k}}} A^{i_1} e_1 \wedge A^{i_2} e_2 \wedge \cdots \wedge A^{i_n} e_n \\
&= (-1)^{n-k} \underbrace{\sum_{\substack{i_1, i_2, \ldots, i_n \in \{0,1\}; \\ i_1 + i_2 + \cdots + i_n = n-k}} A^{i_1} e_1 \wedge A^{i_2} e_2 \wedge \cdots \wedge A^{i_n} e_n}_{\substack{= \Phi(e_1 \wedge e_2 \wedge \cdots \wedge e_n) \\ \text{(by the definition of } \Phi)}} \\
&= (-1)^{n-k} \Phi\left(e_1 \wedge e_2 \wedge \cdots \wedge e_n\right). \qquad (24)
\end{aligned}
$$

Now, let $w_1, w_2, \ldots, w_n \in V$ be arbitrary. Then, there exists some $\lambda \in \mathbb{K}$ such that $w_1 \wedge w_2 \wedge \cdots \wedge w_n = \lambda \cdot e_1 \wedge e_2 \wedge \cdots \wedge e_n$ (since the 1-tuple $(e_1 \wedge e_2 \wedge \cdots \wedge e_n)$ is a

basis of the $\mathbb{K}$-module $\Lambda^n V$). Consider this λ. Now,

$$\begin{aligned}
a_k \cdot \underbrace{w_1 \wedge w_2 \wedge \cdots \wedge w_n}_{=\lambda \cdot e_1 \wedge e_2 \wedge \cdots \wedge e_n} &= \lambda \cdot \underbrace{a_k \cdot e_1 \wedge e_2 \wedge \cdots \wedge e_n}_{\substack{=(-1)^{n-k}\Phi(e_1 \wedge e_2 \wedge \cdots \wedge e_n) \\ \text{(by (24))}}} \\
&= \lambda \cdot (-1)^{n-k} \Phi\left(e_1 \wedge e_2 \wedge \cdots \wedge e_n\right) \\
&= (-1)^{n-k} \underbrace{\lambda \cdot \Phi\left(e_1 \wedge e_2 \wedge \cdots \wedge e_n\right)}_{\substack{=\Phi(\lambda \cdot e_1 \wedge e_2 \wedge \cdots \wedge e_n) \\ \text{(since the map } \Phi \text{ is } \mathbb{K}\text{-linear)}}} \\
&= (-1)^{n-k} \Phi\left(\underbrace{\lambda \cdot e_1 \wedge e_2 \wedge \cdots \wedge e_n}_{=w_1 \wedge w_2 \wedge \cdots \wedge w_n}\right) \\
&= (-1)^{n-k} \underbrace{\Phi\left(w_1 \wedge w_2 \wedge \cdots \wedge w_n\right)}_{\substack{= \sum\limits_{\substack{i_1, i_2, \ldots, i_n \in \{0,1\}; \\ i_1 + i_2 + \cdots + i_n = n-k}} A^{i_1} w_1 \wedge A^{i_2} w_2 \wedge \cdots \wedge A^{i_n} w_n \\ \text{(by the definition of } \Phi)}} \\
&= (-1)^{n-k} \sum_{\substack{i_1, i_2, \ldots, i_n \in \{0,1\}; \\ i_1 + i_2 + \cdots + i_n = n-k}} A^{i_1} w_1 \wedge A^{i_2} w_2 \wedge \cdots \wedge A^{i_n} w_n.
\end{aligned}$$

Thus, Proposition 1.19 is proven again. □

1.11. Appendix: Proof of Proposition 1.17

We have yet to prove Proposition 1.17. This is much easier than proving its converse, which we have already done. We begin by stating a trivial consequence of Theorem 1.12:

Corollary 1.36. Let $\mathbb{K}$ be a commutative ring. Let $\mathbb{L}$ be a commutative $\mathbb{K}$-algebra. Let $n \in \mathbb{N}$. Let $u \in \mathbb{L}$. Assume that there exists a monic polynomial $f \in \mathbb{K}[t]$ of degree n such that $f(u) = 0$. Then, $\mathbb{K}[u] = \left\langle u^0, u^1, \ldots, u^{n-1} \right\rangle_{\mathbb{K}}$.

Proof of Corollary 1.36. Assertion $\mathcal{A}$ of Theorem 1.12 holds (since there exists a monic polynomial $f \in \mathbb{K}[t]$ of degree n such that $f(u) = 0$). Hence, Assertion $\mathcal{D}$ of Theorem 1.12 holds as well (since Theorem 1.12 yields that Assertions $\mathcal{A}$ and $\mathcal{D}$ are equivalent). In other words, $\mathbb{K}[u] = \left\langle u^0, u^1, \ldots, u^{n-1} \right\rangle_{\mathbb{K}}$. This proves Corollary 1.36. □

Proof of Proposition 1.17. Let $\mathbb{M}$ be the $\mathbb{K}$-subalgebra of $\mathbb{L}$ generated by the coefficients of the characteristic polynomial $\chi_A \in \mathbb{L}[t]$. Let $u_1, u_2, \ldots, u_m$ be these coefficients. These coefficients $u_1, u_2, \ldots, u_m$ are integral over $\mathbb{K}$ (since we assumed that each coefficient of the characteristic polynomial $\chi_A \in \mathbb{L}[t]$ is integral over $\mathbb{K}$), and generate $\mathbb{M}$ as a $\mathbb{K}$-algebra (by the definition of $\mathbb{M}$); thus, in particular, they are

elements of $\mathbb{M}$. Hence, Theorem 1.15 (applied to $\mathbb{M}$ instead of $\mathbb{L}$) yields that the $\mathbb{K}$-module $\mathbb{M}$ is finitely generated. Thus, the $\mathbb{K}$-module $\mathbb{M}^n$ is finitely generated as well.

Clearly, $\mathbb{L}$ is a commutative $\mathbb{M}$-algebra (since $\mathbb{M}$ is a subring of $\mathbb{L}$). Also, the $\mathbb{M}$-algebra $\mathbb{M}[A]$ is commutative (since it is generated by a single element A over the commutative ring $\mathbb{M}$).

On the other hand, the Cayley–Hamilton theorem yields $\chi_A(A) = 0$. But the characteristic polynomial χ_A is a monic polynomial of degree n; furthermore, all its coefficients belong to $\mathbb{M}$ (since $\mathbb{M}$ was defined to be the $\mathbb{K}$-subalgebra of $\mathbb{L}$ generated by these coefficients). Thus, the polynomial χ_A belongs to $\mathbb{M}[t]$. Hence, there exists a monic polynomial $f \in \mathbb{M}[t]$ of degree n such that $f(A) = 0$ (namely, $f = \chi_A$). Hence, Corollary 1.36 (applied to $\mathbb{M}$, $\mathbb{M}[A]$ and A instead of $\mathbb{K}$, $\mathbb{L}$ and u) yields $\mathbb{M}[A] = \left\langle A^0, A^1, \ldots, A^{n-1} \right\rangle_{\mathbb{M}}$. Thus, each element of $\mathbb{M}[A]$ is an $\mathbb{M}$-linear combination of the powers $A^0, A^1, \ldots, A^{n-1}$. Therefore, the $\mathbb{K}$-module homomorphism

