Записки научных
семинаров ПОМИ
Том 470, 2018 г.

N. A. Vavilov

TOWARDS THE REVERSE DECOMPOSITION OF UNIPOTENTS

Abstract. Decomposition of unipotents gives short polynomial expressions of the *conjugates* of elementary generators as products of elementaries. It turns out that with some minor twist the decomposition of unipotents can be read backwards, to give very short polynomial expressions of elementary generators themselves in terms of elementary conjugates of an arbitrary matrix and its inverse. For absolute elementary subgroups of *classical* groups this was recently observed by Raimund Preusser. I discuss various generalisations of these results for *exceptional* groups, specifically those of types E_6 and E_7, and also mention further possible generalisations and applications.

Jesus said to them, "Have you never read in the Scriptures: "The stone that the builders rejected has become the cornerstone"; this was the Lord's doing, and it is marvelous in our eyes."
Matthew 21:42

§1. Introduction

Decomposition of unipotents [32] was first proposed by Alexei Stepanov [29] for $\mathrm{GL}(n, R)$ in 1987, immediately generalised to other split classical groups by the present author [40], and soon afterwards announced also for exceptional Chevalley groups [41, 53]. It was then extensively developed in other contexts by a number of authors, see the recent papers [24, 31, 45] for many further references. In its simplest form it can be viewed as an effective/constructive version of the **normality of elementary subgroups**, see [23, 33–35].

Namely, let Φ be an irreducible root system of rank $\geqslant 2$, R be an arbitrary commutative ring with 1, and $G(\Phi, R)$ be the simply connected Chevalley group of type Φ over R. Further, fix a split maximal torus

---

Key words and phrases: classical groups, Chevalley groups, normal structure, elementary subgroups, decomposition of unipotents, reverse decomposition of unipotents.
This publication is supported by Russian Science Foundation grant 17-11-01261.

$T(\Phi, R)$ of $G(\Phi, R)$ and the corresponding elementary generators $x_\alpha(\xi)$, where $\alpha \in \Phi$, $\xi \in R$. Let $E(\Phi, R)$ be the elementary subgroup spanned by all these elementary generators. Then the above normality theorem asserts that $E(\Phi, R)$ is normal in $G(\Phi, R)$.

In its simplest form, **decomposition of unipotents** provides explicit polynomial formulae expressing the conjugate $gx_\alpha(\xi)g^{-1}$ of an elementary generator by an arbitrary matrix $g \in G(\Phi, R)$ as a product of elementaries. Such formulae are especially straightforward and uniform for simply laced root systems admitting microweights, those of types $\mathrm{A}_l, \mathrm{D}_l, \mathrm{E}_6$ and E_7.

Specifically, denote by $E^L(\Phi, R)$ the subset of $E(\Phi, R)$ consisting of products of $\leqslant L$ elementary generators. Then [53] (the details of proofs were published in [32, 41, 42]) implies in particular that $gx_\alpha(\xi)g^{-1} \in E^L(\Phi, R)$, where L takes the following values: $4(l+1)l$ for A_l; $L = 4 \cdot 2l \cdot 2(l-1)$ for D_l; $4 \cdot 27 \cdot 16$ for E_6; and $4 \cdot 56 \cdot 27$ for E_7, respectively [21, 42]

Another keynote classical result in the structure theory of Chevalley groups is **description of E-normalised subgroups**, i.e. subgroups of $G(\Phi, R)$ normalised by the elementary group $E(\Phi, R)$. Roughly, such a description asserts that under some mild assumptions for any such subgroup H there exists a unique ideal $I \trianglelefteq R$ such that $E(\Phi, R, I) \leqslant H \leqslant C(\Phi, R, I)$, where $E(\Phi, R, I)$ and $C(\Phi, R, I)$ are the relative elementary subgroup and the full congruence subgroups of level I, respectively.

It would be difficult to mention even the most important historical steps towards this result. The pathbreaking contribution was due to Hyman Bass [7], who created the setting for such a description, and established it for finite dimensional rings. There was the whole history of generalisations to other classical groups, starting with the work of Anthony Bak [5, 8], see [6, 22] for a detailed survey.

The next fantastic breakthrough, completely unexpected at the time, came with the work of John Wilson [55], who described E-normalised subgroups in $\mathrm{GL}(n, R)$, $n \geqslant 4$, over an *arbitrary* commutative ring R. In 1973–1975 Igor Golubchik improved the bound to $n \geqslant 3$ and generalised this description to some non-commutative rings, and to classical groups. However, his outstanding results [14, 15] were never properly published in a form accessible to Western readers [16, 17], and are largely ignored[1]. For

[1] As an aside: in 1981 I was a *Gutachter* for Golubchik's Thesis [16]. I can certify that complete proofs of the results announced in [14, 15], and, in fact, more general results, were there at that time. Unlike all subsequent proofs, including those by Vaserstein and Preusser, they are based on truly *non-commutative* localisations. In terms of the class

[almost] *commutative* rings, several short and transparent proofs of these results based on different ideas were then proposed by Leonid Vaserstein, Zenon Borewicz, myself, Alexei Stepanov and others, see, for instance, [9, 32, 36].

An [almost] complete generalisations to Chevalley groups was obtained by Eiichi Abe, Kazuo Suzuki and Leonid Vaserstein [1–3, 37]. The last dots for symplectic groups and G_2 were put by Douglas Costa and Gordon Keller [12, 13]. Then further proofs with very clear geometrical structure were proposed, inculding [30, 31, 46, 47, 50], etc.

Now, it is natural to ask what would be an *effective/constructive* version of that? Until very recently, this was only known in some very special cases. Thus, for $\mathrm{SL}(n, \mathbb{Z})$, $n \geqslant 3$, Joel Brenner [11] established that for an arbitrary non-central matrix $g \in \mathrm{SL}(n, \mathbb{Z})$ there is a *bounded* product of conjugates of g and g^{-1} that is a non-trivial elementary transvection $t_{ij}(\xi)$, $\xi \neq 0$. Brenner's proof used the theory of elementary divisors, and even generalisations to other groups over PID were not immediate. And of course, there was *whatsoever* no hope to write such similar formulae for arbitrary commutative rings.

