European Journal of Mathematics (2023) 9:53
https://doi.org/10.1007/s40879-023-00627-y

RESEARCH ARTICLE

Bounded generation and commutator width of Chevalley groups: function case

Boris Kunyavskiĭ[1] · Eugene Plotkin[1] · Nikolai Vavilov[2]

In memory of Irina Suprunenko, our dear friend and colleague

Received: 24 August 2022 / Revised: 5 November 2022 / Accepted: 25 March 2023 /
Published online: 29 June 2023

Abstract

We prove that Chevalley groups over polynomial rings $\mathbb{F}_q[t]$ and over Laurent polynomial $\mathbb{F}_q[t, t^{-1}]$ rings, where $\mathbb{F}_q$ is a finite field, are boundedly elementarily generated. Using this we produce explicit bounds of the commutator width of these groups. Under some additional assumptions, we prove similar results for other classes of Chevalley groups over Dedekind rings of arithmetic rings in positive characteristic. As a corollary, we produce explicit estimates for the commutator width of affine Kac–Moody groups defined over finite fields. The paper contains also a broader discussion of the bounded generation problem for groups of Lie type, some applications and a list of unsolved problems in the field.

Keywords Chevalley groups · Kac–Moody groups · Bounded generation · Polynomial rings · First order rigidity

Mathematics Subject Classification 20G07

Research of Boris Kunyavskiĭ and Eugene Plotkin was supported by the ISF Grants 1623/16 and 1994/20. Nikolai Vavilov thanks the “Basis” Foundation Grant N. 20-7-1-27-1 “Higher Symbols in Algebraic K-Theory”.

✉ Eugene Plotkin
plotkin.evgeny@gmail.com

Boris Kunyavskiĭ
kunyav@gmail.com

Nikolai Vavilov
nikolai-vavilov@yandex.ru

[1] Department of Mathematics, Bar-Ilan University, Ramat Gan 5290002, Israel

[2] Department of Mathematics and Computer Science, St. Petersburg State University, St. Petersburg, Russia

1 Introduction

In the present paper, we consider Chevalley groups $G = G(\Phi, R)$ and their elementary subgroups $E(\Phi, R)$ over various classes of rings, primarily over Dedekind rings of arithmetic type. In some special cases these groups are closely related to various Kac–Moody type groups, and we can derive some non-trivial corollaries in this situation.

Primarily, we are interested in the classical problems of estimating the width of $G(\Phi, R)$ and $E(\Phi, R)$ with respect to the following two paradigmatic generating sets:

- The elementary generators $x_\alpha(\xi)$, $\alpha \in \Phi$, $\xi \in R$. We say that a group G is *boundedly elementarily generated* if it has finite width $w_{\mathrm{E}}(G)$ with respect to elementary generators.
- Commutators $[x, y] = xyx^{-1}y^{-1}$, where $x, y \in G$. In this case we say that G has *finite commutator width* $w_{\mathrm{C}}(G)$.

(To treat the cases where G is not perfect, we define its commutator width as supremum of the commutator lengths of the elements of the commutator subgroup $[G, G]$; abusing notation, we still denote it by $w_{\mathrm{C}}(G)$ and keep this notation and convention throughout the paper; see Sects. 3.1, 3.2 where the arising subtleties are discussed in some detail.)

In the proofs we work also with other related generating sets, such as elements in the unipotent radicals of various parabolic subgroups, which are closely related but better behaved with respect to stability maps.

For Chevalley groups of rank $\geqslant 2$, bounded generation in terms of elementary generators and bounded generation in terms of commutators are essentially equivalent. Indeed, in this case the Chevalley commutator formula readily implies that every elementary generator of G lying in $[G, G]$ can be presented as a product of a bounded number of commutators. Conversely, a very deep result by Alexei Stepanov and others (see in particular [66, 77], and in final form [76]) implies that given any commutative ring R, every commutator in $E(\Phi, R)$ is a product of not more than L elementary generators, with the bound $L = L(\Phi)$ depending on Φ alone. But of course the actual estimates of $w_{\mathrm{E}}(G)$ and $w_{\mathrm{C}}(G)$ can be very different.

Both problems have attracted considerable attention over the last 40 years or so. Roughly, the situation is as follows. Bounded elementary generation always holds with obvious bounds for 0-dimensional rings and usually fails for rings of dimension $\geqslant 2$. But for 1-dimensional rings it is problematic.

Thus, from the existence of arbitrary long division chains in Euclidean algorithm it follows that $\mathrm{SL}(2, \mathbb{Z})$ and $\mathrm{SL}(2, \mathbb{F}_q[t])$ are not boundedly elementarily generated. But this could be attributed to the exceptional behaviour of rank 1 groups. Much more surprisingly, Wilberd van der Kallen [83] established that bounded generation fails even for $\mathrm{SL}(3, \mathbb{C}[t])$, a group of Lie rank 2 over a Euclidean ring!

An emblematic example of 1-dimensional rings are Dedekind rings of arithmetic type $R = \mathcal{O}_S$, for which bounded elementary generation of $G(\Phi, R)$ is intrinsically related to the positive solution of the congruence subgroup problem in that group. This connection was first noted by Vladimir Platonov and Andrei Rapinchuk, see [56, 60, 61].

For the *number case* the situation is well understood, even for rank 1 groups. After the initial breakthrough by Douglas Carter and Gordon Keller [9, 10], later expanded by Oleg Tavgen [78] and many others, we now know bounded generation with excellent bounds depending on the type of Φ and the class number of R for *all Chevalley groups* of rank $\geqslant 2$. Apart from the rings $R = \mathcal{O}_S$, $|S| = 1$, with finite multiplicative group, similar results are even available for $\mathrm{SL}(2, R)$, see a detailed survey in Sect. 3.

However, the *function case* turned out to be much more recalcitrant, and is not solved up to now, apart from some important but isolated results, such as the works by Clifford Queen [59] and Bogdan Nica [51], which treat the group $\mathrm{SL}(2, R)$ over *some* arithmetic function rings with infinite multiplicative groups, and the groups $\mathrm{SL}(n, \mathbb{F}_q[t])$, $n \geqslant 3$, respectively.[1]

Here we expand these results to all Chevalley groups, obtaining explicit bounds. The first major new result of the present paper establishes bounded elementary generation for all Chevalley groups of rank at least 2 over the most classical, and in a sense the most difficult example, polynomial rings $\mathbb{F}_q[t]$ with coefficients in finite fields.[2]

Theorem A *Let $G(\Phi, R)$ be a simply connected Chevalley group of type Φ, $\mathrm{rk}(\Phi) \geqslant 2$ over $R = \mathbb{F}_q[t]$. Then the width of $G(\Phi, R)$ with respect to elementary generators is bounded by a constant not depending on q.*

The proof of this result constitutes about half of the paper. *Some* bound in the bounded generation for all Chevalley groups can be easily derived from the case of rank two systems by a version of the usual Tavgen's trick [78, Theorem 1], described in [68, 89].

- For A_2 bounded generation of $\mathrm{SL}(3, \mathbb{F}_q[t])$ is precisely the main result of Nica [51].
- A large part of the present paper is the analysis of the most difficult case of $\mathrm{Sp}(4, \mathbb{F}_q[t])$, which is the Chevalley group of type C_2. Again, we take the proof in Tavgen's paper [78, Section 4], as a prototype. But there is a substantial difference, since now we have to verify some arithmetic properties that are well known in the number case, but for which we could not find any reference in the function case.
- Luckily, we do not have to imitate Tavgen's proof [78, Section 5], for the remaining case of the Chevalley group of type G_2. Instead of a difficult direct calculation, we show that this case can be derived from the case of A_2 by the usual stability arguments.

For $\mathrm{SL}(n, R)$ there is a realistic bound of the width in elementary generators, in terms of stability conditions, taking into account the fact that for Dedekind rings $\mathrm{sr}(R) = 1.5$. The aforementioned proof of Theorem A gives us an occasion to return to the stability

[1] The difference between the number and function cases is subtle enough and may be overlooked when approaching from outside. We quote from page 2 of the memoir [28]: '…G is known to be boundedly generated by X only in a few cases, namely, when R is a finite extension of $\mathbb{Z}$ or $F[t]$, with F a finite field.' In a sense, the present paper, along with [51], can be viewed as a first step along the long and painful road to justification of this brave claim.

[2] After the preliminary version of the present paper has been finished, there appeared a preprint of Alexander Trost [81] where the statement of our Theorem A was established for the ring of integers R of an arbitrary global function field K, with a bound of the form $L(d, q) \cdot |\Phi|$, where the factor L depends on q and of the degree d of K. His method is similar to Morris' approach in [49]. The subsequent preprint [82] contains further improvements for the group $\mathrm{SL}(n, R)$, $n \geqslant 3$, over an arbitrary global function ring R.

arguments for all Chevalley groups, and obtain bounds which are substantially better than the ones that could be obtained via Tavgen's trick.

Alternatively, Theorem A can be restated in the following equivalent form. The difference is that in this case the computations of many authors, subsumed and expanded by Andrei Smolensky [67], allow one to produce very reasonable bounds, usually at most 6, 7 or 8 commutators.

Theorem B *Let $G(\Phi, R)$ be a simply connected Chevalley group of type Φ, $\mathrm{rk}(\Phi) \geqslant 2$ over $R = \mathbb{F}_q[t]$. Then $G(\Phi, R)$ is of finite commutator width.*

Remark 1.1 The commutator width of a Chevalley group of type Φ depends on the lattice $\mathcal{P}$ determining it. For example, $w_C(\mathrm{PSL}(2, \mathbb{Q})) = 1$ while $w_C(\mathrm{SL}(2, \mathbb{Q})) = 2$ (the matrix $-I$ is not representable as a single commutator and is a product of two commutators, see [79]). So, if the lattice is not stated explicitly, under $w_C(G(\Phi, R))$ we always mean maximum, i.e., the commutator width of the simply connected group.

See Sect. 3.2 for the discussion of subtleties arising in the cases where G is not perfect.

In fact, for applications to Kac–Moody groups we do not need the full force of Theorem A. We only need a similar result for the equally classical but *much easier* example of *Laurent* polynomial rings $\mathbb{F}_q[t, t^{-1}]$ with coefficients in finite fields.

For Chevalley groups over such rings bounded generation can be derived from Theorem A. Yet, the bounds thus obtained will not be the best possible ones. However, the multiplicative group of the ring $R = \mathbb{F}_q[t, t^{-1}]$ is *infinite*. This means that alternatively bounded generation can be derived—with much better bounds!—from the result by Clifford Queen [59]. Let us state the most spectacular finiteness result in terms of unitriangular factors obtained along this route.

Theorem C *Let $R = \mathcal{O}_S$ be the ring of S-integers of K, a function field of one variable over $\mathbb{F}_q$ with S containing at least two places. Assume that at least one of the following holds:*

- *either at least one of these places has degree one, or*
- *the class number of R, as a Dedekind domain, is prime to $q - 1$.*

Then any simply connected Chevalley group $G = G(\Phi, R)$ admits the following decompositions:

$$G = UU^-UU^-U = U^-UU^-UU^-.$$

Such a sharp bound was quite unexpected for us. In particular, Chevalley groups over such *arithmetic* rings have the same commutator width as Chevalley groups over rings of stable rank 1, see [67].

In particular, we can now give the same bounds for affine Kac–Moody groups.

Theorem D *The commutator width of an affine elementary untwisted Kac–Moody group $\widetilde{E}_{\mathrm{sc}}(A, \mathbb{F}_q)$ over a finite field $\mathbb{F}_q$ is $\leqslant L'$, where*

- *$L' = 5$ for $\Phi = \mathrm{F}_4$ and $\Phi = \mathrm{A}_l$, $l = 2k + 1$, $k = 0, 1, \ldots$;*

- $L' = 6$ *for* $\Phi = \mathrm{A}_l, l = 2k, k = 1, 2, \dots, \Phi = \mathrm{B}_l, \mathrm{C}_l, \mathrm{D}_l,$ *for* $l \geqslant 3$ *or* $\Phi = \mathrm{E}_7, \mathrm{E}_8,$ *or, finally,* $\Phi = \mathrm{C}_2, \mathrm{G}_2$ *under the additional assumption that* 1 *is the sum of two units in* R (*which is automatically the case provided* $q \neq 2$);
- $L' = 7$ *for* $\Phi = \mathrm{E}_6$.

The paper is organised as follows. In Sect. 2 we recall the necessary notation and preliminaries and in Sect. 3 provide background and historical survey. The next four sections constitute the technical core of the paper. Namely, in Sect. 4 we sketch the scheme of the proof of Theorem A, of which Theorem B is an immediate corollary, and reduce its proof to the rank 2 groups. This reduction is a variation of Tavgen's rank reduction trick, a further slight improvement of the rank reduction results in [68, 89]. In Sect. 5 we revisit surjective stability for K_1 modeled on Chevalley groups, with explicit bounds, and, in particular, reduce the case of the group $\mathrm{G}_2(R)$ to the known case of $\mathrm{SL}(3, R)$. In Sect. 6 we prove Theorem A for the group $\mathrm{Sp}(4, R)$, which is the most exciting case of all, and requires rather difficult algebraic and arithmetic considerations. Section 7 contains an alternative argument based on reducing to rank 3 groups and separate consideration of the types B_3 and C_3. Incidentally, this gives estimates with better constants. After that, in Sect. 8 we develop an alternative approach to bounded elementary generation, based on Queen's result, that gives sharper bounds for some classes of rings R with infinite multiplicative groups, including Laurent polynomial rings, thus proving Theorem C. The next section is devoted to applications. In Sect. 9.1 we discuss applications to Kac–Moody groups over finite fields and prove Theorem D, and in Sect. 9.2 we obtain some applications of bounded generation in model theory. Finally, in Sect. 10 we present some relevant concluding remarks and open problems.

2 Notation and preliminaries

In this section we briefly recall the notation that will be used throughout the paper. For more details on Chevalley groups over rings see [87, 88], where one can find many further references.

2.1 Chevalley groups

Let Φ be a reduced irreducible root system of rank $\geqslant 2$, and $W = W(\Phi)$ be its Weyl group. Choose an order on Φ and let Φ^+, Φ^- and $\Pi = \{\alpha_1, \dots, \alpha_l\}$ be the corresponding sets of positive, negative and fundamental roots, respectively. Further, we consider a lattice $\mathcal{P}$ intermediate between the root lattice $\mathcal{Q}(\Phi)$ and the weight lattice $\mathcal{P}(\Phi)$. Finally, let R be a commutative ring with 1, with the multiplicative group R^*.

These data determine the Chevalley group $G = G_{\mathcal{P}}(\Phi, R)$, of type $(\Phi, \mathcal{P})$ over R. It is usually constructed as the group of R-points of the Chevalley–Demazure group scheme $G_{\mathcal{P}}(\Phi, -)$ of type $(\Phi, \mathcal{P})$. In the case $\mathcal{P} = \mathcal{P}(\Phi)$ the group G is called simply connected and is denoted by $G_{\mathrm{sc}}(\Phi, R)$. In another extreme case $\mathcal{P} = \mathcal{Q}(\Phi)$ the group G is called adjoint and is denoted by $G_{\mathrm{ad}}(\Phi, R)$. Many results do not depend on the lattice $\mathcal{P}$ and hold for all groups of a given type Φ. In all such cases, or when $\mathcal{P}$ is determined by the context, we omit any reference to $\mathcal{P}$ in the notation and denote by $G(\Phi, R)$ *any* Chevalley group of type Φ over R. Usually, we assume that $G(\Phi, R)$ is simply connected.

In what follows, we also fix a split maximal torus $T = T(\Phi, R)$ in $G = G(\Phi, R)$ and identify Φ with $\Phi(G, T)$. This choice uniquely determines the unipotent root subgroups, X_α, $\alpha \in \Phi$, in G, elementary with respect to T. As usual, we fix maps $x_\alpha \colon R \mapsto X_\alpha$, so that $X_\alpha = \{x_\alpha(\xi) \mid \xi \in R\}$, and require that these parametrisations are interrelated by the Chevalley commutator formula with integer coefficients, see [13, 75]. The above unipotent elements $x_\alpha(\xi)$, where $\alpha \in \Phi$, $\xi \in R$, elementary with respect to $T(\Phi, R)$, are also called [elementary] unipotent root elements or, for short, simply root unipotents.

Further,

$$E(\Phi, R) = \langle x_\alpha(\xi),\ \alpha \in \Phi,\ \xi \in R \rangle$$

denotes the *absolute* elementary subgroup of $G(\Phi, R)$, spanned by all elementary root unipotents, or, what is the same, by all [elementary] root subgroups X_α, $\alpha \in \Phi$.

Since we are interested in the bounded generation, we also consider the *subset* $E^L(\Phi, R)$, consisting of products of $\leqslant L$ root unipotents. Since $E^L(\Phi, R)$ contains all generators of $E(\Phi, R)$, it is not a subgroup of $E(\Phi, R)$, unless $E^L(\Phi, R) = E(\Phi, R)$.

2.2 Root elements

Further, let $\alpha \in \Phi$ and $\varepsilon \in R^*$. As usual, we set

$$w_\alpha(\varepsilon) = x_\alpha(\varepsilon) x_{-\alpha}(-\varepsilon^{-1}) x_\alpha(\varepsilon), \quad h_\alpha(\varepsilon) = w_\alpha(\varepsilon) w_\alpha(1)^{-1}.$$

The elements $h_\alpha(\varepsilon)$ are called semisimple root elements.

By definition, $h_\alpha(\varepsilon)$ is a product of *six* elementary unipotents—well, actually if you look inside, *five* of them. However, it is classically known that $h_\alpha(\varepsilon)$ is a product of *four* elementary unipotents[3]. To somewhat improve some of the ulterior bounds we need a still more precise form of this classical observation, asserting that the first/last of these four factors can be chosen either lower, or upper, with an *arbitrary* invertible parameter. After that the remaining three factors are uniquely determined.

The following fact is obvious, but we could not find an explicit reference.

Lemma 2.1 *Let R be any commutative ring. Then for any $\varepsilon, \eta \in R^*$ the matrix $h_\alpha(\varepsilon)$ can be represented as the product of the form*

[3] On the other hand, since $B \cap U^- = e$, it is never a product of *three* such unipotents, unless $\varepsilon = 1$.

$$\begin{aligned}
h_\alpha(\varepsilon) &= x_{-\alpha}(\eta)\,x_\alpha\big(-\eta^{-1}(1-\varepsilon^{-1})\big)\,x_{-\alpha}(-\varepsilon\eta)\,x_\alpha\big(\varepsilon^{-1}\eta^{-1}(1-\varepsilon^{-1})\big)\\
&= x_{-\alpha}\big(\varepsilon^{-1}\eta^{-1}(1-\varepsilon^{-1})\big)\,x_\alpha(-\varepsilon\eta)\,x_{-\alpha}\big(-\eta^{-1}(1-\varepsilon^{-1})\big)\,x_\alpha(\eta)\\
&= x_\alpha(\varepsilon\eta(1-\varepsilon))\,x_{-\alpha}(-\varepsilon^{-1}\eta^{-1})\,x_\alpha(-\eta(1-\varepsilon))\,x_{-\alpha}(\eta^{-1})\\
&= x_\alpha(\eta^{-1})\,x_{-\alpha}(-\eta(1-\varepsilon))\,x_\alpha(-\varepsilon^{-1}\eta^{-1})\,x_{-\alpha}(\varepsilon\eta(1-\varepsilon)).
\end{aligned}$$

Proof Verify one of these formulae by a direct calculation in $\mathrm{SL}(2, R)$, then transpose, invert and transpose-invert it. □

Corollary 2.2 *Let R be any commutative ring. Then for any $\varepsilon, \lambda \in R^*$ the matrix $h_\alpha(\varepsilon)$ can be transformed to $h_\alpha(\lambda)$ by four elementary moves.*

Proof By Lemma 2.1, $h_\alpha(\varepsilon\lambda^{-1}) = h_\alpha(\varepsilon)(h_\alpha(\lambda))^{-1}$ can be transformed to 1 by four elementary moves, whence the statement. □

Next, let $N = N(\Phi, R)$ be the algebraic normaliser of the torus $T = T(\Phi, R)$, i.e. the subgroup, generated by $T = T(\Phi, R)$ and all elements $w_\alpha(1)$, $\alpha \in \Phi$. The factor-group N/T is canonically isomorphic to the Weyl group W, and for each $w \in W$ we fix its preimage $n_w \in N$. Clearly, such a preimage can be taken in $E(\Phi, R)$. Indeed, for a root reflection w_α one can take $w_\alpha(1) \in E(\Phi, R)$ as its preimage, any element w of the Weyl group can be expressed as a product of root reflections.

In particular, we get the following classical result, which is crucial in reduction to smaller ranks.

Lemma 2.3 *The elementary Chevalley group $E(\Phi, R)$ is generated by unipotent root elements $x_\alpha(\xi)$, $\alpha \in \pm\Pi$, $\xi \in R$, corresponding to the fundamental and negative fundamental roots.*

Further, let $B = B(\Phi, R)$ and $B^- = B^-(\Phi, R)$ be a pair of opposite Borel subgroups containing $T = T(\Phi, R)$, standard with respect to the given order. Recall that B and B^- are semidirect products $B = T \leftthreetimes U$ and $B^- = T \leftthreetimes U^-$, of the torus T and their unipotent radicals

$$\begin{aligned}
U &= U(\Phi, R) = \langle x_\alpha(\xi),\ \alpha \in \Phi^+,\ \xi \in R\rangle,\\
U^- &= U^-(\Phi, R) = \langle x_\alpha(\xi),\ \alpha \in \Phi^-,\ \xi \in R\rangle.
\end{aligned}$$

Here, as usual, for a subset X of a group G one denotes by $\langle X\rangle$ the subgroup in G generated by X. Semidirect product decomposition of B amounts to saying that $B = TU = UT$, and at that $U \trianglelefteq B$ and $T \cap U = 1$. Similar facts hold with B and U replaced by B^- and U^-. Sometimes, to speak of both subgroups U and U^- simultaneously, we denote $U = U(\Phi, R)$ by $U^+ = U^+(\Phi, R)$.

2.3 Levi decomposition

The main role in the reduction to smaller ranks is played by Levi decomposition for elementary parabolic subgroups. In general, one can associate a subgroup $E(S) = E(S, R)$ to any closed set $S \subseteq \Phi$. Recall that a subset S of Φ is called *closed*, if for any two roots $\alpha, \beta \in S$ the fact that $\alpha + \beta \in \Phi$, implies that already $\alpha + \beta \in S$. Now, we define $E(S) = E(S, R)$ as the subgroup generated by all elementary root unipotent subgroups $X_\alpha, \alpha \in S$:

$$E(S, R) = \langle x_\alpha(\xi), \ \alpha \in S, \ \xi \in R \rangle.$$

In this notation, U and U^- coincide with $E(\Phi^+, R)$ and $E(\Phi^-, R)$, respectively. The groups $E(S, R)$ are particularly important in the case where S is a *special* (= *unipotent*) set of roots; in other words, where $S \cap (-S) = \varnothing$. In this case $E(S, R)$ coincides with the *product* of root subgroups $X_\alpha, \alpha \in S$, in some/any fixed order.

