Записки научных
семинаров ПОМИ
Том 484, 2019 г.

N. Vavilov

TOWARDS THE REVERSE DECOMPOSITION OF UNIPOTENTS. II. THE RELATIVE CASE

Abstract. Recently Raimund Preusser displayed very short polynomial expressions of elementary generators in classical groups over an arbitrary commutative ring as products of conjugates of an arbitrary matrix and its inverse by absolute elementary matrices. In particular, this provides very short proofs for description of normal subgroups. In [27] I discussed various generalisations of these results to exceptional groups, specifically those of types E_6 and E_7. Here, I produce a further variation of Preusser's wonderful idea. Namely, in the case of $GL(n,R)$, $n \geqslant 4$, I obtain similar expressions of elementary transvections as conjugates of $g \in GL(n,R)$ and g^{-1} by *relative* elementary matrices $x \in E(n,J)$ and then $x \in E(n,R,J)$, for an ideal $J \trianglelefteq R$. Again, in particular, this allows to give very short proofs for the description of subgroups normalised by $E(n,J)$ or $E(n,R,J)$ – and thus also of subnormal subgroups in $GL(n,R)$.

What bird has done yesterday, the man may do next year, be it fly, be it moult, be it hatch, be it nest.
James Joyce, Finnegans wake

§1. Introduction

Let R be an commutative ring with 1, $GL(n,R)$ be the general linear group of degree n over R. As usual, e denotes the identity matrix, whereas e_{ij}, $1 \leqslant i,j \leqslant n$, denotes a standard matrix unit. For $\xi \in R$ and $1 \leqslant i \neq j \leqslant n$ one denotes by $t_{ij}(\xi) = e + \xi e_{ij}$ the corresponding elementary transvection. To an ideal $I \trianglelefteq R$, one assigns the elementary subgroup

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

In turn, 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 group $E(n,R)$.

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.

Consider the reduction homomorphism $\rho_I : \mathrm{GL}(n,R) \longrightarrow \mathrm{GL}(n,R/I)$ modulo I. Then the *principal* congruence subgroup $\mathrm{GL}(n,R,I)$ is the kernel of the reduction homomorphism ρ_I, whereas the *full* congruence subgroup $C(n,R,I)$ is the full pre-image of the centre of $\mathrm{GL}(n,R/I)$, with respect to ρ_I.

We need also the less familiar *brimming* congruence subgroup $G(n,R,I)$, which is the full preimage of the diagonal subgroup $D(n,R/I) \leqslant \mathrm{GL}(n,R/I)$. In the terminology of Zenon Borewicz, $G(n,R,I) = G(\sigma)$ is the net subgroup corresponding to the D-net $\sigma = (\sigma_{ij})$, $1 \leqslant i \neq j \leqslant n$, such that $\sigma_{ij} = I$ for all $i \neq j$, while $\sigma_{ii} = R$ as they should be, for D-nets, see [6,7].

In the present paper we are interested in description of subgroups of $\mathrm{GL}(n,R)$, normalised either by the elementary subgroup $E(n,J)$, or by the relative elementary subgroup $E(n,R,J)$, for some ideal $J \trianglelefteq R$.

As discovered by J. Wilson [31], for subgroups normalised by $E(n,R,J)$ the answer may be stated as follows: there exists an integer m, with the following property. For any subgroup $H \leqslant \mathrm{GL}(n,R)$ normalised by $E(n,R,J)$ there exists an ideal $I \trianglelefteq R$ such that

$$E(n,R,J^mI) \leqslant H \leqslant C(n,R,I).$$

This line of research was then pursued in many further papers, see below.

For subgroups $H \leqslant \mathrm{GL}(n,R)$ normalised by $E(n,J)$ the answer is similar, but is stated in terms of different subgroups, the lower layer being smaller, while the upper layer being larger, than for subgroups normalised by the larger group $E(n,R,J)$, namely

$$E(n,J^mI) \leqslant H \leqslant G(n,R,I).$$

This is precisely what was established inside the proofs of results on subgroups normalised by $E(n,R,J)$, even if it was not stated this way.

Unlike the absolute case, here the ideal I is not unique as such. But it is unique up to the equivalence relation $\Diamond_J$, which is described as follows. Let A and B be two ideals of the ring R. We set $A\Diamond_J B$ if there exist such integers r,s that $J^rA \leqslant B$ and $J^sB \leqslant A$, see [1,15].

By a standard argument due to John Wilson [31], this description implies also the description of subnormal subgroups in $\mathrm{GL}(n,R)$. Namely, let $H \trianglelefteq_d \mathrm{GL}(n,R)$ be a subnormal subgroup of depth d. Then

$$E(n,R,I^r) \leqslant H \leqslant C(n,R,I),$$

for some ideal I and some $r \leqslant (m^d-1)/(m-1)$.

Technically, the main issue is to find the smallest possible value of m, in the above answer. Historically, the published estimates in chronological order were as follows[1]:

- $m \leqslant 7$ for $n \geqslant 4$, John Wilson, 1972 [31],
- $m \leqslant 24$ (under some stability conditions), Anthony Bak, 1982 [1],
- $m \leqslant 6$, Leonid Vaserstein, 1986 [23],
- $m \leqslant 40$, Li Fuan and Liu Mulan, 1987 [15],
- $m \leqslant 5$, the present author, 1990 [26],
- $m \leqslant 4$, Vaserstein, 1990 [24].

Since there are examples where the above inclusions do not hold with $m = 2$, see [23], it only remains to ascertain, whether the correct bound is $m = 3$ or $m = 4$.