Thus, we were seriously perplexed, when we've first seen the preprints of [26, 27] by Raimund Preusser in Summer 2017. The calculations in [26] start in exactly the same way as in [32], so predictably our assessment of these papers came through the following three stages:

- There must be nothing new as compared with [32].
- Gosh, why is it true at all?
- It is a fantastic breakthrough in the structure theory of algebraic-like groups!

Technically, the twist introduced by Raimund Preusser in the decomposition of unipotents seems to be minor. It consists in expressing a conjugate of an elementary generator not as a product of factors sitting in proper parabolics of certain types, but rather sitting in the products of these parabolics by something small in the unipotent radicals of the opposite parabolics. We were aware of the idea itself [31] (in fact, it was implicit already in [7, 38]), but never appreciated the whole significance of this apparently small variation.

of rings they were never superceded and their publication would make much sense even today.

In fact, it allows to reduce degree of the resulting polynomials, and thus both to completely avoid the cumbersome "main lemma" (see, for instance, [44] and discussion there), establishing that the coefficients of the occuring polynomials generate the unit ideal, and drastically lower the depth of commutators. In particular, Preusser's idea allows to prove analogues of Brenner's lemma for groups of all types over arbitrary commutative rings, and *much more*.

In other words, decomposition of unipotents conveys *explicit short* polynomial factorisations of the root elements $gx_\alpha(\xi)g^{-1}$ — which are the obvious generators of $E(\Phi,R)^g$ — in terms of elementary generators. For *classical* groups Preusser's trick achieves also the opposite. It provides similar *explicit short* polynomial expressions of the elementary generators [of the lower level] of $g^{E(\Phi,R)}$, in terms of the obvious generators of this last group, in other words, of the elementary conjugates of g and g^{-1}. Thus, Preusser's papers are the first major advance in the direction of what can be dubbed the **reverse decomposition of unipotents**.

Immediately after understanding this idea, I and Zuhong Zhang were able to generalise it to *exceptional* groups as well, to subgroups normalised by the relative elementary subgroups $E(\Phi,R,I)$, corresponding to an ideal $I \trianglelefteq R$, and to some other situations. In particular, for Chevalley groups a slight refinement of the methods of [41, 42], and subsequent works gives us the following results. For an arbitrary commutative ring R and an arbitrary matrix $g \in G(\Phi,R)$ one can write explicit formulae, expressing an elementary generator $x_\alpha(\xi)$, where ξ belongs to the level of g, as products of at most $8 \cdot \dim(G)$ elementary conjugates of g and g^{-1}.

In § 2 we explain the original idea by Preusser in the simplest case of $\mathrm{GL}(n,R)$. In § 3 we sketch how with the methods of [41, 42] the same idea works — with the same length bounds! — for Chevalley groups of types E_6 and E_7. However, even in these (easy!) cases detailed proofs require some heed, and will be published elsewhere. Finally, in § 4 we mention further imminent applications of these ideas.

§2. Preusser's idea for $\mathrm{GL}(n,R)$

In this section, we explain the main idea of [26, 27] in the simplest example of the general linear group $G = \mathrm{GL}(n,R)$ of degree $n \geqslant 3$ over a commutative ring R. For this, we have to recall some notation.

We use some basic commutator calculus in groups. Our commutators are always left-normed, $[x,y] = xyx^{-1}y^{-1}$. Further, ${}^x y = xyx^{-1}$ and $y^x =$

$x^{-1}yx$ denote the left and the right conjugates of y by x, respectively. We use obvious commutator identities such as $[x,yz]=[x,y]\cdot {}^y[x,z]$ without any specific reference.

For an ideal $I\trianglelefteq R$ denote by $E(n,I)$ the corresponding elementary subgroup, generated by the elementary transvections of level I:

$$E(n,I)=\big\langle t_{ij}(\xi),\xi\in I,1\leqslant i\neq j\leqslant n\big\rangle.$$

Recall, that an [elementary] transvection $t_{ij}(\xi)$, corresponding to $\xi\in R$ and $1\leqslant i\neq j\leqslant n$, equals $t_{ij}(\xi)=e+\xi e_{ij}$. Here, as usual, e is the identity matrix and e_{ij} is a standard matrix unit. Further, the relative elementary subgroup $E(n,R,I)$ of level I is defined as the normal closure of $E(n,I)$ in the absolute elementary subgroup $E(n,R)$.

Let $g\in \mathrm{GL}(n,R)$ be an invertible matrix. It is written in terms of its entries as $g=(g_{ij})$, $1\leqslant i,j\leqslant n$. Entries of the inverse matrix $g^{-1}=(g'_{ij})$, $1\leqslant i,j\leqslant n$, are denoted by g'_{ij}. A matrix of the form $g^x=x^{-1}gx$, where $x\in E(n,R)$, is called an *elementary conjugate* of g.

By R^n we denote the free *right* R-module, consisting of columns of height n with components in R. The standard base in R^n (consisting of the columns of identity matrix e) is denoted by $e_1,\dots,e_n$. The group $G=\mathrm{GL}(n,R)$ acts on R^n by *left* multiplication. The stabiliser of the coordinate subspace $\langle e_1,\dots,e_m\rangle$ is called a [standard] parabolic [subgroup] and is denoted $P_m=\mathrm{Stab}_G\big(\langle e_1,\dots,e_m\rangle\big)$. Its conjugates are called parabolics of type P_m. In the field case it is indeed a maximal subgroup.

The subgroup of P_m generated by $t_{ij}(\xi)$, where $\xi\in R$, $1\leqslant i\leqslant m$, $m+1\leqslant j\leqslant n$, is denoted by U_m and is called the *unipotent radical* of P_m. Obviously, U_m is an *abelian* normal subgroup of P_m.

Further, consider the reduction homomorphism

$$\rho_I:\ \mathrm{GL}(n,R)\longrightarrow \mathrm{GL}(n,R/I)$$

modulo the ideal I. By definition, the principal congruence subgroup $\mathrm{GL}(n,R,I)$ is the kernel of ρ_I, whereas the full congruence subgroup $C(n,R,I)$ is the preimage of the centre of $\mathrm{GL}(n,R/I)$ under ρ_I.