Let again $S \subseteq \Phi$ be a closed set of roots. Then S can be decomposed into a disjoint union of its *reductive* (= *symmetric*) part S^{r}, consisting of those $\alpha \in S$, for which $-\alpha \in S$, and its *unipotent* part S^{u}, consisting of those $\alpha \in S$, for which $-\alpha \notin S$. The set S^{r} is a closed root subsystem, whereas the set S^{u} is special. Moreover, S^{u} is an *ideal* of S, in other words, if $\alpha \in S$, $\beta \in S^{\mathrm{u}}$ and $\alpha + \beta \in \Phi$, then $\alpha + \beta \in S^{\mathrm{u}}$. *Levi decomposition* asserts that the group $E(S, R)$ decomposes into semidirect product $E(S, R) = E(S^{\mathrm{r}}, R) \rightthreetimes E(S^{\mathrm{u}}, R)$ of its *Levi subgroup* $E(S^{\mathrm{r}}, R)$ and its *unipotent radical* $E(S^{\mathrm{u}}, R)$.

Especially important is the case of elementary subgroups corresponding to the maximal parabolic subschemes. Denote by $m_k(\alpha)$ the coefficient of α_k in the expansion of α with respect to the fundamental roots:

$$\alpha = \sum_{k=1}^{l} m_k(\alpha)\alpha_k.$$

Now, fix an $r = 1, \dots, l$—in fact, in the reduction to smaller rank it suffices to employ only terminal parabolic subgroups, even only the ones corresponding to the first and the last fundamental roots, $r = 1, l$. Denote by

$$S = S_r = \{\alpha \in \Phi : m_r(\alpha) \geqslant 0\}$$

the r-th standard parabolic subset in Φ. As usual, the reductive part $\Delta = \Delta_r$ and the special part $\Sigma = \Sigma_r$ of the set $S = S_r$ are defined as

$$\Delta = \{\alpha \in \Phi : m_r(\alpha) = 0\}, \quad \Sigma = \{\alpha \in \Phi : m_r(\alpha) > 0\}.$$

The opposite parabolic subset and its special part are defined similarly

$$S^- = S_r^- = \{\alpha \in \Phi : m_r(\alpha) \leqslant 0\}, \quad \Sigma^- = \{\alpha \in \Phi : m_r(\alpha) < 0\}.$$

Obviously, the reductive part S_r^- equals Δ.

Denote by P_r the *elementary* [maximal] parabolic subgroup of the elementary group $E(\Phi, R)$. By definition,

$$P_r = E(S_r, R) = \langle x_\alpha(\xi), \ \alpha \in S_r, \ \xi \in R \rangle.$$

Now Levi decomposition asserts that the group P_r can be represented as the semidirect product

$$P_r = L_r \leftthreetimes U_r = E(\Delta, R) \leftthreetimes E(\Sigma, R)$$

of the elementary Levi subgroup $L_r = E(\Delta, R)$ and the unipotent radical $U_r = E(\Sigma, R)$. Recall that

$$L_r = E(\Delta, R) = \langle x_\alpha(\xi),\ \alpha \in \Delta,\ \xi \in R \rangle,$$

whereas

$$U_r = E(\Sigma, R) = \langle x_\alpha(\xi),\ \alpha \in \Sigma,\ \xi \in R \rangle.$$

A similar decomposition holds for the opposite parabolic subgroup P_r^-, whereby the Levi subgroup is the same as for P_r, but the unipotent radical U_r is replaced by the opposite unipotent radical $U_r^- = E(-\Sigma, R)$.

As a matter of fact, we use Levi decomposition in the following form. It will be convenient to slightly change the notation and write $U(\Sigma, R) = E(\Sigma, R)$ and $U^-(\Sigma, R) = E(-\Sigma, R)$.

Lemma 2.4 *The group $\langle U^\sigma(\Delta, R), U^\rho(\Sigma, R) \rangle$, where $\sigma, \rho = \pm 1$, is the semidirect product of its normal subgroup $U^\rho(\Sigma, R)$ and the complementary subgroup $U^\sigma(\Delta, R)$.*

In other words, it is asserted here that the subgroups $U^\pm(\Delta, R)$ normalise each of the groups $U^\pm(\Sigma, R)$, so that, in particular, one has the following four equalities for products:

$$U^\pm(\Delta, R) U^\pm(\Sigma, R) = U^\pm(\Sigma, R) U^\pm(\Delta, R),$$

and, furthermore, the following four obvious equalities for intersections hold:

$$U^\pm(\Delta, R) \cap U^\pm(\Sigma, R) = 1.$$

In particular, one has the following decompositions:

$$U(\Phi, R) = U(\Delta, R) \leftthreetimes U(\Sigma, R), \quad U^-(\Phi, R) = U^-(\Delta, R) \leftthreetimes U^-(\Sigma, R).$$

3 Bounded generation. State of art

To put the results of the present paper in context, here we briefly recall what is known concerning the finite elementary width and the finite commutator width of Chevalley groups over rings. This will give us an occasion to explain some basic ideas behind our proof.

3.1 Length and width

Let G be a group and X be a set of its generators. Usually one considers symmetric sets, for which $X^{-1} = X$.

- The *length* $l_X(g)$ of an element $g \in G$ with respect to X is the minimal k such that g can be expressed as the product $g = x_1 \dots x_k$, $x_i \in X$.
- The *width* $w_X(G)$ of G with respect to X is the supremum of $l_X(g)$ over all $g \in G$.

We say that a group G has *bounded generation* with respect to X if the width $w_X(G)$ is finite.[4] In the case when $w_X(G) = \infty$, one says that G does not have bounded word length with respect to X.

The problem of calculating or estimating $w_X(G)$ has attracted a lot of attention, especially when G is one of the classical-like groups over skew-fields.

There are *hundreds* of papers which address this problem in the case when G is a classical group such as $\mathrm{SL}(n, R)$ or $\mathrm{Sp}(2l, R)$ or its large subgroup, whereas X is a natural set of its generators.

- Classically, over fields and other small-dimensional rings one would think of elementary transvections, all transvections, or Eichler–Siegel–Dickson (ESD)-transvections, reflections, pseudo-reflections, or other small-dimensional transformations.
- Other common choice would be a class of matrices determined by their eigenvalues such as the set of all involutions, a non-central conjugacy class, or the set of all commutators.
- More exotic choices include matrices distinct from the identity matrix in one column, symmetric matrices, etc.

In many classical cases exact values or at least sharp estimates of $w_X(G)$ are available. Sometimes there are even more precise results, explicitly calculating the length of *individual* elements in terms of certain geometric invariants such as, e.g., the dimension of its residual space, or the like.

More generally, oftentimes one considers any subset $X \subseteq G$ and looks at the width $w_X(\langle X \rangle)$. For instance, one calls the width of the commutator subgroup $[G, G]$ with respect to the set of all commutators the *commutator width* of G itself, regardless of whether the group G is perfect. This is a prototypical example of what is called the *word length* problems, when one tries to calculate or estimate the width of the verbal subgroup of G with respect to a word w with respect to the set of values of w in G.

[4] In the literature, expressions *bounded generation* and *finite width* are used in several related but significantly different contexts. Oftentimes one calls a group G boundedly generated if it has bounded generation with respect to the powers of some *finite* generating set. This amounts to the group being a finite product of several *cyclic* subgroups. In many situations it is equally meaningful to consider groups which are finite products of *abelian* subgroups. Finally, one calls families G_i of finitely presented groups boundedly generated if they can be presented in such a way that the sums of the number of generators and relations of G_i are uniformly bounded.

3.2 Elementary width and commutator width

In the present paper we focus on the much less studied case, where $G = G(\Phi, R)$ is a Chevalley group or its elementary subgroup $E(\Phi, R)$ over a commutative ring R, and on the closely related case of Kac–Moody groups. In this setting exact calculations of $w_X(G)$ with respect to most of the generating sets are usually beyond reach.

In the present paper we are primarily interested in the two following candidates for the generating set X for $E(\Phi, R)$:

- The set of *elementary* root unipotents

$$\Omega = \{x_\alpha(\xi) \mid \alpha \in \Phi, \xi \in R\}$$

  relative to the choice of a split maximal torus T.
- The set of commutators

$$C = \Big\{[x, y] = xyx^{-1}y^{-1} \mid x \in G(\Phi, R),\ y \in E(\Phi, R)\Big\}.$$

It is a classical theorem due to Suslin, Kopeiko and Taddei that for $\mathrm{rk}(R) \geqslant 2$ one indeed has $C \subseteq E(\Phi, R)$.

The width $w_\Omega(E(\Phi, R))$ is usually denoted $w_{\mathrm{E}}(G(\Phi, R))$ and is called the *elementary width* of $G(\Phi, R)$. Clearly, $w_{\mathrm{E}}(G(\Phi, R))$ is the smallest L such that $E(\Phi, R) = E^L(\Phi, R)$.

Similarly, the width $w_{\mathrm{C}}(E(\Phi, R))$ is oftentimes called the *commutator width* of $G(\Phi, R)$ itself.

Remark 3.1 Notice the subtleties related to the necessity to distinguish the Chevalley group $G(\Phi, R)$ itself, its commutator, the elementary subgroup $E(\Phi, R)$, etc. In the arithmetic situation they usually all coincide in ranks $\geqslant 2$, even in the relative case, this is precisely the [almost] positive solution of the congruence subgroup problem. But for the group $\mathrm{SL}(2, R)$ (and occasionally for some groups of rank 2) one will have to impose additional restrictions.

Anyway, in the arithmetic case for simply connected groups of rank at least two it follows from [46] that $G_{\mathrm{sc}}(\Phi, R) = E_{\mathrm{sc}}(\Phi, R)$. This means that the above set C equals the set of all commutators in $G_{\mathrm{sc}}(\Phi, R)$. That said, one sees that Theorem B is indeed equivalent to Theorem A.

One has to mention the exceptional cases where G is not perfect. For groups of rank at least 2, this happens if and only if R has $\mathbb{F}_2$ among its residue fields and Φ is of type C_2 or G_2. In the case where R is a field, this was first noticed by Robert Steinberg [74]. For more general rings, this was proved by Michael Stein [71, Corollary 4.4]. Note that no additional exceptions arise even for reductive groups, see [44].

Inside the proofs we have to consider some other related generating sets, such as, for instance:

- the set of all unitriangular elements

$$U(\Phi, R) \cup U^-(\Phi, R); \quad \text{or}$$

- the set of all root unipotents

$$\Omega^G = \left\{ x_\alpha(\xi)^g \mid \alpha \in \Phi,\ \xi \in R,\ g \in G(\Phi, R) \right\},$$

which are better behaved with respect to reduction to smaller ranks.

3.3 The case of 0-dimensional rings

Finiteness of the elementary width is a very rare and extremely significant phenomenon which has repercussions everywhere in the structure theory of the group. It is obvious, and classically known that Chevalley groups over fields and semi-local rings have finite elementary width. In fact, the groups over 0-dimensional rings rejoice short factorisations such as Bruhat decomposition or Gauß decomposition. Such factorisations are essentially tantamount to bounded elementary generation with very sharp bounds.

In fact, Bruhat decomposition immediately implies that over a field the elementary width of $G(\Phi, K)$ does not exceed $2N + 4l$ (here and below $N = |\Phi^+|$, $l = \mathrm{rk}(\Phi)$). It immediately follows that the commutator width of $G(\Phi, K)$ is also finite.

But determining the precise value of the commutator width turned out to be a very challenging problem—for finite fields it was the famous Ore conjecture. Without trying to follow the whole tortuous path, we just mention the two definitive contributions. For fields containing $\geqslant 8$ elements Erich Ellers and Nikolai Gordeev [EG] using Gauß decomposition with prescribed semi-simple part have proven that $w_C(G_{\mathrm{ad}}(\Phi, R)) = 1$, while $w_C(G_{\mathrm{sc}}(\Phi, R)) \leqslant 2$. This was then extended to the groups over small fields $\mathbb{F}_q$, $q = 2, 3, 4, 5, 7$, by Martin Liebeck, Eamonn O'Brien, Aner Shalev, and Pham Huu Tiep [40, 41], using explicit information about their maximal subgroups and very delicate character estimates.

Similarly, Gauß decomposition which holds over arbitrary *semi-local* rings—or even over rings of stable rank 1, see [68][5] in particular—implies that the elementary width of $G(\Phi, R)$ does not exceed $3N + 4l$. Actually, [89] gives another estimate in terms of unitriangular decomposition, $4N$, which is usually better for groups of very small ranks, say, up to 4 or 5. What seems to not have been noted in the literature, is that the LUP-decomposition of Chevalley groups over *local* rings provides the same upper bound on their width as for fields, $2N + 4l$.

As above, bounded elementary generation implies finite commutator width. However, providing sharp bounds for this width turned out to be a difficult problem. Skipping a detailed description of the early work by Keith Dennis, Leonid Vaserstein, You Hong, and others, pertaining to the classical groups [3, 23, 24, 86], we just mention a recent paper by Andrei Smolensky [67], where such an estimate is obtained for all Chevalley groups. The commutator width $w_C(E(\Phi, R))$ does not exceed 3 for $\Phi = \mathrm{A}_l$ and F_4, does not exceed 4 for all other types, except maybe for E_6, and does not exceed 5 for $G(\mathrm{E}_6, R)$. [We strongly believe that the commutator width does not

[5] In the literature, three of four completely different commodities are merchandised under the common name of Gauß decomposition: (1) the big cell decomposition LU, as in the affine group schemes textbooks, (2) the Birkhoff LPU-decomposition as in Ellers–Gordeev [25], (3) the LUP-decomposition, as in the computational linear algebra textbooks. Here we speak of (4) the DULU-decomposition, consult [68] for the historical background.

exceed 4 also for E_6, but we were discouraged by the extent of calculations necessary to improve the bound in this remaining case.]

Note that so far there are no examples of matrices from $\mathrm{SL}(n, \mathbb{Z})$, $\mathrm{SL}(n, \mathbb{F}_q[t])$, $\mathrm{SL}(n, \mathbb{F}_q[t, t^{-1}])$ ($n \geqslant 3$), not representable as a single commutator.

3.4 Counter-examples

The groups of rank 1 only occasionally can have finite elementary width, or finite commutator width, for that matter. Over a Euclidean ring R elementary expressions in $\mathrm{SL}(2, R)$ correspond to continued fractions.

In fact, the existence of arbitrarily long division chains in $\mathbb{Z}$ and in $K[t]$ implies that the groups $\mathrm{SL}(2, \mathbb{Z})$ and $\mathrm{SL}(2, \mathbb{F}_q[t])$ cannot be boundedly generated. The most classical example are the Fibonacci matrices

$$\begin{pmatrix} F_{m+1} & F_m \\ F_m & F_{m-1} \end{pmatrix}$$

which for even m require precisely m elementary factors.

Remark 3.2 For an odd m a similar matrix looks as

$$\begin{pmatrix} F_m & F_{m+1} \\ F_{m-1} & F_m \end{pmatrix},$$

which strongly suggests that while considering the width problems in $\mathrm{GL}(2, R)$ it might be more expedient to switch to Cohn's generators

$$\begin{pmatrix} x & 1 \\ 1 & 0 \end{pmatrix}.$$

Of course, the same holds for $\mathrm{SL}(2, \mathbb{F}_q[t])$, where instead of two consecutive Fibonacci numbers one should take two sufficiently generic polynomials of two consecutive degrees m and $m - 1$, placing the one of the higher degree into the NW or NE corner, depending on the parity of m. Many such similar examples were constructed by Paul Cohn [16] and others starting with the mid-1960s.

What came as a shock, though, was that the elementary width of rank $\geqslant 2$ groups over a Euclidean ring can be infinite too. Indeed, using methods of higher algebraic K-theory Wilberd van der Kallen [83] has proven that $\mathrm{SL}(3, \mathbb{C}[t])$ has infinite elementary width. Later Igor Erovenko came up with a somewhat more elementary proof [26]. On the other hand, soon thereafter Dennis and Vaserstein [24] noticed that $\mathrm{SL}(3, \mathbb{C}[t])$ does not even have finite commutator width.

3.5 Dedekind rings of arithmetic type, groups of rank $\geqslant 2$

For rings of dimension $\geqslant 2$ one cannot in general expect bounded generation. An extremely interesting borderline case are 1-dimensional rings, especially the classical example of the Dedekind rings of arithmetic type. Below, K is a global field, i.e. a finite extension of $\mathbb{Q}$ in characteristic 0, or a finite extension of $\mathbb{F}_q(t)$, $q = p^m$, in positive characteristic p. Further, S is a finite set of valuations of K, containing all Archimedean ones in the number case, and $R = \mathcal{O}_S$.

The *number* case is well understood. The initial breakthrough was due to David Carter and Gordon Keller who have proven that $\mathrm{SL}(n, R)$, $n \geqslant 3$, is boundedly elementary generated and gave explicit bounds on in terms of n and the discriminant[6] of K, see [9]. The proof in this paper is essentially an *effectivisation* of the usual verification of the familiar properties of Mennicke symbols.

Actually, their published proof is based on explicit rank reduction in terms of the stable rank, see below. It remains to verify bounded generation of $\mathrm{SL}(3, R)$. One of the key calculations in that paper, Lemma 1, can be described as follows. Let $A \in \mathrm{SL}(2, R)$ be a matrix with the first row (a, b). Then A^m can be transformed to a matrix in $\mathrm{SL}(2, R)$ with the first row (a^m, b) by a sequence of not more than 16 elementary transformations in $\mathrm{SL}(3, R)$—sic!

However, Carter and Keller mention that their original approach was based on model theory. To elucidate the connection, recall that van der Kallen [83] observed that the obstruction to bounded elementary generation of the group $E(\Phi, R)$ is the quotient $E(\Phi, R)^{\infty}/E(\Phi, R^{\infty})$ (countably many copies). This establishes connection with ultraproducts and non-standard models. Namely, it can be interpreted as the equivalence of the bounded generation of $E(\Phi, R)$ and the [almost] positive solution of the congruence subgroup problem for $G(\Phi, {}^*R)$ for non-standard models ${}^*R$ of R.

Carter and Keller came up with such a proof for the group $\mathrm{SL}(n, R)$, initially for $n \geqslant 3$, see [11]. Dave Witte Morris [49] gave an exposition of this proof in a somewhat more traditional logical language (first-order properties, compactness theorem, etc.). Unfortunately, this proof is not much easier than a direct algebraic proof[7] and it gives no bound whatsoever on the elementary width.

In [10] Carter and Keller have given a separate elementary proof specifically for the [easier] case of $\mathrm{SL}(n, \mathbb{Z})$, $n \geqslant 3$, in terms of direct matrix manipulations, mimicking the verification of the properties of Mennicke symbols (but without explicitly mentioning the work of Mennicke and/or of Bass–Milnor–Serre [7]). In particular, they have proven that the elementary width of $\mathrm{SL}(3, \mathbb{Z})$ does not exceed 48,[8] later this bound was reduced by Nica [51] to 37.

Soon thereafter, Oleg Tavgen invented a different, purely elementary approach to rank reduction, which allowed him to reduce the proof of bounded generation for all Chevalley groups to groups of rank 2. After that he succeeded in settling the cases of rank 2 groups, $\mathrm{Sp}(4, R)$ and the Chevalley group of type G_2 (and, in fact, also twisted Chevalley groups of rank 2) by direct matrix calculations. These important advances

[6] Or, actually, the number of its prime divisors. Later, Loukanidis and Murty [43, 50] obtained bounds that depended on n and the degree $|K : \mathbb{Q}|$ of K, not the discriminant.

[7] Well, explicit use of infinitesimals does make life easier, sometimes. Say, in R itself there are no non-obvious ideals such that $I^2 = I$, whereas in ${}^*R$ there is such an ideal $\mathcal{I}$ consisting of all infinitesimal elements, which can be very handy. But these simplifications are mostly relevant in the [difficult] case $n = 2$, see the next subsection.

[8] The proof from [10] with several successful deteriorations, without reference to [10], and with a worse bound 73 was subsequently reproduced in [2].

sum up to his main result, the bounded elementary generation of Chevalley groups of rank $\geqslant 2$ over arithmetic Dedekind rings in the *number* case.

3.6 Dedekind rings of arithmetic type, groups of rank 1

There is a critical difference in behaviour of $\mathrm{SL}(2, R)$, depending on whether $|S| = 1$, in which case R^* is finite, and $|S| \geqslant 2$, when R^* is infinite. As we know, for the case $|S| = 1$ the answer to the question on bounded elementary generation is negative, so in the rest of the subsection we assume that R^* is infinite.

Again in the *number* case the situation is well understood. Elementary generation of $\mathrm{SL}(2, R)$ is closely related to generalisations of Euclidean algorithm. Important early inroads in this direction were suggested [apparently independently!] by Timothy O'Meara [54], who simultaneously considered the number case and the *function* case, and by Paul Cohn [16], who proposed vast [non-commutative] generalisations.

About a decade later, George Cooke and Peter Weinberger [19] systematically studied the length of division chains [17, 18] in the number case. For the case, where R^* is infinite, their main results implied that modulo some form of the Generalised Riemann Hypothesis (GRH), any matrix in $\mathrm{SL}(2, R)$ is a product of $\leqslant 9$ elementary transvections.

The results of Hendrik Lenstra on the Generalised Artin Conjecture [39]—again conditional on GRH—imply that whenever S contains at least one real valuation, the bound here can be reduced to $\leqslant 7$. Observe that the best possible bound here is[9] $\leqslant 5$. However, Cooke and Weinberger proposed an example of a matrix over a *totally imaginary* ring R of degree 4 which cannot be expressed as a product of less than six elementary matrices.

It has taken quite some time to get rid of the dependence on the GRH and to improve bounds here. Modulo the GRH, Bruce Jordan and Yevgeny Zaytman [35] have slightly remodelled the Cooke–Weinberger argument and improved the bound to five elementary transvections if K is *not* totally imaginary, to six elementary transvections when S contains at least one non-Archimedean place, and to seven elementary transvections for the integers of a totally imaginary field.

One of the first unconditional results was obtained by Bernhard Liehl [42], but he imposed some additional restrictions on the number field K, and his proof does not give good bounds. Almost simultaneously Carter and Keller, jointly with Eugene Paige, came up with the first general *logical* proof [12], somewhat refashioned in [49]. But, as we already mentioned, this proof gives no bounds whatsoever. About a decade later Loukanidis and Murty [43, 50] proposed an unconditional *analytic* argument, but it only works provided S is sufficiently large, say $|S| \geqslant \max(5, 2|K : \mathbb{Q}| - 3)$.

