ABSTRACT. We give a brief survey of some problems related to the solvability of word equations in matrix groups and polynomial equations in matrix algebras, focusing on the cases where the matrix entries belong to a global field or its ring of integers.

What is needed is a corpus of explicit concrete cases and a middlebrow arithmetic theory which would provide both a practicable means to obtain them and a framework to understand any unexpected regularities.

J. W. S. Cassels, E. V. Flynn, Prolegomena to a Middlebrow Arithmetic of Curves of Genus 2, Cambridge, 1996

1. INTRODUCTION

In the present survey, by a noncommutative Diophantine problem we mean the question on the existence of a solution to an equation of the form

$$F(A_1,\ldots,A_m,X_1,\ldots,X_d) = 0,$$  (1.1)

where the $A_i$ are fixed elements of an associative or Lie algebra $\mathcal{A}$, or of a group $G$, $F$ is an associative or Lie polynomial $P$, or a group word $w$, and solutions are $d$-uples $(X_1,\ldots,X_d)$ sought in $\mathcal{A}^d$ or $G^d$, respectively (in the group case we change 0 to 1 in the right-hand side of (1.1)). We restrict our attention to a particular class of equations:

$$F(X_1,\ldots,X_d) = A,$$  (1.2)
where $A$ is fixed and only scalar constants are permitted in the left-hand side of (1.2).

If the algebra $A$ or the group $G$ admits a faithful matrix representation over a field $k$ or a ring $R$, then we can regard the matrix entries as variables and thus reduce our noncommutative Diophantine problem to a commutative one (in the group case, this reduction may require either cancelling denominators arising from negative powers of $X_i$'s in group words, or using representations in SL rather than in GL, which can easily be achieved by embedding $GL_n$ into $SL_{n+1}$).

This “abelianization” has several drawbacks. First, we replace a single (though noncommutative) polynomial equation with a system of polynomial equations which may be really huge (though commutative), require computer investigation, and go far beyond software limitations. Second, this abelianization process is somehow forgetful: essential symmetries built in the original algebra or group are hidden in the arising polynomial system, and it is not an easy task to reveal them.

However, sometimes one should be ready to pay this price, having in mind a potential reward: full access to a mill driven by arithmetic-geometric power, allowing one to grind hard grain. Such a reward is particularly achievable in the finite group case, when one can use representations over finite fields and apply to the resulting problem of the existence of rational points on algebraic varieties over finite fields the whole strength of high-tech homological algebra developed over the twentieth century (Weil–Grothendieck–Deligne). Perhaps the first, quite impressive instance of such tour de force is due to Bombieri [Bom1] (see also a follow-up in [Bom2]), who ingeniously applied elimination techniques to polynomial systems over finite fields which arise in classification of finite simple groups of Ree type (note that computer calculations played a critically important role). Over the past decade, some new applications of arithmetic-geometric methods appeared, which resulted in solving long-standing problems in the theory of finite [BGGKPP1]–[BGGKPP2] and infinite [BS1]–[BS2] groups; see surveys [Sh1]–[Sh3], [GKP], [BGaK] for more details (including the discussion of the famous Ore problem [Mall], where the impressive final stop was recently put in [LOST]).

In the present paper, our focus is, however, somewhat different: we are going to discuss the situation when matrix representations described above are defined over some infinite field (or ring) of arithmetic nature. Our primary interest is addressed to the field $\mathbb{Q}$ of rational numbers and the ring $\mathbb{Z}$ of rational integers. Here, comparing to the finite field case, achievements are much more modest, to put it mildly. We will give a brief survey of some of those, trying to put the relevant problems in a general context of local-global considerations, traditional for classical “commutative” Diophantine geometry.

Note that perhaps the only existing general result, Borel’s theorem [Bor], guaranteeing, in the case where $G$ is a connected semisimple linear algebraic group over an algebraically closed field and $w$ is any non-identity group word, the solvability of equation (1.2) with a “general” (in the sense of the Zariski topology) right-hand side,
EQUATIONS IN MATRIX GROUPS AND ALGEBRAS

gives very little for smaller fields: A. Thom showed that even over reals, if \( G \) is compact, the solutions of the relevant equation may, for an appropriately chosen \( w \), be included in an arbitrarily small (in the sense of real topology) neighbourhood of the identity [Th]. Once again, it is worth quoting the book of Cassels and Flynn [CF], the title of which inspired (and was shamelessly plagiarised by) the author of this survey: ‘Modern geometers take account of the ground field, but regard a field extension as a cheap manoeuvre. For practical computations it is desperately expensive: the theory we seek must avoid them at almost all costs.’