$$\begin{aligned} \pi : \mathbb{M}^n &\to \mathbb{M}[A], \\ (m_0, m_1, \ldots, m_{n-1}) &\mapsto m_0 A^0 + m_1 A^1 + \cdots + m_{n-1} A^{n-1} \end{aligned}$$

is surjective. Hence, Lemma 1.11 (applied to $M = \mathbb{M}^n$, $N = \mathbb{M}[A]$ and $f = \pi$) shows that the $\mathbb{K}$-module $\mathbb{M}[A]$ is finitely generated. This $\mathbb{K}$-module $\mathbb{M}[A]$ clearly satisfies $A \cdot \mathbb{M}[A] \subseteq \mathbb{M}[A]$ (since it is a $\mathbb{K}$-algebra and contains A). Moreover, every $v \in \mathbb{M}[A]$ satisfying $v \cdot \mathbb{M}[A] = 0$ satisfies $v = 0$ (since it satisfies $v = v \cdot \underbrace{1}_{\in \mathbb{M}[A]} \in v \cdot \mathbb{M}[A] = 0$). Hence, Corollary 1.13 (applied to $\mathbb{M}[A]$, A, $\mathbb{M}[A]$ and $\mathbb{M}[A]$ instead of $\mathbb{L}$, u, C and U) yields that $A \in \mathbb{M}[A]$ is integral over $\mathbb{K}$. In other words, $A \in \mathbb{L}^{n \times n}$ is integral over $\mathbb{K}$. This proves Proposition 1.17. $\square$

2. Dynamical Behaviour of Linear and Additive Cellular Automata

A *discrete time dynamical system* (DTDS) is a pair $(\mathcal{X}, \mathcal{F})$, where $\mathcal{X}$ is any set equipped with a distance function d and $\mathcal{F} : \mathcal{X} \to \mathcal{X}$ is a map that is continuous on $\mathcal{X}$ according to the topology induced by d (see [KatHas95, LinMar95] for a background on discrete time dynamical systems). The main goal of this section is to prove that several important properties of discrete time dynamical systems are decidable for additive cellular automata over a finite abelian group. First of all, we will prove that the following properties are decidable for a nontrivial subclass of additive cellular automata, namely, the linear cellular automata over $\mathbb{K}^n$ (see [LebMar95] for an introduction to them), where $\mathbb{K}$ is the finite ring $\mathbb{Z}/m\mathbb{Z}$:

- sensitivity to the initial conditions and equicontinuity;
- equicontinuity;

- topological transitivity and other mixing properties;
- ergodicity and other ergodic properties.

Then, we will extend the decidability results about the above mentioned properties and other ones (as injectivity and surjectivity) from linear cellular automata to additive cellular automata over a finite abelian group.

2.1. Background on DTDS and Cellular Automata

We begin by reviewing some general notions about discrete time dynamical systems and cellular automata.

Let $(\mathcal{X}, \mathcal{F})$ be a DTDS.

We say that $(\mathcal{X}, \mathcal{F})$ is *surjective*, resp., *injective*, if $\mathcal{F}$ is *surjective*, resp., *injective*. The DTDS $(\mathcal{X}, \mathcal{F})$ is *sensitive to the initial conditions* (or simply *sensitive*) if there exists $\varepsilon > 0$ such that for any $x \in \mathcal{X}$ and any $\delta > 0$ there is an element $y \in \mathcal{X}$ such that $0 < d(y, x) < \delta$ and $d(\mathcal{F}^k(y), \mathcal{F}^k(x)) > \varepsilon$ for some $k \in \mathbb{N}$. The system $(\mathcal{X}, \mathcal{F})$ is said to be *equicontinuous* if $\forall \varepsilon > 0$ there exists $\delta > 0$ such that for all $x, y \in \mathcal{X}$, $d(x, y) < \delta$ implies that $\forall k \in \mathbb{N}$, $d(\mathcal{F}^k(x), \mathcal{F}^k(y)) < \varepsilon$. As dynamical properties, sensitivity and equicontinuity represent the main features of unstable and stable dynamical systems, respectively. The former is the well-known basic component and essence of the chaotic behavior of discrete time dynamical systems, while the latter is a strong form of stability.

The DTDS $(\mathcal{X}, \mathcal{F})$ is *topologically transitive* (or, simply, *transitive*) if for all nonempty open subsets U and V of $\mathcal{X}$ there exists a natural number h such that $\mathcal{F}^h(U) \cap V \neq \emptyset$, while it is said to be *topologically mixing* if for all nonempty open subsets U and V of $\mathcal{X}$ there exists a natural number h_0 such that the previous intersection condition holds for every $h \geq h_0$. Clearly, topological mixing is a stronger condition than transitivity. Moreover, $(\mathcal{X}, \mathcal{F})$ is *topologically weakly mixing* if the DTDS $(\mathcal{X} \times \mathcal{X}, \mathcal{F} \times \mathcal{F})$ is topologically transitive, while it is *totally transitive* if $(\mathcal{X}, \mathcal{F}^h)$ is topologically transitive for all $h \in \mathbb{N}$.

Let $(\mathcal{X}, \mathcal{M}, \mu)$ be a probability space and let $(\mathcal{X}, \mathcal{F})$ be a DTDS where $\mathcal{F}$ is a measurable map which preserves μ, i.e., $\mu(E) = \mu(\mathcal{F}^{-1}(E))$ for every $E \in \mathcal{M}$. The DTDS $(\mathcal{X}, \mathcal{F})$, or, the map $\mathcal{F}$, is *ergodic* with respect to μ if for every $E \in \mathcal{M}$

$$\left(E = \mathcal{F}^{-1}(E)\right) \Rightarrow \mu(E)(1 - \mu(E)) = 0$$

It is well known that $\mathcal{F}$ is ergodic iff for any pair of sets $A, B \in \mathcal{M}$ it holds that

$$\lim_{h \to \infty} \frac{1}{h} \sum_{i=0}^{h-1} \mu(\mathcal{F}^{-i}(A) \cap B) = \mu(A)\mu(B)$$

The DTDS $(\mathcal{X}, \mathcal{F})$ is *(ergodic) mixing*, if for any pair of sets $A, B \in \mathcal{M}$ it holds that

$$\lim_{h \to \infty} \mu(\mathcal{F}^{-h}(A) \cap B) = \mu(A)\mu(B) \ ,$$

while it is *(ergodic) weak mixing*, if for any pair of sets $A, B \in \mathcal{M}$ it holds that

$$\lim_{h\to\infty} \frac{1}{h} \sum_{i=0}^{h-1} |\mu(\mathcal{F}^{-i}(A) \cap B) - \mu(A)\mu(B)| = 0$$

We now recall some general notions about cellular automata.

Let S be a finite set. A configuration over S is a map from $\mathbb{Z}$ to S. We consider the following *space of configurations*

$$S^{\mathbb{Z}} = \{\boldsymbol{c} |\ \boldsymbol{c} \colon \mathbb{Z} \to S\}\,.$$

Each element $\boldsymbol{c} \in S^{\mathbb{Z}}$ can be visualized as an infinite one-dimensional cell lattice in which each cell $i \in \mathbb{Z}$ contains the element $\boldsymbol{c}_i \in S$.

Let $r \in \mathbb{N}$ and $\delta \colon S^{2r+1} \to S^n$ be any map. We say that r is the radius of δ.