In the present paper I start chasing the *constructive* versions of the above results, with the best possible bounds. With this end, we have to start clearly distinguishing $E(n, J)$ and $E(n, R, J)$ in our results. For a matrix $g \in \mathrm{GL}(n, R)$ we denote by $\mathrm{lev}(g)$ the *upper level* of g, generated by its entries g_{ij}, $1 \leqslant i \neq j \leqslant n$ outside of the principal diagonal *and* by the pairwise differences $g_{ii} - g_{jj}$, $i \neq j$, of its diagonal entries. As opposed to that, the *outer diagonal upper level* of g, denoted by $\mathrm{leo}(g)$, is generated by its outer diagonal entries g_{ij}, $1 \leqslant i \neq j \leqslant n$, *alone*, see § 2 for notation and precise definitions of all requisite concepts.

The following result is a partial relative analogue of [27], Theorem 1. The full such analogue will be established in Theorem 4 below.

Theorem 1. *Let R be a commutative ring, $J \trianglelefteq R$ be an ideal of R, $n \geqslant 4$. Further, let $g \in \mathrm{GL}(n, R)$. Then for any $\xi \in \mathrm{leo}(g)$, any five elements $a_1, \ldots, a_5 \in J$, and all $1 \leqslant i \neq j \leqslant n$ the elementary transvection $t_{ij}(a_1 \ldots a_5 \xi)$ is a product of $\leqslant 32(n^2 - n)$ conjugates of g and g^{-1} by elements of $E(n, J)$.*

Of course, one would wish to give an absolute bound that applies to all ideals $J \trianglelefteq R$ and to all $\xi \in J^5 \,\mathrm{leo}(g)$. However, this is not possible. In the absolute case the ideal R is principal, here the answer depends on the number of generators of J. Let J be generated by d elements $b_1, \ldots, b_d$,

[1]Clearly, [1] and [15] drop out of the mainstream. The reason is that [1] was published some 15 years after completion, and [15] relied upon [1]. Nevertheless, these papers are very pertinent in what concerns discussion of the relative commutator formulae and the equivalence relation $\Diamond_J$.

then J^m is generated by at most $\binom{d+m-1}{m}$ products with repetitions of m among $b_1, \dots, b_d$. In particular, it follows from Theorem 1 that for an ideal J generated by d elements, $t_{ij}(\xi)$, $\xi \in J^5 \operatorname{leo}(g)$, is the product of $\leqslant 32\binom{d+4}{5}(n^2-n)$ elementary conjugates of g and g^{-1}. However, for applications the exponent of J is much more important than the exact number of factors, which may be unbounded, when R is not Noetherian.

This theorem is, in fact, a very powerful constructive version of the following result, which was proven in [26], though not stated there this way, see [34] for details.

Theorem 2. *Let R be a commutative ring with 1, $n \geqslant 4$. Further, let $J \trianglelefteq R$ be an ideal of R, and $H \leqslant \mathrm{GL}(n, R)$ be a subgroup normalised by the elementary subgroup $E(n, J)$. Then there exists an ideal $I \trianglelefteq R$ such that*
$$E(n, J^5 I) \leqslant H \leqslant G(n, R, I).$$
This ideal I is unique up to the equivalence relation $\Diamond_J$.

The proof of Theorem 1 is presented in § 3. For the most part it is an easy remoulding of the brilliant observation by Raimund Preusser [17, 18], see also another exposition of this result in [27]. Namely, he proposed to express a conjugate of an elementary generator not as a product of factors sitting in proper parabolics of certain types, as in [7] or [21], but as 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 [20] – in fact, it was implicit already in [4, 24] – but never appreciated the whole significance of this apparently small variation. Actually, there is another extremely important feature, which greatly facilitates analysis of the relative case. Namely, with respect to the entries of our matrix g, the degrees of parameters of the elementary matrices engaged at the first move are 1, whereas they were equal to 2 in the decomposition of unipotents.

In turn, Theorem 2 is an immediate corollary of Theorem 1. Most of the auxiliary results in § 3 hold for $n \geqslant 3$, the stronger assumption $n \geqslant 4$ is only invoked at the very end.

In §§ 4, 5 we prove Theorems 3 and 4, with $E(n, J)$ in the above results replaced by $E(n, R, J)$. These are essentially minor variations of Theorem 2 and Theorem 1, respectively, imposing somewhat stronger conditions, and arriving at somewhat stronger conclusions.

Next, in § 6 we start discussing how by to relax the bound on m in the above results from $m = 5$ to $m = 4$, which would then provide constructive version of [24] in the commutative case. Eventually, we plan to relax it to $m = 3$ for $n \geqslant 4$, but that would require much fancier calculations, than the ones in the present paper.

Finally, in § 7 we discuss the current situation for other groups, such as Chevalley groups or unitary groups, and suggest several related problems.

§2. Notation and preliminary facts

2.1. Basic 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.

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. Transvections are subject to the usual elementary relations, such as additivity, and the Chevalley commutator formula $[t_{ij}(\xi), t_{jh}(\zeta)] = t_{ih}(\xi\zeta)$. In particular, additivity implies that for a fixed pair $1 \leqslant i \neq j \leqslant n$ of distinct indices all elementary transvection $t_{ij}(\xi)$, $\xi \in R$, form a subgroup X_{ij} called an elementary root subgroup.

As in the introduction, for an ideal $I \trianglelefteq R$ we denote by $E(n, I)$ the corresponding elementary subgroup, generated by the elementary transvections of level I. 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. We consider the following three congruence subgroups of level I:

• the principal congruence subgroup $\mathrm{GL}(n,R,I)$ is the kernel of ρ_I,

• the full congruence subgroup $C(n,R,I)$ is the full preimage of the centre of $\mathrm{GL}(n,R/I)$ under ρ_I.

• The brimming congruence subgroup $G(n,R,I)$ is the full preimage of the diagonal subgroup of $\mathrm{GL}(n,R/I)$ under ρ_I.