The author has to apologise that the corpus of presented explicit examples, as well as the theory brought for explaining some regularities, seem, at the current stage, too poor. What is badly needed is some replacement of the missing algebraic chain in Manin’s triad appearing in the title of [Man]; by now, we are not aware of any reasonable analogues of Galois-cohomological machinery and descent methods, which proved their power in clarifying ties between geometric and arithmetic properties of varieties considered in [Man] and [CF] (and in many other sources). This explains why the term “middlebrow” has been downgraded in the title of this essay.

2. LOCAL OBSTRACTIONS

First, it should be emphasised once again that in the present text we restrict our attention to considering word equations in simple algebraic groups, as well as to polynomial equations in simple matrix algebras (and close to such). Apart from the reason mentioned in the introduction (that we want to use algebraic-geometric machinery), there are other arguments in favour of such limitations: roughly, in a “random” group a “random” word equation has little chance to be solvable; even if we drop one of the assumptions on the group (either simplicity, or algebraicity), even the simplest word equations, such as the commutator equation \( [x, y] = g \), may turn out to be nonsolvable; see [KBKP] for relevant discussions and counter-examples.

So throughout below all word (resp. polynomial) equations will be considered in \( G(k) \), where \( G \) is a connected simple linear algebraic \( k \)-group (resp. in the matrix algebra \( M_n(k) \)).

As our focus is on arithmetic properties of matrix equations, \( k \) will typically stand for a number field (usually \( k = \mathbb{Q} \)) or its ring of integers.

We are going to discuss the traditional local-global approach to solving matrix equations over a number field or its ring of integers, so the first question is the existence of local obstructions. Here we notice the first unusual phenomenon: even an innocent looking matrix equation may have no solutions over \( \mathbb{C} \). Below are the simplest instances.

2.1. Power maps. Let \( w: G \to G \) be defined by \( w(x) = x^m \) where \( m \geq 2 \) is an integer. Looking at the equation
in $G = \text{SL}_2(\mathbb{C})$, we immediately notice that it has no solutions (apply Jordan’s theorem). This obstruction can easily be removed by factoring out the centre: indeed, one can show that in $\text{PSL}_2(\mathbb{C})$ any power equation $x^m = a$ is solvable. However, this observation is somewhat misleading because for simple algebraic groups of types other than $A$ the obstruction over $\mathbb{C}$ does exist:

**Theorem 2.1** ([Ch1], [Ch2], [Ste]). The map $x \mapsto x^m$ is surjective on $\mathcal{G}(k)$ ($k$ is an algebraically closed field of characteristic exponent $p$, $\mathcal{G}$ is a connected semisimple algebraic $k$-group) if and only if $m$ is prime to $p r z$, where $z$ is the order of the centre of $\mathcal{G}$ and $r$ is the product of “bad” primes.

So even in characteristic zero (when $p = 1$), even for groups of adjoint type (when $z = 1$), the power map $x \mapsto x^m$ is not surjective whenever $m$ is divisible by a bad prime $r$, which can be 2, 3 or 5, depending on the type of the Dynkin diagram of $\mathcal{G}$.

Other local obstructions ($p$-adic and real) to the surjectivity of power maps are also completely described, see [Ch3], [Ch4].

Let us now go over to more complicated maps.

### 2.2. Group case.

As above, here $\mathcal{G}$ stands for a connected semisimple algebraic $k$-group, $G = \mathcal{G}(k)$, $w$ is a word in $d$ letters. We denote by the same letter the map $w : G^d \to G$, $(g_1, \ldots, g_d) \mapsto w(g_1, \ldots, g_d)$.

Assume that $\mathcal{G}$ is of adjoint type, so $G$ has no centre. Here is a brief account of what is known by now.

(i) For $k = \mathbb{C}$ it is a challenging open question whether or not there exist a nonpower word $w$ (i.e., $w$ cannot be represented as a proper power of another word $v$) and a group $\mathcal{G}$ such that the map $w$ is not surjective; see [KBKP] for a survey of known surjectivity results.

(ii) For $k = \mathbb{R}$, A. Thom [Th] constructed words $w$ in two variables such that the image of the corresponding word map on the unitary group lies within a neighbourhood of the identity matrix, and this neighbourhood can be made as small as we wish by an appropriate choice of $w$.

(iii) This construction can easily be extended to an arbitrary compact simple real group and also (using a theorem of Bruhat–Tits–Rousseau, see [Pr] for a short proof) to an arbitrary anisotropic $p$-adic simple group (which is necessarily of type $A$).