Definition 2.1 (Cellular Automaton)**.** A *one-dimensional CA based on a radius r local rule* δ is a pair $(S^{\mathbb{Z}}, F)$, where

$$F \colon S^{\mathbb{Z}} \to S^{\mathbb{Z}},$$

is the *global transition map* defined as follows:

$$\forall \boldsymbol{c} \in S^{\mathbb{Z}},\ \forall i \in \mathbb{Z}, \quad F(\boldsymbol{c})_i = \delta\left(\boldsymbol{c}_{i-r}, \ldots, \boldsymbol{c}_{i+r}\right). \tag{25}$$

In other words, the content of cell i in the configuration $F(\boldsymbol{c})$ is a function of the content of cells $i-r, \ldots, i+r$ in the configuration $\boldsymbol{c}$.

We stress that the local rule δ completely determines the global rule F of a CA.

In order to study the dynamical properties of one-dimensional CA, we introduce a distance over the space of the configurations. Namely, $S^{\mathbb{Z}}$ is equipped with the Tychonoff distance d defined as follows

$$\forall \boldsymbol{c}, \boldsymbol{c}' \in S^{\mathbb{Z}}, \quad d(\boldsymbol{c}, \boldsymbol{c}') = \frac{1}{2^{\ell}} \quad \text{where } \ell = \min\{i \in \mathbb{N} \,:\, \boldsymbol{c}_i \neq \boldsymbol{c}'_i \text{ or } \boldsymbol{c}_{-i} \neq \boldsymbol{c}'_{-i}\}\ .$$

It is easy to verify that metric topology induced by d coincides with the product topology induced by the discrete topology on $S^{\mathbb{Z}}$. With this topology, $S^{\mathbb{Z}}$ is a compact and totally disconnected space and the global transition map F of any CA $(S^{\mathbb{Z}}, F)$ turns out to be (uniformly) continuous. Therefore, any CA itself is also a discrete time dynamical system. Moreover, any map $F : S^{\mathbb{Z}} \to S^{\mathbb{Z}}$ is the global transition rule of a CA if and only if F is (uniformly) continuous and $F \circ \sigma = \sigma \circ F$, where $\sigma : S^{\mathbb{Z}} \to S^{\mathbb{Z}}$ is the *shift map* defined as $\forall \boldsymbol{c} \in S^{\mathbb{Z}}$, $\forall i \in \mathbb{Z}$, $\sigma(\boldsymbol{c})_i = \boldsymbol{c}_{i+1}$. From now, when no misunderstanding is possible, we identify a CA with its global rule. Moreover, whenever an ergodic property is considered for CA, μ is the well-known Haar measure over the collection $\mathcal{M}$ of measurable subsets of $S^{\mathbb{Z}}$, i.e., the one defined as the product measure induced by the uniform probability distribution over S.

2.2. Additive and Linear Cellular Automata

Let us introduce the background of additive CA. The alphabet S will be a finite abelian group G, with group operation $+$, neutral element 0, and inverse operation $-$. In this way, the configuration space $G^{\mathbb{Z}}$ turns out to be a finite abelian group, too, where the group operation of $G^{\mathbb{Z}}$ is the component-wise extension of $+$ to $G^{\mathbb{Z}}$. With an abuse of notation, we denote by the same symbols $+$, 0, and $-$ the group operation, the neutral element, and the inverse operation, respectively, both of G and $G^{\mathbb{Z}}$. Observe that $+$ and $-$ are continuous functions in the topology induced by the metric d. A configuration $\boldsymbol{c} \in G^{\mathbb{Z}}$ is said to be *finite* if the number of positions $i \in \mathbb{Z}$ with $\boldsymbol{c}_i \neq 0$ is finite.

Definition 2.2 (Additive Cellular Automata)**.** An *additive CA* over a abelian finite group G is a CA $(G^{\mathbb{Z}}, F)$ where the global transition map $F : G^{\mathbb{Z}} \to G^{\mathbb{Z}}$ is an endomorphism of $G^{\mathbb{Z}}$.

The *sum of two additive CA* F_1 and F_2 over G is naturally defined as the map on $G^{\mathbb{Z}}$ denoted by $F_1 + F_2$ and such that

$$\forall \boldsymbol{c} \in G^{\mathbb{Z}}, \quad (F_1 + F_2)(\boldsymbol{c}) = F_1(\boldsymbol{c}) + F_2(\boldsymbol{c})$$

Clearly, $F_1 + F_2$ is an additive CA over G.

We now recall the notion of linear CA, an important subclass of additive CA. We stress that, whenever the term *linear* is involved, the alphabet S is $\mathbb{K}^n$, where $\mathbb{K} = \mathbb{Z}/m\mathbb{Z}$ for some positive integer m. Both $\mathbb{K}^n$ and $(\mathbb{K}^n)^{\mathbb{Z}}$ become $\mathbb{K}$-modules in the obvious (i.e., entrywise) way.

A local rule $\delta\colon (\mathbb{K}^n)^{2r+1} \to \mathbb{K}^n$ of radius r is said to be linear if it is defined by $2r+1$ matrices $A_{-r}, \ldots, A_0, \ldots, A_r \in \mathbb{K}^{n \times n}$ as follows:

$$\forall (x_{-r}, \ldots, x_0, \ldots, x_r) \in (\mathbb{K}^n)^{2r+1}, \quad \delta(x_{-r}, \ldots, x_0, \ldots, x_r) = \sum_{i=-r}^{r} A_i \cdot x_i \ .$$

Definition 2.3 (Linear Cellular Automata (LCA))**.** A *linear CA (LCA)* over $\mathbb{K}^n$ is a CA based on a linear local rule.

We remark that for a linear one-dimensional CA, equation (25) becomes

$$F(\boldsymbol{c})_i = A_{i-r} \cdot \boldsymbol{c}_{i-r} + \ldots + A_{i+r} \cdot \boldsymbol{c}_{i+r}$$

Let $\mathbb{K}[X, X^{-1}]$ denote the set of Laurent polynomials with coefficients in $\mathbb{K}$. Before proceeding, let us recall the *formal power series* (fps) which have been successfully used to study the dynamical behaviour of LCA in the case $n = 1$ [ItoOsa83, ManMan99]. The idea of this formalism is that configurations and global rules are represented by suitable polynomials and the application of the global rule turns

into multiplications of polynomials. In the more general case of LCA over $\mathbb{K}^n$, a configuration $\boldsymbol{c} \in (\mathbb{K}^n)^{\mathbb{Z}}$ can be associated with the fps

$$\boldsymbol{P_c}(X) = \sum_{i\in\mathbb{Z}} \boldsymbol{c}_i X^i = \begin{bmatrix} c^1(X) \\ \vdots \\ c^n(X) \end{bmatrix} = \begin{bmatrix} \sum_{i\in\mathbb{Z}} c_i^1 X^i \\ \vdots \\ \sum_{i\in\mathbb{Z}} c_i^n X^i \end{bmatrix}$$

Then, if F is the global rule of a LCA defined by $A_{-r}, \ldots, A_0, \ldots, A_r$, one finds

$$\boldsymbol{P}_{F(\boldsymbol{c})}(X) = A \cdot \boldsymbol{P_c}(X)$$

where

$$A = \sum_{i=-r}^{r} A_i X^{-i} \in \mathbb{K}[X, X^{-1}]^{n\times n}$$

is the *finite fps*, or, *the matrix, associated with the LCA F*. In this way, for any integer $k > 0$ the fps associated with F^k is A^t, and then $\boldsymbol{P}_{F^k(\boldsymbol{c})}(X) = A^k \cdot \boldsymbol{P_c}(X)$.