Some 10 years ago Maxim Vsemirnov and Sury [90] considered the key example of $\mathrm{SL}(2, \mathbb{Z}[1/p])$, obtaining the bound $\leqslant 5$ *unconditionally*. This was a key inroad to the first complete unconditional solution of the general case with a good bound, in the work of Alexander Morgan, Andrei Rapinchuk and Sury [48]. The bound they gave is $\leqslant 9$, but for the case when S contains at least one real or non-Archimedean

[9] See [89], where it is [essentially] proven for $R = \mathbb{Z}[1/p]$, again modulo GRH.

valuation[10] it was almost immediately improved [with the same ideas] to $\leqslant 8$ by Jordan and Zaytman [35].

3.7 Reduction to smaller ranks

Let us explain, what do the width bounds obtained for ranks 1 or 2 imply for higher ranks.

There are two basic ways to reduce the problem of bounded generation for a Chevalley groups to similar problems for groups of smaller ranks. We will start with Tavgen's reduction theorem, which came later historically, but is both more elementary and more general, than the reduction based on stability conditions. On the other hand, explicit factorisations resulting from stability conditions are not always available, but when they are, they give sharper bounds.

To present Tavgen's idea in its simplest form, let us consider not the width in elementary generators, but a coarser problem of determining the width of $G(\Phi, R)$ in terms of the elements belonging to the unipotent subgroups U and U^-. As far as we know, this problem was first systematically considered by Dennis and Vaserstein in the context of the closely related problem of estimating the commutator width for $\mathrm{SL}(n, R)$, see [23, 24]. In other words, we are interested in finding the smallest m such that

$$G(\Phi, R) = UU^-UU^-\dots U^\pm, \quad m \text{ factors},$$

where the last factor equals U or U^- depending on whether m is odd or even.

Essentially, Tavgen observed that if there are root subsystems $\Psi_1, \dots, \Psi_t \subseteq \Phi$ which contain all fundamental roots, and such that each of the Chevalley groups $G(\Psi_1, R), \dots, G(\Psi_t, R)$ admits a similar decomposition with m factors, then $G(\Phi, R)$ itself admits such a decomposition with m factors. In this [and in fact slightly more general] form this reduction is described in [68, 89]. Modulo the Levi decomposition of parabolic subgroups and the Chevalley commutator formula it is undergraduate group theory, see the next section for precise statements, somewhat broader discussion, and a proof.

Since every element of U is a product of not more than $N = |\Phi^+|$ elementary generators, Tavgen's theorem suffices to give plausible bounds for the elementary width of large rank groups in terms of the elementary widths of their rank 1 or rank 2 subgroups. However, these bounds tend to be somewhat exaggerated.

Actually, for small dimensional rings there is a more precise form of reduction in terms of the stability conditions. For $\mathrm{GL}(n, R)$ such a reduction in terms of the usual stable rank $\mathrm{sr}(R)$ was first proposed by Hyman Bass in 1964, and then improved by Vaserstein, Dennis, Kolster, and others. Namely, for $n \geqslant \mathrm{sr}(R)$ the usual proof of the surjective stability for SK_1 grants the following decomposition:

$$\mathrm{SL}(n+1, R) = \mathrm{SL}(n, R)\, U_n U_n^- U_n U_n^-.$$

[10] Recall that our standing assumption $|S| \geqslant 2$ excludes the problematic case $R = \mathbb{Z}$.

It follows that if $\mathrm{SL}(n, R)$ has the elementary width $\leqslant s$, then $\mathrm{SL}(n+1, R)$ has the elementary width $\leqslant s + 4n$ — and in fact $\leqslant s + 3n + \mathrm{sr}(R)$, if you look inside the proof.

For *Dedekind rings* this bound was slightly improved by Carter and Keller [9], who noticed that one can do slightly better by observing that $\mathrm{sr}(R) \leqslant 1.5$. This means that for $n \geqslant 2$ one needs just one addition instead of two, to get a shorter unimodular row. This gives for the elementary width of $\mathrm{SL}(n, R)$ the estimate $s + \frac{3}{2}n^2 - \frac{1}{2}n - 5$, where s is the elementary width of $\mathrm{SL}(2, R)$.

Surjective stability of K_1 in terms of various stability conditions—the usual stable rank $\mathrm{sr}(R)$, the absolute stable rank $\mathrm{asr}(R)$, or the like—is known for all relevant embeddings of other Chevalley groups. For the usual stability embeddings of classical groups of the same type, it is indeed classical, starting with the work of Anthony Bak and Leonid Vaserstein. For cross-type and exceptional emdeddings such similar results were established by Michael Stein and Eugene Plotkin, see in particular [57, 58, 72, 73]. However, at least in the exceptional cases it was not stated in the form of such precise decompositions as above.

As a result, the explicit bounds for other groups—let alone their improvements for Dedekind rings—were never mentioned in the available literature. Even in the number case Tavgen only states finiteness, without providing any specific bound. In Sect. 5 below, as part of the proof of Theorem A, we return to this problem, and procure such explicit bounds.

Let us mention yet another extremely pregnant generalisation, *bounded reduction*. In fact, even below the usual stability conditions and even in the absence of the bounded generation for $G(\Psi, R)$, it makes sense to speak of the number of elementary generators necessary to reduce an element g of $G(\Phi, R)$ to an element of $G(\Psi, R)$, for a subsystem $\Psi \subseteq \Phi$.

One such prominent example are polynomial rings $R[t_1, \dots, R_m]$, where bounded reduction holds starting with a rank depending on R alone, not on the number of indeterminates. For the case of $\mathrm{SL}(n, R[t_1, \dots, R_m])$ this is essentially an effectivisation of Suslin's solution of the K_1-analogue of Serre's problem, explicit bounds were obtained in the remarkable paper by Leonid Vaserstein [84], which unfortunately remained unpublished. For other split classical groups such bounds were recently obtained by Pavel Gvozdevsky [32].

3.8 The function case

In the function case, until now much less was known concerning the bounded generation of Chevalley groups. On the one hand, an analogue of Riemann's Hypothesis was known in this case for quite some time. On the other hand, in the positive characteristic additional arithmetic difficulties occur, that have no obvious counterparts in characteristic 0. They reflect in particular in the structure of arithmetic subgroups in the function case. For instance, it is well known that the group $\mathrm{SL}(2, K[t])$ is not even finitely generated, whereas the groups $\mathrm{SL}(2, K[t, t^{-1}])$ and $\mathrm{SL}(3, K[t])$ are finitely generated but not finitely presented.

Until very recently the only published result was that by Clifford Queen [59]. We discuss this and related work in much more detail in Sect. 8. Queen's main result implies that under some additional assumptions on R—which hold, for instance, for Laurent polynomial rings $\mathbb{F}_q[t, t^{-1}]$ with coefficients in a finite field—the elementary width of the group $\mathrm{SL}(2, R)$ is $\leqslant 5$. As we know, this implies, in particular, bounded elementary generation of all Chevalley groups $G(\Phi, R)$.

The case of the groups over the usual polynomial ring $\mathbb{F}_q[t]$ long remained open. Only in 2018 has Bogdan Nica established the bounded elementary generation of $\mathrm{SL}(n, \mathbb{F}_q[t])$, $n \geqslant 3$. Part of the problem is that in characteristic $p > 0$ bounded elementary generation is not the same as bounded generation in terms of cyclic subgroups. For instance, the groups $\mathrm{SL}(n, \mathbb{F}_q[t])$ do not have bounded generation in this abstract sense, see [1].

This is exactly where we jump in. As already stated in the introduction, in the present paper we prove bounded *elementary* generation for all Chevalley groups of rank $\geqslant 2$ over the usual polynomial rings $\mathbb{F}_q[t]$ and—with better bounds—for Chevalley groups of rank $\geqslant 1$ over a class of function rings with infinite multiplicative group, including the Laurent polynomial rings $\mathbb{F}_q[t, t^{-1}]$.

3.9 Further prospects

The historical description is already rather long, we cannot mention many further aspects. A systematic survey should include at least:

- Partial positive results, such as bounded expressions of elementary conjugates and commutators in terms of elementary generators—decomposition of unipotents, Stepanov's universal localisation, and the like.
- Connection with the prestability kernel, bounded generation of SL_2 in terms of Vaserstein prestability generators, [85], etc.
- Connection of the bounded generation with the congruence subgroup problem, Kazhdan's property (T), finite presentation, super-rigidity, etc.
- Implications for the bounded generation by cyclic/abelian subgroups, including actions, etc.
- Extension of known bounds for word width (such as in [6]) to the function case.

We intend to return to [some of] these subjects in an expected sequel to the present paper.

4 Outline of the proof of Theorem A and reduction to rank 2

In this section we sketch the main ideas of the proof and implement the rank reduction. Together with the result by Nica [51], this already suffices to establish Theorem A for simply laced types and type F_4.

4.1 Outline of the proof

The proofs of bounded generation for the rings of integers of an algebraic number field, see [2, 9, 10, 78], deploy similar ideas. Let

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

be a matrix from $\mathrm{SL}(2, R)$ nested either in $G = \mathrm{SL}(3, R)$ or in $G = \mathrm{Sp}(4, R)$. Observe that in the second case there are two natural embeddings of $\mathrm{SL}(2, R)$, on short roots and on long roots, and that is a major aspect of the quest. We also provide an approach based on the reduction to Chevalley groups of rank 3. This approach has some advantages and makes use of embeddings of the Chevalley group of type $G(\mathrm{A}_2, R)$ into either $G(\mathrm{C}_3, R)$ or $G(\mathrm{B}_3, R)$. The Chevalley groups of type G_2 are to be treated separately anyway, but they do not occur in the analysis of higher rank cases.

The goal is to reduce A to the identity matrix by elementary transformations in G in such a way that the number of elementary factors does not depend on A. The guideline of the proof can be summarised as follows:

- Eventually, one has to transform A to a matrix with an invertible entry by a bounded number of elementary transformations.
- One way to do that is to use a version of Little Fermat's Theorem. So we need some entry of A in an appropriate power.
- Hence, we need to produce an elementary descendant B of A with some entry, say the first one, to be a^k, where k is an appropriate power. This is achieved by Lemmas 1 in [9, 10], [2, Lemma 2], [78, Proposition 3].
- The proof follows from the miraculous fact that the matrix

  $$A^k = \begin{pmatrix} a & b \\ c & d \end{pmatrix}^k$$

  coincides modulo elementary matrices with the matrix

  $$B = \begin{pmatrix} a^k & b \\ * & * \end{pmatrix}.$$

- This miracle is none other than the multiplicative property of Mennicke symbols, so this is not a surprise at all modulo a tricky proof of this property (see [45, 47], etc).
- It remains to use a combination of analytic tools such as Dirichlet's theorem on primes in arithmetic progressions and, if needed, reciprocity laws to obtain by elementary transformations a matrix of the form

  $$B = \begin{pmatrix} a^k & p \\ q & * \end{pmatrix}$$

where the pair (a^k, p) satisfies $a^k - 1 = ps$ for some s.

Note that Nica [51] modified the proof using the so-called "swindling lemma". We shall discuss this trick in more detail in Sect. 6. Actually, "swindling" is merely a weaker version of the multiplicativity of Mennicke symbols. However, the advantage is that this weaker form is cheaper in terms of the number of elementary moves, and here we generalise this approach to the symplectic case as well.

Remark 4.1 One of the points of the present work is that, unlike the proofs based on model theory, here we get *efficient* realistic estimates for the number of elementary factors, with bounds that depend on Φ alone. In some cases, like for the bounded reduction to smaller rank, our bounds are [very close to] the best possible ones. For small ranks, there might be still some gap between the counter-examples and the estimates we obtain, but our upper bounds are still reasonably close to the theoretically best possible ones. The lower bounds in such similar problems are usually quite difficult to obtain, anyway.

4.2 Tavgen's reduction theorem

Here we reproduce with minor variations the *elementary* reduction procedure due to Tavgen, in the form mentioned in [68, 89]. This procedure suffices to reduce Theorem A for groups of rank $\geqslant 3$ to the groups $\mathrm{SL}(3, R)$ and $\mathrm{Sp}(4, R)$. It of course works also for reduction to $\mathrm{Sp}(6, R)$ and $\mathrm{SO}(7, R)$ used in Sect. 7. Moreover, the bounds it gives are quite reasonable, though clearly not the best possible ones. In Sect. 5 we work out the *stable* reduction, based on the fact the stable rank of Dedekind rings equals 1.5. This approach gives much better bounds for reduction, sometimes the sharp ones, and for exceptional groups it is new even in the number case.

Tavgen's approach works more smoothly for *unitriangular factorisations*, in other words, for expressions of elementary subgroup $E(\Phi, R)$ as a product of subgroups $U(\Phi, R)$ and $U^-(\Phi, R)$,

$$E(\Phi, R) = U(\Phi, R)U^-(\Phi, R) \cdots U^{\pm}(\Phi, R).$$

Later on in [68] it was applied to *triangular factorisations*, where also the toral factor is admitted.[11]

The leading idea of Tavgen's proof is very general and beautiful, and works in many other related situations. It relies on the fact that for systems of rank $\geqslant 2$ every fundamental root falls into the subsystem of smaller rank obtained by dropping either the first or the last fundamental root. However, as was pointed out by the referee of [68], the argument applies without any modification in a much more general setting. Namely, it suffices to assume that the required decomposition holds for *some* subsystems $\Delta = \Delta_1, \ldots, \Delta_t$, whose union contains all fundamental roots of Φ. These subsystems do not have to be terminal, or even irreducible, for that matter!

[11] As we know from Sect. 4, this does not influence boundedness or lack thereof, but may affect the actual bounds.

Theorem 4.2 *Let Φ be a reduced irreducible root system of rank $l \geqslant 2$, and R be a commutative ring. Further, let $\Delta = \Delta_1, \dots, \Delta_t$ be some subsystems of Φ, whose union contains all fundamental roots of Φ. Suppose that for all $\Delta = \Delta_1, \dots, \Delta_l$, the elementary Chevalley group $E(\Delta, R)$ admits a unitriangular factorisation*

$$E(\Delta, R) = U(\Delta, R)U^-(\Delta, R) \cdots U^{\pm}(\Delta, R)$$

of length L. Then the elementary Chevalley group $E(\Phi, R)$ itself admits unitriangular factorisation

$$E(\Phi, R) = U(\Phi, R)U^-(\Phi, R) \cdots U^{\pm}(\Phi, R)$$

of the same length L.

Let us reproduce the details of the argument. By definition,

$$Y = U(\Phi, R)U^-(\Phi, R) \cdots U^{\pm}(\Phi, R)$$

is a *subset* in $E(\Phi, R)$. Usually, the easiest way to prove that a subset $Y \subseteq G$ coincides with the whole group G consists in the following

Lemma 4.3 *Assume that $Y \subseteq G$, $Y \neq \varnothing$, and let $X \subseteq G$ be a symmetric generating set. If $XY \subseteq Y$, then $Y = G$.*

Now, we are all set to finish the proof of Theorem 4.2.

Proof By Lemma 2.3, the group $E(\Phi, R)$ is generated by the fundamental root elements

$$X = \{x_\alpha(\xi) \mid \alpha \in \pm\Pi,\ \xi \in R\}.$$

Thus, by Lemma 4.3 is suffices to prove that $XY \subseteq Y$.

Fix a fundamental root unipotent $x_\alpha(\xi)$. Since $\mathrm{rk}(\Phi) \geqslant 2$, the root α belongs to at least one of the subsystems $\Delta = \Delta_r$, where $r = 1, \dots, t$. Set $\Sigma = \Sigma_r$ and express $U^{\pm}(\Phi, R)$ in the form

$$U(\Phi, R) = U(\Delta, R)U(\Sigma, R), \quad U^-(\Phi, R) = U^-(\Delta, R)U^-(\Sigma, R).$$

Using Lemma 2.4, we see that

$$Y = U(\Delta, R)U^-(\Delta, R) \cdots U^{\pm}(\Delta, R) \cdot U(\Sigma, R)U^-(\Sigma, R) \cdots U^{\pm}(\Sigma, R).$$

Since $\alpha \in \Delta$, one has $x_\alpha(\xi) \in E(\Delta, R)$, so that the inclusion $x_\alpha(\xi)Y \subseteq Y$ immediately follows from the assumption. □

4.3 Proof of Theorem A for simply laced systems and in the case of F_4

In [68] the authors commented that they do not see immediate applications of the more general form of Tavgen's reduction theorem, as stated above. Here, we notice that it is in fact *surprisingly* strong, since it allows one to pass from *some* smaller rank subsystems to the whole system, without looking at any other subsystems, including those of intermediate ranks! Indeed, it may happen that for those other subsystems bounded generation holds with some larger bound, or bluntly fails.

Of course, the easiest case is when the group $\mathrm{SL}(2, R)$ itself has bounded elementary generation.

Corollary 4.4 *Let any element of* $\mathrm{SL}(2, R)$ *be a product of* $\leqslant L$ *elementaries. Then any simply connected Chevalley group* $G = G(\Phi, R)$ *admits unitriangular factorisation*

$$G = UU^{-}U \cdots U^{\pm}$$

of length L*.*

However, this is very seldom the case, so one should start looking at larger rank subsystems. Recall that in the A_2 case Theorem A was proven by Nica [51]. His main new result can be stated as follows.

Proposition 4.5 *Any element of* $\mathrm{SL}(3, \mathbb{F}_q[t])$ *is a product of* $\leqslant 41$ *elementary transvections.*

We are not contending for the best possible bounds in terms of unitriangular matrices at this stage, since later we improve the bounds anyway. Interestingly, the main arithmetic ingredient of his proof is the Kornblum–Artin functional version of Dirichlet's theorem on primes in arithmetic progressions. In Sect. 6 below we shall see how it works in the parallel example of $\mathrm{Sp}(4, \mathbb{F}_q[t])$.

Now, together with Theorem 4.2 this result by Nica implies Theorem A for the two following cases:

- Chevalley groups of *simply laced* type Φ of rank $\geqslant 2$. Indeed, in this case Π is covered by the fundamental copies of A_2 spanned by all pairs of adjacent fundamental roots.
- Chevalley group of type F_4. Indeed, in this case Π is covered by *two* fundamental copies of A_2 — the long one A_2, spanned by the two fundamental *long* roots, and the short one $\widetilde{\mathrm{A}}_2$, spanned by the two fundamental *short* roots.

Observe that in the second case it is neither assumed, nor does it follow that the group $\mathrm{Sp}(4, R)$ is boundedly generated! Even more amazingly, the same applies to the subgroups of types B_3 and C_3.

However, for root systems of types B_l and C_l there are short/long roots that cannot be embedded into any irreducible rank 2 subsystem other than C_2. Thus, to be able to apply Theorem 4.2 we have to explicitly dismantle elements of $\mathrm{Sp}(4, \mathbb{F}_q[X])$ into elementary factors. This is exactly what is achieved in Sect. 6.

However, since we are interested in actual bounds, before treating this case, we have to recall an alternative approach to rank reduction, based on stability conditions.

In the next section we recall the stability conditions themselves and illustrate how they work for Chevalley groups of type G_2. Later, in Sects. 5 and 6 we produce similar arguments for groups of types C_2, C_3, and B_3.

5 Proof of Theorem A in the case of G_2

The purpose of this section is two-fold. As a first objective, here we provide the proof of Theorem A for the Chevalley group of type G_2. This is done by virtue of surjective stability for the embedding $\mathrm{A}_2 \longrightarrow \mathrm{G}_2$. Using this opportunity, we revisit stability for Dedekind rings also for other embeddings, and obtain accurate bounds for reduction in this case. For exceptional groups such explicit bounds are new even in the number case.

5.1 Stability conditions

Traditionally, stability results are stated in terms of stability conditions. The first such condition, stable rank, was introduced by Hyman Bass back in 1964. However, *surjective* stability results for K_1 for embeddings other than the simplest stability embeddings

$$\mathrm{SL}(n, R) \longrightarrow \mathrm{SL}(n+1, R) \quad \text{and} \quad \mathrm{Sp}(2l, R) \longrightarrow \mathrm{Sp}(2(l+1), R)$$

usually require *stronger* stability conditions, such as the *absolute* stable rank, etc.

Modulo some small additive constants, all these ranks are bounded by the Krull dimension $\dim(R)$ or even the Jacobson dimension $\dim(\mathrm{Max}(R))$ of the ring R. On the other hand, arithmetic rings, such as Dedekind rings and their kin, usually satisfy even stronger stability conditions than the ones that would follow from their dimension. Here we very briefly recall some of these conditions, limiting ourselves only to those that are actually used in the sequel.

A row $(a_1, \ldots, a_n) \in {}^nR$ is called *unimodular* if its components $a_1, \ldots, a_n$ generate R as a right ideal,

$$a_1 R + \cdots + a_n R = R,$$

or, what is the same, if there exist $b_1, \ldots, b_n \in R$ such that

$$a_1 b_1 + \cdots + a_n b_n = 1.$$

A row $(a_1, \ldots, a_{n+1}) \in {}^{n+1}R$ of length $n+1$ is called *stable* if there exist $b_1, \ldots, b_n \in R$ such that the ideal generated by

$$a_1 + a_{n+1} b_1,\ a_2 + a_{n+1} b_2,\ \ldots,\ a_n + a_{n+1} b_n$$

coincides with the ideal generated by $a_1, \ldots, a_{n+1}$.

The *stable rank* $\mathrm{sr}(R)$ of the ring R is the *smallest* n such that every unimodular row $(a_1, \dots, a_{n+1})$ of length $n+1$ is *stable*. In other words, there exist $b_1, \dots, b_n \in R$ such that the row

$$(a_1 + a_{n+1}b_1, a_2 + a_{n+1}b_2, \dots, a_n + a_{n+1}b_n)$$

of length n is unimodular. If no such n exists, one writes $\mathrm{sr}(R) = \infty$.

Bass himself denoted stability of unimodular rows of length $n+1$ by $\mathrm{SR}_{n+1}(R)$. It is easy to see that condition $\mathrm{SR}_m(R)$ implies condition $\mathrm{SR}_n(R)$ for all $n \geqslant m$, so that the stable rank is defined correctly: if $n > \mathrm{sr}(R)$, then every unimodular row of length n is stable. Clearly, this means that when $n > \mathrm{sr}(R) + 1$ one can iterate the process of shortening a unimodular row and eventually reduce any unimodular row to a unimodular row of length $\mathrm{sr}(R)$.

For representations other than the vector representations of SL_n and Sp_{2l}, the stock of available elementary transformations is limited, so that one has to work with pieces of unimodular rows, that are not themselves unimodular. However, stability of all non-unimodular rows is an *exceedingly* restrictive condition — though Dedekind rings satisfy precisely something of the sort!

The most familiar variation of stable rank, that works for other classical groups, is the absolute stable rank. For commutative rings this condition was introduced by David Estes and Jack Ohm [29], whereas Michael Stein [73] discovered its relevance in the study of orthogonal groups and exceptional groups.

For a row $(a_1, \dots, a_n) \in {}^nR$ let us denote by $J(a_1, \dots, a_n)$ the intersection of the maximal ideals of the ring R containing $a_1, \dots, a_n$. In particular, a row is unimodular if and only if $J(a_1, \dots, a_n) = R$.