If examples as in (i) above do exist, one can ask whether the surjectivity can be tested on real points:

**Question 2.2.** Do there exist a word $w$ and a simple algebraic $\mathbb{R}$-group $\mathcal{G}$ such that the map $w_{\mathbb{R}} : (\mathcal{G}(\mathbb{R}))^d \to \mathcal{G}(\mathbb{R})$ is surjective but the map $w_{\mathbb{C}} : (\mathcal{G}(\mathbb{C}))^d \to \mathcal{G}(\mathbb{C})$ is not?
Similar questions can be asked for justification of “global-to-local” surjectivity principles:

**Question 2.3.** Do there exist a word $w$ and a simple algebraic $\mathbb{Q}$-group $G$ such that the map $w_\mathbb{Q} : (G(\mathbb{Q}))^d \to G(\mathbb{Q})$ is surjective but the map $w_\mathbb{R} : (G(\mathbb{R}))^d \to G(\mathbb{R})$ is not?

**Question 2.4.** Do there exist a word $w$, a simple algebraic $\mathbb{Q}$-group $G$ and a prime number $p$ such that the map $w_\mathbb{Q} : (G(\mathbb{Q}))^d \to G(\mathbb{Q})$ is surjective but the map $w_p : (G(\mathbb{Q}_p))^d \to G(\mathbb{Q}_p)$ is not?

The author is not aware of any results in this direction.

### 2.3. Algebra case

For the solvability of polynomial equations (1.2) in matrix algebras there are several obstructions: obvious, such as the case where $F$ is an identical relation of the algebra under consideration, or the case where $F$ is a Lie polynomial and hence its image consists of trace zero matrices; less obvious but known for decades, such as the case where all values of $F$ lie in the centre of the algebra; more recent, described in [KBMR1]–[KBMR3], [Sp] (for matrices over an algebraically closed field) and [Male] (for matrices over a real closed field); see these papers, as well as the survey [KBKP] for more details and some open problems.

The case of Lie polynomials on simple Chevalley Lie algebras was treated in some detail in [BGKP], where among other things, one can find counter-examples to the surjectivity as well. However, the cases of real and $p$-adic algebras remain almost unexplored, for both associative and Lie algebras. In particular, the author does not know anything regarding analogues of Questions 2.2–2.4.

The next questions refer to an eventual analogue of Thom’s construction described in 2.2(ii):

**Questions 2.5.** Do there exist a simple real Lie algebra $\mathfrak{g}$ and a Lie polynomial $P$ such that the image of the map $P : \mathfrak{g}^d \to \mathfrak{g}$ lies within an open (in the real topology) neighbourhood of zero $U$ (different from $\mathfrak{g}$)? Can one choose $U$ as small as we wish by an appropriate choice of $P$? Can one arrange this for $\mathfrak{g} = \mathfrak{su}(n)$, the algebra of skew-hermitian trace zero matrices? Can one do the same for some simple $p$-adic Lie algebra?

### 3. Local-global principles

In light of Questions 2.2–2.4, it seems reasonable to formulate local-global principles for the surjectivity of word (or polynomial) maps as follows. Let $G$ denote a simple linear algebraic $\mathbb{Q}$-group. For $p \leq \infty$ we denote by $\mathbb{Q}_p$ the completion of $\mathbb{Q}$ at $p$ where $\mathbb{Q}_\infty := \mathbb{R}$.

**Question 3.1.** Let $w(x_1, \ldots, x_d) = g$ (resp. $P(X_1, \ldots, X_d) = a$) be a word (resp. polynomial) equation in the group $G(\mathbb{Q})$ (resp. in the associative algebra $M_n(\mathbb{Q})$). Suppose that for every $g \in G(\mathbb{Q})$ (resp. for every $a \in M_n(\mathbb{Q})$) the equation is solvable
in $G(\mathbb{Q}_p)$ (resp. in $M_n(\mathbb{Q}_p)$) for all $p \leq \infty$. Is it true that for every $g \in G(\mathbb{Q})$ (resp. for every $a \in M_n(\mathbb{Q})$) it is solvable in $G(\mathbb{Q})$ (resp. in $M_n(\mathbb{Q})$)?

If the answer to this question is “yes”, we say that the pair $(w,G)$ (resp. $(P,M_n)$) satisfies weak Grunwald–Wang principle (see the next section for justification of this terminology).

One can also consider a variant of this principle, where the local requirements are weakened so that the global solvability is implied by the local solvability at all but finitely many $p$. In this case we say that the pair under consideration satisfies strong Grunwald–Wang principle.

Naturally, both versions can be discussed when $\mathbb{Q}$ is replaced with any number field (or, more generally, any global field).

As in the case of obstructions over $\mathbb{C}$, let us start with power maps.