A matrix $A \in \mathbb{K}[X, X^{-1}]^{n\times n}$ is in *Frobenius normal form* if

$$A = \begin{bmatrix} 0 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 1 & \ddots & 0 & 0 \\ 0 & 0 & 0 & \ddots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 0 & 1 \\ \mathbb{a}_0 & \mathbb{a}_1 & \mathbb{a}_2 & \ldots & \mathbb{a}_{n-2} & \mathbb{a}_{n-1} \end{bmatrix} \tag{26}$$

where each $\mathbb{a}_i \in \mathbb{K}[X, X^{-1}]$. Recall that the coefficients of χ_A are just the elements $\mathbb{a}_i$ of the n-th row of A (up to sign).

Definition 2.4 (Frobenius LCA)**.** A LCA $((\mathbb{K}^n)^{\mathbb{Z}}, F)$ is said to be a *Frobenius LCA* if the fps $A \in \mathbb{K}[X, X^{-1}]^{n\times n}$ associated with F is in Frobenius normal form.

2.3. Decidability Results about Linear CA

We now deal with sensitivity and equicontinuity for LCA over $\mathbb{K}^n$. First of all, we remind that a dichotomy between sensitivity and equicontinuity holds for LCA. Moreover, these properties are characterized by the behavior of the powers of the matrix associated with a LCA.

Proposition 2.5 ([Dennun19]). Let $\left((\mathbb{K}^n)^{\mathbb{Z}}, F\right)$ be a LCA over $\mathbb{K}^n$ and let A be the matrix associated with F. The following statements are equivalent:

1. F is sensitive to the initial conditions;
2. F is not equicontinuous;
3. $\left|\{A^1, A^2, A^3, \ldots\}\right| = \infty$.

Note that the statement "$\left|\{A^1, A^2, A^3, \ldots\}\right| = \infty$" here is clearly equivalent to "$\left|\{A^0, A^1, A^2, \ldots\}\right| = \infty$", which is the kind of statement discussed in Theorem 1.2. An immediate consequence of Proposition 2.5 is that any decidable characterization of sensitivity to the initial conditions in terms of the matrices defining LCA over $\mathbb{K}^n$ would also provide a characterization of equicontinuity. In the sequel, we are going to show that such a characterization actually exists. First of all, we remind that the a decidable characterization of sensitivity and equicontinuity was provided for the class of Frobenius LCA in [Dennun19]. In particular, the following result holds.

Theorem 2.6 (Theorem 31 in [Dennun19]). Sensitivity and equicontinuity are decidable for Frobenius LCA over $\mathbb{K}^n$.

By means of the main result from Section 1 we are now able to prove the following

Theorem 2.7. Sensitivity and equicontinuity are decidable for LCA over $\mathbb{K}^n$.

Proof. Let $\left((\mathbb{K}^n)^{\mathbb{Z}}, G\right)$ be any LCA over $\mathbb{K}^n$ and let A be the matrix associated with G. Consider the Frobenius LCA $\left((\mathbb{K}^n)^{\mathbb{Z}}, F\right)$ such that $\chi_A = \chi_B$, where B is the matrix (in Frobenius normal form) associated with F. By Theorem 1.2 and Proposition 2.5 the former LCA is equicontinuous if and only if the latter is. Theorem 2.6 concludes the proof. □

For a sake of completeness, we recall that injectivity and surjectivity are decidable for LCA over $\mathbb{K}^n$. This result follows from a characterization of these properties in terms of the determinant of the matrix associated with a LCA and from the fact that injectivity and surjectivity are decidable for LCA over $\mathbb{K}$ (for the latter, see [ItoOsa83]).

Theorem 2.8 ([LebMar95, Kari00]). Injectivity and equicontinuity are decidable for LCA over $\mathbb{K}^n$. In particular, a LCA over $\mathbb{K}^n$ is injective (resp., surjective) if and only if the determinant of the matrix associated with it is the fps associated with an injective (resp., surjective) LCA over $\mathbb{K}$.

The decidability of topologically transitivity, ergodicity, and other mixing and ergodic properties for LCA over $\mathbb{K}^n$ has been recently proved in [Dennun19b]. In particular, authors showed the equivalence of all the mixing and ergodic properties for additive CA over a finite abelian group and the decidability for LCA over $\mathbb{K}^n$ (see also [Dennun19c]).

Theorem 2.9 ([Dennun19b])**.** Let F be any additive CA over a finite abelian group. The following statements are equivalent:

1. F is topologically transitive;
2. F is ergodic;
3. F is surjective and for every $k \in \mathbb{N}$ it holds that $F^k - I$ is surjective;
4. F is topologically mixing;
5. F is weak topologically transitive;
6. F is totally transitive;
7. F is weakly ergodic mixing;
8. F is ergodic mixing.

Moreover, all the previously mentioned properties are decidable for LCA over $\mathbb{K}^n$.

2.4. From Linear to Additive CA

In this section we are going to prove that sensitivity, equicontinuity, injectivity, surjectivity, topological transitivity, and all the properties equivalent to the latter are decidable also for additive CA over a finite abelian group. For each of them we will reach the decidability result by exploiting the analogous one obtained for LCA and extending it to the wide class of additive CA over a finite abelian group.

We recall that the local rule $\delta : G^{2r+1} \to G$ of an additive CA of radius r over a finite abelian group G can be written as

$$\forall (x_{-r}, \dots, x_r) \in G^{2r+1}, \qquad \delta(x_{-r}, \dots, x_r) = \sum_{i=-r}^{r} \delta_i(x_i) \tag{27}$$

where the functions δ_i are endomorphisms of G.

The fundamental theorem of finite abelian groups states that every finite abelian group G is isomorphic to $\bigoplus_{i=1}^{h} \mathbb{Z}/k_i\mathbb{Z}$ where the numbers $k_1, \dots, k_h$ are powers of (not necessarily distinct) primes and $\oplus$ is the direct sum operation. Hence, the global rule F of an additive CA over G splits into the direct sum of a suitable number h' of additive CA over subgroups $G_1, \dots, G_{h'}$ with $h' \leq h$ and such that

$\gcd(|G_i|, |G_j|) = 1$ for each pair of distinct $i, j \in \{1, \ldots, h'\}$. Each of them can be studied separately and then the analysis of the dynamical behavior of F can be carried out by combining together the results obtained for each component.

In order to make things clearer, consider the following example. If F is an additive CA over $G \cong \mathbb{Z}/4\mathbb{Z} \times \mathbb{Z}/8\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/25\mathbb{Z}$ then F splits into the direct sum of 3 additive CA F_1, F_2, and F_3 over $\mathbb{Z}/4\mathbb{Z} \times \mathbb{Z}/8\mathbb{Z}$, $\mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z}$ and $\mathbb{Z}/25\mathbb{Z}$, respectively. Therefore, F will be sensitive to initial conditions (resp., topologically transitive) iff at least one (resp., each) F_i is sensitive to initial conditions.