In other words, the elements $g \in \mathrm{GL}(n,R,I)$ are congruent to e modulo I, meaning that $g_{ij} \in I$ for all $i \neq j$, whereas $g_{ii} \in 1+I$. The elements of the full congruence subgroups are not necesssarily congruent to 1 modulo I, but $g_{ii} \equiv g_{jj} \pmod I$ for all $i \neq j$. On the other hand, $g \in G(n,R,I)$ means simply that $g_{ij} \in I$ for $i \neq j$, no further condition is imposed on the diagonal entries. Clearly,

$$\mathrm{GL}(n,R,I) \leqslant C(n,R,I) \leqslant G(n,R,I).$$

2.2. Levels. Next, we discuss the four natural notions of level, two upper levels and two lower levels. The reason why we usually do not see the difference between all these levels is that when we consider $E(n,R)$-normalised subgroups the two upper levels coincide, as do the two lower levels. When the standard description of $E(n,R)$-normalised subgroup holds, upper levels coincide with lower levels. But for $E(n,J)$-normalised subgroups they may be all different, and relations among them are more complicated.

• 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 $g \in C(n,R,I)$. Such an ideal is generated by the off-diagonal entries g_{ij}, $1 \leqslant i \neq j \leqslant n$, and by the pair-wise 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 $i = 1,\dots,n-1$.

• Similarly the *outer diagonal upper level* of a matrix $g = (g_{ij}) \in \mathrm{GL}(n,R)$ is the *smallest* ideal $I = \mathrm{leo}(g)$ such that $g \in G(n,R,I)$. Clearly, such an ideal is generated by the off-diagonal entries g_{ij}, $1 \leqslant i \neq j \leqslant n$, and by the pair-wise differences of its diagonal entries $g_{ii} - g_{jj}$, $1 \leqslant i \neq j \leqslant n$.

Again, it suffices to consider only the *fundamental* differences $g_{i+1,i+1}-g_{ii}$, where $i=1,\dots,n-1$.

Thus, the upper levels $\mathrm{lev}(g)$ and $\mathrm{leo}(g)$ are generated by n^2-1 and by n^2-n elements, respectively. 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 these bounds 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 $E(n,R,I)\leqslant g^{E(n,R)}$.

• Similarly, the *outer diagonal lower level* $I=\mathrm{loo}(g)$ of a matrix $g\in\mathrm{GL}(n,R)$ is the *largest* ideal such that $E(n,I)\leqslant g^{E(n,R)}$.

Clearly, $\mathrm{lol}(g)\leqslant\mathrm{loo}(g)\leqslant\mathrm{leo}(g)\leqslant\mathrm{lev}(g)$. 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.

2.3. Geometry of transvections. For a pair (X,Y) of elementary root subgroups $X=X_{ij}$, $Y=X_{hk}$ there are very few possibilities up to simultaneous conjugation corresponding to the interrelations of roots in the root system of type A_{n-1}. We list them below, according to the growing angle between the roots.

• Angle 0, the two corresponding groups are *equal*.

• Angle $\pi/3$, there are 3 distinct indices among i,j,h,k and at that $i\neq k$, $j\neq h$. In this case the root subgroups will be called *commuting*. For $n\geqslant 3$ there are two orbits of commuting root subgroups with representatives (X_{12},X_{13}) and (X_{21},X_{31}) respectively.

• Angle $\pi/2$, the indices i,j,h,k are all distinct, in which case the subgroups are called *orthogonal*. For $n\geqslant 4$ there is just one orbit of pairs of orthogonal root subgroups under conjugation.

• Angle $2\pi/3$, there are 3 distinct indices among i,j,h,k and at that $i\neq h$, $j\neq k$. In this case the root subgroups will be called *non-commuting*. Actually in this case the mixed commutator $[X,Y]$ is again a root subgroup. For $n\geqslant 3$ there are two orbits of such pairs with representatives (X_{12},X_{23}) and (X_{23},X_{12}) respectively.

• Angle π, which is the most interesting case, $(i,j) = (k,h)$, when the subgroups are called *opposite*. For $n \geqslant 2$ every pair of opposite root subgroups is conjugate to the pair (X_{12}, X_{21}).

We will freely apply this terminology to the transvections $t_{ij}(\xi)$ and $t_{hk}(\zeta)$ sitting in these root subgroups, and to the matrix positions (i,j) and (h,k) themselves.

§3. Subgroups normalised by $E(n,J)$

In this section we prove Theorems 1 and 2.

In [27], Theorem 1, we have already discussed the following result by Raimund Preusser [17], and its proof. Of course, in the absolute case we do not have to distinguish $E(n,R)$ and $E(n,R,R)$. Below we explain why with the naive approach 8 in the absolute case becomes 32 in the relative case.

Lemma 1. *Let R be a commutative ring, $n \geqslant 3$, 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 $t_{ij}(\xi)$ belongs to the $E(n,R)$-normalised subgroup $g^{E(n,R)}$ generated by g, and in fact is a product of $\leqslant 8(n^2-1)$ elementary conjugates of g and its inverse g^{-1}.*

Let us start with the outer diagonal level. In fact, [17] proves the following fact.

Lemma 2. *Let R be commutative, $n \geqslant 3$, and $g \in \mathrm{GL}(n,R)$. Then for any $1 \leqslant r \neq s \leqslant n$ and all $1 \leqslant i \neq j \leqslant n$ the elementary transvection $t_{ij}(g_{rs})$ belongs to $g^{E(n,R)}$, and in fact is a product of $\leqslant 8$ elementary conjugates of g and g^{-1}.*

The proof of this result consists of three elementary moves, when one passes from g to the commutator $[g,x]$, for some $x \in E(n,R)$, to get $t_{hk}(g_{rs}) \in g^{E(n,R)}$ in *some* position (h,k). After that it is concluded by the reference to the fact that $E(n,R)$ contains preimages of all permutations, so that it does not matter in which position outside of the principal diagonal we landed.