One says that a commutative ring R satisfies condition ASR_{n+1} if for any row $(a_1, \dots, a_{n+1})$ of length $n+1$ there exist $b_1, \dots b_n \in R$ such that

$$J(a_1 + a_{n+1}b_1, \dots, a_n + a_{n+1}b_n) = J(a_1, \dots, a_{n+1}).$$

It is obvious that condition $\mathrm{ASR}_m(R)$ implies condition $\mathrm{ASR}_n(R)$ for all $n \geqslant m$. The *absolute stable rank* $\mathrm{asr}(R)$ of the ring R is the smallest natural n for which condition $\mathrm{ASR}_{n+1}(R)$ holds. Clearly, $\mathrm{sr}(R) \leqslant \mathrm{asr}(R)$.

The classical theorem of Estes and Ohm [29] asserts that for commutative rings one has

$$\mathrm{asr}(R) \leqslant \dim(\mathrm{Max}(R)) + 1,$$

a similar estimate for $\mathrm{sr}(R)$ follows from a classical theorem of Bass. Thus, in particular, any Dedekind ring satisfies $\mathrm{ASR}_3(R)$ — and, as we recall below, a much stronger condition.

5.2 Surjective stability for K_1 and bounded reduction.

Recall that the K_1-functor modelled on a Chevalley group $G(\Phi, R)$ is defined as

$$\mathrm{K}_1(\Phi, R) = G(\Phi, R)/E(\Phi, R).$$

For [irreducible] root systems of rank $\geqslant 2$ the elementary subgroup $E(\Phi, R)$ is a normal subgroup of $G(\Phi, R)$, so that in this case $\mathrm{K}_1(\Phi, R)$ is a group.

Now, by the homomorphism theorem every embedding of root systems $\Delta \subset \Phi$ gives rise to the stability map

$$\nu = \nu_{\Delta \to \Phi} \colon \mathrm{K}_1(\Phi, \Delta) \longrightarrow \mathrm{K}_1(\Phi, R),$$

and one of the archetypical classical problems of the algebraic K-theory, whose study was initiated by Hyman Bass in the early 1960s, is to find conditions under which this map is surjective or injective.

Clearly, *surjective stability* for the embedding $\Delta \subset \Phi$ amounts to the equality

$$G(\Phi, R) = G(\Delta, R)E(\Phi, R).$$

In other words, any matrix $g \in G(\Phi, R)$ can be expressed as a product of a matrix from $G(\Delta, R)$ and elementary unipotents.

However, *in the stable range*, that is when $\mathrm{rk}(\Delta)$ is large with respect to $\dim(R)$, one can use the above stability conditions and establish rather more. In this setup, all customary proofs of surjective stability afford not just elementary reduction to smaller rank, but *bounded* elementary reduction. In other words, they establish an equality of the type

$$G(\Phi, R) = G(\Delta, R)E^L(\Phi, R),$$

for some constant L depending on the dimension of the ring R and the embedding $\Delta \subset \Phi$. This means that we have *bounded reduction*: any matrix $g \in G(\Phi, R)$ can be expressed as a product of a matrix from $G(\Delta, R)$ and not more than L elementary unipotents, where L does not depend on g.

When Δ is the reductive part of a parabolic subset S of Φ, the actual value of L is estimated in terms of the order of the unipotent part Σ of S. Thus, as we have already mentioned in Sect. 4, for the embedding $\mathrm{A}_{n-1} \subset \mathrm{A}_n$ the original Bass's proof furnishes the following classical decomposition

$$\mathrm{SL}(n+1, R) = \mathrm{SL}(n, R)\, U_n U_n^- U_n U_n^-,$$

which implies that in this case L is at most $4n$. Actually, since one needs only $\mathrm{sr}(R)$ additions to shorten a unimodular row, this bound immediately reduces to $3n + \mathrm{sr}(R)$.

However, for all other embeddings, apart from $\mathrm{C}_{n-1} \subset \mathrm{C}_n$, and especially for exceptional groups and for root subsystems that are not reductive parts of parabolic subsets, it is not that immediate. Even in the classical cases, not to mention the exceptional ones, the exact number of elementary unipotents used in the reduction was not explicitly tracked.

Indeed, the existing proofs of surjective stability do not bother about explicit bounds. At the moment, one could invoke a previously known stability result with the same or weaker stability condition, one would do that, without actually reproducing the reduction procedure, or worrying for the shortest elementary expressions. For anyone familiar with the proofs of surjective stability in, say [31, 57, 58, 73], it is clear that they afford bounded reduction with *some* L. Note that these bounds are valid in the case of any base ring of Krull dimension 1 and hence for any Dedeking ring. But any such bounds are not explicit there, and one should go over all proofs in these papers once again even to produce *some* bounds (not the best possible ones!).

Additional features of the exceptional cases are that—with the sole exception of G_2—their minimal representations are too large for manual matrix computations, and even in these representations the elementary unipotents are significantly more complicated. Thus, instead of matrices one should use some tools from representation theory, as do [31, 57, 58, 73]. It would take quite a few pages to describe these tools, and adjust them to our needs. To establish Theorem A with some [reasonable] bound, we do not need that. Actually, we intend to return to this issue in the sequel to this paper, and come up with *sharp* bounds. In the next section we limit ourselves with the proof specifically for the long root embeddings $\mathrm{A}_1 \subset \mathrm{A}_2 \subset \mathrm{G}_2$.

5.3 Proof of Theorem A for G_2

In his pathbreaking paper [73] Michael Stein proves, in particular, that under the absolute stable range condition $\mathrm{ASR}_3(R)$ one has

$$G(\mathrm{G}_2, R) = G(\mathrm{A}_1, R)E(\mathrm{G}_2, R) = G(\mathrm{A}_2, R)E(\mathrm{G}_2, R),$$

[long root embeddings], this is his Theorem 4.1.m. Below, we go through the proof of that theorem, to come up with an actual bound.

Theorem 5.1 *Under the assumption* $\mathrm{ASR}_3(R)$ *one has*

$$G(\mathrm{G}_2, R) = G(\mathrm{A}_1, R)E^{24}(\mathrm{G}_2, R) = G(\mathrm{A}_2, R)E^{24}(\mathrm{G}_2, R).$$

Clearly, this result together with the main theorem of [51] immediately implies the claim of Theorem A for the case of G_2. Indeed, $\mathrm{SL}(3, R)$ is boundedly elementary generated, and since Dedekind rings have dimension $\leqslant 1$ and thus satisfy condition ASR_3, it follows from the above result that

$$w_{\mathrm{E}}(G(\mathrm{G}_2, R)) \leqslant w_{\mathrm{E}}(G(\mathrm{A}_2, R)) + 24.$$

Proof Our proof closely follows that in [73], pages 102–104, and we essentially preserve the notation thereof. Let α_1, α_2 be the fundamental roots of G_2, with α_2 long.

Further, consider the short roots

$$\alpha = -\alpha_1, \quad \beta = 2\alpha_1 + \alpha_2, \quad \gamma = -\alpha_1 - \alpha_2,$$

which clearly sum to zero, $\alpha + \beta + \gamma = 0$.

Consider the 7-dimensional short root representation of $G(\mathrm{G}_2, R)$, with the highest weight $\mu = \beta$, its weights are the short roots $\pm\alpha$, $\pm\beta$, $\pm\gamma$ and 0. Order the weights by height, $\mu = \beta, -\gamma, -\alpha, 0, \alpha, \gamma, -\beta$.

As usual, the entries of matrices $g \in G(\mathrm{G}_2, R)$ are indexed by pairs of weights, $g = (g_{\lambda,\mu})$, where $\lambda, \mu = \beta, \dots, -\beta$.

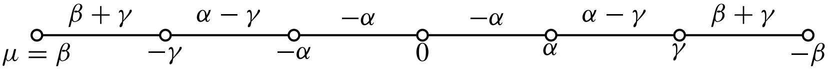
Initially, we concentrate on the first column $g_{*\beta}$ of this matrix, which is the image of the highest weight vector under the action of g. For typographical reasons, we denote this column by

$$(x_\beta, x_{-\gamma}, x_{-\alpha}, x_0, x_\alpha, x_\gamma, x_{-\beta}).$$

It is our intention to reduce this column to the form $(1, *, *, *, *, *, *)$ by elementary unipotents.

This can be done as follows. Not to proliferate indices in this *and further stability calculations*, we will not *rename* [as mathematicians would do], but *reset* [as is typical in programming] our variables g and x, still denoting them by the same letters after each successive transformation.

In order to make the action of elementary unipotents visible, below we present the weight diagram of the 7-dimensional short root representation of $G(\mathrm{G}_2, R)$:



As usual, the action of the elementary unipotent $x_\gamma(t)$ on the first column $g_{*\beta}$ can be viewed by looking for pairs of weights on the weight diagram connected by the root γ.

- Using condition $\mathrm{SR}_3(R)$, we can find $a_1, a_2 \in R$ such that the shorter column

  $$(x_\beta + a_1 x_0, x_{-\gamma}, x_{-\alpha} + a_2 x_0, \_, x_\alpha, x_\gamma, x_{-\beta}),$$

  where the blank indicates the position of the component x_0 that we drop, is unimodular. Reset g to $x_\beta(a_1)x_{-\alpha}(a_2)g$—this requires two elementary operations. After this step we may assume that $(x_\beta, x_{-\gamma}, x_{-\alpha}, x_\alpha, x_\gamma, x_{-\beta})$ is unimodular.
- Observe that every elementary long root unipotent $x_\delta(\xi)$ adds one of the components $x_\beta, x_\alpha, x_\gamma$ to another one of them, acts in the opposite direction on the components $x_{-\beta}, x_{-\alpha}, x_{-\gamma}$, and fixes x_0. This corresponds to the decomposition of the 7-dimensional representation of $G(\mathrm{G}_2, R)$ into two 3-dimensional and one 1-dimensional invariant subspaces, when restricted to $G(\mathrm{A}_2, R)$.

Thus, we consider the ideal I generated by the components $x_\beta, x_\alpha, x_\gamma$. As we just observed, this ideal is not changed by the action of any element of $E(\mathrm{A}_2, R)$. However, under the condition $\mathrm{SR}_3(R/I)$ transitivity of the action $\mathrm{SL}(3, R)$ in the 3-dimensional vector representation is well known from the work of Bass. For this we need two additions to shorten a unimodular column over R/I of length 3 to two positions, then two additions to get 1 in the third position, and, finally, two additions to clear the components in the remaining two positions. This is six elementary operations altogether.

This means that further multiplying g by six factors of the form $x_{\pm(\beta-\alpha)}(*)$ and $x_{\pm(\beta-\gamma)}(*)$ we obtain a column of height 6

$$(x_\beta + a_1 x_0, x_{-\gamma}, x_{-\alpha} + a_2 x_0, \_, x_\alpha, x_\gamma, x_{-\beta}),$$

subject to the extra condition that

$$x_{-\beta} \equiv 1 \pmod{I}, \qquad x_{-\alpha}, x_{-\gamma} \equiv 0 \pmod{I}.$$

In other words, already the following column of height 4

$$(x_\beta, \_, \_, \_, x_\alpha, x_\gamma, x_{-\beta}).$$

is unimodular.

So far, we only invoked the usual stable rank condition SR_3. Next, the tricky part comes, which requires the use of ASR_3.

- Using condition $\mathrm{ASR}_3(R)$ we can find $b_1, b_2 \in R$ such that the ideal J generated by $x_\alpha + b_1 x_\beta, x_\gamma + b_2 x_\beta$ is contained in the same maximal ideals that the [a priori larger] ideal I generated by $x_\beta, x_\alpha, x_\gamma$.
  This means that resetting g to $x_{\alpha-\beta}(b_1) x_{\gamma-\beta}(b_2) g$—that is further two elementary operations—we may assume that the following column of height 3

  $$(\_, \_, \_, \_, x_\alpha, x_\gamma, x_{-\beta})$$

  is unimodular.
- Now, using condition $\mathrm{SR}_3(R)$ once more we can find $c_1, c_2 \in R$ such that the following column of height 2

  $$(\_, \_, \_, \_, x_\alpha + c_1 x_{-\beta}, x_\gamma + c_2 x_{-\beta}, \_)$$

  is unimodular.
  As usual, we reset g to $x_{-\gamma}(c_1) x_{-\alpha}(c_2) g$—that is two more elementary operations.
- After the previous step we may assume that

  $$(\_, \_, \_, \_, x_\alpha, x_\gamma, \_)$$

  is unimodular, and we are done. It remains to express

  $$1 - x_\beta = d_1 x_\alpha + d_2 x_\gamma,$$

and to reset g to $x_{\beta-\alpha}(d_1)x_{\beta-\gamma}(d_2)g$—that is two more elementary operations—to achieve our intermediate goal $x_\beta = 1$.

- Up to now we have used 14 elementary operations in $E(\mathrm{G}_2, R)$. On the other hand, a matrix g with 1 in the diagonal position corresponding to the highest weight can be readily reduced to smaller rank, in our case,

$$g \in G(\mathrm{A}_1, R)U_2U_2^-.$$

This is exactly the celebrated Chevalley–Matsumoto decomposition theorem, see, for instance [46, 73, 87] (the same argument was used in [78]). But $\dim(U_2) = 5$, which consumes $\leqslant 10$ more elementary unipotents, not more 24 elementary factors altogether, as claimed. □

We are in possession of similar reduction results, with pretty sharp bounds, also for all other exceptional cases. But calculations with columns of height 26, 27, 56 and 248 are quite a bit more involved. In the present paper we limit ourselves with *some* explicit bounds, resulting from Tavgen's approach. We intend to come up with much sharper bounds in the sequel to this paper.

5.4 Improvements for Dedekind rings

As is well known, for Dedekind rings the constants in the reduction can be slightly improved. This is based on the well-known property that the ideals I in Dedekind rings are not just 2-generated, but rather 1.5-generated. In other words, one of the generators can be an arbitrary non-zero element of I.

More precisely, let $I \trianglelefteq R$ be an ideal of a Dedekind ring R. Then for any $a \in I$, $a \neq 0$, there exists $b \in I$ such that $aR + bR = I$. This translates into the following stability condition, weaker than $\mathrm{sr}(R) = 1$, but strictly stronger than $\mathrm{sr}(R) = 2$.

Lemma 5.2 *Let R be a Dedekind ring, and $I \trianglelefteq R$ be its ideal. Then for any three elements $a, b, c \in R$ generating I there exists $d \in R$ such that $a, b + dc$ or $a + dc, b$ generate I.*

In particular, *one* addition, instead of two suffices to shorten a unimodular colum of height 3. This property was used by Carter and Keller to get a sharp bound for $\mathrm{SL}(n, R)$, since to reduce a matrix from $\mathrm{SL}(3, R)$ to a matrix from $\mathrm{SL}(2, R)$ one now needs seven elementary operations instead of eight that are expected for general rings with $\mathrm{sr}(R) = 2$.

Here we illustrate this idea by slightly improving the bound in the result of the previous section pertaining to groups of type G_2.

Proposition 5.3 *For a Dedekind ring R one has*

$$G(\mathrm{G}_2, R) = G(\mathrm{A}_2, R)E^{20}(\mathrm{G}_2, R).$$

Proof In each one of the first, third and fourth steps of the procedure described in the proof of Theorem 5.1 one now needs only 1 elementary operation instead of 2.

Further, at the second step inside $\mathrm{SL}(3, R)$ one now needs five elementary operations instead of 6. □

We have a similar improvement for all other exceptional cases, which is new in the number case, and allows one to improve all known bounds. However, its proof requires a painstaking tracking of elementary operations in their minimal representations, and we postpone it to the sequel of this paper.

6 Proof of Theorem A in the case of C_2

6.1 Notation and stability calculations for C_2

Let $G = G(\mathrm{C}_2, R)$, where $R = \mathbb{F}_q[t]$. Fix an order on Φ, and let as usual Φ^+ and Π be the sets of positive and fundamental roots, respectively. Then $\Pi = \{\alpha = \epsilon_1 - \epsilon_2, \beta = 2\epsilon_2\}$ and

$$\Phi^+ = \big\{\alpha = \epsilon_1 - \epsilon_2, \beta = 2\epsilon_2, \alpha + \beta = \epsilon_1 + \epsilon_2, 2\alpha + \beta = 2\epsilon_1\big\}.$$

We fix a representation with the highest weight $\mu = \epsilon_1$. So the other weights are

$$\mu - \alpha = \epsilon_2, \quad \mu - (\alpha + \beta) = -\epsilon_2, \quad \mu - (2\alpha + \beta) = -\epsilon_1.$$

Then $G(\mathrm{C}_2, R)$ is the symplectic group $\mathrm{Sp}(4, R)$ of 4×4-matrices preserving the form

$$B(x, y) = (x_l y_{-1} - x_{-1} y_1) + (x_2 y_{-2} - x_{-2} y_2).$$

Finally, α and $\alpha + \beta$ are short roots while β and $2\alpha + \beta$ are long ones.

Take an arbitrary matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix} \in \mathrm{Sp}(4, R).$$

As we know, any embedding of root systems $\Delta \subset \Phi$ induces a group homomorphism $G(\Delta, R) \to G(\Phi, R)$. Its image will be denoted by $G(\Delta \subset \Phi, R)$. This can be applied to the special case $\Delta = \{\pm\gamma\}$, γ is a root of Φ. We get an embedding φ_γ of the group $G(\Delta, R)$, which is isomorphic to $\mathrm{SL}(2, R)$, into the Chevalley group $G(\Phi, R)$. In this case the image of this embedding will be denoted by $G^\gamma = G^\gamma(R)$.

Thus, for every root $\gamma \in \mathrm{C}_2$ we have the subgroup $G^\gamma(R) = \mathrm{Sp}^\gamma(4, R)$. In particular,

$$x_\gamma(\xi) = \varphi_\gamma \begin{pmatrix} 1 & \xi \\ 0 & 1 \end{pmatrix}, \quad x_{-\gamma}(\xi) = \varphi_\gamma \begin{pmatrix} 1 & 0 \\ \xi & 1 \end{pmatrix}.$$

In this notation, set

$$A' = \varphi_\beta \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad \widetilde{A}' = \varphi_\alpha \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

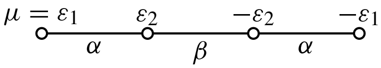
so that the regular embedding $\mathrm{A}_1 \subset \mathrm{C}_2$ on the long roots β and $-\beta$ gives rise to the matrix

$$A' = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & a & b & 0 \\ 0 & c & d & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

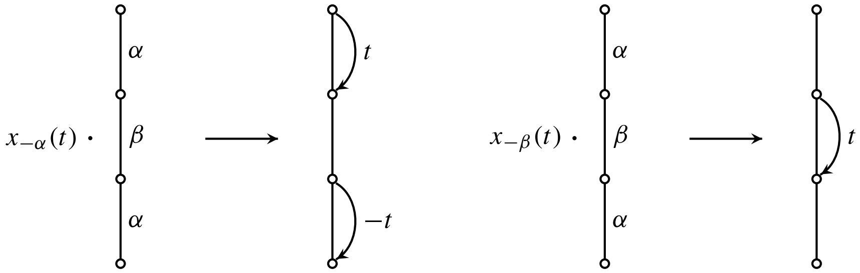
and the regular embedding $\widetilde{\mathrm{A}}_1 \subset \mathrm{C}_2$ on the short roots α and $-\alpha$ gives rise to the matrix

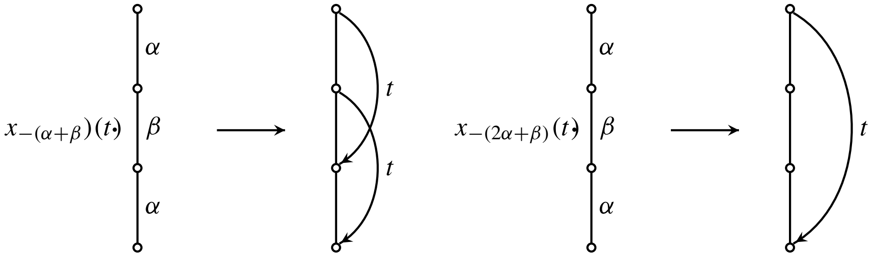
$$\widetilde{A}' = \begin{pmatrix} a & b & 0 & 0 \\ c & d & 0 & 0 \\ 0 & 0 & a & -b \\ 0 & 0 & -c & d \end{pmatrix}.$$

Since we need a bunch of calculations with matrices from $\mathrm{Sp}(4, R)$, we start with some visualization of these calculations. Our main tool is the technique of weight diagrams (see [73, 88]). We work with representations with some highest weight μ. In our case the weight diagram of C_2 type is quite simple:



The entries of matrices $g \in G(\mathrm{C}_2, R)$ are indexed by pairs of weights, $g = (g_{\lambda_1,\lambda_2})$. We concentrate on the first column $g_{*\mu}$ of this matrix, which is the image of the highest weight vector under the action of g. The action of elementary unipotents on the first column of g is depicted on the following self-explaining picture:





Lemma 6.1 *A matrix A in $G(\mathrm{C}_2, R)$ can be moved to A' in $G(\mathrm{A}_1 \subset \mathrm{C}_2, R)$ by $\leqslant 10$ elementary transformations.*

Proof Recall that A' in $G(\mathrm{A}_1 \subset \mathrm{C}_2, R)$ is the image of $A \in G(\mathrm{A}_1, R)$ embedded into $G(\mathrm{C}_2, R)$ on long roots. We denote elements of the first column of $A \in \mathrm{Sp}(4, R)$ lexicographically, $x = (x_1, x_2, x_3, x_4)$. Let $I = \langle x_1, x_2, x_3\rangle$ be the ideal generated by the first three entries of x. Thus, $I + \langle x_4\rangle = R$.

Since R is a Dedekind ring, there exists $t \in R$ such that in $x' = x_{-\alpha}(t)x$ the ideal I is generated by two entries, $I = \langle x_2', x_3'\rangle$, see Lemma 5.2. Note that $\langle x_4\rangle \equiv \langle x_4'\rangle \pmod{I}$. The first column of $x' = x_{-\alpha}(t)x$ is unimodular, and we have

$$\langle x_2', x_3', x_4'\rangle = I + \langle x_4'\rangle = R.$$

Then there exist $t_1, t_2, t_3 \in R$ such that in $x_\alpha(t_1)x_{\alpha+\beta}(t_2)x_{2\alpha+\beta}(t_3)x'$ we obtain the first column of the form $x = (1, *, *, *)$ (cf. [73]).

Having an invertible element in the NW corner of the matrix, it remains to make three elementary moves downstairs and three left-to-right elementary moves to get zeros in the first column and the first row. Thus we transformed A to A' by $10 = 1+3+3+3$ elementary transformations in total. □

6.2 Extracting roots of Mennicke symbols

Our goal is to prove, in the function field case, that one can extract m^{th} roots of Mennicke symbols. This is an essential ingredient in performing elementary operations below, see Lemmas 6.14 and 6.15, which can only be applied when one of the matrix entries is a square.