The above considerations lead us to three distinct scenarios:

1) $G \cong \mathbb{Z}/p^k\mathbb{Z}$. Then, G is cyclic and we can define each δ_i simply assigning the value of δ_i applied to the unique generator of G. Moreover, every pair δ_i, δ_j commutes, i.e., $\delta_i \circ \delta_j = \delta_j \circ \delta_i$, and this makes it possible a detailed analysis of the global behavior of F. Indeed, additive cellular automata over $\mathbb{Z}/p^k\mathbb{Z}$ are nothing but LCA over $\mathbb{Z}/p^k\mathbb{Z}$ and almost all dynamical properties, including sensitivity to the initial conditions, equicontinutity, injectivity, surjectivity, topological transitivity and so on are well understood and characterized (see [ManMan99]).

2) $G \cong (\mathbb{Z}/p^k\mathbb{Z})^n$. In this case, G is not cyclic anymore and has n generators. We can define each δ_i assigning the value of δ_i for each generator of G. This gives rise to the class of linear CA over $(\mathbb{Z}/p^k\mathbb{Z})^n$. Now, δ_i and δ_j do not commute in general and this makes the analysis of the dynamical behavior much harder. Nevertheless, in Section 2 we have proved that sensitivity and equicontinuity are decidable by exploiting the main result of Section 1. As pointed out in [Dennun19], we also recall that linear CA over $(\mathbb{Z}/p^k\mathbb{Z})^n$ allow the investigation of some classes of non-uniform CA over $\mathbb{Z}/p^k\mathbb{Z}$ ((for these latter see [Cattan09, Dennun12, Dennun13]).

3) $G \cong \bigoplus_{i=1}^{n} \mathbb{Z}/p^{k_i}\mathbb{Z}$. In this case ($\mathbb{Z}/4\mathbb{Z} \times \mathbb{Z}/8\mathbb{Z}$ in the example), G is again not cyclic and turns out to be a subsystem of a suitable LCA. Then, the analysis of the dynamical behavior of F is even more complex than in **2)**. We do not even know easy checkable characterizations of basic properties like surjectivity or injectivity so far. We will provide them in the sequel as we stated at the beginning of this section.

Therefore, WLOG in the sequel we can assume that $G = \mathbb{Z}/p^{k_1}\mathbb{Z} \times \ldots \times \mathbb{Z}/p^{k_n}\mathbb{Z}$ with $k_1 \geq k_2 \geq \ldots \geq k_n$ in order to reach our goal.

For any $i \in \{1, \ldots, n\}$ let us denote by $\boldsymbol{e}^{(i)} \in G^{\mathbb{Z}}$ the bi-infinite configuration such that $\boldsymbol{e}_0^{(i)} = e_i$ and $\boldsymbol{e}_j^{(i)} = 0$ for every integer $j \neq 0$.

Definition 2.10. Let $(G^{\mathbb{Z}}, F)$ be an additive CA over G. We say that $\boldsymbol{e}^{(i)} \in G^{\mathbb{Z}}$ *spreads under* F if for every $\ell \in \mathbb{N}$ there exists $k \in \mathbb{N}$ such that $F^k(\boldsymbol{e}^{(i)})_j \neq 0$ for some integer j with $|j| > \ell$.

Remark 2.11. Whenever we consider $\boldsymbol{P}_{\boldsymbol{e}^{(i)}}(X) \in G[X, X^{-1}]$, we will say that $\boldsymbol{P}_{\boldsymbol{e}^{(i)}}(X)$ spreads under F if for every $\ell \in \mathbb{N}$ there exists $k \in \mathbb{N}$ such that $\boldsymbol{P}_{F^k(\boldsymbol{e}^{(i)})}(X)$ has at least one component with a non null monomial of degree which is greater than ℓ in absolute value. Clearly, $\boldsymbol{P}_{\boldsymbol{e}^{(i)}}(X)$ spreads under F if and only if $\boldsymbol{e}^{(i)}$ spreads under F.

Let $\hat{G} = \mathbb{Z}/p^{k_1}\mathbb{Z} \times \ldots \times \mathbb{Z}/p^{k_1}\mathbb{Z}$. Define the map $\psi : G \to \hat{G}$ as follows

$$\forall i = 1, \ldots, n, \quad \forall h \in G, \quad \psi(h)^i = h^i \, p^{k_1 - k_i} \ .$$

Definition 2.12. We define the function $\Psi : G^{\mathbb{Z}} \to \hat{G}^{\mathbb{Z}}$ as the component-wise extension of ψ, i.e.,

$$\forall \boldsymbol{c} \in G^{\mathbb{Z}}, \quad \forall j \in \mathbb{Z}, \quad \Psi(\boldsymbol{c})_j = \psi(\boldsymbol{c}_j) \ .$$

It is easy to check that Ψ is continuous and injective, but not surjective. Since every configuration $\boldsymbol{c} \in G^{\mathbb{Z}}$ (or $\hat{G}^{\mathbb{Z}}$) is associated with the fps $\boldsymbol{P}_{\boldsymbol{c}}(X) \in G[X, X^{-1}]$ (or $\hat{G}[X, X^{-1}]$), with an abuse of notation we will sometimes consider Ψ as map from $G[X, X^{-1}]$ to $\hat{G}[X, X^{-1}]$ with the obvious meaning.

For any additive CA over G, we are now going to define a LCA over $(\mathbb{Z}/p^{k_1}\mathbb{Z})^n$ associated to it. With a further abuse of notation, in the sequel we will write p^{-m} with $m \in \mathbb{N}$ even if this quantity might not exist in $\mathbb{Z}/p^k\mathbb{Z}$. However, we will use it only when it multiplies $p^{m'}$ for some integer $m' > m$. In such a way $p^{m'-m}$ is well defined in $\mathbb{Z}/p^k\mathbb{Z}$ and we will note it as product $p^{-m} \cdot p^{m'}$.

Definition 2.13. Let $(G^{\mathbb{Z}}, F)$ be any additive CA and let $\delta : G^{2r+1} \to G$ be its local rule defined, according to (27), by $2r+1$ endomorphisms $\delta_{-r}, \ldots, \delta_r$ of G . For each $z \in \{-r, \ldots, r\}$, we define the matrix $A_z = (a_{i,j}^{(z)})_{1 \le i \le n, 1 \le j \le n} \in (\mathbb{Z}/p^{k_1}\mathbb{Z})^{n \times n}$ as

$$\forall i, j \in \{1, \ldots, n\}, \qquad a_{i,j}^{(z)} = p^{k_j - k_i} \cdot \delta_z(e_j)^i$$

The *LCA associated with the additive CA* $(G^{\mathbb{Z}}, F)$ is $(\hat{G}^{\mathbb{Z}}, L)$, where L is defined by $A_{-r}, \ldots, A_r$ or, equivalently, by $A = \sum_{z=-r}^{r} A_z X^{-z} \in \hat{G}[X, X^{-1}]^{n \times n}$.

Remark 2.14. Since every δ_z is an endomorphism of G, by construction A turns out to be well defined.

Remark 2.15. The following diagram commutes

$$\begin{array}{ccc} G^{\mathbb{Z}} & \xrightarrow{F} & G^{\mathbb{Z}} \\ \Psi \big\downarrow & & \big\downarrow \Psi \\ \hat{G}^{\mathbb{Z}} & \xrightarrow[L]{} & \hat{G}^{\mathbb{Z}} \end{array},$$

i.e., $L \circ \Psi = \Psi \circ F$. This is the reason why we also say that $(\hat{G}^{\mathbb{Z}}, L)$ is the LCA associated with $(G^{\mathbb{Z}}, F)$ *via the embedding* Ψ.