### 3.1. Power maps.

Variants of Question 3.1 can be posed for $G$ other than simple linear algebraic groups, say, for commutative algebraic groups. In this context, the case of power words $w = x^m$ was thoroughly investigated. For the multiplicative group $\mathbb{G}_m$, the answer follows from the Grunwald–Wang theorem [AT, Chapter X], [NSW, Chapter IX], [Co, Appendix A] (see [Roq] for historical details). Given a global field $k$ and a natural number $m$, it says whether there is an element of $k$ which is an $m^{th}$ power everywhere (or almost everywhere) locally but not globally. In particular, if $m$ is divisible by 8, there are counter-examples to strong Grunwald–Wang principle for $(x^m, \mathbb{G}_m, \mathbb{Q})$. For $(x^m, \mathbb{G}_m, \mathbb{Q}(\sqrt{7}))$ there are well-known counter-examples to weak Grunwald–Wang principle: 16 is not an $8^{th}$ power in $\mathbb{Q}(\sqrt{7})$ though it is an $8^{th}$ power everywhere locally.

For power maps on more general commutative algebraic groups, perhaps the first counter-example to weak Grunwald–Wang principle was constructed by Cassels and Flynn [CF, 6.9]. In their example, $m = 2$ and $G$ is an abelian $\mathbb{Q}$-surface. Over the past decade, both versions of Grunwald–Wang principles (weak and strong) for power maps have been studied in detail, which led to almost conclusive results for algebraic tori [DZ1], [Il], elliptic curves [Wo], [DZ1]–[DZ3], [Pa1]–[Pa3], [PRV], [Cr2], and higher-dimensional abelian varieties [Wo], [Cr1]. The interested reader is referred to [Cr1], [Cr2] for a detailed account of these results as well as for various ramifications of local-global principles for higher cohomology.

These considerations lead to counter-examples to weak Grunwald–Wang principle for power maps on simple linear algebraic groups defined over number fields such as $\mathbb{Q}(\sqrt{7})$: indeed, the diagonal $2 \times 2$ matrix with diagonal entries 16 and $1/16$ cannot be an $8^{th}$ power of a matrix from $\text{SL}_2(\mathbb{Q})$ though it is an $8^{th}$ power everywhere locally.

As to more general words, the author is not aware of any counter-examples to Grunwald–Wang principles.

### 3.2. Digression on relations with rationality and approximation problems.

It is well known that counter-examples to Grunwald–Wang principles are closely related to counter-examples to E. Noether’s problem on the rationality of the field of
invariants \( \mathbb{Q}(x_1, \ldots, x_n)^G \) with respect to the permutation action of a finite group \( G \) on the \( x_i \). Perhaps this relation was first noticed by Saltman [Sal1], [Sal2] (see also [Sw] for a survey of these results), who provided an argument based on the existence (or, more precisely, non-existence) of generic \( G \)-extensions and described relations with lifting and approximation problems. There is another approach, where a bridge between the Grunwald–Wang and Noether problems goes through birational invariants of certain algebraic tori.

Namely, according to Sansuc [San, §§1–2], the cohomological obstruction to the classical strong Grunwald–Wang principle is of purely algebraic nature: the deviation from this principle is measured by the finite abelian group

\[
\mathbb{H}_1^{\mathrm{e}}(k, \mu_m) = \ker \left( H^1(\Gamma, \mu_m) \to \bigoplus_{\langle \gamma \rangle \subset \Gamma} H^1(\langle \gamma \rangle, \mu_m) \right),
\]

(3.1)

where \( \mu_m \) is the group of \( m \)th roots of unity, \( \Gamma = \text{Gal}(K/k) \) is the Galois group of the splitting field \( K \) of \( \mu_m \), and the sum in the right-hand side is taken over all cyclic subgroups \( \langle \gamma \rangle \) of \( \Gamma \). Note that the deviation from the weak Grunwald–Wang principle is measured by the subgroup \( \mathbb{H}_1^{\mathrm{e}}(k, \mu_m) \) of \( \mathbb{H}_1^{\mathrm{e}}(k, \mu_m) \), which is of arithmetic nature, as well as their quotient, measuring the gap between the weak and strong principles (see [San, Lemme 1.4 and Sec. 2a]).