Recall that the *upper level* of a matrix $g=(g_{ij})\in \mathrm{GL}(n,R)$ is the *smallest* ideal $I=\mathrm{lev}(g)$ such that $g\in C(n,R,I)$. Clearly, such an ideal is generated by the off-diagonal entries g_{ij}, $1\leqslant i\neq j\leqslant n$, and by the pairwise differences of its diagonal entries $g_{ii}-g_{jj}$, $1\leqslant i\neq j\leqslant n$. Clearly, it suffices to consider only the *fundamental* differences $g_{i+1,i+1}-g_{ii}$, where $i=1,\dots,n-1$. Thus, the upper level $\mathrm{lev}(g)$ is generated by n^2-1 elements, and by looking at the generic invertible matrix with commuting entries (say in the structure ring $\mathbb{Z}[\mathrm{GL}_n]$ of the affine group scheme GL_n) one immediately sees that this bound cannot be improved in general.

Further, denote by $g^{E(n,R)}$ the smallest $E(n,R)$-normalised subgroup of $\mathrm{GL}(n,R)$ containing g. The *lower level* $I = \mathrm{lol}(g)$ of a matrix $g \in \mathrm{GL}(n,R)$ is the *largest* ideal such that $E(n,R,I) \leqslant g^{E(n,R)}$. The standard description of $E(n,R)$-normalised subgroups (which holds, in particular, when R is commutative and $n \geqslant 3$) is *equivalent* to the claim that for any matrix $g \in \mathrm{GL}(n,R)$ its lower and the upper level coincide, $\mathrm{lol}(g) = \mathrm{lev}(g)$. This ideal is usually called simply the **level** of g.

The proof of the following result in [26], Theorem 12, does not depend on the standard description and can be regarded as its effective — stronger! — version.

Theorem 1. *Let R be commutative, $n \geqslant 3$, and $g \in \mathrm{GL}(n,R)$. Then for any $\xi \in \mathrm{lev}(g)$ and all $1 \leqslant i \neq j \leqslant n$ the elementary transvection $t_{ij}(\xi)$ is a product of $\leqslant 8(n^2-1)$ elementary conjugates of g and g^{-1}.*

In view of the above definition of level, this astonishing result asserts that any elementary generator $t_{ij}(g_{hk})$, where $1 \leqslant i \neq j \leqslant n$, $1 \leqslant h \neq k \leqslant n$, is a product of not more than 8 elementary conjugates of g and g^{-1}. Similarly, $t_{ij}(g_{hh} - g_{kk})$ require not more than 8 such elementary conjugates, modulo the previous generators. Of course, here we express the elementary generators of $E(n, \mathrm{lev}(g))$, rather than those of the relative subgroup $E(n,R,\mathrm{lev}(g))$ itself. However, passage to the $E(n,R)$-normalised subgroup simply amounts to another elementary conjugation, and does not change the number of such factors.

Recall that after [7] virtually all proofs of the standard description relied on the two following reductions[2]. First, the **level reduction** that asserts that to establish the standard description it suffices to prove that the lower level of any non-central matrix $g \in \mathrm{GL}(n,R)$ is $\neq 0$. Second, the **parabolic reduction** that asserts that the lower level of any non-central matrix g contained in a proper parabolic is $\neq 0$. Thus, one only non-trivial step was to prove that for any non-central g the intersection of $g^{E(n,R)}$ with some proper parabolic is non-central.

[2]This does not apply to the proofs by Igor Golubchik [16, 17], though. He worked in the situation where the standard commutator formulae were not known at the time, and had to rely on more oblique and sophisticated forms of level reduction. This alone would justify publication of updated versions of his proofs.

Now, the proof of the above result in [26] starts in essentially the same way, as the proofs of the standard description in [9, 32]. However, there is an important twist.

Recall that the proofs in [9, 32] both start as follows. Let i, j, h be any pair-wise distinct indices and r be any index between 1 and n. Set $x_r = t_{ij}(g'_{hr})t_{ih}(-g'_{jr})$. Then $x_r g^{-1}$ has the same r-th column as g^{-1} and, thus, $gx_r g^{-1}$ hits the parabolic subgroup $P = \mathrm{Stab}_G(\langle e_r\rangle)$ of type P_1. This shows that in the case $r \neq j, h$ the commutator $[x_r^{-1}, g] = x_r^{-1} \cdot gx_r g^{-1}$ sits in the same parabolic P.

If this commutator is non-central, we are done. If all such commutators are central, it imposes certain equations on the entries of our initial matrix g. For $\mathrm{GL}(n, R)$ it is easy to analyse these equations directly to conclude that then g itself was in a proper parabolic, this is exactly what is done in [9]. Entries of the above commutators are polynomials of degree 3 or 4 in the entries of g and g^{-1}. One had to do some work to lower the degree of the resulting equations, to eventually verify that certain entries of the matrix g were 0 from the very start.

The approach of [32] was more systematic. Namely, there it is observed that in the above calculation one can replace x_r by another elementary matrix $y_r = t_{ij}(g_{rh}g'_{hr})t_{ih}(-g_{rh}g'_{jr})$, whose parameters are multiples of those of x_r, and that (since R is commutative!) the product of y_r over $r = 1, \ldots, n$ equals $t_{ij}(1)$. Thus, for a matrix g, such that all $[y_r^{-1}, g]$ are central, also the commutator $[t_{ij}(1), g]$ is central, and (since $E(n, R)$ is perfect), g commutes with $t_{ij}(1)$. Again, this means that g itself was in a proper parabolic.

In these calculations we had to circumvent the cases $r \neq j, h$, where the commutator $[x_r^{-1}, g]$ was not sitting in a proper parabolic. However, retrospectively, this means exactly *rejecting the cornerstone*. Below I replicate the proof from [26], with minimal changes. To simplify the notation, we interchange g and g^{-1} in the above recapitulation. Thus, from now on $x_r = t_{ij}(g_{hr})t_{ih}(-g_{jr})$. The following argument by Preusser *in particular* provides a new proof of the standard description. This is the first such proof after Igor Golubchik [16, 17] that does not directly invoke standard commutator formulae. Unlike *all* preceding proofs — *including* those by Golubchik! — it does not hinge on *any* form of level reduction[3], but directly proves the equality $\mathrm{lol}(g) = \mathrm{lev}(g)$.