2.4.1. Sensitivity and Equicontinuity for Additive Cellular Automata

Let us start with the decidability of sensitivity and equicontinuity.

Lemma 2.16. Let $(G^{\mathbb{Z}}, F)$ be any additive CA. If for some $i \in \{1, \dots, n\}$ the configuration $\boldsymbol{e}^{(i)} \in G^{\mathbb{Z}}$ spreads under F then $(G^{\mathbb{Z}}, F)$ is sensitive to the initial conditions.

Proof. We prove that F is sensitive with constant $\varepsilon = 1$. Let $\boldsymbol{e}^{(i)} \in G^{\mathbb{Z}}$ be the configuration spreading under F. Choose arbitrarily an integer $\ell \in \mathbb{N}$ and a configuration $\boldsymbol{c} \in G^{\mathbb{Z}}$. Let $t \in \mathbb{N}$ and $j \notin \{-\ell, \dots, \ell\}$ be the integers such that $F^t(\boldsymbol{e}^{(i)})_j \neq 0$. Consider the configuration $\boldsymbol{c}' = \boldsymbol{c} + \sigma^j(\boldsymbol{e}^{(i)})$. Clearly, it holds that $d(\boldsymbol{c}, \boldsymbol{c}') < 2^{-\ell}$ and $F^t(\boldsymbol{c}') = F^t(\boldsymbol{c}) + F^t(\sigma^j(\boldsymbol{e}^{(i)})) = F^t(\boldsymbol{c}) + \sigma^j(F^t(\boldsymbol{e}^{(i)}))$. So, we get $d(F^t(\boldsymbol{c}'), F^t(\boldsymbol{c})) = 1$ and this concludes the proof. □

In order to prove the decidability of sensitivity, we need to deal with the degree of a polynomial. We stress that the notions we are going to introduce apply to finite Laurent polynomials.

Definition 2.17. Given any finite polynomial $\mathbb{p}(X) \in \mathbb{Z}/p^{k_1}\mathbb{Z}\left[X, X^{-1}\right]$, the *positive* (resp., *negative*) *degree of* $\mathbb{p}(X)$, denoted by $deg^+[\mathbb{p}(X)]$ (resp., $deg^-[\mathbb{p}(X)]$) is the maximum (resp., minimum) degree among those of the monomials having both positive (resp., negative) degree and coefficient which is not multiple of p. If there is no monomial satisfying both the required conditions, then $deg^+[\mathbb{p}(X)] = 0$ (resp., $deg^-[\mathbb{p}(X)]$=0).

Lemma 2.18. Let $(\hat{G}^{\mathbb{Z}}, L)$ be a LCA and let $A \in \mathbb{Z}/p^{k_1}\mathbb{Z}\left[X, X^{-1}\right]^{n \times n}$ be the matrix associated to it. If $(\hat{G}^{\mathbb{Z}}, L)$ is sensitive then for every integer $m \geq 1$ there exists an integer $k \geq 1$ such that at least one entry of A^k has either positive or negative degree with absolute value which is greater than m.

Proof. We can write $A = B + p \cdot C$ for some $A, B \in \mathbb{Z}/p^{k_1}\mathbb{Z}\left[X, X^{-1}\right]^{n\times n}$, where the monomials of all entries of A have coefficient which is not multiple of p. Assume that there exists a bound $b \geq 1$ such that for every $k \geq 1$ all entries of A^k have degree less than b in absolute value. Therefore, it holds that $\left|\{A^k, k \geq 1\}\right| < \infty$ and so $(\hat{G}^{\mathbb{Z}}, L)$ is not sensitive. □

We are now able to prove the following important result.

Theorem 2.19. Let $(G^{\mathbb{Z}}, F)$ be any additive CA over G and let $(\hat{G}^{\mathbb{Z}}, L)$ be the LCA associated to it via the embedding Ψ. Then, the CA $(G^{\mathbb{Z}}, F)$ is sensitive to the initial conditions if and only if $(\hat{G}^{\mathbb{Z}}, L)$ is.

Proof. $\Longrightarrow$: Assume that $(\hat{G}^{\mathbb{Z}}, L)$ is not sensitive. Then, by Proposition 2.5, there exist two integers $k \in \mathbb{N}$ and $m > 0$ such that $L^{k+m} = L^k$. Therefore, we get $\Psi \circ F^{k+m} = L^{k+m} \circ \Psi = L^k \circ \Psi = \Psi \circ F^k$. Since Ψ is injective, it holds that $F^{k+m} = F^k$ and so $(G^{\mathbb{Z}}, F)$ is not sensitive.
$\Longleftarrow$: Assume that $(\hat{G}^{\mathbb{Z}}, L)$ is sensitive and for any natural k let $A^k = (a_{i,j}^{(k)})_{1\leq i\leq n, 1\leq j\leq n}$ be the k-th power of $A \in \mathbb{Z}/p^{k_1}\mathbb{Z}\left[X, X^{-1}\right]^{n\times n}$, where A is the matrix associated to $(\hat{G}^{\mathbb{Z}}, L)$. We are going to show that at least one configuration among $\boldsymbol{e}^{(1)}, \ldots, \boldsymbol{e}^{(n)}$ spreads under F. Choose arbitrarily $\ell \in \mathbb{N}$. By Lemma 2.18, there exist an integer $m \geq 1$ and one entry (i, j) such that either $deg^-[a_{i,j}^{(m)}] < -\ell$ or $deg^+[a_{i,j}^{(m)}] > \ell$. WLOG suppose that $deg^+[a_{i,j}^{(m)}] > \ell$. The i–th component of $\boldsymbol{P}_{F^m(\boldsymbol{e}^{(j)})}(X)$ is the well defined polynomial $p^{k_i-k_1} \cdot p^{k_1-k_j} \cdot a_{i,j}^{(m)}$. Since $deg^+[a_{i,j}^{(m)}] > \ell$, we can state that $\boldsymbol{e}^{(j)}$ spreads under F. By Lemma 2.16, it follows that $(G^{\mathbb{Z}}, F)$ is sensitive. □

As immediate consequence of Theorem 2.19 we can state that the dichotomy between sensitivity and equicontinuity also holds for additive CA.

Corollary 2.20. Any additive CA over a finite abelian group is sensitive to the initial conditions if and only if it is not equicontinuous.

The following decidability result follows from Theorem 2.19 and the decidability of sensitivity for LCA.

Corollary 2.21. Equicontinuity and sensitivity to the initial conditions are decidable for additive CA over a finite abelian group.

Proof. Use Theorem 2.7 and 2.19. □

2.4.2. Surjectivity and Injectivity for Additive Cellular Automata

We now study injectivity and surjectivity for additive CA.

Lemma 2.22. Let $(\hat{G}^{\mathbb{Z}}, L)$ be any LCA over $\hat{G}$. For any configuration $\boldsymbol{b} \in \hat{G}^{\mathbb{Z}}$ with $\boldsymbol{b} \neq 0$ and $L(\boldsymbol{b}) = 0$ there exists a configuration $\boldsymbol{b}' \in \Psi(G^{\mathbb{Z}})$ such that $\boldsymbol{b}' \neq 0$ and $L(\boldsymbol{b}') = 0$. In particular, if $\boldsymbol{b}$ is finite then $\boldsymbol{b}'$ is too.