The previous stability argument was quite general. Below, we restrict our attention to the particular case of the base ring $R = \mathcal{O} = \mathbb{F}_q[t]$.

Actually, we only need the case $m = 2$. We shall proceed along a more general way of reasoning. Namely, we shall first establish the statement in the case $m = q - 1$. If q is odd, the case $m = 2$ follows: after extracting an m^{th} root, we can then raise to the $(m/2)^{\text{th}}$ power to get a square root. So we first assume that q is odd and $m = q - 1$, leaving the problem of extracting square roots in the case of characteristic 2 for separate consideration.

Let us fix some notation. Denote $K = \mathbb{F}_q(t)$. Let $\mathfrak{p}_\infty$ be the infinite place of K, it corresponds to the valuation v_∞ of $\mathcal{O}$ given by

$$v_\infty(f) = -\deg f.$$

This valuation naturally extends to K by setting $v_\infty(f/g) = \deg g - \deg f$. For the completion of K at this place we have

$$K_{v_\infty} = \mathbb{F}_q((1/t)),$$

the field of Laurent series in $1/t$. For brevity, we denote this field by K_∞. Let $\mathcal{O}_\infty = \mathbb{F}_q[[1/t]]$ denote its ring of integers, it is a discrete valuation ring with maximal ideal $\mathfrak{p}_\infty = (1/t)$ and residue field $\mathbb{F}_q$. The residue of $f_0 = a_0 + a_{-1}/t + \cdots \in \mathcal{O}_\infty$ equals a_0. If $f, g \in \mathbb{F}_q[t]$ are polynomials of the same degree, the residue of f/g is equal to the ratio of their leading coefficients.

As mentioned above, we first consider the case where q is odd and $m = q - 1$.

We start with the following observation on extracting roots in K_∞.

Observation 6.2 (cf. [70]) *Given $f \in K_\infty$ with leading term $a_M x^M$, f is an m^{th} power if and only if M is divisible by m and a_M has an m^{th} root in $\mathbb{F}_q$.*

Indeed, suppose that $f = g^m$, then $v_\infty(f) = m v_\infty(g)$, so that $-\deg(f) = -m\deg(g)$ and m divides M. Write $f = x^M f_0$ where $f_0 = a_M + a_{M-1}/t + \cdots$, then $f_0 = g_0^m$ for some $g_0 \in \mathcal{O}_\infty$, $g_0 = b_0 + b_{-1}/t + \cdots$. Taking residues modulo $\mathfrak{p}_\infty$, we get $a_M = b_0^m$.

Conversely, write $f = x^M f_0$, where $f_0 = a_M + a_{M-1}/t + \cdots$, and suppose that m divides M and a_M is an m^{th} power in $\mathbb{F}_q$. Then the polynomial $x^m - a_M \in \mathbb{F}_q[x]$ has a root in $\mathbb{F}_q$, and as $m = q - 1$ is prime to the characteristic of $\mathbb{F}_q$, this root is simple. Hence by Hensel's lemma, it lifts to a root of the polynomial $x^m - f_0 \in \mathcal{O}_\infty[x]$, which belongs to K_∞. Therefore f_0 is an m^{th} power in K_∞, hence so is f.

The subsequent arguments are mainly based on combining two powerful classic tools: algebraic, the m^{th} power reciprocity law, and analytic, (generalised) Dirichlet's theorem on primes in arithmetic progressions, as in [9] (and also [7, 47]).

More precisely, we use the Kornblum–Artin version of Dirichlet's theorem:

Theorem 6.3 (Kornblum–Artin, [63, Theorem 4.8]) *Let a, b be relatively prime polynomials in $\mathcal{O} = \mathbb{F}_q[t]$, $\deg a > 0$. Then there are infinitely many monic irreducible polynomials b' congruent to b modulo $a\mathcal{O}$. Moreover, such b' can be of arbitrary degree N, provided N is sufficiently large.*

The reciprocity law we use in our set-up can be formulated as a product formula for local residue m^{th} power symbols

$$\prod_{\mathfrak{p}} \left(\frac{\alpha, \beta}{\mathfrak{p}}\right)_m = 1. \tag{1}$$

Here $\alpha, \beta \in K^*$ are fixed, and $\mathfrak{p}$ runs over all places of K. For computations below, we use an explicit formula by Hermann Ludwig Schmid, see, e.g. formula (27) in [62]:

$$\left(\frac{\alpha, \beta}{\mathfrak{p}}\right)_m = N_{\mathfrak{p}}\left((-1)^{ab} \frac{\alpha^b}{\beta^a}(\mathfrak{p})\right)^{\frac{q-1}{m}}, \tag{2}$$

where $a = v_{\mathfrak{p}}(\alpha)$, $b = v_{\mathfrak{p}}(\beta)$, $f(\mathfrak{p})$ stands for the image of $f \in K$ in the residue field $\kappa(\mathfrak{p})$ of $\mathfrak{p}$, and $N_{\mathfrak{p}}$ is the norm map from $\kappa(\mathfrak{p})$ to $\mathbb{F}_q$. (The expression raised to the power $(q-1)/m$ in formula (2) is usually called tame symbol.)

The power residue symbol takes values in the group of m^{th} roots of 1 (which is clear from the right-hand side of formula (2)). From the same formula it is clear that for all but finitely many $\mathfrak{p}$ these values are equal to 1 (namely, for those with $v_{\mathfrak{p}}(\alpha) = v_{\mathfrak{p}}(\beta) = 0$). It is well known that this symbol is bimultiplicative.

We are now ready to state and prove an arithmetic lemma which allows us to perform the needed elementary transformations below. It is completely parallel (in the statement and in the proof) to Lemma 3 of [9].

Lemma 6.4 *Let $G = \mathrm{SL}_2(\mathcal{O})$. Let m be either $q-1$ or 2. Then for any $A = \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix} \in G$ there exists $A' = \begin{pmatrix} a^m & b \\ c & d \end{pmatrix} \in G$ elementarily equivalent to A.*

Proof We combine the proof of Lemma 3 in [9] with some facts from [7]. We closely follow the arguments and the notation of [9].[12]

As mentioned above, we first assume that q is odd and $m = q-1$.

As in [9], we may assume that the elements of the first row of A are nonzero. Indeed, if, say, $a_1 = 0$, then b_1 is a nonzero constant, and hence A is elementarily equivalent to A' with $a = 1$ (note that 1 is an m^{th} power in $\mathbb{F}_q$.)

For reader's convenience, we break the proof into several short steps and emphasize the conclusive part of each step by putting in it italic.

Step 1. *One can choose $u, w \in K_\infty^*$ so that the m^{th} local residue at $\mathfrak{p}_\infty$*

$$\zeta = \left(\frac{u, w}{\mathfrak{p}_\infty}\right)_m$$

is a primitive m^{th} root of 1.

This follows from the fact that the residue symbol is non-degenerate, see, e.g. the proof of Case 1 of Theorem 3.5 in [7]. In our set-up, one can argue in a more straightforward way, using formula (2). Under our assumptions, this formula reduces to

$$\left(\frac{u, w}{\mathfrak{p}_\infty}\right)_m = (-1)^{\deg u \deg w} \frac{u^{-\deg w}}{w^{-\deg u}}(\mathfrak{p}_\infty). \tag{3}$$

(Note that the numerator and denominator of the fraction appearing in formula (3) are polynomials of the same degree, hence its residue is well defined and equals the ratio of their leading coefficients.)

[12] There is an exception: in [9] the term 'local units' is used for calling nonzero elements of the local field K_v, where v is a place of K. We avoid using such a terminology because 'local unit' commonly refers to an invertible element of the valuation ring O_v.

Hence one can choose degree one polynomials $u = u_0 + u_1 t$ and $w = w_0 + w_1 t$ such that $-w_1/u_1$ is a primitive element of $\mathbb{F}_q$. Say, let us choose $w = -1 + t$ and u_1 a primitive element of $\mathbb{F}_q$.

Step 2. Consider the arithmetic progression $\{a_1 + b_1\mathcal{O}\}$. By Theorem 6.3, it contains a monic irreducible polynomial $a_2 = t^d + \alpha_{d-1}t^{d-1} + \cdots$ of sufficiently large degree d such that

$$d \equiv 1 \pmod{m}. \tag{4}$$

With our choice of $w = -1 + t$, we have

$$\frac{1}{w} = \frac{1}{-1+t} = \frac{1}{t(1-t^{-1})} = \frac{1}{t}\,(1 + t^{-1} + t^{-2} + \cdots),$$

so that

$$\frac{a_2}{w} = t^{d-1}(1 + \alpha_{d-1}t^{-1} + \cdots)(1 + t^{-1} + t^{-2} + \cdots).$$

Combining congruence (4) with Observation 6.2 and noticing that 1 is an m^{th} power in $\mathbb{F}_q$ for $m = q - 1$, we conclude that *a_2/w is an m^{th} power in K_∞*.

Step 3. We have

$$\left(\frac{u, a_2}{\mathfrak{p}_\infty}\right)_m = \left(\frac{u, a_2/w}{\mathfrak{p}_\infty}\right)_m \cdot \left(\frac{u, w}{\mathfrak{p}_\infty}\right)_m = 1 \cdot \left(\frac{u, w}{\mathfrak{p}_\infty}\right)_m = \zeta.$$

The first equality follows from the multiplicativity of the power residue symbol, and the second equality is a consequence of the choice of a_2 made at Step 2. (Recall that if one of the components of the symbol is an m^{th} power, the symbol equals 1.)

Thus, $\left(\frac{u,a_2}{\mathfrak{p}_\infty}\right)_m$ *is a primitive m^{th} root of* 1 (see Step 1).

Step 4. Since by Step 3 the symbol $\left(\frac{u,a_2}{\mathfrak{p}_\infty}\right)_m$ is a primitive m^{th} root of 1, its powers take all nonzero values in $\mathbb{F}_q$. Hence there exists k such that

$$\left(\frac{u, a_2}{\mathfrak{p}_\infty}\right)_m^k = \left(\frac{b_1, a_2}{a_2\mathcal{O}}\right)_m^{-1},$$

i.e. we have

$$\left(\frac{b_1, a_2}{a_2\mathcal{O}}\right)_m \cdot \left(\frac{u, a_2}{\mathfrak{p}_\infty}\right)_m^k = 1. \tag{5}$$

Note that if necessary, we can replace k by any larger integer k' congruent to k modulo m, and equality (5) will remain valid. *So we set $s = u^k$ and assume that k is large enough.*

Step 5. Using Theorem 6.3 once again, choose an irreducible polynomial b of degree $k = \deg s$ such that

$$b \equiv b_1 \pmod{a_2\mathcal{O}}. \tag{6}$$

On multiplying b by a nonzero constant, we can equalize the leading coefficients of the polynomials b and s. Thus in the sequel we may and shall assume that *b and s have the same degree and the same leading coefficient.*

Step 6. Since the polynomials b and a_2 are irreducible, the m^{th} power reciprocity law reduces to the equality

$$\left(\frac{b, a_2}{b\mathcal{O}}\right)_m \cdot \left(\frac{b, a_2}{a_2\mathcal{O}}\right)_m \cdot \left(\frac{b, a_2}{\mathfrak{p}_\infty}\right)_m = 1 \tag{7}$$

(all other symbols $\left(\frac{b,a_2}{\mathfrak{p}}\right)_m$ are equal to 1 because $v_{\mathfrak{p}}(b) = v_{\mathfrak{p}}(a_2) = 0$).

Let us show that the product of the second and third factors equals 1.

Looking at the second factor, we note that by congruence (6),

$$\left(\frac{b, a_2}{a_2\mathcal{O}}\right)_m = \left(\frac{b_1, a_2}{a_2\mathcal{O}}\right)_m$$

(use formula (2)). As to the third factor, it is equal to $\left(\frac{s,a_2}{\mathfrak{p}_\infty}\right)_m$ because the polynomials b and s are chosen at Step 5 so that they have the same degree and the same leading coefficient, and hence by formula (3) the corresponding symbols coincide. By the choice of s made at Step 4, we have $s = u^k$, so that by the multiplicativity of the residue symbol we have

$$\left(\frac{s, a_2}{\mathfrak{p}_\infty}\right)_m = \left(\frac{u, a_2}{\mathfrak{p}_\infty}\right)_m^k,$$

and we finish by applying (5).

Thus (7) gives $\left(\frac{b,a_2}{b\mathcal{O}}\right)_m = 1$. Swapping components of the symbol inverts its value, hence also

$$\left(\frac{a_2, b}{b\mathcal{O}}\right)_m = 1. \tag{8}$$

Step 7. As both b and a_2 are irreducible, from (8) we conclude that a_2 is an m^{th} power modulo b, i.e. there exists a such that

$$a^m \equiv a_2 \pmod{b}. \tag{9}$$

Step 8. The choices made for a_2 at Step 2 and for b at Step 5, together with congruence (9) obtained at Step 7, allow one to prove the lemma by three elementary operations:

$$\begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix} \to \begin{pmatrix} a_2 & b_1 \\ * & * \end{pmatrix} \to \begin{pmatrix} a_2 & b \\ * & * \end{pmatrix} \to \begin{pmatrix} a^m & b \\ * & * \end{pmatrix}.$$

This finishes the proof in the case where q is odd.

Suppose now that q is a power of 2. In this case, extracting $(q-1)^{\text{th}}$ roots of Mennicke symbols can be done in exactly the same way.

So we only have to consider the problem of extracting square roots. In characteristic 2, this is easy. Indeed, if a polynomial $f \in \mathbb{F}_q[t]$ is irreducible, any $g \in \mathbb{F}_q[t]$ is a square modulo f because its image $\bar{g}$ in the field $\mathbb{F}_q[t]/(f)$ of characteristic 2 is a square, as any other element of a finite field of characteristic 2. Thus it is enough to implement Steps 2 and 5 of the first part of the proof, only taking care of the irreducibility of a_2 and b. □

Remark 6.5 If needed, one can arrange the m^{th} power in the NE corner of A' instead of the upper-left one, without additional elementary operations.

Remark 6.6 If needed, one can arrange an irreducible polynomial not only in the NE corner of A' but also in the lower-left one (at the expense of the fourth elementary operation). Indeed, as the matrix A' is unimodular, the entries a^m and c of its left column are coprime, and one can apply the Kornblum–Artin theorem to the arithmetic progression $\{c + a^m \mathcal{O}\}$ to find an irreducible c' congruent to c modulo a^m. On adding an appropriate multiple of the first row to the second one provides the needed irreducible polynomial c' in the SW corner.

6.3 Swindling lemma for $\widetilde{\mathrm{A}}_1 \subset \mathrm{C}_2$

The following lemmas are headed towards Proposition 6.10, which is a symplectic analogue of the swindling lemma by Nica [51] for the *short root* embedding of a matrix $A \in \mathrm{SL}(2, R)$ into $\mathrm{Sp}(4, R)$.

We start with the following symplectic analogue of the swindling lemma for the *long root* embedding. It is weaker than what we actually need, since here we can only move squares. These calculations are purely formal, here R is an arbitrary commutative ring.

Lemma 6.7 *Let $a, b, c, d, s \in R$, $ad - bcs^2 = 1$ and, moreover, $a \equiv d \equiv 1 \pmod{s}$. Then*

$$\varphi_\beta \begin{pmatrix} a & b \\ cs^2 & d \end{pmatrix} \text{ can be moved to } \varphi_{2\alpha+\beta} \begin{pmatrix} d & -c \\ -bs^2 & a \end{pmatrix}$$

by eight elementary transformations.

Proof Specifically, let $a = 1 + st$, for some $t \in R$. Start with a matrix

$$A = \varphi_\beta \begin{pmatrix} a & b \\ cs^2 & d \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & a & b & 0 \\ 0 & cs^2 & d & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \in \mathrm{SL}(2, R) \leqslant \mathrm{Sp}(4, R).$$

Step 1.

$$A = Ax_{-(\alpha+\beta)}(s) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ bs & a & b & 0 \\ ds & cs^2 & d & 0 \\ 0 & s & 0 & 1 \end{pmatrix}.$$

Step 2.

$$A = x_{\alpha}(cs)A = \begin{pmatrix} 1+bcs^2 & acs & bcs & 0 \\ bs & a & b & 0 \\ ds & 0 & d & -cs \\ 0 & s & 0 & 1 \end{pmatrix}.$$

Step 3.

$$A = x_{\alpha+\beta}(-t)A = \begin{pmatrix} d & acs & bcs-dt & cst \\ bs & 1 & b & -t \\ ds & 0 & d & -cs \\ 0 & s & 0 & 1 \end{pmatrix}.$$

Step 4.

$$A = Ax_{-\alpha}(-bs) = \begin{pmatrix} d-abcs^2 & acs & bcs-dt+bcs^2t & cst \\ 0 & 1 & b-bst & -t \\ ds & 0 & d-bcs^2 & -cs \\ -bs^2 & s & bs & 1 \end{pmatrix}.$$

Step 5.

$$A = Ax_{\beta}(-b+bst) = \begin{pmatrix} d-abcs^2 & acs & -dt+abcs^2t & cst \\ 0 & 1 & 0 & -t \\ ds & 0 & d-bcs^2 & -cs \\ -bs^2 & s & bs^2t & 1 \end{pmatrix}.$$

Step 6.

$$A = Ax_{\alpha+\beta}(t) = \begin{pmatrix} d-abcs^2 & acs & 0 & (1+a)cst \\ 0 & 1 & 0 & 0 \\ ds & 0 & 1 & -cs \\ -bs^2 & s & 0 & a \end{pmatrix}.$$

Step 7.

$$A = x_{2\alpha+\beta}(-ac)A = \begin{pmatrix} d & 0 & 0 & -c \\ 0 & 1 & 0 & 0 \\ ds & 0 & 1 & -cs \\ -bs^2 & s & 0 & a \end{pmatrix}.$$

Step 8.

$$A = x_{-(\alpha+\beta)}(-s)g = \begin{pmatrix} d & 0 & 0 & -c \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -bs^2 & 0 & 0 & a \end{pmatrix}.$$

□

The following lemma is an explicit version of Bass–Milnor–Serre, [7, Lemma 13.3]. It expresses one of the [various!] multiplicativity properties of Mennicke symbols in the symplectic case. We use it here, since it is cheaper than other such multiplicativity properties, in terms of the number of elementary moves.

Lemma 6.8 *Let $a, b, c, d, x, y, z \in R$, $ad - bc = 1$ and $az - xy = 1$. Then*

$$\varphi_\alpha \begin{pmatrix} a & b \\ c & d \end{pmatrix} \varphi_\beta \begin{pmatrix} a & x \\ y & z \end{pmatrix} = \begin{pmatrix} a & b & 0 & 0 \\ c & d & 0 & 0 \\ 0 & 0 & a & -b \\ 0 & 0 & -c & d \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & a & x & 0 \\ 0 & y & z & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

can be moved to

$$\varphi_{2\alpha+\beta} \begin{pmatrix} a & b^2x \\ c^2y & d(1-bc)+b^2c^2z \end{pmatrix} = \begin{pmatrix} a & 0 & 0 & b^2x \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ c^2y & 0 & 0 & d(1-bc)+b^2c^2z \end{pmatrix}$$

by six elementary transformations.

Proof The product we start with equals

$$A = \begin{pmatrix} a & ab & bx & 0 \\ c & ad & dx & 0 \\ 0 & ay & az & -b \\ 0 & -cy & -cz & d \end{pmatrix}.$$

Step 1.

$$A = Ax_\alpha(-b) = \begin{pmatrix} a & 0 & bx & b^2x \\ c & 1 & dx & bdx \\ 0 & ay & 1+xy & bxy \\ 0 & -cy & -cz & d-bcz \end{pmatrix}.$$

Step 2.

$$A = Ax_{-\alpha}(-c) = \begin{pmatrix} a & 0 & abdx & b^2x \\ 0 & 1 & ad^2x & bdx \\ -acy & ay & 1+adxy & bxy \\ c^2y & -cy & -cdxy & d-bcz \end{pmatrix}.$$

Step 3.

$$A = Ax_\beta(-ad^2x) = \begin{pmatrix} a & 0 & abdx & b^2x \\ 0 & 1 & 0 & bdx \\ -acy & ay & 1-abcdxy & bxy \\ c^2y & -cy & bc^2dxy & d-bcz \end{pmatrix}.$$

Step 4.

$$A = Ax_{\alpha+\beta}(-bdx) = \begin{pmatrix} a & 0 & 0 & b^2x \\ 0 & 1 & 0 & 0 \\ -acy & ay & 1 & -b^2cxy \\ c^2y & -cy & 0 & d(1-bc)+b^2c^2z \end{pmatrix}.$$

Step 5.

$$A = x_{-(\alpha+\beta)}(cy)A = \begin{pmatrix} a & 0 & 0 & b^2x \\ 0 & 1 & 0 & 0 \\ 0 & ay & 1 & 0 \\ c^2y & 0 & 0 & d(1-bc)+b^2c^2z \end{pmatrix}.$$

Step 6.

$$A = x_\beta(-ay)A = \begin{pmatrix} a & 0 & 0 & b^2x \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ c^2y & 0 & 0 & d(1-bc)+b^2c^2z \end{pmatrix}.$$

□

Lemma 6.9 *Let $a, b, c, d \in R$, $ad - bc = 1$. Then*

$$\varphi_\alpha\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b & 0 & 0 \\ c & d & 0 & 0 \\ 0 & 0 & a & -b \\ 0 & 0 & -c & d \end{pmatrix}$$

can be moved to

$$\varphi_{(2\alpha+\beta)}\begin{pmatrix} a & b^2 \\ -c^2 & d(1-bc) \end{pmatrix} = \begin{pmatrix} a & 0 & 0 & b^2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -c^2 & 0 & 0 & d(1-bc) \end{pmatrix}$$

by not more than nine elementary transformations.

Proof In the previous lemma, take

$$\begin{pmatrix} a & x \\ y & z \end{pmatrix} = \begin{pmatrix} a & 1 \\ -1 & 0 \end{pmatrix}.$$

This last matrix is a product of three elementary transformations in SL_2,

$$\begin{pmatrix} a & 1 \\ -1 & 0 \end{pmatrix} = t_{21}(-1)t_{12}(1)t_{21}(a-1) = t_{12}(1-a)t_{21}(-1)t_{12}(1),$$

summing up to $6+3=9$. □

Now, we are all set to derive from Lemmas 6.7 and 6.9 a life-size symplectic analogue of the *swindling lemma* by Nica [51] for *short roots*.

Proposition 6.10 *Let $a, b, c, d, s \in R$, $ad - bcs = 1$ and, moreover, $a \equiv d \pmod{s}$. Then*

$$\varphi_\alpha \begin{pmatrix} a & b \\ cs & d \end{pmatrix} \text{ can be moved to } \varphi_\alpha \begin{pmatrix} d & c \\ bs & a \end{pmatrix}$$

by not more than 26 *elementary transformations.*

Proof By Lemma 6.9

$$\varphi_\alpha \begin{pmatrix} a & b \\ cs & d \end{pmatrix} \text{ can be moved to } \varphi_{2\alpha+\beta} \begin{pmatrix} a & b^2 \\ -c^2s^2 & d(1-bcs) \end{pmatrix}$$

by not more than nine elementary operations.

