Iterated polynomials are dense

Pascal Autissier Jean-Philippe Furter Egor Yasinsky

Abstract

For any infinite field $\mathbf{k}$ and any positive integer r, we show constructively that the map sending each polynomial $P \in \mathbf{k}[x]$ to its r-th iterate $P^{\circ r}$ is dominant in various inductive limit topologies on the space of all polynomials.

Contents

1 Equations in groups

1.1 The word map

In what follows, by a *word* we mean an element $w = w(x_1, \dots, x_N)$ of the free group $\mathcal{F}_N$ on N generators $x_1, \dots, x_N$, for some $N \geqslant 1$. Let G be a group. The *word map* on G defined by w is the map

$$\mathbf{w}\colon G^N \to G, \quad (g_1, \dots, g_N) \mapsto w(g_1, \dots, g_N).$$

Assuming that $\mathbf{w} \neq \mathrm{id}$, what can be said about the image of $\mathbf{w}$? Or, in other words, what can be said about the solutions $(g_1, \ldots, g_N) \in G^N$ of the equation

$$w(g_1, \ldots, g_N) = g \tag{$*$}$$

when one varies $g \in G$? This problem has a long and remarkable history, and we refer to [Sha13, BGK14, GKP18] for some excellent surveys. Clearly, one cannot expect the map $\mathbf{w}$ to be surjective already for power words x^r, $r \geqslant 2$. For example, taking $G = \mathrm{SL}_2(\mathbf{k})$, where $\mathbf{k} = \mathbb{R}$ or $\mathbf{k} = \mathbb{C}$, we easily find matrices which do not admit square roots in G, and thus $w = x^2$ induces a non-surjective word map.

1.2 Borel's dominance theorem and weakly exponential Lie groups

Still, one can ask: how "large" is the image $\mathbf{w}(G^N)$ of a word map? For instance, if the group G is endowed with a reasonable topology, is the image of $\mathbf{w}$ dense in G? Here are two prototypical results in this direction which inspire our main theorem, presented in Section 1.4 below.

A morphism $f\colon X \to Y$ of topological spaces is called *dominant* if its image $f(X)$ is dense in Y. When X and Y are algebraic varieties, equipped with the Zariski topology, Chevalley's theorem implies that a morphism $f\colon X \to Y$ is dominant if and only if $f(X)$ contains a non-empty Zariski open subset. The following remarkable theorem is due to A. Borel.

Theorem 1.1 ([Bor83])**.** *If $\mathbf{k}$ is a field, G is a connected semisimple linear algebraic $\mathbf{k}$-group, and $w \neq \mathrm{id}$, then the corresponding word map $\mathbf{w}\colon G^N \to G$ is dominant.*

The second result concerns complex Lie groups and power maps $\mathbf{w}\colon g \mapsto g^r$ defined on them. Standard tools from Lie theory, such as the exponential map $\exp\colon \mathfrak{g} \to G$, provide an easy way to extracting r-th roots. Specifically, given an element $g \in G$, assume there exists $v \in \mathfrak{g}$ such that $g = \exp(v)$. Then $\exp(v/r)$ is an r-th root of g. Unfortunately, the exponential map of a Lie group is not necessarily surjective — this issue has been the subject of extensive study (see the references in the surveys mentioned above). However, for complex Lie groups the following can be said[1].

Theorem 1.2 ([HM78, Theorem 2.11])**.** *Every complex connected Lie group G is weakly exponential, i.e. the image of the exponential map $\exp\colon \mathfrak{g} \to G$ is dense.*

Hence, for $w = x^r$ and "typical" g, the equation ($*$) admits a solution.

[1] In the semisimple case, this theorem is also due to A. Borel.

1.3 Endomorphisms of the affine space and some ind-topologies

This note was motivated by a question posed in the surveys [GKP18, BGK14], which deal with word equations in simple matrix groups. Let $\mathbf{k}$ be any field. Noting J.-P. Serre's observation [Ser10] that the group $\mathrm{Bir}(\mathbb{P}^2_{\mathbf{k}})$ of birational transformations of the projective plane resembles simple linear algebraic groups, the authors of these surveys asked whether word maps are dominant for the Cremona group $G = \mathrm{Bir}(\mathbb{P}^2_{\mathbf{k}})$ and any known "interesting" topology on it, see [BGK14, Problem 7.11] and [GKP18, Question 3.11].