Proof. The commutator $[x_r^{-1}, g^{-1}] \in g^{E(n,R)}$ is the product of *two* elementary conjugates of g and g^{-1}. When $r = j$, the r-th column of this commutator differs from the column e_r of the identity matrix in exactly one position. Namely, its entry in the position (i,j) equals $-g_{hj}$. This means that even not being in the above parabolic $P = \mathrm{Stab}_G(\langle e_j\rangle)$, this commutator has the form $t_{ij}(-g_{hj})x$, for some $x \in P$.

Next, observe that for any $s \neq i, j$ the elementary transvection $t_{js}(1)$ sits in the unipotent radical U of the parabolic subgroup P. Obviously, $[xy, z]^x = [y, z] \cdot [x, z]^x = [y, z] \cdot [z, x^{-1}]$. Thus,

$$y = [t_{ij}(-g_{hj})x, t_{js}(-1)]^{t_{ij}(-g_{hj})} = [x, t_{js}(-1)] \cdot [t_{js}(-1), t_{ij}(g_{hj})] \in g^{E(n,R)}$$

is the product of *four* elementary conjugates of g and g^{-1}. In the above expression of y the first commutator $z = [x, t_{js}(-1)]$ belongs to the unipotent radical U, while the second commutator equals $t_{is}(g_{hj})$.

Finally, since $t_{ji}(1) \in U$ and U is abelian, one can conclude that

$$[t_{ji}(1), y] = [t_{ji}(1), z t_{is}(g_{hj})] = t_{js}(g_{hj}) \in g^{E(n,R)}$$

is the product of *eight* elementary conjugates of g and g^{-1}. Since j and h here are arbitrary, and the group $g^{E(n,R)}$ is normalised by the absolute elementary subgroup $E(n,R)$, which contains classes of all permutation matrices modulo the diagonal subgroup $\mathrm{diag}(\pm 1, \dots, \pm 1)$, this proves the claim for the off-diagonal entries.

To finish the proof of the theorem it only remains to observe that the entry of the matrix $g^{t_{hk}(1)} \in g^{E(n,R)}$ in the position (h,k) equals $g_{hk} + g_{hh} - g_{kk} - g_{kh}$, it follows from the above that $t_{ij}(g_{hk} + g_{hh} - g_{kk} - g_{kh}) \in g^{E(n,R)}$ is the product of *eight* elementary conjugates of g and g^{-1}. Since $t_{ij}(g_{hk})$ and $t_{ij}(g_{kh})$ are already accounted for, it follows that we need at most *eight* further elementary conjugates of g and g^{-1} to express $t_{ij}(g_{hh} - g_{kk}) \in g^{E(n,R)}$. $\square$

[3] Almost simultaneously Alexei Stepanov [31] proposed another *drastic* simplification of the known proofs for standard description, kind of a *universal* level reduction [30]. It was one of the major tools in the proof of normal structure for isotropic reductive groups in the recent joint paper by Stavrova and Stepanov [28]. However, Stepanov's idea seems to me to be an upswing in a completely different direction. As of today, it only works for algebraic groups. Also, when Preussers approach can be effectuated, it gives more specific results.

§3. Reverse decomposition of unipotents in Chevalley groups of types E_6 and E_7

Now I briefly sketch how this idea works, without any major remoulding, for Chevalley groups of simply laced type Φ. For $\Phi = \mathrm{D}_l$ this is already done by Preusser himself in [26], Theorem 27. Remarkably, for $\Phi = \mathrm{E}_6$ and $\Phi = \mathrm{E}_7$ we have to use the initial A_5-proof and A_7-proof, respectively, see [41, 42], rather than the subsequent technically less demanding A_2-proofs and their descendants, [44, 46, 47]. Reverse decomposition of unipotents works also for the type $\Phi = \mathrm{E}_8$. In fact in this case it is based on the original D_8-proof, and is *easier* than direct decomposition of unipotents. However, many details of proofs are rather different from the cases E_6 and E_7, and will be described separately.

I use the same notation as in [41, 42, 45]. We fix a reduced irreducible root system Φ of rank $l = \mathrm{rk}(\Phi)$ and a commutative ring R. As above, $G(\Phi, R)$ denotes the simply connected Chevalley group of type Φ over R. We choose an order on Φ and let $\Pi = \{\alpha_1, \dots, \alpha_l\}$ be the corresponding set of fundamental roots, while $s_1, \dots, s_l$ are the correspodning fundamental reflections in the Weyl group $W(\Phi)$. Our numbering of the fundamental roots complies with [10]. In the sequel, δ denotes the maximal root corresponding to this choice of Π.

Further, we fix a split maximal torus $T(\Phi, R)$ in $G(\Phi, R)$ and parametrisations of the root subgroups X_α, $\alpha \in \Phi$, elementary with respect to this torus. The elements $x_\alpha(\xi)$, $\alpha \in \Phi$, $\xi \in R$, are called *elementary root unipotents*. For an ideal $I \trianglelefteq R$ we consider the elementary subgroup

$$E(\Phi, I) = \big\langle x_\alpha(\xi) \mid \alpha \in \Phi,\ \xi \in I \big\rangle$$

and denote by $E(\Phi, R, I)$ its normal closure in the absolute elementary subgroup $E(\Phi, R)$.

In the sequel, we view the groups $G = G(\Phi, R)$ as linear groups acting in a rational representation (V, π). Usually, we identify an element $g \in G$ with its image $\pi(g) \in \mathrm{GL}(V)$ in this representation. The most handy for $\Phi = \mathrm{E}_6, \mathrm{E}_7$ are the fundamental representations $V = V(\varpi_1)$ and $V = V(\varpi_7)$, respectively. These are faithful microweight representations of dimensions 27 and 56, respectively, and we denote by $\Lambda(\pi)$ the corresponding sets of weights, which are all extremal, and thus of multiplicity one.

We fix a crystal base v^λ, $\lambda \in \Lambda(\pi)$, in V, which is a positive admissible base. With respect to this base an element $g \in G(\Phi, R)$ may be identified with its matrix $g = (g_{\lambda\mu})$, $\lambda, \mu \in \Lambda(\pi)$, in $\mathrm{GL}(27, R)$ or $\mathrm{GL}(56, R)$, depending on whether $\Phi = \mathrm{E}_6$ or E_7. The μ-th column of g will be denoted by $g_{*,\mu}$.