Now we can apply Lemma 6.7 to transform the latter matrix to the matrix of the form $\varphi_{-\beta}$

$$\varphi_\beta \begin{pmatrix} d(1-bcs) & c^2 \\ -b^2s^2 & a \end{pmatrix} = \varphi_{-\beta} \begin{pmatrix} a & b^2s^2 \\ -c^2 & d(1-bcs) \end{pmatrix}$$

by eight elementary operations.

Note that switching the first column with the second one is nothing else than replacing β by $-\beta$. Inside SL_2, such a replacement amounts to the conjugation by w_β. However, with respect to the embedding of SL_2 into Sp_4, it is just a different parametrisation, which gives *the same* matrix in Sp_4, so that no additional elementary moves are needed.

The angle between α and $2\alpha+\beta$ is the same as the angle between $-\beta$ and α. Thus, we can apply Lemma 6.9 once more, and get the desired matrix by not more than nine further elementary operations:

$$\varphi_{-\beta} \begin{pmatrix} a & b^2s^2 \\ -c^2 & d(1-bcs) \end{pmatrix} \longrightarrow \varphi_\alpha \begin{pmatrix} d & c \\ bs & a \end{pmatrix}.$$

Altogether, we have expended not more than $8+9+9=26$ elementary moves. □

Remark 6.11 The above lemmas allow numerous releases.

- To replace columns by rows, one transposes all factors: the transpose of an elementary move is again an elementary move.

- More interestingly, one can switch a and b in, say, Lemma 6.9, thus reducing

$$\varphi_\gamma \begin{pmatrix} a & b \\ c & d \end{pmatrix} \text{ to the form } \varphi_\beta \begin{pmatrix} a^2 & b \\ c(1+ad) & d^2 \end{pmatrix}.$$

However, this is *not* a conjugation. It amounts to a multiplication by a Weyl group element on the left, and by *another* Weyl group element on the right! Such transformations are still elementary, of course, but they may affect the length.

Remark 6.12 We believe that the estimate in this lemma might be *grossly* exaggerated. In the above proof we switched between the short root and the long root positions. We would expect that by implementing swindling in place, the number of elementary operations here could be reduced to something like 7, 8 or 9.

6.4 Bounded elementary generation for Sp$(4, \mathbb{F}_q[t])$

We start with a matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2, R)$$

embedded into Sp$(4, R)$ on the *long root* position $A \in G(\mathrm{A}_1 \subset \mathrm{C}_2, R)$, as in Lemma 6.1:

$$A = \varphi_\beta \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & a & b & 0 \\ 0 & c & d & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

We argue as follows. First, we need to get a matrix in SL$(2, R)$ with a square entry, to be able to move it to a *short root* position.

Lemma 6.13 *Any matrix*

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

from SL$(2, R)$ *can be moved by three elementary transformations in* SL$(2, R)$ *to a matrix of the form*

$$A = \begin{pmatrix} * & b_1^2 \\ * & * \end{pmatrix}.$$

Proof See Lemma 6.4. □

Lemma 6.14 *Let $A \in \mathrm{SL}(2, R)$ be of the form*

$$A = \begin{pmatrix} a & b^2 \\ c' & d' \end{pmatrix}.$$

Then it can be transformed to the matrix of the form

$$A = \begin{pmatrix} a & b^2 \\ -c^2 & d \end{pmatrix}$$

by one elementary transformation.

Proof The argument we produce below is in fact a minor conversion of [7, Lemma 5.3]. Indeed, let

$$A = \begin{pmatrix} a & b^2 \\ c' & d' \end{pmatrix}.$$

Then (a, b^2) is unimodular, and there exist $x, y \in R$ such that $ax + yb^2 = 1$. Setting $c = -b^2y^2$, $d = x(1 + b^2y)$, we get

$$ad - b^2c = ax + ab^2xy + b^4y^2 = ax + b^2y(ax + b^2y) = 1.$$

Consequently,

$$A_1 = \begin{pmatrix} a & b^2 \\ c & d \end{pmatrix} \in \mathrm{SL}(2, R)$$

and thus

$$AA_1^{-1} = \begin{pmatrix} a & b^2 \\ c' & d' \end{pmatrix} \begin{pmatrix} d & -b^2 \\ -c & a \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ c'd - d'c & 1 \end{pmatrix}.$$

Finally,

$$\begin{pmatrix} 1 & 0 \\ -c'd + d'c & 1 \end{pmatrix} \begin{pmatrix} a & b^2 \\ c' & d' \end{pmatrix} = \begin{pmatrix} a & b^2 \\ c & d \end{pmatrix} = \begin{pmatrix} a & b^2 \\ -b^2y^2 & d \end{pmatrix}..$$

□

Lemma 6.15 *Any matrix of the form*

$$A = \varphi_\beta \begin{pmatrix} a & b^2 \\ c & d \end{pmatrix} \in \mathrm{Sp}(4, R)$$

can be moved to a matrix of the form

$$A_1 = \varphi_\alpha \begin{pmatrix} a & b \\ * & * \end{pmatrix}$$

by not more than 10 *elementary transformations in* $\mathrm{Sp}(4, R)$.

Proof Use Lemma 6.14 to get square in the SW corner of A by one elementary move. We get a matrix of the form

$$A' = \varphi_\beta \begin{pmatrix} a & b^2 \\ -c^2 & * \end{pmatrix}.$$

Use Lemma 6.9 to transform A' to

$$A_1 = \varphi_\alpha \begin{pmatrix} a & b \\ * & * \end{pmatrix}$$

by not more than nine elementary moves. □

Remark 6.16 The above amounts to saying that we need at most nine elementary transformations to move a fundamental short root SL_2 to a fundamental long root SL_2, but we might spend up to 10 elementary transformations to move in the opposite direction.

- Summarising the above, we managed to move the original matrix A to a matrix of the form

  $$\varphi_\alpha \begin{pmatrix} * & * \\ * & * \end{pmatrix} \in \mathrm{Sp}_4^\alpha(R),$$

  the fundamental $\mathrm{SL}(2, R)$ in the short root embedding. The total number of elementary transformations to that stage is $3 + 10 = 13$.
  The symplectic swindling lemma for the short root embedding $\widetilde{\mathrm{A}}_1 \to \mathrm{C}_2$ was established in Proposition 6.10. At this point, we can follow the proof by Nica for the $\mathrm{SL}(3, R)$ case almost verbatim. [Alternatively, we *could* follow Carter–Keller's approach, but Nica's approach furnishes a somewhat better bound.] For the sake of self-completeness, we reproduce all details (see [51] for the original exposition).
  We start with a matrix

  $$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2, R)$$

  and proceed as follows.

- Using the Kornblum–Artin version of Dirichlet's theorem (see Theorem 6.3), make b and c in the above matrix irreducible of coprime degrees $\deg(b)$ and $\deg(c)$. Then

  $$\delta(b) = \frac{q^{\deg(b)} - 1}{q - 1} \quad \text{and} \quad \delta(c) = \frac{q^{\deg(c)} - 1}{q - 1}$$

are also coprime. In other words, there exist $u, v \in \mathbb{N}$ such that

$$u\delta(b) - v\delta(c) = 1.$$

This requires not more than two elementary moves.

- It follows that

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}^{u\delta(b)} \cdot \begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-v\delta(c)}.$$

We reduce the factors independently.

- To this end, recall that by the Cayley–Hamilton theorem, $A^2 = \mathrm{tr}(A)A - I$ and $A^m = x(\mathrm{tr}(A))I + y(\mathrm{tr}(A))A$, where I stands for the identity matrix and x, y are polynomials in $\mathbb{Z}[t]$ (see Remark 6.17 below). For an arbitrary m one has

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^m = x\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + y\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} x + ya & yb \\ yc & x + yd \end{pmatrix}.$$

By explicit calculations we get

$$x + ya \equiv a^m \pmod{b} \quad \text{and} \quad x + ya \equiv a^m \pmod{c}.$$

Remark 6.17 In fact, x and y are explicitly known, morally they are the values of two consecutive Chebyshev polynomials U_{m-1} and U_m at $\mathrm{tr}(A)/2 = (a + d)/2$, which allows one to argue differently, *without swindling*. But we do not use it here because this approach would require more elementary moves.

• Now, using swindling on short roots embedding provided by Proposition 6.10 we reduce

$$A = \begin{pmatrix} x + ya & yb \\ yc & x + yd \end{pmatrix}$$

[in the short root position!] to either

$$B = \begin{pmatrix} x + ya & y^2 b \\ c & x + yd \end{pmatrix}$$

or

$$C = \begin{pmatrix} x + ya & b \\ y^2 c & x + yd \end{pmatrix}$$

depending on whether we argue modulo c or modulo b.

• Taking $m = v\delta(c)$, we see that the first matrix is triangular modulo c and that $x + ya \mod c \in \mathbb{F}_q^*$. Since $x + ya \equiv a^m \pmod{c}$, for the latter inclusion we shall check that $z := a^{\delta(c)} \bmod c$ lies in $\mathbb{F}_q^*$. Denote $M = \deg c$. Let $\mathbb{F}_{q'} = \mathbb{F}_{q^M}$ be the extension of degree M of the field $\mathbb{F}_q$, and set $e := a \bmod c \in \mathbb{F}_{q'}$. We shall prove $z^q = z$, i.e., $z^{q-1} = 1$. We have

$$\begin{aligned} z^{q-1} = (a^{\delta(c)} \bmod c)^{q-1} &= ((a \bmod c)^{\delta(c)})^{q-1} = (e^{\delta(c)})^{q-1} \\ &= (e^{(q^M-1)/(q-1)})^{q-1} = e^{q^M-1} = e^{q'} = 1. \end{aligned}$$

Denote $u := x + ya = u \mod c \in \mathbb{F}_q^*$. Applying the same arguments to the matrix B^{-1}, we conclude that $x + yd \mod c = u^{-1} \in \mathbb{F}_q^*$. We have

$$\begin{pmatrix} x+ya & y^2b \\ c & x+yd \end{pmatrix} = \begin{pmatrix} u+cr & y^2b \\ c & u^{-1}+cq \end{pmatrix} \longrightarrow \begin{pmatrix} u & 0 \\ c & u^{-1} \end{pmatrix} \longrightarrow \begin{pmatrix} u & 0 \\ 0 & u^{-1} \end{pmatrix} = h_2$$

in $3 = 2 + 1$ elementary moves (the element in the NE corner of the penultimate matrix is automatically zero because the determinant of the matrix is equal to one).

- Similarly, taking $m = u\delta(b)$, we see that the second matrix is triangular modulo b, and we have $v := x + ya \mod b \in \mathbb{F}_q^*$, $x + yd \mod b = v^{-1} \in \mathbb{F}_q^*$. Accordingly, it can be reduced the matrix of the form

$$\begin{pmatrix} v & 0 \\ 0 & v^{-1} \end{pmatrix} = h_1$$

in three elementary moves.

- By Corollary 2.2,

$$\varphi_\alpha(h_1)\varphi_\alpha(h_2) = \begin{pmatrix} v & 0 & 0 & 0 \\ 0 & v^{-1} & 0 & 0 \\ 0 & 0 & v^{-1} & 0 \\ 0 & 0 & 0 & v \end{pmatrix} \begin{pmatrix} u & 0 & 0 & 0 \\ 0 & u^{-1} & 0 & 0 \\ 0 & 0 & u^{-1} & 0 \\ 0 & 0 & 0 & u \end{pmatrix}$$

can be reduced to the identity matrix in four moves.

Calculating the total number of all elementary transformations used so far one gets the following result.

Theorem 6.18 *The elementary width of* $\mathrm{Sp}(4, \mathbb{F}_q[t])$ *is finite and, moreover,*

$$w_{\mathrm{E}}(\mathrm{Sp}(4, \mathbb{F}_q[t])) \leqslant 79.$$

Proof We have to apply Lemmas 6.1 (10 moves), 6.13 (3 moves), 6.15 (10 moves), Proposition 6.10 (twice) ($2 \cdot 26 = 52$ moves), and Corollary 2.2 (4 moves), which gives 79 moves, as claimed. □

7 Proof of Theorem A via the reduction to rank 3 case

Let $G(\Phi, R)$ be a Chevalley group of rank $\geqslant 3$. Then by stable calculations we can reduce the question of bounded elementary generation of $G(\Phi, R)$ to the root systems of rank 3 rather than those of rank 2. This approach allows us to obtain somewhat better estimates for the elementary width of $G(\Phi, R)$. With this end we have to consider $\Phi = \mathrm{C}_3$ and $\Phi = \mathrm{B}_3$ separately.

7.1 Proof of Theorem A for C$_3$ case

Recall that $G(\mathrm{C}_3, R)$ is the symplectic group $\mathrm{Sp}(6, R)$ of 6×6-matrices preserving the form

$$B(x, y) = (x_l y_{-1} - x_{-1} y_1) + (x_2 y_{-2} - x_{-2} y_2) + (x_3 y_{-3} - x_{-3} y_3).$$

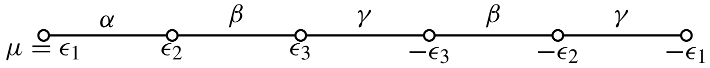
In this case,

$$\Pi = \{\alpha = \epsilon_1 - \epsilon_2, \beta = \epsilon_2 - \epsilon_3, \gamma = 2\epsilon_3\}.$$

We fix a representation with the highest weight $\mu = \epsilon_1$ — the vector representation. Other weights of the vector representation are

$$\mu - \alpha = \epsilon_2, \quad \mu - (\alpha + \beta) = \epsilon_3, \quad \mu - (\alpha + \beta + \gamma) = -\epsilon_3,$$
$$\mu - (\alpha + 2\beta + \gamma) = -\epsilon_2, \quad \mu - (2\alpha + 2\beta + \gamma) = -\epsilon_1.$$

The corresponding weight diagram looks as follows:



Take an arbitrary matrix

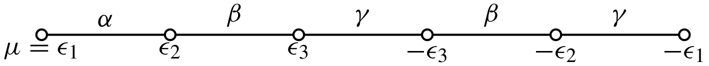
$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & a_{36} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} & a_{46} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} & a_{56} \\ a_{61} & a_{62} & a_{63} & a_{64} & a_{65} & a_{66} \end{pmatrix} \in \mathrm{Sp}(6, R).$$

The embedding $\mathrm{C}_2 \subset \mathrm{C}_3$ gives rise to

$$A' = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & a_{22} & a_{23} & a_{24} & a_{25} & 0 \\ 0 & a_{32} & a_{33} & a_{34} & a_{35} & 0 \\ 0 & a_{42} & a_{43} & a_{44} & a_{45} & 0 \\ 0 & a_{52} & a_{53} & a_{54} & a_{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \in G(\mathrm{C}_2 \subset \mathrm{C}_3).$$

Lemma 7.1 *A matrix A in $G(\mathrm{C}_3, R)$ can be moved to A' in $G(\mathrm{C}_2 \subset C_3, R)$ by $\leqslant 16$ elementary transformations.*

Proof Let the fundamental roots of C_3 be $\alpha = \epsilon_1 - \epsilon_2$, $\beta = \epsilon_2 - \epsilon_3$, $\gamma = 2\epsilon_3$. We fix a representation with the highest weight $\mu = \epsilon_1$. The corresponding weight diagram is as follows:



Let x be the first column of $A \in \mathrm{Sp}(6, R)$,

$$x = (x_1, x_2, x_3, x_{-3}, x_{-2}, x_1).$$

We need to reduce it by elementary transformations to

$$x = (1, 0, 0, 0, 0, 0).$$

- Since R is a Dedekind ring, there exists $t \in R$ such that $x_{-\alpha}(t)x$ is unimodular, see Lemma 5.2.
- Then there exist $t_1, t_2, t_3, t_4, t_5 \in R$ such that in

  $$x_\alpha(t_1)x_{\alpha+\beta}(t_2)x_{\alpha+\beta+\gamma}(t_3)x_{\alpha+2\beta+\gamma}(t_4)x_{\alpha+2\beta+2\gamma}(t_5)x$$

  we obtain the first column of the form

  $$x = (1, *, *, *, *, *)$$

  (cf. [73]).
- Having 1 in the NW corner of the matrix, it remains to apply five downward elementary moves to get

  $$x = (1, 0, 0, 0, 0, 0).$$

  Other five elementary moves allow to make the first row $x = (1, 0, 0, 0, 0, 0)$ as well.

Summarising the above, we see that at most $16 = 1 + 5 + 5 + 5$ moves are needed to reduce $A \in \mathrm{Sp}(6, R)$ to A' in $G(\mathrm{C}_2 \subset \mathrm{C}_3, R)$. □

Using Lemma 6.1, the matrix A' can be moved to $G(\mathrm{A}_1 \subset \mathrm{C}_2 \subset \mathrm{C}_3, R)$ by not more than 10 elementary moves.

Similarly, using Lemma 6.4 + the usual stability for $\mathrm{SL}(3, R)$ the matrix A' can be moved to $A'' \in G(\widetilde{\mathrm{A}}_1 \subset \mathrm{C}_2 \subset \mathrm{C}_3, R)$ by at most $3 + 9 = 12$ elementary moves.

The matrix A'' is of the form

$$A'' = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & a_{22} & a_{23} & 0 & 0 & 0 \\ 0 & a_{32} & a_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & a_{44} & a_{45} & 0 \\ 0 & 0 & 0 & a_{54} & a_{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & B & 0 & 0 \\ 0 & 0 & B^{-1} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

where $B \in \mathrm{SL}(2, R)$.

Now look at the matrix

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & a_{22} & a_{23} \\ 0 & a_{32} & a_{33} \end{pmatrix} \in \mathrm{SL}(2, R) \leqslant \mathrm{SL}(3, R).$$

According to Nica's Theorem it can be moved to the identity matrix in not more than 34 elementary transformations [51].

Summing up all elementary moves above we get

Theorem 7.2 *The elementary width of* $\mathrm{Sp}(6, \mathbb{F}_q[x])$ *is finite and, moreover,*

$$w_{\mathrm{E}}(\mathrm{Sp}(6, \mathbb{F}_q[t])) \leqslant 72.$$

Proof $16 + 10 + 12 + 34 = 72$. □

7.2 Proof of Theorem A for B_3 case

In this case,

$$\Pi = \{\alpha = \epsilon_1 - \epsilon_2, \beta = \epsilon_2 - \epsilon_3, \gamma = \epsilon_3\}.$$

We fix the 7-dimensional orthogonal representation with the highest weight $\mu = \epsilon_1$—the vector representation. Other weights of the vector representation are

$$\mu - \alpha = \epsilon_2,\ \mu - (\alpha + \beta) = \epsilon_3,\ \mu - (\alpha + \beta + \gamma) = 0,\ \mu - (2\alpha + \beta + 2\gamma) = -\epsilon_3,$$
$$\mu - (\alpha + 2\beta + 2\gamma) = -\epsilon_2,\ \mu - (2\alpha + 2\beta + 2\gamma) = -\epsilon_1.$$

Take an arbitrary matrix

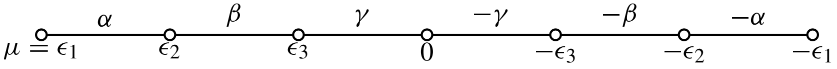
$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} & a_{17} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} & a_{27} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & a_{36} & a_{37} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} & a_{46} & a_{47} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} & a_{56} & a_{57} \\ a_{61} & a_{62} & a_{63} & a_{64} & a_{65} & a_{66} & a_{67} \\ a_{71} & a_{72} & a_{73} & a_{74} & a_{75} & a_{76} & a_{77} \end{pmatrix} \in \mathrm{SO}(7, R).$$

The embedding $\mathrm{B}_2 \subset \mathrm{B}_3$ gives rise to

$$A' = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} & 0 \\ 0 & a_{32} & a_{33} & a_{34} & a_{35} & a_{36} & 0 \\ 0 & a_{42} & a_{43} & a_{44} & a_{45} & a_{46} & 0 \\ 0 & a_{52} & a_{53} & a_{54} & a_{55} & a_{56} & 0 \\ 0 & a_{62} & a_{63} & a_{64} & a_{65} & a_{66} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \in G(\mathrm{B}_2 \subset \mathrm{B}_3).$$

Lemma 7.3 *A matrix A in $G(\mathrm{B}_3, R)$ can be moved to A' in $G(\mathrm{B}_2 \subset \mathrm{B}_3, R)$ by $\leqslant 21$ elementary transformations.*

Proof As usual, we focus on the first column $A_{*\mu}$ of A. The action of elementary unipotents on the first column of A can be viewed via the weight diagram



Denote the first column by

$$x = (x_1, x_2, x_3, x_0, x_{-3}, x_{-2}, x_{-1}).$$

We need to get the column

$$(1, 0, 0, 0, 0, 0, 0)$$

by elementary transformations. The adapt the proof from [73, Theorem 2.1], with some minor improvements for Dedekind rings.

- Consider the ideal $I = \langle x_{-3}, x_{-2}, x_{-1} \rangle$. Then the column (x_1, x_2, x_3, x_0) is unimodular in R/I. By Lemma 5.2, there exists t_0 such that in $x_\gamma(t_0)x$ the column (x_1, x_2, x_3) is unimodular in R/I.
- There are t_1, t_2, t_3 such that the first component of $x_{-\alpha}(t_1)x_\beta(t_2)x_\alpha(t_3)x$ is a unit in R/I.

- Then there are t_4, t_5 such that in $x_{-\alpha}(t_4)x_{-\beta}(t_5)x$ we have

$$x_1 \equiv 1 \pmod{I}, \quad x_2 \equiv x_3 \equiv 0 \pmod{I}.$$

  Hence the column

$$(x_1, -, -, -, x_{-3}, x_{-2}, x_{-1})$$

  is unimodular in R.
- Then there exists t_6 (Lemma 5.2) such that in $x_\alpha(t_6)x$ the column

$$(x_1, -, -, -, x_{-3}, x_{-2}, -)$$

  is unimodular in R.
- Then there is t_7 such that in either $x_\beta(t_7)x$ or in $x_{\alpha+2\beta+2\gamma}(t_7)x$ the column

$$(x_1, -, -, -, x_{-3}, -, -)$$

  is unimodular.
- Then there exist t_8 and t_9 such that in $x_{-(\alpha+2\beta+2\gamma)}(t_9)x_{-\beta}(t_8))x$ we obtain the column

$$(x_1, -, -, -, x_{-3}, 1, -).$$

- One more elementary transformation provides the column

$$(1, -, -, -, -, -, -).$$

- Finally, we need five more unipotents acting downstairs to get the first column

$$(1, 0, 0, 0, 0, 0, 0).$$

  The total number of elementary unipotents used in the process is 16.
- We need five more transformations to bring the first row to the same shape.