We consider the closely related automorphism group of the affine space $\mathrm{Aut}(\mathbb{A}^n_{\mathbf{k}})$ and, more generally, the monoid $\mathrm{End}(\mathbb{A}^n_{\mathbf{k}})$ of algebraic endomorphisms. Recall that an endomorphism of $\mathbb{A}^n_{\mathbf{k}}$ is given by

$$f\colon x = (x_1, \ldots, x_n) \mapsto (f_1(x), \ldots, f_n(x)),$$

where $f_1, \ldots, f_n \in \mathbf{k}[x_1, \ldots, x_n]$ are polynomials. To simplify the notation, we often write $(f_1, \ldots, f_n)$ in what follows. The *degree* $\deg f$ of an endomorphism $f = (f_1, \ldots, f_n)$ is defined as $\deg(f) = \max\{\deg(f_1), \ldots, \deg(f_n)\}$. The subset of invertible elements of the monoid $\mathcal{E} = \mathrm{End}(\mathbb{A}^n_{\mathbf{k}})$ is the group $\mathrm{Aut}(\mathbb{A}^n_{\mathbf{k}})$ of automorphisms of $\mathbb{A}^n_{\mathbf{k}}$.

For each $d \geqslant 0$, we identify the set

$$\mathcal{E}_{\leqslant d} = \left\{ f \in \mathrm{End}(\mathbb{A}^n_{\mathbf{k}})\colon \deg f \leqslant d \right\}$$

with the vector space $\mathbf{k}^N$, where $N = \binom{d+n}{n} n$, in an obvious way. Following I. R. Šafarevich [Sha66], we view $\mathrm{End}(\mathbb{A}^n_{\mathbf{k}})$ as an *ind-monoid*.

Definition 1.3. Consider the filtration

$$\mathcal{E}_{\leqslant 1} \subseteq \mathcal{E}_{\leqslant 2} \subseteq \ldots \subseteq \mathcal{E}_{\leqslant d} \subseteq \mathcal{E}_{\leqslant d+1} \subseteq \ldots, \quad \mathcal{E} = \bigcup_{d=1}^{\infty} \mathcal{E}_{\leqslant d}. \tag{1}$$

1. A set $S \subseteq \mathcal{E}$ is called *Zariski closed* if $S \cap \mathcal{E}_{\leqslant d}$ is Zariski closed in $\mathcal{E}_{\leqslant d}$ for all $d \geqslant 1$. The corresponding topology on $\mathrm{End}(\mathbb{A}^n_{\mathbf{k}})$ is called the *Zariski ind-topology*.

2. Let $(\mathbf{k}, |\cdot|)$ be a valued field with a non-trivial absolute value $|\cdot|$. The map $(x, y) \mapsto |x - y|$ is a metric on $\mathbf{k}$ which yields a topology on $\mathbf{k}$ in the usual way. The sets $\mathcal{E}_{\leqslant d}$, being identified with a $\mathbf{k}$-vector space $\mathbf{k}^N$, where $N = \binom{d+n}{n} n$, can be endowed with a uniform norm $\|x_1 e_1 + \ldots x_N e_N\|_\infty = \max_{1 \leqslant i \leqslant N} |x_i|$ for a given choice of basis $\langle e_1, \ldots, e_N \rangle = \mathcal{E}_{\leqslant d}$. Furthermore, any two such uniform norms with respect to two different bases are equivalent, and hence they define the same topology on $\mathcal{E}_{\leqslant d}$. If $(\mathbf{k}, |\cdot|)$ is complete, then all norms on $\mathcal{E}_{\leqslant d}$ are equivalent to the uniform norm and

therefore define the same topology. In any case, by an abuse of terminology, we call this topology *Euclidean*. A set $S \subseteq \mathcal{E}$ is called *Euclidean closed* if $S \cap \mathcal{E}_{\leqslant d}$ is closed in this topology on $\mathcal{E}_{\leqslant d}$ for all $d \geqslant 1$. The corresponding topology on $\mathrm{End}(\mathbb{A}^n_{\mathbf{k}})$ is called the *Euclidean ind-topology*. Note that this topology is still Hausdorff.

1.4 Main result

The main result of this note is the following.

Theorem. *Let* $\mathbf{k}$ *be a field,* $r \geqslant 1$ *be an integer and* $w = x^r$ *be a power word. Consider the corresponding power map on* $\mathrm{End}(\mathbb{A}^1_{\mathbf{k}}) \simeq \mathbf{k}[x]$*, which sends every polynomial to its* r*-th iterate:*

$$\mathbf{w}\colon \mathbf{k}[x] \to \mathbf{k}[x], \quad P \mapsto P^{\circ r}.$$

Then the following holds.

1. *If* $(\mathbf{k}, |\cdot|)$ *is a valued field with a non-trivial absolute value, then* $\mathbf{w}$ *is dominant in the Euclidean ind-topology.*
2. *The map* $\mathbf{w}$ *is dominant in the Zariski ind-topology for any infinite field* $\mathbf{k}$*.*

Furthermore, the statement stays true when w *is replaced with any non-trivial word* $w \in \mathcal{M}_N$ *in the free monoid* $\mathcal{M}_N$ *on* N *generators.*

Note that, unlike the case of dominant morphisms between algebraic varieties, the image of a dominant map in the ind-topology *does not*, in general, contain a Zariski-open subset.

Remark 1.4.

