

M. A. Buryakov, N. A. Vavilov

RELATIVE DECOMPOSITION OF TRANSVECTIONS: EXPLICIT BOUNDS

ABSTRACT. Let R be a commutative associative ring with 1, and let $G = \mathrm{GL}(n, R)$ be the general linear group of degree $n \geq 3$ over R . Further, let $I \trianglelefteq R$ be an ideal of R . In the present note, which is a marginalia to the paper of Alexei Stepanov and the second named author(2000), we obtain explicit expressions of the elementary transvection $gt_{ij}(\xi)g^{-1}$, where $1 \leq i \neq j \leq n$, $\xi \in I$ and $g \in G$, as products of the Stein–Tits–Vaserstein generators of the relative elementary group $E(n, R, I)$.

INTRODUCTION

Let R be a commutative ring with 1, $G = \mathrm{GL}(n, R)$ be the general linear group of degree $n \geq 3$ over R . For an ideal $I \trianglelefteq R$ denote by $E(n, I)$ the elementary subgroup generated by the elementary transvections of level I :

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

The corresponding *relative* elementary subgroup $E(n, R, I)$ is defined as the normal closure of $E(n, I)$ in the absolute elementary subgroup $E(n, R)$.

One of the pivotal results in the structure theory of linear groups, Suslin’s normality theorem [18], asserts that relative elementary subgroups $E(n, R, I)$ are in fact normal in the whole general linear group $\mathrm{GL}(n, R)$. In other words, $gt_{ij}(\xi)g^{-1} \in E(n, R, I)$ for all $1 \leq i \neq j \leq n$, $\xi \in I$, and all $g \in \mathrm{GL}(n, R)$. Later, many further proofs were proposed, using various versions of localisation and geometric methods, see, in particular, [1, 5, 17, 19, 21], and many further references in [2, 6, 9].

To state an *effective* version of this result, for a natural $L \in \mathbb{N}$ denote by $E^L(n, R)$ the subset of $E(n, R)$ consisting of products of $\leq L$ elementary transvections.

Key words and phrases: general linear group, congruence subgroups, elementary subgroups, standard commutator formulae.

The work of the second named author is supported by the Russian Science Foundation grant 17-11-01261.

In the *absolute* case, Suslin's initial proof provided an explicit factorisation which in particular implied that for all $1 \leq i \neq j \leq n$, $\xi \in R$, and $g \in \mathrm{GL}(n, R)$ one has $gt_{ij}(\xi)g^{-1} \in E^L(n, R)$ for

$$L \leq n(n-1)(n+2),$$

cubic in n .

Later, **decomposition of transvections**, developed by Alexei Stepanov and the second author, see in particular [17] and references there, furnished other such factorisations, improving the above bound to a better one,

$$L \leq 4n(n-1),$$

quadratic in n . [Of course, original Suslin's approach applies to a broader class of linear transvections, not just conjugates of the elementary ones, and thus proves *stronger* results, than just normality of $E(n, R)$.]

This bound was crucial in obtaining sharp estimates of the width of commutators in $\mathrm{GL}(n, R)$ in terms of transvections in the work of Alexander Sivatsky and Stepanov [14], and then in some later applications, [8, 10, 11].

In connection with similar applications at the relative level, see [8, 16], it is natural to ask, what are the explicit bounds in the effective versions of Suslin's normality theorem in the *relative* case, for the groups $E(n, R, I)$, $I \trianglelefteq R$. Amazingly, in the general case, without some strong additional assumptions on the ring R , no such explicit bounds seem to be available in the existing literature.

In the present paper, we slightly adapt the proof from [17] to obtain such a bound, *cubic* in n , in terms of the Stein–Tits–Vaserstein generators of $E(n, R, I)$, see [15, 20, 23]. Of course, now we have the advantage of the hindsight, provided by the papers by the second author and Zhang Zuhong, where very similar calculations were performed at the birelative level, see [27, 28] and references there.

Namely, as a group $E(n, R, I)$ is generated by the elements of the form

$$z_{ij}(\xi, \zeta) = t_{ji}(\zeta)t_{ij}(\xi)t_{ji}(-\zeta),$$

where $1 \leq i \neq j \leq n$, while $\xi \in I$ and $\zeta \in R$. For a natural $L \in \mathbb{N}$ denote by $E^L(n, R, I)$ the subset of $E(n, R, I)$ consisting of products of $\leq L$ such elementary generators $z_{ij}(\xi, \zeta)$.

Now the main result of the present note can be stated as follows.