The first part of that argument transcribes verbatim to $E(n,J)$-normalised subgroups, but now it does matter, where we end up. We might need one, two, or in the exceptional case of $n = 3$ even three additional moves to return the entry from the position where we landed to all other positions.

So now we have to carefully monitor not just the number of moves, but also the position of g_{rs} after each move.

Lemma 3. *Let R be commutative, $n \geqslant 3$, and $g \in \mathrm{GL}(n, R)$. Then for any $1 \leqslant r \neq s \leqslant n$, any $a_1, a_2, a_3 \in J$ one has $t_{sr}(a_1 a_2 a_3 g_{rs}) \in g^{E(n,J)}$. These transvections are products of $\leqslant 8$ elementary $E(n, J)$-conjugates of g and g^{-1}.*

Proof. Below, we reproduce essentially the same argument as in the proofs of Lemmas 1 and 2, but now for $E(n, J)$-normalised subgroups, instead of $E(n, R)$-normalised subgroups, tracing the position where g_{rs} arrives after each move.

Move 1. Let i, j, h be any three pair-wise distinct indices. Setting

$$x_r = t_{ij}(a_1 g_{hr}) t_{ih}(-a_1 g_{jr})$$

for some $a_1 \in J$ we see that the commutator $[x_r^{-1}, g^{-1}] \in g^{E(n,J)}$ is the product of *two* $E(n, J)$-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 $-a_1 g_{hj}$. This means that even not being in the above parabolic $P = \mathrm{Stab}_G(\langle e_j \rangle)$, this commutator has the form $t_{ij}(-a_1 g_{hj})x$, for some $x \in P$.

Move 2. Next, observe that for any $s \neq i, j$ and any $a_2 \in J$ the elementary transvection $t_{js}(a_2)$ sits in $U_P(J)$ of the parabolic subgroup P. Obviously, $[xy, z]^x = [y, z] \cdot [x, z]^x = [y, z] \cdot [z, x^{-1}]$. Thus,

$$\begin{aligned} y &= [t_{ij}(-a_1 g_{hj})x, t_{js}(-a_2)]^{t_{ij}(-a_1 g_{hj})} \\ &= [x, t_{js}(-a_2)] \cdot [t_{js}(-a_2), t_{ij}(a_1 g_{hj})] \in g^{E(n,J)} \end{aligned}$$

is the product of *four* elementary conjugates of g and g^{-1}. In the above expression of y the first commutator $z = [x, t_{js}(-a_2)]$ belongs to the unipotent radical U_P, while the second commutator equals $t_{is}(a_1 a_2 g_{hj})$.

Naive move 3. Next, take any $a_3 \in J$. Since $t_{ji}(a_2) \in U_P(J)$ and U_P is abelian, one can conclude that

$$[t_{ji}(a_3), y] = [t_{ji}(a_3), z t_{is}(a_1 a_2 g_{hj})] = t_{js}(a_3 a_1 a_2 g_{hj}) \in g^{E(n,J)}$$

is the product of *eight* $E(n, J)$-elementary conjugates of g and g^{-1}. Since j and h here are arbitrary distinct, in the case $n = 3$ there is a unique choice for the third index $i \neq j, h$, and then again the unique choice of the third index $s \neq i, j$, namely, $s = h$. Thus, in this case after three elementary moves we can only get inclusion $t_{sr}(J^3 g_{rs}) \leqslant g^{E(3,J)}$ in the *opposite* position. $\square$

For $n \geqslant 4$ even with a similar naive argument we arrive at a stronger conclusion.

Lemma 4. *Let R be commutative, $n \geqslant 4$, and $g \in \mathrm{GL}(n, R)$. Then for any $1 \leqslant r \neq s \leqslant n$, any $a_1, a_2, a_3 \in J$ one has $t_{sj}(a_1 a_2 a_3 g_{rs}) \in g^{E(n,J)}$ and $t_{ir}(a_1 a_2 a_3 g_{rs}) \in g^{E(n,J)}$, for all $j \neq s$ and all $i \neq r$. As above, these transvections are products of $\leqslant 8$ elementary $E(n, J)$-conjugates of g and g^{-1}.*

Proof. Since $n \geqslant 4$, for any $s \neq j$ we can always choose $i \neq j, h, s$ in the above argument to conclude that $t_{sj}(J^3 g_{rs}) \leqslant g^{E(n,J)}$ for all $j \neq s$. Similarly, replacing in the above argument columns by rows (switching in Move 1 matrices g and g^{-1} and taking $x_s = t_{ih}(a_1 g_{sj}) t_{jh}(-a_1 g_{si})$ instead of the initial x_r), we get also the inclusion $t_{ir}(J^3 g_{rs}) \leqslant g^{E(n,J)}$ for all $i \neq r$. Thus, for $n \geqslant 4$, we also get the desired inclusions with $m = 3$ in the *non-commuting* positions. $\square$

In the absolute case, this was already the end of line, since $E(n, R)$ contains monomial matrix w_π corresponding to any permutation $\pi \in S_n$. Conjugating by these matrices (which is just another elementary conjugation in $E(n, R)$) we conclude that all $t_{ij}(g_{rs})$, $1 \leqslant i \neq j \leqslant n$, $1 \leqslant r \neq s \leqslant n$, are products of $\leqslant 8$ elementary $E(n, R)$-conjugates of g and g^{-1}.

Unfortunately, this does not work this way in the relative case. The product $w_\pi x$, where $x \in E(n, J)$, does not belong to $E(n, J)$. This means that one needs two additional commutators with elementary transvections of level J to put g_{rs}, $r \neq s$, into any position for $n \geqslant 4$.