- Whenever it is defined, the Euclidean ind-topology on $\mathrm{End}(\mathbb{A}^n_{\mathbf{k}})$ is finer than the Zariski ind-topology and therefore, for $\mathbf{k}$ valued, statement (1) is stronger than statement (2).
- The "Furthermore" part of the theorem follows from its main part. Indeed, if $w = x_{i_1}^{m_1} x_{i_2}^{m_2} \dots x_{i_s}^{m_s} \in \mathcal{M}_N$ is any word (here $m_1, \dots, m_s > 0$, as we work in the monoid), then the image of $\mathbf{w}$ contains $\mathbf{w}(P, \dots, P) = P^{\circ(m_1 + \dots + m_s)}$ for all $P \in \mathbf{k}[x]$, hence it is dense.
- Let $|\cdot|_1, \dots, |\cdot|_n$ be pairwise non-equivalent non-trivial absolute values on $\mathbf{k}$ (so in particular they induce different topologies on $\mathbf{k}$). The proof of our main result and

the Artin-Whaples approximation theorem [AW45, Theorem 1] imply that, given any polynomials $Q_1, \dots, Q_n \in \mathbf{k}[x]$, $r \geqslant 1$ and $\eta > 0$, there exists $P \in \mathbf{k}[x]$ such that

$$\|P^{\circ r} - Q_i\|_i < \eta,$$

for all $i \in \{1, \dots, n\}$. Here, P is a polynomial of degree at most

$$d = \max\{1, \deg Q_1, \dots, \deg Q_n\} + r$$

(see the key proposition in Section 2.2) and $\|\cdot\|_i$ denotes a norm on $\mathcal{E}_{\leqslant d^r}$ induced by the absolute value $|\cdot|_i$, as in Definition 1.3.

1.5 Finitary case

Let $\mathbf{k} = \mathbf{F}_q$ be a finite field. Then, the filtration (1) in Definition 1.3 is a filtration by finite sets. Fix $r \geqslant 2$. For each $d \geqslant 1$ let

$$\mathcal{I}(d, r) = \big\{P \in \mathcal{E}_{\leqslant d} \colon P = Q^{\circ r} \text{ for some } Q \in \mathbf{k}[x]\big\}.$$

Notice that the "asymptotic density" of r-th iterates is zero:

$$\lim_{d \to +\infty} \frac{|\mathcal{I}(d, r)|}{|\mathcal{E}_{\leqslant d}|} = 0.$$

Indeed, suppose that $P = Q^{\circ r}$. Then $\deg P \leqslant d$ if and only if $\deg Q \leqslant d^{1/r}$. Hence there is a surjection from $\mathcal{E}_{\leqslant d^{1/r}}$ to $\mathcal{I}(d, r)$, and we deduce $|\mathcal{I}(d, r)| \leqslant q^{d^{1/r}+1}$. Therefore,

$$\frac{|\mathcal{I}(d, r)|}{|\mathcal{E}_{\leqslant d}|} = \frac{|\mathcal{I}(d, r)|}{q^{d+1}} \leqslant q^{d^{1/r}-d} \to 0 \quad \text{when} \quad d \to +\infty.$$

Nevertheless, the following questions seem interesting to us.

Question 1.5. Fix a finite base field $\mathbf{k} = \mathbf{F}_q$ and an integer $r \geqslant 2$. What can be said about the asymptotic of the integer sequence $|\mathcal{I}(d, r)|$ as d grows? Does there exist a limit of the rational sequence $q^{-d-1}|\mathcal{I}(d^r, r)|$ as $d \to +\infty$?

Acknowledgements

We thank Serge Cantat for his helpful comments and suggestions on the draft of this paper.

2 Power maps on $\mathrm{End}(\mathbb{A}^1_{\mathbf{k}})$

2.1 Hasse derivative

We start with recalling some definitions.

Definition 2.1 (*Hasse derivative*, see e.g. [Gol03, §1.3])**.** Let A be a commutative ring. For each polynomial $P(x) = p_n x^n + p_{n-1}x^{n-1} + \cdots + p_1 x + p_0 \in A[x]$ and non-negative integer $j \leqslant n$, define the j-*th Hasse derivative* of P by

$$P^{[j]}(x) = \sum_{k=j}^{n} p_k \binom{k}{j} x^{k-j}.$$

Note that the j-th Hasse derivative of P satisfies $P^{(j)} = j!P^{[j]}$, where $P^{(j)}$ denotes the usual j-th derivative. In what follows, we will use the following two properties of Hasse derivatives.

Proposition 2.2. *Let A be a commutative ring, and $P, Q \in A[x]$. Then the following holds.*

1. **Taylor's formula:**
$$P(a+b) = \sum_{j \geqslant 0} P^{[j]}(a)\, b^j$$
for any $a, b \in A$.