On the other hand, one can show (see [CTS]) that the same finite abelian group \( \mathbb{H}_1^{\mathrm{e}}(k, \mu_m) \) provides a Galois-cohomological obstruction to the rationality of the field of invariants in Noether’s problem, and give another interpretation of this obstruction. Namely, following Voskresenskiı (see [Vo1], [Vo2, Chapter VII]), consider, for any finite abelian group \( C \) (say, for \( C = \mathbb{Z}/m \)), the \( k \)-torus \( \mathbb{Q}_C = S/C \) where \( S \) is a quasitrivial torus (say, for \( C = \mathbb{Z}/m \) we have \( S = R_{K/k} \mathbb{G}_m \) where \( K \) is the splitting field of the character group \( \hat{C} = \mu_m \), i.e., the fixed field with respect to the action of \( \text{Gal}(k) \) on \( \hat{C} \)). Then the field of invariants in Noether’s problem for \( C \) is none other than the field of rational functions \( k(\mathbb{Q}_C) \). From the exact sequence of \( \Gamma \)-modules

\[
0 \to \hat{Q}_C \to \hat{S} \to \hat{C} \to 0,
\]

using the fact that \( \hat{S} \) is a permutation \( \Gamma \)-module and hence the group

\[
\mathfrak{H}_1^{\mathrm{e}}(k, \hat{S}) := \ker \left( H^2(\Gamma, \hat{S}) \to \bigoplus_{\langle \gamma \rangle \subset \Gamma} H^2(\langle \gamma \rangle, \hat{S}) \right),
\]

is trivial [San, Lemme 1.9], we obtain that \( \mathbb{H}_1^{\mathrm{e}}(k, \hat{C}) \cong \mathfrak{H}_1^{\mathrm{e}}(k, \hat{Q}_C) \). This latter group is a birational invariant of the \( k \)-torus \( \mathbb{Q}_C \). It is isomorphic to \( \text{Br}(X)/\text{Br}(k) \), where \( X \) is a smooth projective \( k \)-variety containing \( \mathbb{Q}_C \) as an open subset [CTS, Proposition 9.5(ii)] (the unramified Brauer group of \( k(\mathbb{Q}_C)/k \), using an alternative terminology), so whenever this invariant does not vanish, one can conclude that the
field in Noether’s problem for $C$ is not rational. Thus any counter-example to strong Grunwald–Wang principle yields a counter-example to Noether’s problem.

It is worth noting that the abelian group $\mathbb{III}_1^1(k,\mu_m)/\mathbb{III}_1^1(k,\mu_m)$, mentioned above (or, more precisely, its Pontryagin dual), carries somewhat different arithmetical information, which is in a sense complementary to weak Grunwald–Wang principle. Namely, whenever this group vanishes, the classical version of the Grunwald–Wang theorem for the field $k$ and the cyclic group $C = \mathbb{Z}/m$ holds: for any finite set $S$ of places of $k$ and any collection of local field extensions $K^{(v)}/k_v$ with Galois group $\Gamma_v$ ($v \in S$) embeddable into $C$, there exists an extension $K/k$ with Galois group $C$ such that for every $w \mid v$ one has $K_w \cong K^{(v)}$ (see [NSW, IX.2], [Co], [Ne] for other versions). This property of the triple $(k, C, S)$ is sometimes called weak approximation with respect to $S$. The pair $(k, C)$ is said to satisfy weak (resp. very weak) approximation if the triple $(k, C, S)$ satisfies weak approximation for every finite $S$ (resp. for every finite $S$ disjoint from some finite set $S_0$ depending from $k$ and $C$). The classical Grunwald–Wang theory establishes very weak approximation for all abelian groups $C$ and all global fields $k$ and describes all cases when weak approximation holds. These properties also make sense for nonabelian finite groups $C$ and are equivalent to weak (resp. very weak) approximation for quotients $\text{SL}_n/C$ (see, e.g., [Ha2]). According to a well-known observation due to Colliot-Thélène, very weak approximation for $(k, C)$ implies that the inverse Galois problem for $C$ over $k$ is solvable; moreover, it is solvable under a weaker assumption of so-called hyper-weak approximation [Ha2]; a recent paper [DG] describes other interrelations between Grunwald–Wang and inverse Galois problems.

For nonabelian groups $C$, conjectural relations of weak approximation properties to the unramified Brauer group are much more tricky comparing with the abelian case. We leave them aside referring the interested reader to [Ha2] and more recent papers [De], [BDH], [LA], [Ne]. In the latter paper there has been defined another cohomological invariant, detecting the failure of weak approximation for some semi-direct products, and there has been introduced a finer subdivision of Grunwald problems into tame (where all primes dividing the order of $C$ are outside $S$) and wild ones. It was shown in [Ne] that for every $C$ there are wild problems with negative solution. For tame problems no such examples are known.

To finish with this digression, let us mention a recent generalization of the Grunwald–Wang theorem, viewed as a statement on characters of $\mathbb{G}_m$, to automorphic representations of semisimple groups [So, §3].