Proof. Let $\boldsymbol{b} \in \hat{G}^{\mathbb{Z}}$ any configuration with $\boldsymbol{b} \neq 0$ and $L(\boldsymbol{b}) = 0$. Set $\boldsymbol{b}^{(1)} = p \cdot \boldsymbol{b}$. If $\boldsymbol{b}^{(1)} = 0$ then for every $i \in \mathbb{Z}$ each component of $\boldsymbol{b}_i$ has p^{k_1-1} as factor. So, $\boldsymbol{b} \in \Psi(G^{\mathbb{Z}})$ and $\boldsymbol{b}' = \boldsymbol{b}$ is just one possible configuration the thesis requires to exhibit. Otherwise, by repeating the same argument, set $\boldsymbol{b}^{(2)} = p \cdot \boldsymbol{b}^{(1)}$. If $\boldsymbol{b}^{(2)} = 0$ then, for every $i \in \mathbb{Z}$, each component of $\boldsymbol{b}_i^{(1)}$ has p^{k_1-1} as factor and so $\boldsymbol{b}^{(1)} \in \Psi(G^{\mathbb{Z}})$. Since $L(\boldsymbol{b}^{(1)}) = 0$, a configuration we are looking for is $\boldsymbol{b}' = \boldsymbol{b}^{(1)}$. After $k-1$ iterations, i.e., once we get $\boldsymbol{b}^{(k_1-1)} = p \cdot \boldsymbol{b}^{(k-2)}$ (with $\boldsymbol{b}^{(k-2)} \neq 0$), if $\boldsymbol{b}^{(k_1-1)} = 0$ holds we conclude that $\boldsymbol{b}' = \boldsymbol{b}^{(k-2)}$ by using the same argument of the previous steps. Otherwise, by definition, for every $i \in \mathbb{Z}$ each component of $\boldsymbol{b}_i^{(k_1-1)}$ itself certainly contains p^{k_1-1} as factor. Therefore, $\boldsymbol{b}^{(k_1-1)} \in \Psi(G^{\mathbb{Z}})$. Moreover, $L(\boldsymbol{b}^{(k_1-1)}) = 0$. Hence, we can set $\boldsymbol{b}' = \boldsymbol{b}^{(k_1-1)}$ and this concludes the proof. □

The following lemma will be useful for studying both surjectivity and other properties.

Lemma 2.23. Let $(G^{\mathbb{Z}}, F)$ and $(\hat{G}^{\mathbb{Z}}, L)$ be any additive CA over G and any LCA over $\hat{G}$, respectively, such that $L \circ \Psi = \Psi \circ F$. Then, the CA $(G^{\mathbb{Z}}, F)$ is surjective if and only if $(\hat{G}^{\mathbb{Z}}, L)$ is.

Proof. $\Longleftarrow$: Assume that F is not surjective. Then, F is not injective on the finite configurations, i.e., there exist two distinct and finite $\boldsymbol{c}', \boldsymbol{c}'' \in G^{\mathbb{Z}}$ with $F(\boldsymbol{c}') = F(\boldsymbol{c}'')$. Therefore, the element $\boldsymbol{c} = \boldsymbol{c}' - \boldsymbol{c}'' \in G^{\mathbb{Z}}$ is a finite configuration such that $\boldsymbol{c} \neq 0$ and $F(\boldsymbol{c}) = 0$. So, we get both $\Psi(\boldsymbol{c}) \neq 0$ and $L(\Psi(\boldsymbol{c})) = \Psi(F(\boldsymbol{c})) = 0$. Since $\Psi(\boldsymbol{c}) \neq 0$, it follows that L is not surjective.
$\Longrightarrow$: Assume that L is not surjective. Then, it is not injective on the finite configurations. Thus, there exist a finite configuration $\boldsymbol{b} \neq 0$ with $L(\boldsymbol{b}) = 0$. By Lemma 2.22, there exists a finite configuration $\boldsymbol{b}' \in \Psi(G^{\mathbb{Z}})$ such that $\boldsymbol{b}' \neq 0$ and $L(\boldsymbol{b}') = 0$. Let $\boldsymbol{c} \in G^{\mathbb{Z}}$ be the finite configuration such that $\Psi(\boldsymbol{c}) = \boldsymbol{b}'$. Clearly, it holds that $\boldsymbol{c} \neq 0$. We get $\Psi(F(\boldsymbol{c})) = L(\Psi(\boldsymbol{c})) = 0$. Since Ψ is injective, it follows that $F(\boldsymbol{c}) = 0$. Therefore, we conclude that F is not surjective. □

Next two theorems state that surjectivity and injectivity behave as sensitivity when looking at an additive CA over G and the associated LCA via the embedding Ψ.

Theorem 2.24. Let $(G^{\mathbb{Z}}, F)$ be any additive CA over G and let $(\hat{G}^{\mathbb{Z}}, L)$ be the LCA associated with it via the embedding Ψ. Then, the CA $(G^{\mathbb{Z}}, F)$ is surjective if and only if $(\hat{G}^{\mathbb{Z}}, L)$ is.

Proof. Use Lemma 2.23. □

Theorem 2.25. Let $(G^{\mathbb{Z}}, F)$ be any additive CA and let $(\hat{G}^{\mathbb{Z}}, L)$ be the LCA associated with it via the embedding Ψ. Then, the CA $(G^{\mathbb{Z}}, F)$ is injective if and only if $(\hat{G}^{\mathbb{Z}}, L)$ is.

Proof. $\Longleftarrow$: Assume that F is not injective. Then, there exist two distinct configurations $\boldsymbol{c}, \boldsymbol{c}' \in G^{\mathbb{Z}}$ with $F(\boldsymbol{c}) = F(\boldsymbol{c}')$. We get $L(\Psi(\boldsymbol{c})) = \Psi(F(\boldsymbol{c})) = \Psi(F(\boldsymbol{c}')) = L(\Psi(\boldsymbol{c}'))$ and, since Ψ is injective, it follows that L is not injective.
$\Longrightarrow$: Assume that L is not injective. Then, there exists a configuration $\boldsymbol{b} \in \hat{G}^{\mathbb{Z}}$ such that $\boldsymbol{b} \neq 0$ and $L(\boldsymbol{b}) = 0$. By Lemma 2.22, there exists a finite configuration $\boldsymbol{b}' \in \Psi(G^{\mathbb{Z}})$ such that $\boldsymbol{b}' \neq 0$ and $L(\boldsymbol{b}') = 0$. Let $\boldsymbol{c} \in G^{\mathbb{Z}}$ be the finite configuration such that $\Psi(\boldsymbol{c}) = \boldsymbol{b}'$. Clearly, it holds that $\boldsymbol{c} \neq 0$. We get $\Psi(F(\boldsymbol{c})) = L(\Psi(\boldsymbol{c})) = 0$. Since Ψ is injective, it follows that $F(\boldsymbol{c}) = 0$. Since $F(0) = 0$, we conclude that F is not injective. □

The decidability of injectivity and surjectivity for LCA combined with the previous two theorems lead to the corresponding decidability result for additive CA.

Corollary 2.26. Surjectivity and injectivity are decidable for additive CA over a finite abelian group.

Proof. It immediately follows from Theorem 2.8, 2.25, and 2.25. □

2.4.3. Topological transitivity and ergodicity

We start by proving that the embedding Ψ also preserves topological transitivity between an additive CA over G and the associated LCA.