Now, the first significant departure from the linear case occurs. There, the Weyl group acted transitively on pairs of distinct weights. This is not the case anymore. In a milder form this additional complication was already visible in the case of the even orthogonal group $\mathrm{SO}(2n, R)$, also considered by Preusser in [26].

In fact, there are three orbits of $W(\mathrm{E}_6)$ on pairs of weights in $\Lambda(\varpi_1)$ and four such orbits of $W(\mathrm{E}_7)$ on pairs of weights in $\Lambda(\varpi_7)$, distinguished by the distance $d(\lambda, \mu)$ between λ and μ in the *weight graph* (not the weight diagram!). Namely, for E_6 two weights are i) equal $\lambda = \mu$, in which case $d(\lambda, \mu) = 0$; ii) *adjacent* $\lambda - \mu \in \Phi$, in which case $d(\lambda, \mu) = 1$; or iii) *distant*, $\lambda - \mu \notin \Phi \cup \{0\}$, which implies that there is a weight ν adjacent to both, in which case $d(\lambda, \mu) = 2$. For E_7 yet another possibility can occur, when λ and μ are iv) *opposite*, $\lambda = -\mu$, in which case $d(\lambda, \mu) = 3$.

However, the *lower* level $I = \mathrm{lol}(g)$ of an element $g \in G(\Phi, R)$ is described exclusively in terms of elementary unipotents $x_\alpha(\xi)$, $\alpha \in \Phi$, $\xi \in R$, belonging to the $E(\Phi, R)$-normalised subgroup generated by g. As above, it is just the largest ideal $I \trianglelefteq R$ such that $E(\Phi, R, I) \leqslant g^{E(\Phi,R)}$.

On the other hand, it *seems* that the *upper* level $I = \mathrm{lev}(g)$ of g has to be defined in terms of *all* pairs of weights (λ, μ). However, this is not the case. As for $\mathrm{GL}(n, R)$ the upper level is fully determined by those matrix positions that actually do occur in the corresponding Lie algebra. In other words, by the positions corresponding to the pairs of weights (λ, μ) at distance $d(\lambda, \mu) \leqslant 1$. In a suitable admissible base the diagonal positions exhibit the elements of the Cartan subalgebra, corresponding to our choice of maximal torus, whereas positions (λ, μ) at distance 1 display respective root elements, associated with the root $\alpha = \lambda - \mu$.

A construction of (generalised) congruence subgroups based on this observation was essentially contained already in our joint paper with Eugene Plotkin [51]. Following Zenon Borewicz, we considered *nets* of ideals $\sigma = (\sigma_\alpha)$, $\alpha \in \Phi$, subject to the condition $\sigma_\alpha \sigma_\beta \subseteq \sigma_{\alpha+\beta}$, for all $\alpha, \beta, \alpha+\beta \in \Phi$. But in that paper no congruences were imposed on diagonal entries. In [51] this simplifying condition was expressed as $\sigma_{\lambda\lambda} = R$, for all diagonal ideals of the corresponding net. Also, in [51] we required that the structure constants are invertible in R. A general construction of congruence subgroups

$G(\Phi, R, A, B)$ and $C(\Phi, R, A, B)$, in terms of such congruences was later given in [20].

Using the constructions from [20, 51] one can check that to define the upper level of $g = (g_{\lambda\mu})$, $\lambda, \mu \in \Lambda(\pi)$, one can take the following elements:

- $g_{\lambda\mu}$ for *any* set of pairs (λ, μ), one with the difference $\alpha = \lambda - \mu$, for each root $\alpha \in \Phi$;

- $g_{\lambda\lambda} - g_{\mu\mu}$ for *any* set of pairs (λ, μ), one with the difference $\alpha_i = \lambda - \mu$, for each fundamental root $\alpha_i \in \Pi$.

Checking correctness of this definition is a rather straightforward but tedious exercise. Now we are all set to state the main new result of the present paper.

Theorem 2. *Let $g \in \mathrm{G}(\Phi, R)$, where $\Phi = \mathrm{E}_6$ or $\Phi = \mathrm{E}_7$, and let $I \trianglelefteq R$ be the level of g. Then for any $\xi \in I$ and all $\alpha \in \Phi$ the elementary root unipotent $x_\alpha(\xi)$ is the product of $\leqslant 8{\cdot}78$ or $\leqslant 8{\cdot}133$ elementary conjugates of g and g^{-1}, respectively.*

Sketch of proof. The strategy of the proof is exactly the same as for the case $\mathrm{GL}(n, R)$ above. But instead of the elements x_r sitting in a P_1 parabolic of the subgroup $\mathrm{SL}(3, R) \leqslant G(\mathrm{E}_6, R)$ of type $\Delta = \mathrm{A}_2$, now we have to consider classical subgroups of much larger rank. Namely, a subgroup of type $\Delta = \mathrm{A}_5$ in $G(\mathrm{E}_6, R)$ and a subgroup of type $\Delta = \mathrm{A}_7$ in $G(\mathrm{E}_7, R)$, see [41, 42, 47].

Since the Weyl group $W(\Phi)$ acts transitively on roots, for the purpose of exposition one may fix a specific choice of Δ. For E_7 it is exactly the same choice as in the above papers, $\Delta = \langle \delta, \alpha_1, \alpha_3, \alpha_4, \alpha_5, \alpha_6, \alpha_7 \rangle$. However, in the meantime I noticed that for E_6 many details of computation become slightly easier for $\Delta = \langle \alpha_1, \alpha_3, \alpha_4, \alpha_5, \alpha_6 \rangle$, whereas for historical reasons all of the above papers used $\Delta = \langle \delta, \alpha_2, \alpha_4, \alpha_5, \alpha_6 \rangle$.