2. **Leibniz rule:**
$$(PQ)^{[j]} = \sum_{\ell=0}^{j} P^{[\ell]}\, Q^{[j-\ell]}.$$

Remark 2.3. For any $a \in \mathbf{k}$ and $T(x) = \sum_{i=0}^{m} t_i x^i \in \mathbf{k}[x]$, we have

$$T(ax)^{[j]} = \sum_{i=j}^{m} t_i a^i \binom{i}{j} x^{i-j} = a^j T^{[j]}(ax).$$

In the sequel we will apply Definition 2.1 with the ring $A = K[\varepsilon, \varepsilon^{-1}]$ of Laurent polynomials in ε.

Definition 2.4. Let $A = K[\varepsilon, \varepsilon^{-1}]$ and let $P, Q \in A[x]$ be any two polynomials. Fix an integer $\ell \in \mathbb{Z}$. We say that P *and* Q *are equivalent modulo* ε^ℓ, and write $P \equiv Q \mod \varepsilon^\ell$, if

$$P - Q \in \varepsilon^\ell \mathbf{k}[\varepsilon][x].$$

Remark 2.5. This is indeed an equivalence relation on the ring $A[x]$. When $(\mathbf{k}, |\cdot|)$ is a valued field with a non-trivial absolute value, one has $P \equiv Q \mod \varepsilon^\ell$ if and only if $P - Q = O(\varepsilon^\ell)$ when $\varepsilon \to 0$. If no confusion arises, we will use both notations below.

2.2 The key proposition

The proof of our main result relies on the following.

Key proposition. *Let $r \geqslant 2$ be an integer and $\mathbf{k}$ be a field with at least r elements. Let $n \geqslant 2$ be an integer and $Q \in \mathbf{k}[x]$ be a polynomial of degree at most $n-1$. Then there exists a family $\varepsilon \mapsto P_\varepsilon$ of polynomials of degree at most $n+r-1$ parametrized by $\mathbb{A}^1_{\mathbf{k}} \setminus \{0\}$, such that the family $\varepsilon \mapsto (P_\varepsilon)^{\circ r}$ extends to a family parametrized by $\mathbb{A}^1_{\mathbf{k}}$ and whose value at $\varepsilon = 0$ is Q.*

The rest of this section is devoted to the proof of this statement. Let $a_1, \ldots, a_{r-1}$ be distinct elements of $\mathbf{k}^\times$. Put $a_r = 0$.

Lemma 2.6. *There exists a unique polynomial $L \in \mathbf{k}[x]$ of degree $\deg L \leqslant n+r-2$ such that*

$$L(0) = a_1, \quad L(a_k) = a_{k+1} \text{ for all } k \in \{1, \ldots, r-1\}, \quad \text{and } L^{[j]}(0) = 0 \text{ for all } j \in \{1, \ldots, n-1\}.$$

Proof. Indeed, let $L(x) = \ell_{n+r-2}x^{n+r-2} + \cdots + \ell_1 x + \ell_0 \in \mathbf{k}[x]$. The condition $L^{[j]}(0) = 0$ for all $1 \leqslant j \leqslant n-1$ implies that $\ell_1 = \ell_2 = \cdots = \ell_{n-1} = 0$, hence L is of the form

$$L(x) = \ell_{n+r-2}x^{n+r-2} + \cdots + \ell_n x^n + a_1.$$

The conditions $L(a_k) = a_{k+1}$ for all $1 \leqslant k \leqslant r-1$ then give a system of $r-1$ linear equations in $r-1$ variables $\ell_{n+r-2}, \ldots, \ell_n$. The matrix of this system is the Vandermonde matrix $V(a_1, \ldots, a_{r-1})$. Since a_i's are assumed pairwise distinct, the Vandermonde determinant is not zero; this achieves the proof. □

Remark 2.7. Another way to prove Lemma 2.6 is to use the Lagrange interpolation formula:

$$L(x) = a_1 + \sum_{k=1}^{r-1} (a_{k+1} - a_1) \frac{x^n}{a_k^n} \prod_{l \neq k} \frac{x - a_l}{a_k - a_l}.$$

This polynomial L coincides with the one obtained in Lemma 2.6.

Now, define the polynomial

$$R(x) = \prod_{k=1}^{r-1} (a_k - x) \in \mathbf{k}[x] \tag{2}$$

and set

$$c = R(0)^{n+1} \prod_{l=1}^{r-1} R'(a_l). \tag{3}$$

Define the polynomial

$$P(x) = \varepsilon^{(r-1)(2n-3)} R(\varepsilon^{2r}x)(\varepsilon^r x^n + c^{-1}Q) + \varepsilon^{-2r} L(\varepsilon^{2r}x) \in \mathbf{k}[\varepsilon, \varepsilon^{-1}][x]. \tag{4}$$

The key proposition is a direct consequence of the following statement.

Proposition 2.8. *One has* $P^{\circ r} \equiv Q \mod \varepsilon$.

To prove it, we will need several lemmas.