Theorem 2.27. Let $(G^{\mathbb{Z}}, F)$ be any additive CA over G and let $(\hat{G}^{\mathbb{Z}}, L)$ be the LCA associated with it via the embedding Ψ. Then, the CA $(G^{\mathbb{Z}}, F)$ is topologically transitive if and only if $(\hat{G}^{\mathbb{Z}}, L)$ is.

Proof. Since $\Psi \circ F = L \circ \Psi$, for every $k \in \mathbb{N}$ it holds that $\Psi \circ (F^k - I) = \Psi \circ F^k - \Psi = L^k \circ \Psi - \Psi = (L^k - I) \circ \Psi$. By Lemma 2.23 , $F^k - I$ is surjective iff $L^k - I$ is. Theorem 2.24 and 2.9 conclude the proof. □

As a final result, we get the decidability of many mixing and ergodic properties for additive CA over any finite abelian group, including topological transitivity and ergodicity.

Corollary 2.28. All the following properties are decidable for additive CA over any finite abelian group:

1. topological transitivity;
2. ergodicity;
3. topological mixing;

4. weak topological transitivity;
5. total transitivity;
6. weak ergodic mixing;
7. ergodic mixing.

Proof. It is an immediate consequence of Theorem 2.9 and 2.27. □

References

[AllKle14] Allen Altman, Steven Kleiman, *A Term of Commutative Algebra*, 1 September 2013.
https://web.mit.edu/18.705/www/13Ed.pdf

[Aluffi09] Paolo Aluffi, *Algebra: Chapter 0*, Graduate Studies in Mathematics #104, AMS 2009.
https://bookstore.ams.org/gsm-104 and https://www.math.fsu.edu/~aluffi/algebraerrata.2009/Errata.html .

[Almkvi73] Gert Almkvist, *Endomorphisms of finetely [sic] generated projective modules over a commutative ring*, Ark. Mat. **11**, Number 1-2 (1973), pp. 263–301.
https://doi.org/10.1007/BF02388522

[Bourba72] Nicolas Bourbaki, *Elements of Mathematics: Commutative Algebra*, Hermann 1972.

[Bourba74] Nicolas Bourbaki, *Algebra I, Chapters 1-3*, Hermann 1974.
https://archive.org/details/ElementsOfMathematics-AlgebraPart1/page/n0

[Cattan09] Gianpiero Cattaneo, Alberto Dennunzio, Enrico Formenti, and Julien Provillard, *Non-uniform cellular automata*, *Linear cellular automata with multiple state variables*, In LATA 2009, volume 5457 of Lecture Notes in Computer Science, 2009, pp 302–313.

[ChaLoi14] Antoine Chambert-Loir, *(Mostly) Commutative Algebra*, October 24, 2014.
https://webusers.imj-prg.fr/~antoine.chambert-loir/enseignement/2014-15/ga/commalg.pdf

[Conrad13a] Keith Conrad, *Exterior powers*, 31 August 2019.
https://kconrad.math.uconn.edu/blurbs/linmultialg/extmod.pdf

[Conrad13b] Keith Conrad, *Base extension and exterior powers*, 2 March 2013.
https://kconrad.math.uconn.edu/blurbs/linmultialg/extmodbaseextn.pdf

[Dennun19] Alberto Dennunzio, Enrico Formenti, Luca Manzoni, Luciano Margara and Antonio E. Porreca, *On the dynamical behaviour of linear higher-order cellular automata and its decidability*, Information Sciences 486, 2019, pp. 73–87.
https://doi.org/10.1016/j.ins.2019.02.023

[Dennun19b] Alberto Dennunzio, Enrico Formenti, Darij Grinberg, and Luciano Margara, *Additive cellular automata over a finite abelian group: dynamical properties, decidability, and applications*, Preprint, 2019.

[Dennun19c] Alberto Dennunzio, Enrico Formenti, Darij Grinberg, and Luciano Margara, *Additive Cellular Automata Over Finite Abelian Groups: Topological and Measure Theoretic Properties*, In 44th International Symposium on Mathematical Foundations of Computer Science, MFCS 2019, volume 138 of LIPIcs, 2019, pp 68:1–68:15.
https://doi.org/10.4230/LIPIcs.MFCS.2019.68

[Dennun12] Alberto Dennunzio, Enrico Formenti, and Julien Provillard, *Non-uniform cellular automata: Classes, dynamics, and decidability*, Information and Computation 215, 2012, pp. 32–46.
https://doi.org/10.1016/j.ic.2012.02.008

[Dennun13] Alberto Dennunzio, Enrico Formenti, and Julien Provillard, *Local rule distributions, language complexity and non-uniform cellular automata*, Theoretical Computer Science 504, 2013, pp. 38–51.
https://doi.org/10.1016/j.tcs.2012.05.013

[Grinbe15] Darij Grinberg, *Notes on the combinatorial fundamentals of algebra*, 10 January 2019.
https://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/2019-01-10 .

[Grinbe19] Darij Grinberg, *Integrality over ideal semifiltrations*, 13 July 2019, arXiv:1907.06125v1.
Also available at
http://www.cip.ifi.lmu.de/~grinberg/algebra/integrality-merged.pdf

[ItoOsa83] Masanobu Ito, Nobuyasu Osato, and Masakazu Nasu, *Linear cellular automata over $\mathbb{Z}_m$*, Journal of Computer and System Sciences 27, 1983,

pp. 125–140.
https://doi.org/10.1016/0022-0000(83)90033-8

[Kari00] Jarkko Kari, *Linear cellular automata with multiple state variables*, In STACS 2000, volume 1770 of Lecture Notes in Computer Science, 2000, pp 110–121.

[KatHas95] Anatole Katok and Boris Hasselblatt, *Introduction to the Modern Theory of Dynamical Systems*, Cambridge University Press 1995.
https://doi.org/10.1017/CBO9780511809187

[LebMar95] Lieven Le Bruyn and Michel Van den Bergh, *Algebraic properties of linear cellular automata*, Linear algebra and its applications 157, 1991, pp 217–234.
https://doi.org/10.1016/0024-3795(91)90116-E

[LinMar95] Douglas Lind and Brian Marcus, *Symbolic Dynamics and Coding*, Cambridge University Press 1995.
https://doi.org/10.1017/CBO9780511626302

[ManMan99] Giovanni Manzini and Luciano Margara, *A Complete and Efficiently Computable Topological Classification of D-dimensional Linear Cellular Automata over $\mathbb{Z}_m$*, Theoretical Computer Science 221, 1999, pp. 157–177.
https://doi.org/10.1016/S0304-3975(99)00031-6

[Pin19] Jean-Éric Pin, *Mathematical Foundations of Automata Theory*, 13 March 2019.
https://www.irif.fr/~jep/PDF/MPRI/MPRI.pdf

[Steinbe16] Benjamin Steinberg, *Representation Theory of Finite Monoids*, Springer 2016.
https://dx.doi.org/10.1007/978-3-319-43932-7

[SwaHun06] Irena Swanson and Craig Huneke, *Integral Closure of Ideals, Rings, and Modules*, London Mathematical Society Lecture Note Series #336, Cambridge University Press, Cambridge, 2006.
https://people.reed.edu/~iswanson/book/SwansonHuneke.pdf

[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.
https://sites.math.rutgers.edu/~zeilberg/mamarim/mamarimhtml/gert.html