In the case of E_6 the elements x_r, $1 \leqslant r \leqslant n$, in the above argument for $\mathrm{GL}(n, R)$ are replaced by

$$x_\rho = x_{\beta_1}(z_1) x_{\beta_2}(z_2) x_{\beta_3}(z_3) x_{\beta_4}(z_4) x_{\beta_5}(z_5),$$

one for each weight $\rho \in \Lambda(\varpi_1)$, where

$$\beta_1 = \alpha_1, \quad \beta_2 = \alpha_1 + \alpha_2, \quad \beta_3 = \alpha_1 + \alpha_2 + \alpha_3,$$
$$\beta_4 = \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4, \quad \beta_5 = \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 + \alpha_5,$$

whereas

$$z_1 = g_{\tau,\rho}, \quad z_2 = \pm g_{\tau+\beta_3,\rho}, \quad z_3 = \pm g_{\tau+\beta_3+\beta_4,\rho},$$
$$z_4 = \pm g_{\tau+\beta_3+\beta_4+\beta_5,\rho}, \quad z_5 = \pm g_{\tau+\beta_3+\beta_4+\beta_5+\beta_6,\rho}.$$

Here $\tau = -\varpi_6 + \alpha_6 + \alpha_5 + \alpha_4 + \alpha_2$, while the signs are choosen in such a way that all additions of the components of the column $g_{*,\rho}$ in positions

$$\tau, \quad \tau + \alpha_3, \quad \tau + \alpha_3 + \alpha_4, \quad \tau + \alpha_3 + \alpha_4 + \alpha_5, \quad \tau + \alpha_3 + \alpha_4 + \alpha_5 + \alpha_6$$

to the ten components of the column $g_{*,\rho}$ in positions between $\varpi_1 - \alpha_1 - \alpha_2 - \alpha_3 - \alpha_4$ and $\varpi_1 - \delta + \alpha_2$ pair-wise cancel. That such a choice is possible is established as part of the proof of [42], Theorem 5. The two other components of $g_{*,\rho}$ affected by x_ρ are those in positions ϖ_1 and $\varpi_1 - \delta$. However, a direct computation shows that the expressions added to these components are

$$\pm g_{\tau+\alpha_3+\alpha_4+\alpha_5+\alpha_6,\rho} g_{-\varpi_6,\rho}$$
$$\pm g_{\tau+\alpha_3+\alpha_4+\alpha_5,\rho} g_{-\varpi_6+\alpha_6,\rho} \pm g_{\tau+\alpha_3+\alpha_4,\rho} g_{-\varpi_6+\alpha_5+\alpha_6,\rho}$$
$$\pm g_{\tau+\alpha_3,\rho} g_{-\varpi_6+\alpha_4+\alpha_5+\alpha_6,\rho} \pm g_{\tau,\rho} g_{-\varpi_6+\alpha_3+\alpha_4+\alpha_5+\alpha_6,\rho}$$

and, respectively,

$$\pm g_{\tau,\rho} g_{\varpi_1-\alpha_1,\rho} \pm g_{\tau+\alpha_3,\rho} g_{\varpi_1-\alpha_1-\alpha_3,\rho} \pm g_{\tau+\alpha_3+\alpha_4,\rho} g_{\varpi_1-\alpha_1-\alpha_3-\alpha_4,\rho}$$
$$\pm g_{\tau+\alpha_3+\alpha_4+\alpha_5,\rho} g_{\varpi_1-\alpha_1-\alpha_3-\alpha_4-\alpha_5,\rho}$$
$$\pm g_{\tau+\alpha_3+\alpha_4+\alpha_5+\alpha_6,\rho} g_{\varpi_1-\alpha_1-\alpha_3-\alpha_4-\alpha_5-\alpha_6,\rho}.$$

Up to signs, these are exactly two of the 27 quadratic equations defining the highest weight orbit in $V(\varpi_1)$, see [42] and further details in [43, 48]. Again, it is verified as part of the proof of [42], Theorem 5, that the above choice of signs is compatible with the signs with which terms occur in these equations. Thus, multiplication by x_ρ does not alter $g_{*,\rho}$. In other words, $g^{-1}x_\rho g$ sits in a proper parabolic subgroup P of type P_1.

Since $W(\mathrm{E}_6)$ acts transitively on pairs of adjacent weights in $\Lambda(\varpi_1)$, one can for the purposes of exposition take $\rho = \varpi - \alpha_1$. Then $P = P_1^{s_1}$ and, exactly as in the case of $\mathrm{GL}(n, R)$ the commutator $[x_{\varpi-\alpha_1}^{-1}, g^{-1}]$ can be written as $x_{\alpha_1}(-g_{\tau,\varpi-\alpha_1})x$, for some $x \in P$.

Now, since the unipotent radical of P is abelian, one can easily emulate the same argument, as in the proof of Preusser's Theorem, to express the elementary unipotent $x_{\alpha_1+\alpha_3}(g_{\tau,\varpi-\alpha_1})$ as the product of 8 elementary conjugates of g and g^{-1}. Since $W(\mathrm{E}_6)$ transitively acts on pairs of weights

(λ,μ) at the same distance, whereas the group $g^{E(\Phi,R)}$ is normalised by $E(\Phi,R)$, this shows that $x_\alpha(g_{\lambda,\mu})$ is a product of 8 elementary conjugates of g and g^{-1}, for any pair of weights (λ,μ) such that $d(\lambda,\mu)=2$.

Now, one could finish the proof of Theorem 2, with the coefficient 8 replaced by 16, in exactly the same way, as the proof of Theorem 1. Namely, passing from g to $g^{x_\gamma(1)}$, one can in the same way conlcude, that $x_\alpha(g_{\lambda,\mu})$ is a product of 16 elementary conjugates of g and g^{-1} for any pair of weights (λ,μ) such that $d(\lambda,\mu)=1$. Similarly, the elements $x_\alpha(g_{\lambda,\lambda}-g_{\mu,\mu})$, where $d(\lambda,\mu)=2$, are expressed by 8 additional elementary conjugates, which then would give 16 additional elementary conjugates to express $x_\alpha(g_{\lambda,\lambda}-g_{\mu,\mu})$ for $d(\lambda,\mu)=1$.

In the case of E_7 to stabilise the ρ-th column of g we again take exactly the same elements $x_\rho=x_{\beta_1}(z_1)x_{\beta_2}(z_2)\dots x_{\beta_7}(z_7)$, as in [41,42,47]. These elements sit in the unipotent radical of a parabolic subgroup of type P_1 in $G(\Delta,R)$, where Δ is a subsystem of type A_7 in E_7. Their parameters z_i are again (up to sign) exactly respective components of $g_{*,\rho}$. Then calculations in $V(\varpi_7)$ allow us to get in a parabolic of type P_7 in $G(\mathrm{E}_7,R)$, see [42]. Again, these calculations rely both on the explicit knowledge of action structure constants, and of the equations on the highest weight orbit, see [42,43,49]. The apposite choice of ρ then allows to extract individual elementary unipotents to conclude that $x_\alpha(g_{\lambda,\mu})$ is a product of 8 elementary conjugates of g and g^{-1}, for any pair of weights (λ,μ) such that $d(\lambda,\mu)=2$.