Lemma 5. *Let R be commutative, $n \geqslant 4$, and $g \in \mathrm{GL}(n, R)$. Then for any $1 \leqslant r \neq s \leqslant n$, any $a_1, \ldots, a_5 \in J$ one has $t_{ij}(a_1 \ldots a_5 g_{rs}) \in g^{E(n,J)}$, for all $i \neq j$. These transvections are products of $\leqslant 32$ elementary $E(n, J)$-conjugates of g and g^{-1}.*

Proof. To return the entry to any position, we need two more commutations with the elementary generators of $E(n, J)$.

Move 4. Take any $a_4 \in J$. Then, in the above notation,

$$t_{ij}(a_1 \ldots a_4 g_{rs}) = [t_{is}(a_4), t_{sj}(a_1 a_2 a_3 g_{rs})] \in g^{E(n,J)},$$

for all $i, j \neq s$ and, similarly,

$$t_{ij}(a_1 \dots a_4 g_{rs}) = [t_{ir}(a_1 a_2 a_3 g_{rs}), t_{rj}(a_4)] \in g^{E(n,J)},$$

for all $i, j \neq r$. Since in the first case $i = r$ is not excluded, and in the second case $j = s$ is not excluded, we get the required inclusions with $m = 4$ in the *orthogonal* and *commuting* positions.

Move 5. One needs one more commutator to return g_{rs} to the initial position. For any $a_5 \in J$, and any $i \neq r, s$,

$$t_{rs}(a_1 \dots a_5 g_{rs}) = [t_{ri}(a_5), t_{is}(a_1 \dots a_4 g_{rs})] \in g^{E(n,J)},$$

which gives the desired inclusion with $m = 5$ also in this last case. □

However, with this naive approach for the exceptional case $n = 3$ after Move 3 we only get g_{rs} in the *opposite* position, and we need *three* more moves from there, Move 4 to a non-commuting position, Move 5 to a commuting position, Move 6 to the initial position, which results in $m = 6$.

§4. Subgroups normalised by $E(n, R, J)$

Description of subgroups normalised by $E(n, R, J)$ with the same exponent m immediately follows from Theorem 2. The corresponding arguments are standard, and were already contained in the works by John Wilson, Leonid Vaserstein, myself, and Zuhong Zhang. However, since further we are interested in the precise number of elementary conjugates, we need details of computations as a model. It is not enough to invoke the corresponding lemmas from the above papers, we should rather go through their proofs.

Theorem 3. *Let R be a commutative ring with 1, $n \geqslant 4$. Further, let $J \trianglelefteq R$ be an ideal of R, and $H \leqslant \mathrm{GL}(n, R)$ be a subgroup normalised by the elementary subgroup $E(n, R, J)$. Then there exists an ideal $I \trianglelefteq R$ such that*

$$E(n, R, J^5 I) \leqslant H \leqslant C(n, R, I).$$

This ideal I is unique up to the equivalence relation $\lozenge_J$.

The following lemma mostly follows [23, 34], modulo correcting some misprints.

Lemma 6. *Let R be a commutative ring with 1, $n \geqslant 3$, and let $J \trianglelefteq R$ be an ideal of R. Suppose there exists an m such that that for any subgroup*

$H \leqslant \mathrm{GL}(n,R)$ normalised by the elementary subgroup $E(n,J)$ for some ideal we have the inclusion

$$E(n,J^m I) \leqslant H \leqslant G(n,R,I).$$

Then for any subgroup H normalised by the relative elementary subgroup $E(n,R,J)$ we have the inclusion

$$E(n,R,J^m I) \leqslant H \leqslant C(n,R,I)$$

with the same m. This ideal I is unique up to the equivalence relation $\lozenge_J$.

Proof. By assumption, $[H,E(n,R,J)] \leqslant H$. Since $E(n,R,J)$ is already normal in $E(n,R)$, it follows that for any $x \in E(n,R)$ one has

$$[H^x, E(n,R,J)] \leqslant H^x.$$

As the first step, we prove that

$$E(n,J^m I) \leqslant H \leqslant C(n,R,I),$$

more precisely, that $E(n,J^m \operatorname{lev}(H)) \leqslant H$.

With this end observe that the entry of the matrix $g^{t_{rs}(1)} \in H^{t_{rs}(1)}$ in the position (r,s) equals $g_{rs} + g_{rr} - g_{ss} - g_{sr}$, it follows from the above that

$$t_{ij}(J^m(g_{rs} + g_{rr} - g_{ss} - g_{sr})) \in H^{t_{rs}(1)}$$

for any $i \neq j$. But whenever $j \neq r$ and $i \neq s$, $t_{ij}(*)$ commutes with $t_{rs}(1)$. This means that in fact already

$$t_{ij}(J^m(g_{rs} + g_{rr} - g_{ss} - g_{sr})) \in H$$

for all such i,j. But since $t_{ij}(J^m g_{rs})$ and $t_{ij}(J^m g_{sr})$ are already accounted for, we can conclude that in this case $t_{ij}(J^m(g_{rr} - g_{ss})) \in H$.

For the remaining pairs (i,j) one could prove the same inclusion by considering $g^{t_{sr}(1)}$, whose entry in the position (s,r) equals $g_{sr} + g_{rr} - g_{ss} - g_{rs}$, so that we get the inclusion $t_{ij}(J^m(g_{rr} - g_{ss})) \in H$ for *all* pairs $i \neq j$, as claimed.