3.3. **Back to general words.** It is not clear to what extent power words can serve as a prototypical toy-model for investigating local-global properties in the general case. It is a common understanding that, say, commutator words are in a certain sense opposite to powers (cf., e.g., [HSVZ, Section 6.5]). Moreover, looking at equidistribution properties of fibres of word maps on simple algebraic groups over finite fields, one can observe that the behaviour of all nonpower words is in striking contrast with the case
of power words [BK1]. However, the author believes that the lack of experimental
evidence does not prevent from raising various questions of local-global nature, such
as Question 3.1, for nonpower words as well. Note that collecting experimental data
is not an easy task: it is not obvious how to produce a nonpower word \( w \) such that
for some simple algebraic \( \mathbb{Q} \)-group \( G \) the map \( w : G(\mathbb{Q})^d \to G(\mathbb{Q}) \) is not surjective.
Here is an example of such a word.

**Proposition 3.2.** For \( w(x, y) = x^2[x^{-2}, y^{-1}]^2 \), the map \( w : SL_2(\mathbb{Q})^2 \to SL_2(\mathbb{Q}) \) is not surjective.

The assertion is an immediate consequence of results obtained by Jambor, Liebeck and O’Brien [JLO]. We revisit their proof, adjusting it to our setting, to use this
opportunity for demonstrating the power of the **trace method**.

This method dates back to classical works of Vogt, Fricke and Klein. The needed
results are quoted below from [Ho] (see [BGrK], [BGG], [BG], [BGaK] for more details
and applications).

Let \( F_d \) denote the free group on \( d \) generators. For \( G = SL_2(k) \) (\( k \) is any commu-
tative ring with 1) and for any \( u \in F_d \) denote by \( tr(u) : G^d \to G \) the trace character,
\( (g_1, \ldots, g_d) \mapsto tr(u(g_1, \ldots, g_d)) \).

**Theorem 3.3.** [Ho] If \( w \) is an arbitrary element of \( F_d \), then the character of \( w \) can be expressed as a polynomial
\( tr(w) = P_w(t_1, \ldots, t_d, t_{12}, \ldots, t_{12\ldots d}) \)
with integer coefficients in the \( 2^d - 1 \) characters \( t_{i_1i_2\ldots i_\nu} = tr(x_{i_1}x_{i_2} \ldots x_{i_\nu}) \), \( 1 \leq \nu \leq d, \)
\( 1 \leq i_1 < i_2 < \cdots < i_\nu \leq d \).

Let \( G = SL_2(k) \), and let \( \pi : G^d \to \mathbb{A}^{2^d-1} \) be defined by
\( \pi(g_1, \ldots, g_d) = (t_1, \ldots, t_d, t_{12}, \ldots, t_{12\ldots d}) \)
in the notation of Theorem 3.3.

Let \( Z_d := \pi(G^d) \subset \mathbb{A}^{2^d-1} \). Let \( w : G^d \to G \) be a word map. It follows from
Theorem 3.3 that for every \( d \) there exists a polynomial map \( \psi : \mathbb{A}^{2^d-1} \to \mathbb{A}^1 \) such that
the following diagram commutes:

\[
\begin{array}{c}
G^d \xrightarrow{w} G \\
\downarrow \pi \quad \downarrow \text{tr} \\
Z_d(k) \xrightarrow{\psi} \mathbb{A}^1.
\end{array}
\]  

(3.2)

Moreover, for small \( d \) we have a more precise information: one can take \( Z_2 = \mathbb{A}^3 \)
and \( Z_3 \subset \mathbb{A}^7 \) an explicitly given hypersurface. This diagram allows one to reduce
the study of the image and fibres of \( w \) to the corresponding problems for \( \psi \), which
may be much simpler. To ease the notation, we denote the coordinates of \( Z_2 = \mathbb{A}^3 \) by \( (s, t, u) \). For any prime \( p \) we denote by \( P_{w,p}(s, t, u) \) the reduction of the integer
polynomial \( P_w(s, t, u) \) modulo \( p \).
**Proof of the proposition.** We will use diagram (3.2) for \( k = \mathbb{Q} \) and \( k = \mathbb{F}_p \). Corollary 3 of [JLO] implies that for infinitely many primes \( p \) the word map \( w: \text{SL}_2(\mathbb{F}_p)^2 \to \text{SL}_2(\mathbb{F}_p) \) induced by the word \( w \) from the statement of the proposition is not surjective. More precisely, looking at the proof of this corollary, one can see that for these primes the image of \( \psi \) does not contain 0. In other words, for any such \( p \) all values of the polynomial \( P_{w,p} \) are nonzero. Let \( \mathcal{P} \) denote the set of all these primes.