Theorem 1. *Let R be a commutative ring with 1, $I \trianglelefteq R$ be an ideal of R , and $n \geq 3$. Then for all $1 \leq i \neq j \leq n$, $\xi \in I$, and $g \in \mathrm{GL}(n, R)$ one has*

$$gt_{ij}(\xi)g^{-1} \in E^L(n, R, I),$$

for

$$L \leq n \left(\frac{3}{2}n^2 - \frac{3}{2}n - 1 \right).$$

Observe that under appropriate stability assumptions a similar result for all classical Chevalley groups follows from the important paper by Sergei Sinchuk and Andrei Smolensky [13]. The feature of the present paper is that our results hold – with a uniform bound! – for ARBITRARY COMMUTATIVE RINGS.

The balance of the present work is organized as follows. In §1 we recall the Theme of [17] which is used in §2 to reduce the proof of Theorem 1 to analysis of the very special case of elements of the form uvu^{-1} , where u and v are taken from the unipotent radicals of the opposite parabolic subgroups P_1 and P_1^- , and, moreover, v is of level I . This key case is then addressed in §3 by an easy induction on rank. Finally, in §4, we state some related open problems.

§1. DECOMPOSITION OF TRANSVECTIONS

As usual, e denotes the identity matrix and e_{ij} is a standard matrix unit. For $\xi \in R$ and $1 \leq i \neq j \leq n$, we denote by $t_{ij}(\xi) = e + \xi e_{ij}$ the corresponding [elementary] transvection. A matrix $g \in \mathrm{GL}(n, R)$ is written as $g = (g_{ij})$, $1 \leq i, j \leq n$, where g_{ij} is its entry in the position (i, j) . Entries of the inverse matrix $g^{-1} = (g'_{ij})$, $1 \leq i, j \leq n$, are denoted by g'_{ij} .

By R^n we denote the free right R -module, consisting of columns of height n with components in R and by nR , we denote the free left R -module consisting of rows of length n with components in R . Standard bases in R^n and nR , are denoted by e_1, \dots, e_n and f_1, \dots, f_n , respectively.

A *transvection* is a matrix of the form $e + uv$, where $u \in R^n$, $v \in {}^nR$ are a column and a row such that $vu = 0$. Classically, [the line spanned by] u is called the *center* of the transvection $e + uv$, while [the hyperplane orthogonal to] v is called its *axis*, see, for instance [26].

Since we are only interested in this factor, for the time being we modify notation as follows: we consider a transvection $e + uv$ of level I in the first standard parabolic subgroup P_1 . Here, as above, $u \in R^n$, $v \in {}^nI$, are such that $vu = 0$ and $v_1 = 0$.

As in the absolute case, our computation starts with the following factorization of $e + uv$ as the product of four factors in the two opposite unipotent radicals U_1 and U_1^- . It is a routine computation based on the properties of transvections. It is known since [3, 4]. For details, see the proof of either [17], Lemma 3, or [26], Lemma 21.

Lemma 2. *Let A be an ideal of R . Further, let $u \in R^n$ and $v \in {}^nI$ be a column and a row such that $vu = 0$ and $v_1 = 0$. Then*

$$e + uv = [e + (u - u_1e_1)f_1, e + e_1v](e + u_1e_1v).$$

Expanding the commutator and blending its fourth term $e - e_1v \in U_1(I)$ with $e + u_1e_1v \in U_1(I)$, we get

$$e + uv = (e + (u - u_1e_1)f_1)(e + e_1v)(e - (u - u_1e_1)f_1) \cdot (e + (u_1 - 1)e_1v).$$

In the absolute case, this supplied an expression of $e + uv$ as the product of $4(n - 1)$ elementary transvections.

This does not work as easily in the relative case. Since $v \in {}^nI$, $v_1 = 0$, the last factor is indeed the product of $n - 1$ elementary transvections from $E(n, I)$. However, the components of u do not have to belong to I . It is the conjugate of $e + e_1v$ by $e + (u - u_1e_1)f_1$ that belongs to $E(n, R, I)$. We are going to express this conjugate as the product of elementary generators $z_{ij}(\xi, \zeta)$ of the relative group $E(n, R, I)$.

With this end, we simplify the notation once more. From now on, we consider a column $u = (u_2, \dots, u_n)^t \in R^{n-1}$ and a row $v = (v_2, \dots, v_n) \in {}^{n-1}I$ such that $vu = 0$, set

$$y_1(u) = \begin{pmatrix} 1 & & & \\ u_2 & 1 & & \\ \vdots & & \ddots & \\ u_n & & & 1 \end{pmatrix} \in U_1^-, \quad x_1(v) = \begin{pmatrix} 1 & v_2 & \dots & v_n \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix} \in U_1(I),$$

and form the [left] conjugate

$$y_1^{(u)}x_1(v) = y_1(u)x_1(v)y_1(-u) \in E(n, R, I).$$

From the above, it follows that if

$$y_1^{(u)}x_1(v) \in E^M(n, R, I),$$

then

$${}^g t_{ij}(\xi) = g t_{ij}(\xi) g^{-1} \in E^L(n, R, I), \quad L = n(M + n - 1).$$

Thus, it only remains to estimate M . This will be done in the next section by an easy induction on n .