To get the theorem with the coefficient 8 one has to work a little harder, by looking at a different column and a different parabolic, allowing more than one addition from the opposite unipotent radical. The details are routine and elementary, but rather lengthy, and will be published elsewhere.

§4. Further variations and final remarks

Similar, but slightly fancier arguments work also for E_8. In fact, in this case the reverse decomposition of unipotents is *easier* than the usual one. The reason is that working in the adjoint representation $V(\varpi_8)$ to get the usual decomposition of unipotents we have to stabilise not just the columns corresponding to roots, but also the columns corresponding to zero weights! In several respects, this is technically more demanding. Firstly, these columns satisfy only part of the equations on the highest weight orbit, see [4,19]. Secondly, their stabilisers are not parabolics, which requires much trickier versions of Whitehead's lemma, and affects the resulting bounds.

Since in the reverse decomposition of unipotents everything is considered up to an elementary conjugation, we can limit ourselves to parabolics of type P_8. Thus, the usual D_8-proof sketched in [41, 52] suffices to obtain the analogues of Theorems 1 and 2. Also, since in this case the unipotent radical is special, and its centre commutes with the commutator of the Levi subgroup, this allows to spare one commutator at the final stage, to give the same overall bound, as in the cases E_6 and E_7.

As another variation, I could mention that description of subnormal subgroups requires essentially the same calculations, but with the absolute elementary subgroup $E(\Phi, R)$ replaced by the *relative* elementary subgroup $E(\Phi, R, I)$, corresponding to an ideal $I \trianglelefteq R$. In other words, now we are interested in the relation between the upper level, and the lower level of the subgroup $g^{E(\Phi,R,I)}$. However, in general $E(\Phi, R, I)$ does not contain elements that act as Weyl group elements.

This means that expressing an elementary unipotent $x_\alpha(g_{\lambda,\mu})$ as the product of a certain number of elementary conjugates of g and g^{-1}, for some root α, does not necessarily imply that the elementary unipotent $x_\beta(g_{\lambda,\mu})$ can be expressed as the product of the same number of such elementary conjugates, for another root β. One has to use the Chevalley commutator formula instead, which with each commutation doubles the number of requisite elementary conjugates. The details are described in our joint papers with Zuhong Zhang, where we establish similar formulae at the relative level.

The author thanks Victor Petrov and Alexei Stepanov for extremely illuminating and productive discussions of Preusser's work. I am very grateful to Alex Lubotzky, who explained me the importance of what they call Brenner's lemma in the study of word maps, and further related areas. Last, but not least, I am very indebted to Zuhong Zhang, for ongoing discussion of the relative case, over the last years.

References

1. E. Abe, *Chevalley groups over commutative rings.* — In: Proc. Conf. Radical Theory (Sendai, 1988), Uchida Rokakuho Publ. Comp., Tokyo, 1989, pp. 1–23.
2. E. Abe, *Normal subgroups of Chevalley groups over commutative rings.* — Contemp. Math. **83** (1989), 1–17.
3. E. Abe, K. Suzuki, *On normal subgroups of Chevalley groups over commutative rings.* — Tôhoku Math. J. **28** (1976), No. 1 185–198.

4. M. M. Atamanova, A. Yu. Luzgarev, *Cubic forms on adjoint representations of exceptional groups.* — J. Math. Sci. **222** (2017), No. 4, 370–379.
5. A. Bak, *The stable structure of quadratic modules,* Ph. D. Thesis, Columbia University, 1969.
6. A. Bak, N. Vavilov, *Structure of hyperbolic unitary groups.* I. *Elementary subgroups.* — Algebra Colloq. **7** (2000), No. 2, 159–196.
7. H. Bass, K-*theory and stable algebra.* — Inst. Hautes Études Sci. Publ. Math. (1964), No. 22, 5–60.
8. H. Bass, *Unitary algebraic K-theory.* — Lecture Notes Math. **343** (1973), 57–265.
9. Z. I. Borewicz, N. A. Vavilov, *The distribution of subgroups in the full linear group over a commutative ring.* — Proc. Steklov Inst. Math. **3** (1985), 27–46.
10. N. Bourbaki, *Groupes et Algébres de Lie.* Chapitres 4–6, Hermann, Paris (1968).
11. J. L. Brenner, *The linear homogeneous group.* III. — Ann. Math. **71** (1960), 210–223.
12. D. L. Costa, G. E. Keller, *Radix redux: normal subgroups of symplectic group.* — J. reine angew. Math. **427** (1992), No. 1, 51–105.
13. D. L. Costa, G. E. Keller, *On the normal subgroups of* $G_2(A)$. — Trans. Amer. Math. Soc. **351**, (1999), No. 12, 5051–5088.
14. I. Z. Golubchik, *On the general linear group over an associative ring.* — Uspehi Mat. Nauk **28** (1973), no. 3, 179–180.
15. I. Z. Golubchik, *On the normal subgroups of orthogonal group over an associative ring with involution.* — Uspehi Mat. Nauk **30** (1975), No. 6, 165.
16. I. Z. Golubchik, *The normal subgroups of linear and unitary groups over rings.* Ph. D. Thesis, Moscow State University, 1981.
17. I. Z. Golubchik, *On the normal subgroups of the linear and unitary groups over associative rings.* — In: Spaces over Algebras and Some Problems in the Theory of Nets, Ufa, 1985, pp. 122–142.
18. A. J. Hahn, O. T. O'Meara, *The classical groups and* K-*theory*, Springer, Berlin et al., 1989.
19. A. Yu. Luzgarev, *Equations determining the orbit of the highest weight vector in the adjoint representation,* `arXiv:1401.0849v1 [math.AG]` 4 Jan 2014.
20. R. Hazrat, V. Petrov, N. Vavilov, *Relative subgroups in Chevalley groups.* — J. K-Theory **5** (2010), No. 3, 603–618.
21. R. Hazrat, A. Stepanov, N. Vavilov, Zuhong Zhang, *Commutator width in Chevalley groups.* — Note di Matematica **33** (2013), No. 1, 139–170.
22. R. Hazrat, N. Vavilov, *Bak's work on the* K*-theory of rings.* — J. K-Theory **4** (2009), No. 1, 1–65.
23. V. I. Kopeiko, *Stabilization of symplectic groups over a ring of polynomials.* — Math. USSR Sb. **34** (1978), No. 5, 655–669.
24. V. A. Petrov, *Decomposition of transvections: An algebro-geometric approach.* — St. Petersburg J. Math. **28** (2017), No. 1, 109–114.
25. R. Preusser, *Structure of hyperbolic unitary groups* II. *Classification of* E*-normal subgroups.* — Algebra Colloq. **24** (2017), No. 2, 195–232.
26. R. Preusser, *Sandwich classification for* $\mathrm{GL}_n(R)$, $O_{2n}(R)$ *and* $U_{2n}(R,\Lambda)$ *revisited.* — J. Group Theory **21** (2018), No. 1, 21–44.