Let us show that the image of the map \( w: \text{SL}_2(\mathbb{Q})^2 \to \text{SL}_2(\mathbb{Q}) \) does not contain matrices with trace zero. Assume to the contrary that there exist \( A, B \in \text{SL}_2(\mathbb{Q}) \) such that the trace of \( w(A, B) \) equals zero. Let \( s = \text{tr}(A) \), \( t = \text{tr}(B) \), \( u = \text{tr}(AB) \). From diagram (3.2) we have \( P_w(s, t, u) = 0 \). Since the set of prime divisors of denominators of rational numbers \( s, t, u \) is finite and \( \mathcal{P} \) is infinite, there exists \( p \in \mathcal{P} \) such that the reductions \( \bar{s}, \bar{t}, \bar{u} \in \mathbb{F}_p \) are well-defined and \( P_{w,p}(\bar{s}, \bar{t}, \bar{u}) = 0 \), contradiction. \( \square \)

### 3.4. Approximation problems.

By analogy with the commutative case, along with local-global problems such as Question 3.1, one can consider various approximation problems. We restrict ourselves to stating them in the group case.

For a word \( w = w(x_1, \ldots, x_d) \) and \( g \in G(\mathbb{Q}) \), where \( G \), as above, stands for a simple linear algebraic \( \mathbb{Q} \)-group embedded into an affine \( \mathbb{Q} \)-space \( A^N \), denote by \( X_{w,g} \) the affine algebraic variety defined over \( \mathbb{Q} \) by the system of polynomial equations in \( A^N \) arising from the matrix equation \( w(x_1, \ldots, x_d) = g \). Assume that for every \( g \in G(\mathbb{Q}) \) the variety \( X_{w,g} \) has a \( \mathbb{Q} \)-point.

**Question 3.4.**

(i) Is it true that for any finite set \( S \) of primes \( p \leq \infty \) and for every \( g \in G(\mathbb{Q}) \) the set \( X_{w,g}(\mathbb{Q}) \) is dense in \( \prod_{p \in S} X_{w,g}(\mathbb{Q}_p) \)?

(ii) Does there exist a finite set \( S_0 \) such that for any finite set \( S \) of primes \( p \leq \infty \) disjoint from \( S_0 \) and for every \( g \in G(\mathbb{Q}) \) the set \( X_{w,g}(\mathbb{Q}) \) is dense in \( \prod_{p \in S} X_{w,g}(\mathbb{Q}_p) \)?

If Question 3.4(i) (resp. (ii)) is answered in the affirmative, we say that the word \( w \) satisfies weak (resp. very weak) approximation in \( G \).

Although Question 3.4 does not seem to be easy for an arbitrary \( G \), there is a hope that the trace map may be helpful in the case \( G = \text{SL}_2 \). Say, for \( d = 2 \) diagram (3.2) then induces a fibration \( \pi: X_{w,g} \to Y_{w,g} \) where the surface \( Y_{w,g} \subset A^3_{\mathbb{Q}} \) is given by the equation \( \psi(s, u, t) = \text{tr}(g) \). For the commutator word \( w = [x, y] \) this approach will hopefully lead to (at least very) weak approximation in \( \text{SL}_2 \); using [Ha1], one can show that in this case for \( X_{w,g} \) the Brauer–Manin obstruction to weak approximation is the only obstruction. Details will follow in a forthcoming paper [BK2].

### 3.5. Strong approximation and local-global principles over integers.

Diophantine questions similar to those discussed above can be posed in the case where \( G \) is a group scheme over \( \mathbb{Z} \) (or, more generally, over the ring of integers \( O_K \) of a global field \( K \)) such that its generic fibre is a simple linear algebraic group over \( \mathbb{Q} \) (or \( K \)).

Little is known in such set-up. Even local obstructions are not completely understood though some recent results give certain hope for conclusive answers, at least
for $G = SL_n$. Say, for the commutator word $w = [x, y]$ the situation is roughly as follows: obstructions for representing an element of $\mathbb{Z}_p$ as a commutator vanish (after excluding necessary congruence obstructions) as soon as $p$ is sufficiently large with respect to the Lie rank of $G$ (see [GS, Theorem 5], [AGKS, Theorem 3.5]). Here is an explicit instance of such a phenomenon [AGKS, Theorem 3.8]: if $n$ is a proper divisor of $p - 1$, then every element of $PSL_n(\mathbb{Z}_p)$ is a commutator. The reader is referred to [AGKS] for more results in this direction, some of those answer the questions raised in [Sh2].

As to the global case, practically nothing is known, even for simplest words and groups. Here is a tempting question raised by Shalev [Sh2], [Sh3]: is every element of $SL_n(\mathbb{Z})$ ($n \geq 3$) a commutator? For $n = 2$ the answer to a similar question is obviously negative. However, the author is not aware of any answer to the following question.

**Question 3.5.** Let $A \in SL_2(\mathbb{Z})$ be representable as a commutator in all groups $SL_2(\mathbb{Z}_p)$. Is $A$ a commutator in $SL_2(\mathbb{Z})$?