Summarising the above, we see that the total number of elementary transformations needed to reduce A in $G(\mathrm{B}_3, R)$ to A' in $G(\mathrm{B}_2 \subset \mathrm{B}_3, R)$ is 21. □

Lemma 7.4 *A matrix A' in $G(\mathrm{B}_2 \subset \mathrm{B}_3, R)$ can be moved to A'' in $G(\mathrm{A}_1 \subset \mathrm{B}_2, R)$ by $\leqslant 10$ elementary transformations.*

Proof Since the groups of types B_2 and C_2 are isomorphic, one can refer to Lemma 6.1. □

Ultimately, reduction of a matrix from $\Gamma(\mathrm{B}_3, R)$ to $G(\mathrm{A}_1, R)$ along the chain of root system embeddings $\mathrm{A}_1 \subset \mathrm{B}_2 \subset B_3$ requires $\leqslant 31$ elementary transformations.

Since we have a commutative diagram of root embeddings

$$\begin{tikzcd}
A_1 \arrow[r, "f"] \arrow[dr, "\varphi"'] & A_2 \arrow[r, "g"] & B_3 \\
& B_2 \arrow[ur, "\psi"'] &
\end{tikzcd},$$

we have the corresponding diagram of homomorphisms of K_1-functors, see [73] or [58, Lemma 3].

Lemmas 7.3 and 7.4 imply that the composition $\psi \circ \varphi$ is an epimorphism. Hence the homomorphism of K_1-functors g corresponding to $\mathrm{A}_2 \to \mathrm{B}_3$ is an epimorphism as well. Thus we obtain

$$G(\mathrm{B}_3, R) = G(\mathrm{A}_2, R) E^{31}(\mathrm{B}_3, R).$$

Combining this with Nica's theorem, that gives additional $\leqslant 34$ elementary transformations, we obtain the following result.

Theorem 7.5 *The elementary width of* $\mathrm{SO}(7, \mathbb{F}_q[x])$ *is finite and, moreover,*

$$w_{\mathrm{E}}(\mathrm{SO}(7, \mathbb{F}_q[t])) \leqslant 65.$$

Remark 7.6 In this section we used the adjoint group of type B_3 and not the simply connected one. As noted in the introduction, this does not affect the finiteness of the elementary width of an arbitrary group of this type.

8 Proof of Theorem C

Actually, for applications to Kac–Moody groups, we mostly need results for Chevalley groups not over the polynomial ring $\mathbb{F}_q[t]$ but rather over the Laurent polynomial rings $\mathbb{F}_q[t, t^{-1}]$. The key difference between these cases is that while the above polynomial ring contains finitely many units, the Laurent polynomial ring has infinitely many of them, namely all at^m, where $m \in \mathbb{Z}$, $a \in \mathbb{F}_q^*$.

As we have already mentioned in Sect. 3, Chevalley groups over rings with finitely many and infinitely many units may behave very differently. This phenomenon is most striking for $\mathrm{SL}(2, R)$. Recall the typical situation occurring in the number case: the group $\mathrm{SL}(2, \mathbb{Z})$ does not have the property of elementary bounded generation whereas the group $\mathrm{SL}(2, R)$, where R is the ring of S-integers in a number field which has infinitely many units, does, see, e.g., [48] for details.

It seems that elementary bounded generation of $\mathrm{SL}(2, R)$ for rings R of S-integers in a global function field which contain infinitely many units, is in general still open. However, the case $R = \mathbb{F}_q[t, t^{-1}]$ can be easily deduced, and at that with rather sharp bounds, from the results of Clifford Queen [59].

Theorem 4.2 reduces the proof of Theorem C to the case of the group $\mathrm{SL}(2, R)$. However, a very short elementary expression in $\mathrm{SL}(2, R)$, for $R = \mathcal{O}_S$ under some additional assumptions on S, was established by [59]. More precisely, Theorem 2 of the above paper [after correction of a minor inaccuracy] amounts essentially to the following result.

Proposition 8.1 *Let $R = \mathcal{O}_S$ be the ring of S-integers of K, a function field of one variable over $\mathbb{F}_q$ with S containing at least two places. Assume that at least one of the following holds:*

- *either at least one of these places has degree one, or*
- *the class number of R, as a Dedekind domain, is prime to $q - 1$.*

Then any matrix $C \in \mathrm{SL}(2, R)$ can be expressed as the product of five elementary transvections.

Proof In follows from [59, Theorem 2] that in this situation any matrix $g \in \mathrm{SL}(2, R)$ can be expressed as the product

$$g = t_{12}(\zeta_1)t_{21}(\zeta_2)t_{12}(\zeta_3)t_{21}(\zeta_4)h_{12}(\epsilon),$$

for some $\zeta_1, \zeta_2, \zeta_3 \in R$ and $\zeta_4, \epsilon \in R^*$, which immediately gives expression of g as a product of *seven* elementary transvections.

However, we can refer to Lemma 2.1, asserting that the first or the last factor in the expression of $h_{12}(\epsilon)$ as a product of elementary transvections can be an arbitrary invertible element of R. Thus, we can start our elementary expression of $h_{12}(\epsilon)$ with the factor $t_{21}(-\zeta_4)$, that cancels with the previous one. After that $t_{12}(\zeta_3)$ can be subsumed into the second factor of the elementary expression of $h_{12}(\epsilon)$, giving us an expression of g as a product of *five* factors of the form UU^-UU^-U.

Implementing the same reduction procedure as in the proof of [59, Theorem 2] for the second column of g instead of the first one, we get a similar expression of g of the form $U^-UU^-UU^-$. □

Remark 8.2 Queen's proof is mainly based on the principles proposed in the seminal paper of Cooke and Weinberger [19] in the number field set-up. Namely, it uses subtle analytic ingredients, such as a function field analogue of Artin's primitive root conjecture, in order to obtain short division chains. In contrast to the number field case where the validity of Artin's conjecture is only known conditionally on the Generalised Riemann Hypothesis (GRH), its function field analogue, developed by Bilharz in the 1930's, became an unconditional theorem after Weil's work. See the paper of Lenstra [39] for more details, as well for some strengthening of Queen's theorem.

In [59] this result is *stated* correctly, in the form to which we referred in our proof, but if you look inside the proof on page 56, it is claimed there that by three multiplications by elementary matrices one can reduce the first column of g to the form $(1, 0)^t$. This is not the case, from Lemma 5 it only follows that it can be reduced to the form $(\epsilon, 0)^t$. Thus, there is no way to express a matrix g as a product of *four* elementary transvections, as would result from the text of the proof of Theorem 2.

One can correct this either as we do above, or, alternatively, by reducing the first column of g to the form $(1, \epsilon)^t$, with $\epsilon \in R^*$, by *three* elementary operations. After that, one needs two more, to remove ϵ, and another one to remove the non-diagonal element in the first row. This gives the same *five* elementary factors.

It follows from [89] that this result is the best possible. The decomposition $E(2, R) = UU^-UU^-$—or, in fact, any such decomposition of length 4 for any Chevalley group—is *equivalent* to $\mathrm{sr}(R) = 1$. Thus, *five* elementary factors is the best bound one can expect in the number case.

Now, precisely the same argument as the proof of Theorem 1 in the work of Smolensky [67] gives us the following estimate of the commutator width.

Corollary 8.3 *Let R be as in Theorem 8.1. Then the commutator width of the simply connected Chevalley group $G = G(\Phi, R)$ is $\leqslant L$, where*

- $L = 3$ *for* $\Phi = \mathrm{A}_l, \mathrm{F}_4$;
- $L = 4$ *for* $\Phi = \mathrm{B}_l, \mathrm{C}_l, \mathrm{D}_l$, *for* $l \geqslant 3$ *or* $\Phi = \mathrm{E}_7, \mathrm{E}_8$, *or, finally,* $\Phi = \mathrm{C}_2, \mathrm{G}_2$ *under the additional assumption that* 1 *is the sum of two units in* R (*which is automatically the case, provided* $q \neq 2$);
- $L = 5$ *for* $\Phi = \mathrm{E}_6$.

Proof In fact, Smolensky proves these bounds for Chevalley groups over rings with $\mathrm{sr}(R) = 1$. The only property of such a ring R that is used in the proof, is the presence of a unitriangular factorisation of length *four*, $E(\Phi, R) = UU^-UU^-$.

However, since the set of commutators is closed under conjugation, the proof in [67] works if not necessarily the matrix $g \in G(\Phi, R)$ itself, but some of its conjugates admits a unitriangular factorisation of length four. However, in our situation this immediately follows from Theorem C, which establishes the unitriangular factorisation of length *five*, $E(\Phi, R) = UU^-UU^-U$. Up to conjugacy the last factor can be carried in front, and subsumed by the first factor. □

Remark 8.4 (i) We believe that for $\Phi = \mathrm{E}_6$ one could also take $L = 4$, but could not prove this.

(ii) We do not know whether one can improve the estimates for non simply connected groups.

On the other hand, the precise bound on the number of elementary generators is somewhat more delicate. Of course, Theorem C immediately implies the following obvious estimate of the elementary width.

Corollary 8.5 *Let R be as in Theorem C. Then the width of the Chevalley group $G(\Phi, R)$ with respect to the elementary unipotents is $\leqslant 5N$, where $N = |\Phi^+|$ is the number of positive roots.*

This bound is quite reasonable, but still not the best possible one. Using the bounded reduction under stability conditions we can get very sharp estimates for the number of elementary factors in other Chevalley groups. For $\mathrm{SL}(n, R)$ such a reduction with the sharpest possible bound is very classical and is implemented already in Carter—Keller [9]. By the same token, from the above proposition we get

Corollary 8.6 *Let R be as in Theorem C. Then any $g \in \mathrm{SL}(n, R)$ can be expressed as a product of $\leqslant \frac{1}{2}(3n^2 - n)$ elementary transvections.*

Proof Immediately follows from the proposition, via improvement of bounded reduction for Dedekind rings. By the contents of Sect. 5.4, reduction of $\mathrm{SL}(n+1, R)$ to $\mathrm{SL}(n, R)$ requires $\leqslant 3n + 1$ elementary operations. □

9 Applications

In this section we briefly discuss two immediate applications of our results. First of all, they imply that Kac–Moody groups of affine type over a finite field have finite commutator width. This problem served as one of the major initial motivations of the present work. As another application, we state several results on bi-interpretability in model theory.

9.1 Applications to Kac–Moody groups

Here we discuss finite commutator width, where there is an especially straightforward connection between the results for the usual Chevalley group $G(\Phi, \mathbb{F}_q[t, t^{-1}])$ over the Laurent polynomial ring and the corresponding affine Kac–Moody group $\widetilde{G}(A, \mathbb{F}_q)$ over the finite field itself.

Let A be an $n \times n$ indecomposable generalised Cartan matrix of (untwisted) affine type, and let K be a field. By an affine Kac–Moody group $\widetilde{G}_{\mathrm{sc}}(A, K)$ we mean the value of the simply connected Tits functor [80], cf. [55], corresponding to the Cartan matrix A. Denote by $\widetilde{E}_{\mathrm{sc}}(A, K)$ its elementary subgroup. The centres $Z(\widetilde{G}_{\mathrm{sc}}(A, K))$ and $Z(\widetilde{E}_{\mathrm{sc}}(A, K))$ coincide. We have a short exact sequence

$$1 \to Z(\widetilde{E}_{\mathrm{sc}}(A, K)) \to \widetilde{E}_{\mathrm{sc}}(A, K) \to G_{\mathrm{ad}}(\Phi, R) \to 1, \qquad (10)$$

the group $G_{\mathrm{ad}}(\Phi, R) \simeq E_{\mathrm{ad}}(\Phi, R) = E_{\mathrm{ad}}(\Phi, K[t, t^{-1}])$ is usually called the *loop group* [30]. So, the elementary affine Kac–Moody group is just a central extension of the loop group. Now we are in a position to prove Theorem D. Recall its statement.

Theorem D *The commutator width of an affine elementary untwisted Kac–Moody group $\widetilde{E}_{\mathrm{sc}}(A, \mathbb{F}_q)$ over a finite field $\mathbb{F}_q$* is *$\leqslant L'$, where*

- *$L' = 5$ for $\Phi = \mathrm{F}_4$ and $\Phi = \mathrm{A}_l, l = 2k + 1, k = 0, 1, \dots$;*
- *$L' = 6$ for $\Phi = \mathrm{A}_l, l = 2k, k = 1, 2, \dots, \Phi = \mathrm{B}_l, \mathrm{C}_l, \mathrm{D}_l,$ for $l \geqslant 3$ or $\Phi = \mathrm{E}_7, \mathrm{E}_8$, or, finally, $\Phi = \mathrm{C}_2, \mathrm{G}_2$ under the additional assumption that 1 is the sum of two units in R (which is automatically the case provided $q \neq 2$);*
- *$L' = 7$ for $\Phi = \mathrm{E}_6$.*

Proof The idea is to get separate estimates for the commutator lengths of the elements of left and right terms of exact sequence (10) and deduce an estimate for the commutator width of the middle term.

For any $g \in \widetilde{E}_{\mathrm{sc}}(A, K)$ denote by $\bar{g} \in G_{\mathrm{ad}}(\Phi, R)$ its projection. Then $\bar{g}$ is a product of L commutators, $\bar{g} = [\bar{a}_1, \bar{b}_1] \dots [\bar{a}_L, \bar{b}_L]$, where L is given by Corollary 8.3. Define

$g' := [a_1, b_1] \dots [a_L, b_L]$. As $\bar{g} = \bar{g}'$, we have $g = g'h$ for some $h \in Z(\widetilde{E}_{\mathrm{sc}}(A, K))$. We will prove that h is a product of two or three commutators, depending on Φ.

Denote by $\Pi = \{\alpha_1, \dots, \alpha_l\}$ the set of fundamental roots of Φ. Then A is determined by the affine root system $\widetilde{\Phi}$ with fundamental roots $\widetilde{\Pi} = \{\alpha_0, \alpha_1, \dots, \alpha_l\}$, see, e.g. [14, 36]. Accordingly, h can be written as $h = h_{\alpha_0}(\lambda_0)h_{\alpha_1}(\lambda_1) \cdots h_{\alpha_l}(\lambda_l)$, cf. [14].[13] Each h_{α_i} lives in $\mathrm{SL}(2, \mathbb{F}_q)$ and has a bounded commutator length. More precisely, suppose that $\widetilde{\Phi} \neq \widetilde{\mathrm{A}}_l$. Then we can represent h as $h_1 h_2$, where $h_1 = h_{\alpha_{i_1}} \cdots h_{\alpha_{i_k}}$, $h_2 = h_{\beta_{j_1}} \cdots h_{\beta_{j_s}}$ such that all the roots α_{i_n} and α_{i_m}, $n \neq m$, as well as β_p and β_t, $p \neq t$, are mutually orthogonal. Every $h_{\alpha_{i_n}}$, $1 \leqslant n \leqslant k$, and $h_{\beta_{jm}}$, $1 \leqslant m \leqslant s$, lies in $\mathrm{SL}(2, \mathbb{F}_q)$, belongs to the centre of this group, and is thus a single commutator, see [79, Theorem 1]. Hence each of h_1 and h_2 belongs to a direct product of $\mathrm{SL}(2, \mathbb{F}_q)$ and is thus a single commutators. As a result, h is a product of two commutators.

The affine Dynkin diagram of type $\widetilde{\mathrm{A}}_l$, $l \geqslant 2$, is a loop. Let $\widetilde{\Phi} = \widetilde{\mathrm{A}}_l$, $l = 2k + 1$, $k \geqslant 1$. Then still $h = h_1 h_2$, as above, and we need two commutators for h. If $\widetilde{\Phi} = \widetilde{\mathrm{A}}_l$, $l = 2k$, $k \geqslant 1$, then there exists a representation $h = h_1 h_2 h_3$ with the properties as above. In this case h is a product of three commutators.

It remains to combine the estimates for the commutator length of g' from Corollary 8.3 with the estimates for the commutator length of h to get the required values of L' for any g. □

Remark 9.1 We do not attempt to state similar results for the bounded elementary generation, in view of the ambiguity of this notion. In fact, elementary generators of $G(\Phi, \mathbb{F}_q[t, t^{-1}])$ correspond to the *spherical roots* of Φ and themselves do not have bounded width with respect to the elementary generators of the affine Kac–Moody group $\widetilde{G}(\Phi, \mathbb{F}_q)$, parametrised in terms of *affine roots*.

Remark 9.2 Let $\overline{G}(A, K)$ be a *complete* affine Kac–Moody group over a field K. Then $\overline{G}(A, K)$ is isomorphic to the Chevalley group of the form $G(\Phi, K((t)))$ where $K((t))$ is the field of formal Laurent series over K.

According to [25], any noncentral element g of $G(\Phi, K((t)))$ is a single commutator. Any central element z is representable as a product of two noncentral elements and hence as a product of two commutators. Thus the commutator width of $\overline{G}(A, K)$ is at most two.

Remark 9.3 It was noticed by Inna Capdeboscq (private correspondence), that the finiteness of the commutator width for Kac–Moody groups can be deduced directly from the polynomial case via Theorem A, using the affine Bruhat decomposition. However this approach yields much worse estimates than the ones from Theorem D.

9.2 Logical applications

Here we state several corollaries of Theorem A related to model theory.

[13] The relevant facts in [14] are formulated for Kac–Moody groups over $\mathbb{C}$. However, the construction remains valid for an appropriate $\mathbb{Z}$-model [30] and hence the needed results from [14] can be extended to groups over $\mathbb{F}_q$.

First note that some of the facts we use in this section require that the group under consideration is finitely generated. In our context, this is guaranteed for Chevalley groups of rank > 1 thanks to the results of Helmut Behr [8].

The notion of bi-interpretability which plays a crucial role in model-theoretic applications can be found in many sources. We refer the reader to [38].

The first important tool is the following Theorem 3.1 of [4]:

Theorem 9.4 ([4]) *Every infinite finitely generated integral domain is bi-interpretable with $\mathbb{Z}$.*

The next lemma can be, in fact, extracted from [37]. Independently, it immediately follows from Theorem 9.4.

Lemma 9.5 *$\mathbb{F}_q[t]$ and $\mathbb{F}_q[t, t^{-1}]$ are bi-interpretable.*

Proof By Theorem 9.4 both rings are bi-interpretable with $\mathbb{Z}$. So they are bi-interpretable with each other. □

Corollary 9.6 *The groups $G(\Phi, \mathbb{F}_q[t])$ and $G(\Phi, \mathbb{F}_q[t, t^{-1}])$, $\mathrm{rk}(\Phi) > 1$, are bi-interpretable with each other and with the rings $\mathbb{F}_q[t]$ and $\mathbb{F}_q[t, t^{-1}]$.*

Proof Follows immediately from [64, Theorem 1.1], which states that if $G(\Phi, R)$, $\mathrm{rk}(\Phi) > 1$, R is an integral domain, has finite elementary width, then R and $G(\Phi, R)$ are bi-interpretable (assuming that for $\Phi = \mathrm{E}_6, \mathrm{E}_7, \mathrm{E}_8, \mathrm{F}_4$ the order of R^* is at least 2). We use also that $\mathbb{F}_q[t]$ and $\mathbb{F}_q[t, t^{-1}]$ are bi-interpretable in view of Lemma 9.5.□

Recall that given a class of groups $\mathcal{C}$, a group $G \in \mathcal{C}$ is *first order rigid* if every group $H \in \mathcal{C}$ which is elementarily equivalent to G is isomorphic to G. We take $\mathcal{C}$ to be the class of finitely generated groups. A group $G \in \mathcal{C}$ is called *finitely axiomatizable* in $\mathcal{C}$ if the elementary theory $Th(G)$ is determined by a single formula φ, that is every group $H \in \mathcal{C}$ which satisfies φ is isomorphic to G. If $\mathcal{C}$ is the class of finitely generated groups, then the property above is used to be called quasi-finite axiomatizability, or QFA-property [52, 53].

Corollary 9.7 *The groups $G(\Phi, \mathbb{F}_q[t])$ and $G(\Phi, \mathbb{F}_q[t, t^{-1}])$, $\mathrm{rk}(\Phi) > 1$, are first order rigid and quasi-finitely axiomatizable.*

Proof Follows from [64, Corollary 1.2]. □

For the following definitions and facts see [15, 38]. A model M of the theory T is called a *prime model* of T if it elementarily embeds in any model of T. A model M of T is atomic if every type realized in M is principal. A model M is *homogeneous* if for every two tuples $\bar{a} = (a_1, \dots, a_n)$, $\bar{b} = (b_1, \dots, b_n)$ in M^n that realize the same types in M there is an automorphism of M that takes $\bar{a}$ to $\bar{b}$. It is known that a model M of T is prime if and only if it is countable and atomic. Furthermore, if M is atomic then it is homogeneous.

The next applications are the consequence of the philosophy of rich groups, i.e., groups where the first-order logic has the same power as the weak second-order logic. This powerful theory is developed by Kharlampovich–Myasnikov–Sohrabi [38]. The crucial observation regarding rich systems is the following

Theorem 9.8 ([38])

- *Any structure bi-interpretable with a rich structure is rich.*
- *The structures $\mathbb{N}$ and $\mathbb{Z}$ are rich.*

The proof is contained in [38, Theorem 4.7 and Lemma 4.14].

Theorem 9.9 *Let $G(\Phi, R)$ be a simply connected Chevalley group, $\operatorname{rk}\Phi > 1$, and let R be an infinite finitely generated integral domain. Assume that $G(\Phi, R)$ is boundedly elementary generated. Assume also that for $\Phi = \mathrm{E}_6, \mathrm{E}_7, \mathrm{E}_8, \mathrm{F}_4$ the order of R^* is at least 2. Then $G(\Phi, R)$ is a rich group.*

Proof By [64, Theorem 1.1], the ring R and the group $G(\Phi, R)$ are bi-interpretable. By Theorem 9.4, R and $\mathbb{Z}$ are bi-interpretable. By Theorem 9.8, $\mathbb{Z}$ is rich. Hence $G(\Phi, R)$ is also rich by Theorem 9.8. □

Corollary 9.10 *Let $G(\Phi, R)$ be a simply connected Chevalley group. Assume the conditions of Theorem 9.9. are fulfilled. Then*

(1) *The group $G(\Phi, R)$ is quasi-finite axiomatizable.*
(2) *The group $G(\Phi, R)$ is first order rigid.*
(3) *The group $G(\Phi, R)$ is prime.*
(4) *The group $G(\Phi, R)$ is atomic.*
(5) *The group $G(\Phi, R)$ is homogeneous.*
(6) *Every finitely generated subgroup of $G(\Phi, R)$ is definable.*

Proof (1) [64, Corollary 1.3], see also [38, Section 4.5.2].
(2) [64, Corollary 1.3], see also [52].
(3) This is a property of rich groups, see [38, Lemma 4.16].
(4) Follows from the previous item, see [34], [38, Section 4.5.1].
(5) See [38, Section 4.5.1].
(6) See [38, Theorem 4.11]. □

Remark 9.11 Theorem 4.11 of [38] states that all finitely generated subgroups of $G(\Phi, R)$ are even uniformly definable, see [38, Definition 4.7].