Lemma 2.9. *For every* $k \in \{1, \dots, r-1\}$, *one has*

$$P'(\varepsilon^{-2r}a_k) \equiv \varepsilon^{3-2n} R'(a_k) a_k^n \mod \varepsilon^{4-2n}.$$

Proof. First, we differentiate P and obtain

$$P'(x) = \varepsilon^{(r-1)(2n-3)} \left[R'(\varepsilon^{2r}x)\varepsilon^{2r}(\varepsilon^r x^n + c^{-1}Q) + R(\varepsilon^{2r}x)(n\varepsilon^r x^{n-1} + c^{-1}Q') \right] + L'(\varepsilon^{2r}x).$$

Notice that $R(a_k) = 0$ and $L'(a_k) \equiv 0 \mod \varepsilon^{4-2n}$ since $4 - 2n \leqslant 0$ and $L'(a_k) \in \mathbf{k}$. Hence

$$P'(\varepsilon^{-2r}a_k) \equiv \varepsilon^{3-2n} R'(a_k) a_k^n + \varepsilon^{(r-1)(2n-3)+2r} R'(a_k) c^{-1} Q(\varepsilon^{-2r}a_k) \mod \varepsilon^{4-2n},$$

and it suffices to show that

$$\varepsilon^{(r-1)(2n-3)+2r} Q(\varepsilon^{-2r}a_k) \equiv 0 \mod \varepsilon^{4-2n}.$$

This is equivalent to saying that $\varepsilon^{(r-1)(2n-3)+2r-4+2n} Q(\varepsilon^{-2r}a_k) \in \mathbf{k}[\varepsilon]$. But, writing $Q = q_{n-1}x^{n-1} + \cdots + q_0$, this Laurent polynomial equals

$$\varepsilon^{2nr-r-1} \sum_{i=1}^{n} q_{n-i} (\varepsilon^{-2r}a_k)^{n-i} = \sum_{i=1}^{n} q_{n-i} a_k^{n-i} \varepsilon^{2ri-r-1}$$

and we observe that $2ri - r - 1 \geqslant 0$ for all $i \geqslant 1$, which finishes the proof. □

Lemma 2.10. *For all* $j \geqslant 1$, *one has*

$$P^{[j]}(\varepsilon^{-2r}a_k) = O\left(\varepsilon^{2r(j-1)-2n+3}\right).$$

Proof. By Remark 2.3, we have

$$\left(\varepsilon^{-2r}L(\varepsilon^{2r}x)\right)^{[j]} = \varepsilon^{2r(j-1)}L^{[j]}(\varepsilon^{2r}x), \text{ hence } \left(\varepsilon^{-2r}L(\varepsilon^{2r}x)\right)^{[j]}(\varepsilon^{-2r}a_k) = O\left(\varepsilon^{2r(j-1)}\right).$$

Next, we will show that if $u \geqslant 0$ and $\deg S \leqslant \nu$, then for any $a \in \mathbf{k}$ one has

$$(T(\varepsilon^u x)S)^{[j]}\left(\varepsilon^{-u}a\right) = O\left(\varepsilon^{(j-\nu)u}\right). \tag{5}$$

By Leibniz rule, we get

$$(T(\varepsilon^u x)S)^{[j]} = \sum_{\ell=0}^{j}\left(T(\varepsilon^u x)\right)^{[\ell]}S^{[j-\ell]} = \sum_{\ell=0}^{j}\varepsilon^{u\ell}T^{[\ell]}(\varepsilon^u x)S^{[j-\ell]}.$$

Now, as $\deg S^{[j-\ell]} \leqslant \deg S - (j-\ell) \leqslant \nu - j + \ell$ and $u \geqslant 0$, we have

$$\varepsilon^{u\ell}T^{[\ell]}(\varepsilon^u x)S^{[j-\ell]}\Big|_{x=\varepsilon^{-u}a} = \varepsilon^{u\ell}O\left(\varepsilon^{-u(\nu-j+\ell)}\right) = O\left(\varepsilon^{(j-\nu)u}\right),$$

as claimed.

We now apply (5) to the first summand of (4). Namely, taking $u = 2r$, $T = R$ and $S = x^n$ we get

$$\left(R(\varepsilon^{2r}x)\varepsilon^r x^n\right)^{[j]}\Big|_{x=\varepsilon^{-2r}a_k} = O\left(\varepsilon^{2r(j-n)+r}\right),$$

while taking $u = 2r$, $T = R$ and $S = c^{-1}Q$, a polynomial of degree $\leqslant n-1$, we get

$$\left(R(\varepsilon^{2r}x)c^{-1}Q\right)^{[j]}\Big|_{x=\varepsilon^{-2r}a_k} = O\left(\varepsilon^{2r(j-n)+2r}\right).$$

Further multiplication by $\varepsilon^{(r-1)(2n-3)}$ gives $O\left(\varepsilon^{2r(j-1)-2n+3}\right)$ and $O\left(\varepsilon^{2r(j-1)+r-2n+3}\right)$, respectively. Since $2r(j-1)-2n+3 < \min\{2r(j-1)+r-2n+3, 2r(j-1)\}$, we are done. □