(Apparently, only $p = 2$ and $p = 3$ provide nontrivial local obstructions.)

Recent developments in understanding strong approximation for varieties admitting nice fibrations [CTX], [CTH] give some hope for successful application of the trace method, as in approximation problems mentioned above.

One can consider analogues of Question 3.5 for other arithmetically interesting cases, such as Bianchi groups $PSL_2(O_d)$, where $O_d$ is the ring of integers of the imaginary quadratic field $\mathbb{Q}(\sqrt{-d})$. Higher-dimensional analogues, such as Picard modular groups $SU(n, 1; O_d)$, look even more challenging.

**3.6. Case of simple algebras.** In the case where $P(X_1, \ldots, X_d)$ is an associative (or Lie) polynomial, the author is not aware of any example of failure of analogues of Grunwald–Wang principles for the map of simple associative (or Lie) algebras $A^d \to A$ induced by $P$ (apart from the power maps discussed above), neither over global fields, nor over their rings of integers. Some natural cases to be considered are, say, the following ones:

(i) $P = X_1^2 + \cdots + X_d^2$;
(ii) additive commutator $P = [X_1, X_2] = X_1X_2 - X_2X_1$.

Some interesting results for case (i) were obtained in [Va1], [KG], [Pu], however, local-global questions as above seem to remain open. For the additive commutator in matrix algebras, the induced map $A^2 \to A_0$, where $A_0$ denotes the collection of all trace zero elements of $A$, is surjective in each of the following cases (see [KBKP], [Sta] for more details):

(i) $A = M_n(D)$, where $D$ is a central division algebra over a field $k$ and $n \geq 2$ [AR];
(ii) for $n = 1$ this is only known in the case where $k$ is a local field and the degree of $D$ is prime [Ros];
A = M_2(R), where R is a principal ideal domain [Li], [Va2], [RR];

(iv) A = M_n(Z) [LR] or, more generally, A = M_n(R), where R is a principal ideal ring [Sta].

Remark 3.6. Regarding case (iv), for A = M_n(R) where n ≥ 3 and R is an arbitrary Dedekind ring, the question on the surjectivity of the commutator map A^2 → A_0 remains open, see a discussion in [Sta].

Remark 3.7. If the commutator map as above is regarded as a map of Lie algebras sl(n, R) × sl(n, R) → sl(n, R), the situation becomes more subtle: it is not clear whether in the Laffey–Reams–Stasinski theorem on representing any trace zero matrix in the form AB − BA one can choose A and B to be trace zero matrices. Say, this is not always possible in the case n = 2, R = Z: if it were the case, it would also be possible in g = sl(2, F_2), but [g, g] ≠ g. (I thank N. Gordeev for this remark.)

It is tempting to use the local-global approach to the surjectivity problem for this map as well. Naturally, similar problems make sense for any Lie polynomial P(X_1, . . . , X_d), when one can ask about the surjectivity of the induced map of Lie rings g(R)^d → g(R), where g is a Chevalley Lie algebra over Z and R is the ring of integers of a global field.

Remark 3.8. Perhaps, the first class of polynomials to which one can try to extend results known for additive commutator are additive Engel polynomials E_n(X_1, X_2) = [[[X_1, X_2], . . . , X_2]]. Nothing is known about the surjectivity of induced maps. Both the associative and Lie case are completely open, even over fields, even for the simplest polynomial E_3(X_1, X_2) = [[[X_1, X_2], X_2]]. Note, however, that for a somewhat similar polynomial, generalised commutator P(X_1, X_2, X_3) = X_1X_2X_3 − X_3X_2X_1, the surjectivity of the induced map A^3 → A is known for any division algebra A [GKKL] and for any matrix algebra A = M_n(R) over a principal ideal domain R [KL].

Remark 3.9. To conclude, it is worth emphasizing the importance of pursuing parallels between matrix equations in groups and algebras, especially in the Lie algebra case that can be viewed as a bridge between groups and associative algebras. Certain considerations in [AGKS] and [BRL] give some hope for making this into a working tool.

Acknowledgements. The author was supported in part by the Israel Science Foundation, grant 1207/12, and by the Minerva Foundation through the Emmy Noether Research Institute for Mathematics; a part of this work was done when he participated in the trimester program “Arithmetic and Geometry” in the Hausdorff Research Institute for Mathematics (Bonn).

The author thanks T. Bandman, N. Gordeev, D. Neftin and A. Shalev for useful discussions and correspondence and T. Bauer for providing reference [Sta].

References


Department of Mathematics, Bar-Ilan University, 5290002 Ramat Gan, ISRAEL
E-mail address: kunyav@macs.biu.ac.il