As we observed immediately after the statement of Lemma 6, if H is normalised by $E(n,R,J)$, then all its elementary conjugates H^x, $x \in E(n,R)$, are normalised by $E(n,R,J)$. In particular, $E(n,J^m \operatorname{lev}(H^x)) \leqslant H^x$ for all elementary x, with the above m. Since $\operatorname{lev}(H^x) = \operatorname{lev}(H)$, it means that $E(n,J^m \operatorname{lev}(H)) \leqslant H^x$, again for all elementary x. Or, what is the same, $E(n,J^m \operatorname{lev}(H))^x \leqslant H$ for all such x. But this means precisely that $E(n,R,J^m \operatorname{lev}(H)) \leqslant H$. $\square$

Alternatively, Theorem 3 immediately follows from Theorem 4 of the next section, but then, of course, the derivation of Theorem 4 itself from Theorem 1 follows the pattern of Lemma 6.

§5. Relative reverse decomposition of unipotents

The analogue of Theorem 1 looks as follows. Here we consider conjugates of g and g^{-1} by elements of $E(n,R,J)$, not just by elements of $E(n,J)$. But then we can express transvections with parameters in $J^5\operatorname{lev}(g)$, not just in $J^5\operatorname{leo}(g)$. Notice also the change of length. You need the additional $32(n-1)$ factors to express parameters corresponding to diagonal differences $g_{rr}-g_{r+1,r+1}$, $1\leqslant r\leqslant n-1$.

Theorem 4. *Let R be commutative, $J\trianglelefteq R$ be an ideal of R, $n\geqslant 4$. Further, let $g\in\mathrm{GL}(n,R)$. Then for any $\xi\in\operatorname{lev}(g)$, any $a_1,\dots,a_5\in J$, and all $1\leqslant i\neq j\leqslant n$ the elementary transvection $t_{ij}((a_1\dots a_5\xi)$ is a product of $\leqslant 32(n^2-1)$ conjugates of g and g^{-1} by elements of $E(n,R,J)$.*

Proving Theorem 1 we have already expressed $t_{ij}(a_1\dots a_5g_{rs})$ as products of 32 conjugates of g and g^{-1} by elements of $E(n,J)$. It only remains to express $t_{ij}((a_1\dots a_5(g_{rr}-g_{ss}))$ as products of $\leqslant 32$ such conjugates modulo the above. Below we reproduce the argument for an arbitrary m, not just for $m=5$.

Lemma 7. *Let R be commutative, $n\geqslant 3$, and $g\in\mathrm{GL}(n,R)$. Further, let $J\trianglelefteq R$ be an ideal of R. Suppose there exists an m such that for any $a_1,\dots,a_m\in J$, any $r\neq s$ and all $1\leqslant i\neq j\leqslant n$ the elementary transvection $t_{ij}(a_1\dots a_mg_{rs})$ is a product of $\leqslant 2^m$ conjugates of g and g^{-1} by elements of $E(n,J)$. Then modulo those transvections any $t_{ij}(a_1\dots a_m(g_{rr}-g_{ss}))$ is a product of $\leqslant 2^m$ further conjugates of g and g^{-1} by elements of $E(n,R,J)$.*

Proof. Observe that the entry of $g^{t_{rs}(1)}$ in the position (r,s) equals $g_{rs}+g_{rr}-g_{ss}-g_{sr}$. Thus, by assumption applied to $g^{t_{rs}(1)}$ any transvection

$$z=t_{ij}(a_1\dots a_m(g_{rs}+g_{rr}-g_{ss}-g_{sr}))$$

is the product of $\leqslant 2^m$ elementary conjugates of $g^{t_{rs}(1)x}$ or $g^{-t_{rs}(1)x}$, where $x\in E(n,J)$. It follows that $z^{t_{rs}(-1)}$ is the product of $\leqslant 2^m$ factors of the form

$$g^{t_{rs}(1)xt_{rs}(-1)}\quad\text{or}\quad g^{-t_{rs}(1)xt_{rs}(-1)},\qquad\text{where}\quad x\in E(n,J).$$

Obviously, $t_{rs}(1)xt_{rs}(-1) \in E(n,R,J)$. Thus, $z^{-t_{rs}(1)}$ is the product of $\leqslant 2^m$ factors of the form g^y and g^{-y}, for $y \in E(n,R,J)$.

When $j \neq r$ and $i \neq s$ one has $[X_{ij}, X_{rs}] = 1$ so that already z itself is such a product. On the other hand, by assumption applied to g itself $t_{ij}(a_1 \dots a_m g_{rs})$ and $t_{ij}(a_1 \dots a_m g_{sr})$ are products of $\leqslant 2^m$ factors of the form g^x and g^{-x}, for $x \in E(n,J)$. It follows that *modulo these transvections* $t_{ij}(a_1 \dots a_m(g_{rr} - g_{ss}))$ is a product of $\leqslant 2^m$ factors of the form g^y and g^{-y}, for $y \in E(n,R,J)$.

In the remaining cases, $j \neq s$ and $i \neq r$ and we would start our argument with the matrix $g^{t_{sr}(1)}$, whose entry in the position (s,r) equals $g_{sr} + g_{rr} - g_{ss} - g_{rs}$, instead. □

§6. Further ramblings

Now, wielding the idea of replacing g by its elementary conjugates, we can start expressing further small elementary matrices, whose parameters are the entries of g with some extra factors from J. Now, we are not as much concerned with the number of elementary conjugates, as with the occurring power of J. Instead of passing to Move 5 we dwell a bit at Moves 3 and 4, however not for H itself, but rather for its elementary conjugates. Of course, such improvements are only possible for subgroups normalised by $E(n,R,J)$.