Lemma 2.11. *One has*

$$P \equiv \varepsilon^{-2r}a_1 + \varepsilon^{(r-1)(2n-3)}Qc^{-1}R(0) \mod \varepsilon^{(r-1)(2n-3)+1}.$$

Proof. We look at each of the two summands in (4). First, we notice that

$$\varepsilon^{-2r}L(\varepsilon^{2r}x) \equiv \varepsilon^{-2r}a_1 \mod \varepsilon^{(r-1)(2n-3)+1}.$$

Indeed, we can write $L(x) = \sum_{i=0}^{r-2}\ell_{n+i}x^{n+i} + a_1$, as in the proof of Lemma 2.6. Therefore,

we just need to verify that

$$\sum_{i=0}^{r-2} \ell_{n+i} x^{n+i} \varepsilon^{2r(n+i)-2r-(r-1)(2n-3)-1}$$

is a polynomial in $\mathbf{k}[x, \varepsilon]$, i.e. has non-negative powers in ε. But for all $i \in \{0, \dots, r-2\}$ we have

$$2r(n+i) - 2r - (r-1)(2n-3) - 1 = 2ri + r + (2n-4) \geqslant r \geqslant 0,$$

since $2n - 4 \geqslant 0$ and $i \geqslant 0$.

Second, we claim that

$$\varepsilon^{(r-1)(2n-3)} Q c^{-1} R(0) \equiv \varepsilon^{(r-1)(2n-3)} R(\varepsilon^{2r} x)(\varepsilon^r x^n + Q c^{-1}) \mod \varepsilon^{(r-1)(2n-3)+1}.$$

This is equivalent to showing

$$Q c^{-1} R(0) \equiv R(\varepsilon^{2r} x)(\varepsilon^r x^n + Q c^{-1}) \mod \varepsilon,$$

which is obvious, as the constant term of $Q c^{-1} R(0) - R(\varepsilon^{2r} x) Q c^{-1}$, viewed as polynomial in ε, equals zero. □

Lemma 2.12. *For all* $k \in \{1, \dots, r\}$, *one has*

$$P^{\circ k} \equiv \varepsilon^{-2r} a_k + \varepsilon^{(r-k)(2n-3)} Q c^{-1} R(0) \prod_{l=1}^{k-1} R'(a_l) a_l^n \mod \varepsilon^{(r-k)(2n-3)+1}. \tag{6}$$

Proof. We proceed by induction on k. The base $k = 1$ is ensured by Lemma 2.11. So, we assume that the statement holds for k and will show it for $k+1$. By the induction hypothesis, we can write

$$P^{\circ k}(x) = \varepsilon^{-2r} a_k + \varepsilon^{(r-k)(2n-3)} \left[Q(x) c^{-1} R(0) \prod_{l=1}^{k-1} R'(a_l) a_l^n + \varepsilon T(\varepsilon, x) \right]$$

with $T(\varepsilon, x) \in \mathbf{k}[\varepsilon, x]$. Let $S(\varepsilon, x) = Q(x) c^{-1} R(0) \prod_{l=1}^{k-1} R'(a_l) a_l^n + \varepsilon T(\varepsilon, x)$, then

$$P^{\circ k}(x) = \varepsilon^{-2r} a_k + \varepsilon^{(r-k)(2n-3)} S(\varepsilon, x), \tag{7}$$

where

$$S \equiv Q(x) c^{-1} R(0) \prod_{l=1}^{k-1} R'(a_l) a_l^n \mod \varepsilon.$$

By Taylor's formula from Proposition 2.2 applied to (7), we get

$$P^{\circ(k+1)}(x) = P(\varepsilon^{-2r}a_k + \varepsilon^{(r-k)(2n-3)}S) = \sum_{j\geqslant 0} P^{[j]}(\varepsilon^{-2r}a_k)\varepsilon^{j(r-k)(2n-3)}S^j. \tag{8}$$

By Lemma 2.10, we have $P^{[j]}(\varepsilon^{-2r}a_k) = O(\varepsilon^{2r(j-1)-2n+3})$, while $S^j \equiv U(x) \mod \varepsilon$ for some $U(x) \in \mathbf{k}[x]$. Therefore,

$$P^{[j]}(\varepsilon^{-2r}a_k)\varepsilon^{j(r-k)(2n-3)}S^j = O\left(\varepsilon^{2r(j-1)-2n+3+j(r-k)(2n-3)}\right).$$

Consider the function

$$h\colon \mathbb{N} \to \mathbb{Z},\ \ j \mapsto 2r(j-1) - 2n + 3 + j(r-k)(2n-3).$$