§3. PROOF OF THEOREM 1

Thus, to finish the proof of Theorem 1 we only have to express

$$y_1(u)x_1(v) = \begin{pmatrix} 1 + vu & v_2 & \dots & v_n \\ u_2vu & 1 + u_2v_2 & \dots & u_2v_n \\ \vdots & \vdots & \ddots & \vdots \\ u_nv u & u_nv_2 & \dots & 1 + u_nv_n \end{pmatrix},$$

where $u = (u_2, \dots, u_n)^t \in R^{n-1}$ and $v = (v_2, \dots, v_n) \in {}^{n-1}I$, as a product of the generators $z_{ij}(\xi, \zeta)$.

Of course, in the previous section it was additionally assumed that $vu = 0$. There this property was needed to ensure that g belongs to the parabolic subgroup P_1 . However, this additional property is not preserved during induction, so that we cannot assume it from the very start.

The following result is the main new contribution of the present note, as compared with [17]. Observe that, unlike Theorem 1, it works uniformly, starting with $n = 2$. The reason why Theorem 1 does not, is *precisely* that for $n = 2$ there are no nontrivial u and v such that $vu = 0$.

Theorem 2. *Let R be a commutative ring with 1, $I \trianglelefteq R$ be an ideal of R , and $n \geq 2$. Then for all $u \in R^{n-1}$ and $v \in {}^{n-1}I$, one has*

$$y_1(u)x_1(v) \in E^M(n, R, I),$$

for

$$M = M(n) \leq \frac{3}{2}n^2 - \frac{5}{2}n.$$

Proof. The proof proceeds by induction on degree n . The base of induction, $n = 2$, is obvious, since in this case

$$y_1(u)x_1(v) = {}^{t_{21}(u_2)}t_{12}(v_2) = z_{12}(v_2, u_2),$$

is itself a single elementary conjugate of level I .

The induction step is quite straightforward. We express $y_1(u)$ and $x_1(v)$ as

$$y_1(u) = t_{n1}(u_n)y_1(\tilde{u}), \quad x_1(v) = x_1(\tilde{v})t_{1n}(v_n),$$

where

$$\tilde{u} = (u_2, \dots, u_{n-1})^t \in R^{n-2}, \quad \tilde{v} = (v_2, \dots, v_{n-1}) \in {}^{n-2}I.$$

Now

$$y_1(u)x_1(v) = t_{n1}(u)y_1(\tilde{u})(x_1(\tilde{v})t_{1n}(v_n)) = t_{n1}(u)(y_1(\tilde{u})x_1(\tilde{v}) \cdot y_1(\tilde{u})t_{1n}(v_n)),$$

and we separately conjugate the two factors by $t_{n1}(u_n)$.

- The first one of them, $y_1(\tilde{u})x_1(\tilde{v})$ is a product of

$$M(n-1) \leq \frac{3}{2}(n-1)^2 - \frac{5}{2}(n-1)$$

elementary generators of level I by induction hypothesis.

Clearly,

$$t_{n1}(u_n)(y_1(\tilde{u})x_1(\tilde{v})) = y_1(\tilde{u})x_1(\tilde{v}) \cdot x_n(w),$$

where

$$w = (u_n \tilde{v} \tilde{u}, u_n v_2, \dots, u_n v_{n-1}) \in {}^{n-1}I.$$

The extra factor $x_n(w)$ costs at most $n-1$ further elementary transvections of level I , on top of the $M(n-1)$ elementary generators of level I requisite to express the first factor.

- On the other hand,

$$y_1(\tilde{u})t_{1n}(v_n) = t_{1n}(v_n) \cdot y_n(qv_n),$$

where

$$q = (0, u_2, \dots, u_{n-1})^t \in R^{n-1}.$$

Further conjugating by $t_{n1}(u_n)$ we get *one* additional elementary conjugate $z_{n1}(v_n, u_n)$ of level I and

$$t_{n1}(u_n)y_n(qv_n) = y_1(-\hat{q}v_n u_n) \cdot y_n(qv_n),$$

where

$$\hat{q} = (u_2, \dots, u_{n-1}, 0)^t \in R^{n-1}.$$

This last factor is of level I and belongs to the unipotent radical of a parabolic subgroup of type P_2 and, thus, is a product of at most $2(n-2)$ elementary transvections of level I .