Move $4\frac{1}{2}$. Let $g \in H$, and let $h \neq r,s$. Then the entries of $g^{t_{rs}(1)} \in H^{t_{rs}(1)}$ in positions (h,r), (h,s) are equal to g_{hr} and $g_{hs} + g_{hr}$, respectively. From Move 4 applied to $g^{t_{rs}(1)} \in H^{t_{rs}(1)}$ we already know $t_{hr}(a_1 \dots a_4(g_{hs} + g_{hr})) \in H^{t_{rs}(1)}$ for all $a_1, \dots, a_4 \in J$. It follows that

$$t_{hr}(a_1 \dots a_4(g_{hs} + g_{hr}))^{t_{rs}(-1)} = t_{hr}(a_1 \dots a_4(g_{hs} + g_{hr}))t_{hs}(a_1 \dots a_4(g_{hs} + g_{hr})) \in H.$$

Since from Move 4 applied to $g \in H$ we also know that $t_{hr}(a_1 \dots a_4 g_{hs}) \in H$ and $t_{hs}(-a_1 \dots a_4 g_{hr}) \in H$, we can conclude that

$$t_{hr}(a_1 \dots a_4 g_{hr})t_{hs}(-a_1 \dots a_4 g_{hs}) \in H.$$

Switching, as in Move 3 above, columns and rows, we also get similar inclusions $t_{rh}(a_1 \dots a_4 g_{rh})t_{sh}(-a_1 \dots a_4 g_{sh}) \in H$, for all $a_1, \dots, a_4 \in J$.

At this point, we can start rethinking also the previous steps, to get further inclusions.

Move $3\frac{1}{2}$. In the notation of Move $4\frac{1}{2}$, from Move 3 applied to $g^{t_{rs}(1)} \in H^{t_{rs}(1)}$ we know that $t_{sj}(a_1a_2a_3(g_{hr}+g_{hs})) \leqslant H^{t_{rs}(1)}$, for any $a_1, a_2, a_3 \in J$ and any $j \neq r, s$. It follows that

$$t_{sj}(a_1a_2a_3(g_{hs}+g_{hr}))^{t_{rs}(-1)} = t_{sj}(a_1a_2a_3(g_{hs}+g_{hr}))t_{rj}(a_1a_2a_3(g_{hs}+g_{hr})) \in H.$$

Since from Move 3 applied to $g \in H$ we also know that

$$t_{sj}(a_1a_2a_3g_{hs}) \in H \quad \text{and} \quad t_{rj}(a_1a_2a_3g_{hr}) \in H,$$

we can conclude that $t_{sj}(a_1a_2a_3g_{hr})t_{rj}(a_1a_2a_3g_{hs}) \in H$. Switching columns and rows, as in Moves 3 and $4\frac{1}{2}$ above, we also get similar inclusions $t_{js}(a_1a_2a_3g_{rh})t_{jr}(a_1a_2a_3g_{sh}) \in H$, for all $a_1, a_2, a_3 \in J$.

However, to really get the necessary inclusions this way is not immediate. At some point you have to perform the tough computation involving two opposite root elements. But then, if you have started the computation as in §3 above, you still get the bound $m = 4$. In the next paper, I intend to rethink the whole computation from scratch, to commit two opposite root elements from the very start, and to eventually obtain for $n \geqslant 4$ the best possible bound $m = 3$. The case $n = 3$ has to be considered separately anyway, and it might well happen that in this case Vaserstein's bound $m = 4$ cannot be improved.

§7. Final remarks

It would be natural to generalise results of the present paper to other groups, such as Chevalley groups, Bak unitary groups, or Petrov odd unitary groups. Analogues of Theorem 3 (and partly of Theorem 2) are known in some cases, but usually with bounds that are far from optimal.

In the following problems we assume that Φ is an irreducible root system of rank $\mathrm{rk}(\Phi) \geqslant 3$; – rank 2 has to be considered separately anyway.

Problem 1. *Describe subgroups of a Chevalley group $G(\Phi, R)$, normalised by the relative elementary subgroup $E(\Phi, J)$, for an ideal $J \trianglelefteq R$.*

Problem 2. *Describe subgroups of a Chevalley group $G(\Phi, R)$, normalised by the relative elementary subgroup $E(\Phi, R, J)$, for an ideal $J \trianglelefteq R$.*

The simply laced case should be comparatively easy. There is no doubt that combining ideas of the present paper with those of [27], one would get analogues of our Theorems 1–4 for $\Phi = \mathrm{D}_l, \mathrm{E}_6, \mathrm{E}_7$ and E_8 by essentially the same arguments and with precisely the same bounds, as for $\mathrm{GL}(n, R)$, $n \geqslant 4$. Presently, the author and Zuhong Zhang are writing up detailed proofs.

However, the situation for multiply laced systems is substantially harder in several respects. Even in the absolute case, the bounds in [17, 18] seem to be grossly exaggerated. On the other hand, there is a genuine difficulty. As we have already mentioned, in the simply laced case the degrees of parameters of the elementary matrices occurring in the first move are 1. In the doubly laced case, the degrees of these parameters with respect to the entries of the initial matrix g are 2. This makes subsequent computations much trickier.

When 2 is not invertible in R the situation becomes even worse. Again even in the absolute case form parameters or admissible pairs occur, in the statement of definitive answers. In our problem this leads to the corresponding complications at the relative level, such as relative form parameters, etc. Thus, solution of the above problems would provide description of subnormal subgroups only when $2 \in R^*$. Otherwise, one should solve similar, but technically more demanding problems stated in terms of birelative elementary subgroups.

We do not try to describe results for generalised unitary groups. Let us cite some of the recent papers, where one can find many further references [9, 10, 32–36].

Over years I worked on localisation and relativisation jointly with Anthony Bak, Roozbeh Hazrat, Victor Petrov, Alexei Stepanov, and Zuhong Zhang. Successive incentive and inspiration that led to the present paper came from discussions with John Wilson in April 2018, with Zuhong Zhang in July 2018, and with Raimund Preusser and Alexei Stepanov in November 2018. I am very grateful to all of them.