Note that h is a strictly increasing affine function, since $2r + (r-k)(2n-3) > 0$. Thus $h(1) + 1 \leqslant h(j)$ for all $j \geqslant 2$, which gives

$$(r-k)(2n-3) - 2n + 4 = (r-(k+1))(2n-3) + 1 \leqslant h(j).$$

We conclude that

$$\sum_{j\geqslant 2} P^{[j]}(\varepsilon^{-2r}a_k)\varepsilon^{j(r-k)(2n-3)}S^j = O\left(\varepsilon^{(r-(k+1))(2n-3)+1}\right),$$

and thus formula (8) implies

$$P^{\circ(k+1)}(x) \equiv P(\varepsilon^{-2r}a_k) + P^{[1]}(\varepsilon^{-2r}a_k)\varepsilon^{(r-k)(2n-3)}S \mod \varepsilon^{(r-(k+1))(2n-3)+1}. \tag{9}$$

To achieve the proof, it remains to show that the right hand side of (9) is equivalent, modulo $\varepsilon^{(r-(k+1))(2n-3)+1}$, to

$$\varepsilon^{-2r}a_{k+1} + \varepsilon^{(r-(k+1))(2n-3)}Qc^{-1}R(0)\prod_{l=1}^{k} R'(a_l)a_l^n.$$

But $P(\varepsilon^{-2r}a_k) = \varepsilon^{-2r}a_{k+1}$ by Lemma 2.6. On the other hand,

$$P^{[1]}(\varepsilon^{-2r}a_k)\varepsilon^{(r-k)(2n-3)}S \equiv \varepsilon^{(r-(k+1))(2n-3)}Qc^{-1}R(0)\prod_{l=1}^{k} R'(a_l)a_l^n \mod \varepsilon^{(r-(k+1))(2n-3)+1}.$$

Indeed, this is equivalent to

$$P^{[1]}(\varepsilon^{-2r}a_k)S \equiv \varepsilon^{3-2n}Qc^{-1}R(0)\prod_{l=1}^{k}R'(a_l)a_l^n \mod \varepsilon^{4-2n},$$

which immediately follows from Lemma 2.9 and the definition of S. □

We are now in position to prove Proposition 2.8.

Proof of Proposition 2.8. Put $k = r$ in the formula (6). Recall that $a_r = 0$, hence

$$P^{\circ r} \equiv Qc^{-1}R(0)\prod_{l=1}^{r-1}R'(a_l)a_l^n \mod \varepsilon.$$

It remains to notice that, by the definitions (2) and (3) of R and c, one has

$$c^{-1}R(0)\prod_{l=1}^{r-1}R'(a_l)a_l^n = \frac{1}{R(0)^n}\prod_{l=1}^{r-1}a_l^n = 1.$$ □

2.3 Proof of the main result

Let $\mathbf{w}\colon \mathrm{End}(\mathbb{A}^1_{\mathbf{k}}) \to \mathrm{End}(\mathbb{A}^1_{\mathbf{k}})$ be the r-th iterate map defined by the word $w = x^r$, and let $\mathcal{W} = \mathbf{w}(\mathrm{End}(\mathbb{A}^1_{\mathbf{k}}))$ be its image. We will show that $\mathcal{W}$ is dense in $\mathrm{End}(\mathbb{A}^1_{\mathbf{k}})$ in the Zariski and Euclidean ind-topologies. So, let $Q \in \mathrm{End}(\mathbb{A}^1_{\mathbf{k}})$ be any polynomial and $\mathcal{U} \ni Q$ be any of its open neighbourhoods. We then need to show that $\mathcal{U} \cap \mathcal{W} \neq \varnothing$. Put $d = \max\{1, \deg Q\} + r$. Then it is sufficient to show that $\mathcal{U} \cap \mathcal{E}_{\leqslant d^r}$ contains an element of $\mathcal{W}$. In what follows, $\mathcal{E}_{\leqslant d^r}$ is identified with the vector space $\mathbf{k}^N$, where $N = d^r + 1$. By Proposition 2.8, there exists $P_\varepsilon \in \mathbf{k}[\varepsilon, \varepsilon^{-1}][x]$ such that $Q = P_\varepsilon^{\circ r} + \varepsilon T(\varepsilon, x)$, where $T(\varepsilon, x) \in \mathbf{k}[\varepsilon, x]$.

Suppose that $(\mathbf{k}, |\cdot|)$ is a valued field with a non-trivial absolute value. Let us first notice that $\mathbf{k}$ contains elements which are arbitrarily close to zero. More precisely, $|\mathbf{k}^\times| = \{|\varepsilon|\colon \varepsilon \in \mathbf{k}^\times\}$ is a subgroup of $\mathbb{R}_{>0}$, which is either dense (if 1 is its accumulation point), or is of the form $\{t^{\mathbb{Z}}\}$ for some $t \in (0,1)$. Now, the set $\mathcal{U} \cap \mathcal{E}_{\leqslant d^r}$ is open in $\mathcal{E}_{\leqslant d^r}$, hence there are $\delta \in \mathbb{R}_{>0}$ and an open δ-ball $B_\delta(Q) \subset \mathcal{U} \cap \mathcal{E}_{\leqslant d^r}$ centred at Q. Then, by choosing $\varepsilon \in \mathbf{k}^\times$ with $|\varepsilon| \cdot \|T\| < \delta$, we obtain $\|Q - P_\varepsilon^{\circ r}\| = |\varepsilon| \cdot \|T\| < \delta$, hence $P_\varepsilon^{\circ r} \in B_\delta(Q) \subset \mathcal{U} \cap \mathcal{E}_{\leqslant d^r} \cap \mathcal{W}$, as required.