27. R. Preusser, *Sandwich classification for $O_{2n+1}(R)$ and $U_{2n+1}(R,\Delta)$ revisited.* — J. Group Theory **21** (2018), No. 4, 539–571.
28. A. Stavrova, A. Stepanov, *Normal structure of isotropic reductive groups over rings*, `arXiv:1801.08748v1 [math.GR]` 26 Jan 2018.
29. A. Stepanov, *Stability conditions in the theory of linear groups over rings*, Ph. D. Thesis, Leningrad State Univ. (1987), 1–112.
30. A. Stepanov, *Structure of Chevalley groups over rings via universal localization.* — J. Algebra **450** (2016), 522–548.
31. A. V. Stepanov, *A new look at the decomposition of unipotents and the normal structure of Chevalley groups.* — St. Petersburg J. Math. **28** (2017), No. 3, 411–419.
32. A. Stepanov, N. Vavilov, *Decomposition of transvections: a theme with variations.* — K-Theory **19** (2000), No. 2, 109–153.
33. A. A. Suslin, *The structure of the special linear group over polynomial rings.* — Math. USSR Izv. **11** (1977), No. 2, 235–253.
34. A. A. Suslin, V. I. Kopeiko, *Quadratic modules and the orthogonal group over polynomial rings.* — J. Soviet Math. **20** (1982), No. 6, 2665–2691.
35. G. Taddei, *Normalité des groupes élémentaires dans les groupes de Chevalley sur un anneau.* — Contemp. Math. **55** (1986), 693–710.
36. L. N. Vaserstein, *On the normal subgroups of the* GL_n *of a ring.* — Algebraic K-Theory, Evanston 1980, Lecture Notes in Math., vol. 854, Springer, Berlin et al., 1981, pp. 454–465.
37. L. N. Vaserstein, *On normal subgroups of Chevalley groups over commutative rings.* — Tôhoku Math. J. **36**, No. 5 (1986), 219–230.
38. L. N. Vaserstein, *The subnormal structure of general linear groups over rings.* — Math. Proc. Cambridge Phil. Soc. **108**, No. 2 (1990), 219–229.
39. L. Vaserstein, Hong You, *Normal subgroups of classical groups over rings.* — J. Pure Appl. Algebra **105** (1995), 93–106.
40. N. A. Vavilov, *Subgroups of split classical groups.* Habilitationsschrift Leningrad State Univ. (1987), 1–334.
41. N. Vavilov, *Structure of Chevalley groups over commutative rings.* — In: Nonassociative algebras and related topics (Hiroshima, 1990), World Sci. Publ., London et al., 1991, pp. 219–335.
42. N. Vavilov, *A third look at weight diagrams.* — Rend. Sem. Mat. Univ. Padova, **104** (2000), 201–250.
43. N. A. Vavilov, *Numerology of square equations.* — St. Petersburg Math. J. **20**, No. 5 (2009), 687–707.
44. N. A. Vavilov, *An* A_3*-proof of structure theorems for Chevalley groups of types* E_6 *and* E_7. II. *Fundamental lemma.* — St. Petersburg Math. J. **23**, No. 6 (2012), 921–942.
45. N. A. Vavilov, *Decomposition of unipotents for* E_6 *and* E_7*: 25 years after.* — J. Math. Sci. **219**, No. 3 (2016), 355–369.
46. N. A. Vavilov, M. R. Gavrilovich, A_2*-proof of structure theorems for Chevalley groups of types* E_6 *and* E_7. — St. Petersburg Math. J. **16**, No. 4 (2005), 649–672.
47. N. A. Vavilov, M. R. Gavrilovich, S. I. Nikolenko, *Structure of Chevalley groups: the proof from the Book.* — J. Math. Sci. **140**, No. 5 (2007), 626–645.

48. N. A. Vavilov, A. Yu. Luzgarev, *Normalizer of the Chevalley group of type* E_6. — St. Petersburg Math. J. **19**, No. 5 (2008), 699–718.
49. N. A. Vavilov, A. Yu. Luzgarev, *Normalizer of the Chevalley group of type* E_7. — St. Petersburg Math. J. **27**, No. 6 (2015), 899–921.
50. N. A. Vavilov, S. I. Nikolenko, A_2*-proof of structure theorems for the Chevalley group of type* F_4. — St. Petersburg Math. J. **20**, No. 4 (2009), 527–551.
51. N. A. Vavilov, E. B. Plotkin, *Net subgroups of Chevalley groups.* — J. Sov. Math. **19**, No. 1 (1982), 1000–1006.
52. N. A. Vavilov, E. B. Plotkin, *Chevalley groups over commutative rings.* I. *Elementary calculations.* — Acta Appl. Math., **45** (1996), 73–113.
53. N. A. Vavilov, E. B. Plotkin, A. V. Stepanov, *Calculations in Chevalley groups over commutative rings.* — Soviet Math. Dokl. **40**, No. 1 (1990), 145–147.
54. N. A. Vavilov, A. V. Stepanov, *Linear groups over general rings* I. *Generalities.* — J. Math. Sci. (2012), 1–107.
55. J. S. Wilson, *The normal and subnormal structure of general linear groups.* — Proc. Camb. Philos. Soc. **71** (1972), 163–177.

St. Petersburg State University
E-mail: `nikolai-vavilov@yandex.ru`

Поступило 12 ноября 2018 г.