References

1. A. Bak, *Subgroups of the general linear group normalized by relative elementary subgroup.* — Algebraic K-theory, Part II (Oberwolfach, 1980), Lecture Notes Math., Vol. 967, Springer, Berlin 1982, pp. 1–22.
2. A. Bak, *Nonabelian K-theory: the nilpotent class of K_1 and general stability.* — K-Theory **4** (1991), pp 363–397.
3. A. Bak, N. Vavilov, *Structure of hyperbolic unitary groups.* I. *Elementary subgroups.* — Algebra Colloq. **7**, No. 2 (2000), 159–196.
4. H. Bass, K-*theory and stable algebra.* — Inst. Hautes Études Sci. Publ. Math., No. 22 (1964), 5–60.
5. H. Bass, *Unitary algebraic K-theory.* — Lecture Notes Math. **343** (1973), 57–265.

6. Z. I. Borewicz, N. A. Vavilov, *Subgroups of the general linear group over a semilocal ring, containing the group of diagonal matrices. — Proc. Steklov. Inst. Math.* **148** (1980), 41–54.
7. Z. I. Borewicz, N. A. Vavilov, *The distribution of subgroups in the general linear group over a commutative ring.* — Proc. Steklov. Inst. Math. **183** (1985), 27–46.
8. J. L. Brenner, *The linear homogeneous group.* III. — Ann. of Math. **71** (1960), 210–223.
9. G. Habdank, *Mixed commutator groups in classical groups and a classification of subgroups of classical groups normalized by relative elementary groups.* Doktorarbeit Uni. Bielefeld, 1987, 1–71.
10. G. Habdank, *A classification of subgroups of* Λ*-quadratic groups normalized by relative elementary groups.* — Adv. Math. **110** (1995), 191–233.
11. A. J. Hahn, O. T. O'Meara, *The classical groups and* K*-theory.* Springer, Berlin et al., 1989.
12. R. Hazrat, N. Vavilov, *Bak's work on the* K*-theory of rings.* — J. K-Theory **4**, No. 1 (2009), 1–65.
13. R. Hazrat, N. Vavilov, Zuhong Zhang, *The commutators of classical groups.* — J. Math. Sci., New York **222**, No.4 (2017), 466–515.
14. R. Hazrat, Zuhong Zhang, *Generalized commutator formula.* — Commun. Algebra, **39**, No. 4 (2011), 1441–1454.
15. Li Fuan, Liu Mulan, *Generalized sandwich theorem.* — K-Theory, **1** (1987), 171–184.
16. R. Preusser, *Structure of hyperbolic unitary groups* II. *Classification of* E*-normal subgroups.* — Algebra Colloq. **24**, No. 2 (2017), 195–232.
17. R. Preusser, *Sandwich classification for* $\mathrm{GL}_n(R)$*,* $O_{2n}(R)$ *and* $U_{2n}(R,\Lambda)$ *revisited.* — J. Group Theory **21**, No. 1 (2018), 21–44.
18. R. Preusser, *Sandwich classification for* $O_{2n+1}(R)$ *and* $U_{2n+1}(R,\Delta)$ *revisited.* — J. Group Theory **21**, No. 4 (2018), 539–571.
19. A. Stepanov, *Structure of Chevalley groups over rings via universal localization.* — J. Algebra **450** (2016), 522–548.
20. A. V. Stepanov, *A new look at the decomposition of unipotents and the normal structure of Chevalley groups.* — St. Petersburg J. Math. **28**, No. 3 (2017), 411–419.
21. A. Stepanov, N. Vavilov, *Decomposition of transvections: a theme with variations.* — K-Theory **19**, No. 2 (2000), 109–153.
22. 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.
23. L. N. Vaserstein, *The subnormal structure of general linear groups.* — Math. Proc. Cambridge Philos. Soc., **99**, No. 3 (1986), 425–431.
24. L. N. Vaserstein, *The subnormal structure of general linear groups over rings.* — Math. Proc. Cambridge Phil. Soc., **108**, No. 2 (1990), 219–229.
25. L. Vaserstein, Hong You, *Normal subgroups of classical groups over rings.* — J. Pure Appl. Algebra **105** (1995), 93–106.
26. N. Vavilov, *Subnormal structure of general linear group.* — Math. Proc. Cambridge Phil. Soc., **107** (1990), 103–106.

27. N. A. Vavilov, *Towards the reverse decomposition of unipotents.* — Zap. Nauchn. Semin. POMI **470** (2018), 21–37.
28. N. A. Vavilov, *Unrelativised standard commutator formula.* — Zap. Nauch. Semin. POMI. **470** (2018), 38–49.
29. N. A. Vavilov, A. V. Stepanov, *Standard commutator formula.* — Vestnik St. Petersburg State Univ., Ser. 1, **41**, No. 1 (2008), 5–8.
30. N. A. Vavilov, A. V. Stepanov, *Standard commutator formulae, revisited.* — Vestnik St. Petersburg State Univ., Ser.1, **43**, No. 1 (2010), 12–17.
31. J. S. Wilson, *The normal and subnormal structure of general linear groups.* — Proc. Camb. Philos. Soc. **71** (1972), 163–177.
32. Hong You, *Subgroups of classical groups normalised by relative elementary groups.* — J. Pure Appl. Algebra, **216** (2012), 1040–1051.
33. Hong You, Xuemei Zhou, *Subnormal structure of classical groups over Banach algebra.* — J. Algebra **357** (2012), 222–230
34. Zuhong Zhang, *Lower K-theory of unitary groups.* Ph. D. thesis Univ. Belfast, 2007, 1–67.
35. Zuhong Zhang, *Stable sandwich classification theorem for classical-like groups.* — Proc. Cambridge Phil. Soc., **143** (2007), 607–619.
36. Zuhong Zhang, *Subnormal structure of non-stable unitary groups over rings.* — J. Pure Appl. Algebra, **214**, No. 5 (2010), 622–628.

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

Поступило 7 октября 2019 г.