In the case of an arbitrary infinite base field $\mathbf{k}$, we proceed as follows. Writing $\mathcal{E}_{\leqslant d^r} \setminus (\mathcal{U} \cap \mathcal{E}_{\leqslant d^r})$ as the zero set of the ideal $(F_1, \ldots, F_m)$, we notice that $F_i(Q) \neq 0$ for some $i \in \{1, \ldots, m\}$; here we identify Q with a point in $\mathbf{k}^N$. Consider $G(\varepsilon) = F_i(P_\varepsilon^{\circ r}) = F_i(Q - \varepsilon T)$ as a polynomial of ε. Note that G is not a zero polynomial, as $G(0) = F_i(Q) \neq 0$. Since

$\mathbf{k}$ is infinite, there exists $\varepsilon \in \mathbf{k}^\times$ such that $G(\varepsilon) = F_i(P_\varepsilon^{or}) \neq 0$, thus $P_\varepsilon^{or} \in \mathcal{U} \cap \mathcal{E}_{\leqslant d^r}$. This achieves the proof.

References

[AW45] Artin, E. ; Whaples, G.: Axiomatic characterization of fields by the product formula for valuations. In: *Bull. Am. Math. Soc.* 51 (1945), S. 469–492. http://dx.doi.org/10.1090/S0002-9904-1945-08383-9. – DOI 10.1090/S0002–9904–1945–08383–9. – ISSN 0002–9904 5

[BGK14] Bandman, Tatiana ; Garion, Shelly ; Kunyavskiĭ, Boris: Equations in simple matrix groups: algebra, geometry, arithmetic, dynamics. In: *Cent. Eur. J. Math.* 12 (2014), Nr. 2, S. 175–211. http://dx.doi.org/10.2478/s11533-013-0335-4. – DOI 10.2478/s11533–013–0335–4. – ISSN 1895–1074 2, 3

[Bor83] Borel, Armand: On free subgroups of semi-simple groups. In: *Enseign. Math. (2)* 29 (1983), S. 151–164. – ISSN 0013–8584 2

[GKP18] Gordeev, N. L. ; Kunyavskiĭ, B. ; Plotkin, E. B.: Geometry of word equations in simple algebraic groups over special fields. In: *Uspekhi Mat. Nauk* 73 (2018), Nr. 5(443), 3–52. http://dx.doi.org/10.4213/rm9838. – DOI 10.4213/rm9838. – ISSN 0042–1316,2305–2872 2, 3

[Gol03] Goldschmidt, David M.: *Grad. Texts Math.*. Bd. 215: *Algebraic functions and projective curves*. New York, NY: Springer, 2003. http://dx.doi.org/10.1007/b97844. http://dx.doi.org/10.1007/b97844. – ISBN 0–387–95432–5 6

[HM78] Hofmann, Karl H. ; Mukherjea, Arunava: On the density of the image of the exponential function. In: *Math. Ann.* 234 (1978), 263–273. http://dx.doi.org/10.1007/BF01420648. – DOI 10.1007/BF01420648. – ISSN 0025–5831 2

[Ser10] Serre, Jean-Pierre: The Cremona group and its finite subgroups. In: *Séminaire Bourbaki. Volume 2008/2009. Exposés 997–1011*. Paris: Société Mathématique de France (SMF), 2010. – ISBN 978–2–85629–291–4, S. 75–100, ex 3

[Sha66] Shafarevich, I. R.: *On some infinite dimensional groups*. Simpos. Int. Geom. Algebr., Roma 1965, 208-212 (1967); Rend. Mat. Appl., V. Ser. 25, 208-212 (1966)., 1966 3

[Sha13] Shalev, Aner: Some results and problems in the theory of word maps. Version: 2013. http://dx.doi.org/10.1007/978-3-642-39286-3_22. In: *Erdös centennial. On the occasion of Paul Erdös 100th anniversary of his birth*. Berlin: Springer; Budapest: János Bolyai Mathematical Society, 2013. – DOI 10.1007/978–3–642–39286–3_22. – ISBN 978–3–642–39285–6; 978–3–642–39286–3, S. 611–649 2

IMB, Université de Bordeaux, 351 Cours de la Libération, 33405 Talence Cedex, France

pascal.autissier@math.u-bordeaux.fr

jean-philippe.furter@math.u-bordeaux.fr

egor.yasinsky@u-bordeaux.fr