All above evidently implies

Corollary 9.12 *The groups $G = G(\Phi, \mathbb{F}_q[t])$, $\operatorname{rk}(\Phi) > 2$, and $G = G(\Phi, \mathbb{F}_q[t, t^{-1}])$, $\operatorname{rk}(\Phi) > 1$, are QFA, first order rigid, prime, atomic, homogeneous. All their finitely generated subgroups are definable.*

Remark 9.13 Many facts from Corollary 9.12 are known for Chevalley groups $G(\Phi, \mathcal{O})$ over different number rings and for various kinds of arithmetic lattices, see [5, 6, 38, 64, 69].

Remark 9.14 For the sake of completeness, we give a straightforward proof of definability of finitely generated subgroups of $G(\Phi, \mathbb{F}_q[t])$ and $G(\Phi, \mathbb{F}_q[t, t^{-1}])$, $\operatorname{rk}(\Phi) > 1$, which is parallel to the one of [5].

Theorem 9.15 *All finitely generated subgroups of* $G(\Phi, \mathbb{F}_q[t])$ *and* $G(\Phi, \mathbb{F}_q[t, t^{-1}])$, $\mathrm{rk}(\Phi) > 1$, *are definable.*

Proof Every finitely generated group is recursively enumerable. So we are interested in recursively enumerable sets over $\mathbb{F}_q[t]$. But every recursively enumerable relation over $\mathbb{F}_q[t]$ is Diophantine over $\mathbb{F}_q[t]$, see [22]. Hence every finitely generated subgroup H of $G(\Phi, \mathbb{F}_q[t])$ is definable. Since $\mathbb{F}_q[t]$ and $\mathbb{F}_q[t, t^{-1}]$ are bi-interpretable [38], every finitely generated subgroup of $G(\Phi, \mathbb{F}_q[t, t^{-1}])$ is definable. □

Remark 9.16 In this section, we took a straightforward approach mainly based on combining our results on elementary bounded generation with the work of Segal and Tent [64]. Actually, one can go beyond that and obtain far more general results, valid for Chevalley groups over arbitrary commutative rings. This would require a thorough revision of the approach taken in [64] and is postponed to our forthcoming work.

10 Final remarks

As mentioned in Sect. 3.9, there are many fascinating topics related to bounded generation, some of them well beyond the theory of algebraic groups. We are not going to discuss them here, referring the interested reader to the introductory parts of [21, 48].

Instead, we mention some [almost] immediate eventual generalisations of the results of the present paper, to which we plan to return in its [expected] sequel.

- Firstly, it is a very challenging problem to perform scrupulous analysis of the proofs in Sects. 5–7 with an aim to reduce the number of elementary moves. We are pretty sure that the obtained bounds are far from being optimal. Even without attempting to get sharp bounds, we believe that we could improve the bounds in the present paper, and other related results.
- Secondly, we plan to produce all details for the stability reduction for the exceptional cases F_4, E_6, E_7, E_8 in the same spirit as we have done here for G_2 and B_l. The goal is obtain new explicit bounds for the elementary width in these cases, which are better than the known ones even in the number case.
  Let us mention also several broader projects on which we are presently working.
- One should be able to extend our results to the cases of twisted Chevalley groups and quasi-split groups, as in [78]. The case of isotropic groups, in the spirit of [27], and of generalised unitary groups, also seem tractable.

It is worth noting here that further generalisations in this direction might be problematic. Namely, the recent results of Pietro Corvaja, Andrei Rapinchuk, Jinbo Ren, and Umberto Zannier [21] show that infinite S-arithmetic subgroups of absolutely almost simple anisotropic algebraic groups over number fields are never boundedly generated. The reason is that anisotropic groups do not contain unipotent elements, and a linear group which is not virtually solvable does not contain enough semisimple elements to guarantee bounded generation (some quantitative properties which describe the extent of the absence of bounded generation by semi-simple elements were announced in the subsequent note of the same authors, joint with Julian Demeio [20]).

- There remains a tempting problem of extending the results of the present paper, in particular Theorems A and C, to Chevalley groups over rings of integers in more general (or even arbitrary) global function fields (of course, for rank one groups one has to assume that the group of units of the ring is infinite). It looks like the most challenging part of such an extension is to generalise the relevant arithmetic ingredients of the proof. Generalised versions of Dirichlet's theorem are readily available (see, e.g., [7, A.12]) but this might not suffice for transferring the whole argument to a broader set-up. Say, Trost's theorem [81] on bounded elementary generation of Chevalley groups of rank at least 2 in the function field case required an analogue of one of arithmetic statements of [49, Section 3] (note that in the subsequent preprint [82] he managed to circumvent this difficulty). In a similar vein, an eventual generalisation for groups of rank 1 would perhaps require a function field counterpart of a subtle fact from additive combinatorics of integers used in [49, Section 5]. An attempt to get an explicit estimate by generalising Queen's approach in [59] looks even more problematic. However, we are moderately optimistic regarding the treatability of these problems taking into account substantial progress in analytic arithmetic of global function fields that can be observed over the past decades.
- Most of the results so far pertain to the *absolute* case alone. However, it makes sense to ask similar questions for the *relative* case, in other words for the congruence subgroups $G(\Phi, R, I)$, and the elementary subgroups $E(\Phi, R, I)$ of level $I \trianglelefteq R$. The expectation is to get similar *uniform* bounds in terms of the elementary conjugates $x_{-\mathfrak{a}}(\eta)x_{\mathfrak{a}}(\xi)x_{-\mathfrak{a}}(-\eta)$, $\mathfrak{a} \in \Phi$, $\xi \in I$, $\eta \in R$. Some results in this direction are contained in the paper by Sinchuk and Smolensky [65]. As a more remote goal one could think of generalisations to birelative subgroups, see [33].
- Finally, there is a broader area of *partial* bounded generation, bounded generation in terms of other sets of generators, etc. When bounded generation in terms of X does not hold for the group G itself, one could ask, whether the width

$$w_X(Y) = \sup l_X(g), \quad g \in Y,$$

is bounded, for certain subsets $Y \subseteq G$. For instance, the results by Stepanov and others that we mentioned in Sect. 3.9, imply that $w_{\mathrm{E}}(C)$ is [uniformly] bounded for the set C of commutators in any Chevalley group of rank $\geqslant$ over an *arbitrary* commutative ring. Recently, the third author and Raimund Preusser established partial results in the same spirit for the set of m^{th} powers. It is natural to expect that some form of this claim holds for arbitrary words, which would (in particular!) infer a negative answer to the problem of finite verbal width.

Acknowledgements Very special thanks go to Inna Capdeboscq. This paper started jointly with her as a discussion of the bounded commutator width of various classes of Kac–Moody groups over finite fields, and for a long time it was supposed to be a work of *four* authors. Our sincere thanks go to Nikolai Gordeev, Olga Kharlampovich, Jun Morita, Alexei Myasnikov, Denis Osipov, Alexei Stepanov, and Igor Zhukov for useful discussions regarding various aspects of this work. We thank the referees for careful reading and meticulous remarks.

References

1. Abért, M., Lubotzky, A., Pyber, L.: Bounded generation and linear groups. Internat. J. Algebra Comput. **13**(4), 401–413 (2003)
2. Adian, S.I., Mennicke, J.: Bounded generation of $\mathrm{SL}(n, Z)$. Internat. J. Algebra Comput. **2**(4), 357–365 (1992)
3. Arlinghaus, F.A., Vaserstein, L.N., You, H.: Commutators in pseudo-orthogonal groups. J. Austral. Math. Soc. Ser. A **59**(3), 353–365 (1995)
4. Aschenbrenner, M., Khélif, A., Naziazeno, E., Scanlon, T.: The logical complexity of finitely generated commutative rings. Int. Math. Res. Not. IMRN **2020**(1), 112–166 (2020)
5. Avni, N., Lubotzky, A., Meiri, C.: First order rigidity of non-uniform higher rank arithmetic groups. Invent. Math. **217**(1), 219–240 (2019)
6. Avni, N., Meiri, C.: Words have bounded width in $\mathrm{SL}(n, \mathbb{Z})$. Compositio Math. **155**(7), 1245–1258 (2019)
7. Bass, H., Milnor, J., Serre, J.-P.: Solution of the congruence subgroup problem for $\mathrm{SL}_n(n \geqslant 3)$ and $\mathrm{Sp}_{2n}(n \geqslant 2)$. Inst. Hautes Études Sci. Publ. Math. **33**, 59–137 (1967)
8. Behr, H.: Arithmetic groups over function fields. I: A complete characterization of finitely presented arithmetic subgroups of reductive algebraic groups. J. Reine Angew. Math. **495**, 79–118 (1998)
9. Carter, D., Keller, G.: Bounded elementary generation of $\mathrm{SL}_n(\mathcal{O})$. Amer. J. Math. **105**(3), 673–687 (1983)
10. Carter, D., Keller, G.: Elementary expressions for unimodular matrices. Commun. Algebra **12**(3–4), 379–389 (1984)
11. Carter, D., Keller, G.E.: Bounded elementary expressions in $\mathrm{SL}(2, \mathcal{O})$. Preprint, University of Virginia, 1–11 (1985)
12. Carter, D., Keller, G.E., Paige, E.: Bounded expressions in $\mathrm{SL}(2, \mathcal{O})$. Preprint, University of Virginia, 1–21 (1985)
13. Carter, R.W.: Simple Groups of Lie Type. Pure and Applied Mathematics, vol. 28. Wiley, London (1972)
14. Carter, R.W., Chen, Y.: Automorphisms of affine Kac–Moody groups and related Chevalley groups over rings. J. Algebra **155**(1), 44–94 (1993)
15. Chang, C.C., Keisler, H.J.: Model Theory. 2nd edn. Studies in Logic and the Foundations of Mathematics, vol. 73. North Holland, Amsterdam (1977)
16. Cohn, P.M.: On the structure of the GL_2 of a ring. Inst. Hautes Études Sci. Publ. Math. **30**, 5–53 (1966)
17. Cooke, G.E.: A weakening of the Euclidean property for integral domains and applications to algebraic number theory. I. J. Reine Angew. Math. **282**, 133–156 (1976)
18. Cooke, G.E.: A weakening of the Euclidean property for integral domains and applications to algebraic number theory. II. J. Reine Angew. Math. **283**(284), 71–85 (1976)
19. Cooke, G., Weinberger, P.J.: On the construction of division chains in algebraic number rings, with applications to SL_2. Commun. Algebra **3**, 481–524 (1975)
20. Corvaja, P., Demeio, J., Rapinchuk, A., Ren, J., Zannier, U.: Bounded generation by semi-simple elements: quantitative results (2022). arXiv:2203.00755
21. Corvaja, P., Rapinchuk, A.S., Ren, J., Zannier, U.M.: Non-virtually abelian anisotropic linear groups are not boundedly generated. Invent. Math. **227**(1), 1–26 (2022)
22. Demeyer, J.: Recursively enumerable sets of polynomials over a finite field are Diophantine. Invent. Math. **170**(3), 655–670 (2007)
23. Dennis, R.K., Vaserstein, L.N.: On a question of M. Newman on the number of commutators. J. Algebra **118**(1), 150–161 (1988)
24. Dennis, R.K., Vaserstein, L.N.: Commutators in linear groups. K-Theory **2**(6), 761–767 (1989)
25. Ellers, E.W., Gordeev, N.: On the conjectures of J. Thompson and O. Ore. Trans. Amer. Math. Soc. **350**(9), 3657–3671 (1998)
26. Erovenko, I.V.: $\mathrm{SL}_n(F[x])$ is not boundedly generated by elementary matrices: explicit proof. Electron. J. Linear Algebra **11**, 162–167 (2004)
27. Erovenko, I.V., Rapinchuk, A.S.: Bounded generation of S-arithmetic subgroups of isotropic orthogonal groups over number fields. J. Number Theory **119**(1), 28–48 (2008)
28. Ershov, M., Jaikin-Zapirain, A., Kassabov, M.: Property (T) for Groups Graded by Root Systems. Memoirs of the American Mathematical Society, vol. 249(1186). American Mathematical Society, Providence (2017)

29. Estes, D., Ohm, J.: Stable range in commutative rings. J. Algebra **7**, 343–362 (1967)
30. Garland, H.: The arithmetic theory of loop groups. Inst. Hautes Études Sci. Publ. Math. **52**, 5–136 (1980)
31. Gvozdevsky, P.: Improved K_1-stability for the embedding D_5 into E_6. Commun. Algebra **48**(11), 4922–4931 (2020)
32. Gvozdevsky, P.: Bounded reduction of orthogonal matrices over polynomial rings. J. Algebra **602**, 300–321 (2022)
33. Hazrat, R., Stepanov, A., Vavilov, N., Zhang, Z.: Commutator width in Chevalley groups. Note Mat. **33**(1), 139–170 (2013)
34. Hodges, W.: Model Theory. Encyclopedia of Mathematics and its Applications, vol. 42. Cambridge University Press, Cambridge (1993)
35. Jordan, B.W., Zaytman, Y.: On the bounded generation of arithmetic SL_2. Proc. Natl. Acad. Sci. USA **116**(38), 18880–18882 (2019)
36. Kac, V.G.: Infinite-Dimensional Lie Algebras, 3rd edn. Cambridge University Press, Cambridge (1990)
37. Kharlampovich, O., Myasnikov, A.: What does a group algebra know about a free group. Ann. Pure Appl. Logic **169**(6), 523–547 (2018)
38. Kharlampovich, O., Myasnikov, A., Sohrabi, M.: Rich groups, weak second-order logic, and applications. In: Kharlampovich, O., Sklinos, R. (eds.) Groups and Model Theory, pp. 127–193. de Gruyter, Berlin (2021)
39. Lenstra, H.W., Jr.: On Artin's conjecture and Euclid's algorithm in global fields. Invent. Math. **42**, 201–224 (1977)
40. Liebeck, M.W., O'Brien, E.A., Shalev, A., Tiep, P.H.: The Ore conjecture. J. Eur. Math. Soc. (JEMS) **12**(4), 939–1008 (2010)
41. Liebeck, M.W., O'Brien, E.A., Shalev, A., Tiep, P.H.: Commutators in finite quasisimple groups. Bull. London Math. Soc. **43**(6), 1079–1092 (2011)
42. Liehl, B.: On the group SL_2 over orders of arithmetic type. J. Reine Angew. Math. **323**, 153–171 (1981)
43. Loukanidis, D., Murty, V.K.: Bounded generation for SL_n $(n \geqslant 2)$ and Sp_n $(n \geqslant 1)$ (1994). Preprint
44. Luzgarev, A.Yu., Stavrova, A.A.: The elementary subgroup of an isotropic reductive group is perfect. St. Petersburg Math. J. **23**(5), 881–890 (2012)
45. Magurn, B.A.: Algebraic Introduction to K-Theory. Encyclopedia of Mathematics and its Applications, vol. 87. Cambridge University Press, Cambridege (2002)
46. Matsumoto, H.: Sur les sous-groupes arithmétiques des groupes semi-simples déployés. Ann. Sci. Éc. Norm. Sup. **2**, 1–62 (1969)
47. Milnor, J.: Introduction to Algebraic K-Theory. Annals of Mathematics Studies, vol. 72. Princeton University Press, Princeton (1971)
48. Morgan, A.V., Rapinchuk, A.S., Sury, B.: Bounded generation of SL_2 over rings of S-integers with infinitely many units. Algebra Number Theory **12**(8), 1949–1974 (2018)
49. Morris, D.W.: Bounded generation of $\mathrm{SL}(n, A)$ (after D. Carter, G. Keller, and E. Paige). New York J. Math. **13**, 383–421 (2007)
50. Murty, V.K.: Bounded and finite generation of arithmetic groups. In: Dilcher, K. (ed.) Number Theory (Halifax, NS, 1994). CMS Conference Proceedings, vol. 15, pp. 249–261. American Mathematical Society, Providence (1995)
51. Nica, B.: On bounded elementary generation for SL_n over polynomial rings. Israel J. Math. **225**(1), 403–410 (2018)
52. Nies, A.: Separating classes of groups by first-order sentences. Internat. J. Algebra Comput. **13**(3), 287–302 (2003)
53. Oger, F., Sabbagh, G.: Quasi-finitely axiomatizable nilpotent groups. J. Group Theory **9**(1), 95–106 (2006)
54. O'Meara, O.T.: On the finite generation of linear groups over Hasse domains. J. Reine Angew. Math. **217**, 79–108 (1965)
55. Peterson, D.H., Kac, V.G.: Infinite flag varieties and conjugacy theorems. Proc. Nat. Acad. Sci. USA **80**(6), 1778–1782 (1983)
56. Platonov, V.P., Rapinchuk, A.S.: Abstract properties of S-arithmetic groups and the congruence problem. Russ. Acad. Sci. Izv. Math. **40**(3), 455–476 (1993)
57. Plotkin, E.B.: Surjective stabilization for K_1-functor for some exceptional Chevalley groups. J. Soviet Math. **64**(1), 751–766 (1993)

58. Plotkin, E.: On the stability of the $\mathcal{K}_1$-functor for Chevalley groups of type E_7. J. Algebra **210**(1), 67–85 (1998)
59. Queen, C.: Some arithmetic properties of subrings of function fields over finite fields. Arch. Math. (Basel) **26**, 51–56 (1975)
60. Rapinchuk, A.S.: The congruence subgroup problem for arithmetic groups of finite width. Soviet Math. Dokl. **42**(2), 664–668 (1991)
61. Rapinchuk, A.S.: Representations of groups of finite width. Soviet Math. Dokl. **42**(3), 816–820 (1991)
62. Roquette, P.: Class field theory in characteristic p, its origin and development. In: Miyake, K. (ed.) Class Field Theory-its Centenary and Prospect (Tokyo, 1998). Advanced Studies in Pure Mathematics, vol. 30, pp. 549–631. Mathematical Society of Japan, Tokyo (2001)
63. Rosen, M.: Number Theory in Function Fields. Graduate Texts in Mathematics, vol. 210. Springer, New York (2002)
64. Segal, D., Tent, K.: Defining R and $G(R)$, J. Europ. Math. Soc. https://doi.org/10.4171/JEMS/1255. arXiv:2004.13407
65. Sinchuk, S., Smolensky, A.: Decompositions of congruence subgroups of Chevalley groups. Internat. J. Algebra Comput. **28**(6), 935–958 (2018)
66. Sivatski, A., Stepanov, A.: On the word length of commutators in $\mathrm{GL}_n(R)$. $\mathcal{K}$-Theory **17**(4), 295–302 (1999)
67. Smolensky, A.: Commutator width of Chevalley groups over rings of stable rank 1. J. Group Theory **22**(1), 83–101 (2019)
68. Smolensky, A., Sury, B., Vavilov, N.: Gauss decomposition for Chevalley groups, revisited. Internat. J. Group Theory **1**(1), 3–16 (2011)
69. Sohrabi, M., Myasnikov, A.G.: Bi-interpretability with Z and models of the complete elementary theories of $SL_n(\mathcal{O})$, $T_n(\mathcal{O})$ and $\mathrm{GL}_n(\mathcal{O})$, $n \geqslant 3$ (2020). arXiv:2004.03585
70. StackExchange discussion. https://math.stackexchange.com/questions/2690954/laurent-series-of-root-of-polynomials
71. Stein, M.R.: Generators, relations, and coverings of Chevalley groups over commutative rings. Amer. J. Math. **93**, 965–1004 (1971)
72. Stein, M.R.: Surjective stability in dimension 0 for K_2 and related functors. Trans. Amer. Math. Soc. **178**, 165–191 (1973)
73. Stein, M.R.: Stability theorems for $\mathcal{K}_1$, $\mathcal{K}_2$ and related functors modeled on Chevalley groups. Japan J. Math. (N.S.) **4**(1), 77–108 (1978)
74. Steinberg, R.: Générateurs, relations et revêtements de groupes algébriques. In: Colloq. Théorie des Groupes Algébriques (Bruxelles, 1962), pp. 113–127. Librairie Universitaire, Louvain (1962)
75. Steinberg, R.: Lectures on Chevalley Groups. University Lecture Series, vol. 66. American Mathematical Society, Providence (2016)
76. Stepanov, A.: Structure of Chevalley groups over rings via universal localization. J. Algebra **450**, 522–548 (2016)
77. Stepanov, A., Vavilov, N.: On the length of commutators in Chevalley groups. Israel J. Math. **185**, 253–276 (2011)
78. Tavgen, O.I.: Bounded generation of Chevalley groups over rings of algebraic S-integers. Math. USSR-Izv. **36**(1), 101–128 (1991)
79. Thompson, R.C.: Commutators in special and general linear groups. Trans. Amer. Math. Soc. **101**, 16–33 (1961)
80. Tits, J.: Uniqueness and presentation of Kac–Moody groups over fields. J. Algebra **105**(2), 542–573 (1987)
81. Trost, A.A.: Bounded generation by root elements for Chevalley groups defined over rings of integers of function fields with an application in strong boundedness (2021). arXiv:2108.12254
82. Trost, A.A.: Elementary bounded generation for SL_n for global function fields and $n \geqslant 3$ (2022). arXiv:2206.13958
83. van der Kallen, W.: $\mathrm{SL}_3(\mathbb{C}[X])$ does not have bounded word length. In: Dennis, R.K. (ed.) Algebraic K-theory, Part I (Oberwolfach 1980). Lecture Notes Mathematics, vol. 966, pp. 357–361. Springer, Berlin (1982)
84. Vaserstein, L.N.: Bounded reduction of invertible matrices over polynomial rings by addition operations. Preprint, Pennsylvania State University (2006). http://www.personal.psu.edu/lxv1/pm2.pdf
85. Vaserstein, L.N.: Polynomial parametrization for the solution of Diophantine equations and arithmetic groups. Ann. Math. **171**(2), 979–1009 (2010)

86. Vaserstein, L.N., Wheland, E.: Commutators and companion matrices over rings of stable rank 1. Linear Algebra Appl. **142**, 263–277 (1990)
87. Vavilov, N.: Structure of Chevalley groups over commutative rings. In: Yamaguti, K., Kawamoto, N. (eds.) Nonassociative Algebras and Related Topics (Hiroshima, 1990), pp. 219–335. World Scientific Publishing, River Edge (1991)
88. Vavilov, N., Plotkin, E.: Chevalley groups over commutative rings. I: Elementary calculations. Acta Appl. Math. **45**(1), 73–113 (1996)
89. Vavilov, N.A., Smolensky, A.V., Sury, B.: Unitriangular factorizations of Chevalley groups. J. Math. Sci. (N.Y.) **183**(5), 584–599 (2012)
90. Vsemirnov, M.: Short unitriangular factorizations of $\mathrm{SL}_2(\mathbb{Z}[1/p])$. Q. J. Math. **65**(1), 279–290 (2014)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.