Summarizing the above, we see that

$$M(n) \leq M(n-1) + (n-1) + 1 + 2(n-2) \leq \frac{3}{2}n^2 - \frac{5}{2}n,$$

as claimed. \square

Remark 1. It is clear from the proof that one requires at most $n - 1$ elementary conjugates of level I , in positions $(1, 2), \dots, (1, n)$, the rest are elementary transvections of level I .

Remark 2. One can reorganize this proof not as an induction on n , but as reduction to two types of parabolic subgroups, of types P_{n-1} and P_2 , the summand 1 in the last display line of the proof should be interpreted as $M(2)$, whereas $n - 1$ and $2(n - 2)$ are the dimensions of the unipotent radicals U_{n-1} and U_2 , respectively.

It is clear that instead of $n - 1 = (n - 2) + 1$ one could consider any partition $n - 1 = k + m$ into two summands, and obtain an estimate

$$M(n) \leq M(k + 1) + M(m + 1) + \dim(U_{k+1}) + \dim(U_{m+1}).$$

As one can expect from the “divide and conquer” arguments, one gets a better bound when k and m are close, either $k = m$, or $k = m \pm 1$. This is indeed the case. The first case, when this makes a difference, is that of $\mathrm{SL}(5, R)$, where our theorem gives $M \leq 25$, whereas $4 = 2 + 2$ improves it to $M \leq 24$. For $\mathrm{SL}(6, R)$, the partition $5 = 3 + 2$ gives $M \leq 37$, instead of $M \leq 39$ that we get in our theorem. For $\mathrm{SL}(7, R)$ our theorem gives $M \leq 56$, whereas the partition $6 = 4 + 2$ improves it to $M \leq 53$, and $6 = 3 + 3$ further improves it to $M \leq 52$.

For $\mathrm{SL}(n, R)$ itself, the resulting improvement is not very essential, and we do not pursue it here. But we intend to elaborate this idea in the next paper dedicated to generalizations to other Chevalley groups.

§4. FINAL REMARKS

A first draft of the present note was written back in 2018. Initially, we attempted two alternative approaches.

- A straightforward way is to express $x_1(v)$ as the product of transvections $t_{1i}(v_i)$, $i = 2, \dots, n$, distribute conjugation by $y_1(u)$ over this product, and estimate the width of an individual factor $y_1(u)t_{1i}(v_i)y_1(-u)$. Clearly, again all such factors have exactly the same structure, up to simultaneous permutation of rows and columns, so we only have to estimate the width of one of them, say $y_1(u)t_{12}(v_2)y_1(-u)$.

- Another way, that gives a slightly better bound, is to use induction on n . In this approach, initially we isolated one factor of the form $y_1(u)t_{1n}(v_n)y_1(-u)$ at a time, which gives one generator $z_{1n}(\xi, \zeta)$, a bunch of elementary transvections in $E(n, I)$, and, finally, a conjugate of exactly the same shape as $y_1(u)x_1(v)y_1(-u)$, but now in $E(n - 1, R, I)$.

Both ways give *cubic* bounds, which are, however, slightly worse, as far as the highest term, than the one exhibited here. Namely, those proofs yield $2n^3$, instead of $3n^3/2$ as in our Theorem 1. Here, we also proceed by induction on n , but single out the conjugation by $t_{n1}(u_n)$, rather than the conjugate of $t_{1n}(v_n)$.

The calculations in the proof of Theorem 2 are somewhat reminiscent of the calculations in the proof of [22], Theorem 1.2. Of course, there Vaserstein operates at the absolute level, but with a localization parameter, which is essentially the same, as working modulo a principal ideal. Since he considers the usual elementary generators, rather than elementary conjugates, his bounds are different. However, a more direct source for us were the recent calculations in the works of the second author with Zhang Zuhong, see, for instance [27, 28].

Amazingly, for $\mathrm{SL}(3, R)$, the relative bound in Theorem 1 coincides with the bound $L = 24$ obtained in [17] for the absolute case! Unfortunately, this is a typical “law of small numbers” phenomenon. For $n \geq 4$ the relative bound invariably exceeds the absolute one.

Problem 1. Improve the bounds in Theorems 1 and 2.

As we already mentioned in the previous section, some *minor* improvements are very possible. However, we do not believe that a cubic bound could be improved to a quadratic one without some entirely new ideas. At present we do not see any inroad leading in this direction.

Problem 2. Generalize Theorems 1 and 2 to Chevalley groups.

In the presence of abelian unipotent radicals such a generalization is straightforward. Lemmas 1 and 2 are known in this case, see [17] and [24]. In turn, Theorem 2 for D_l can be reduced to two parabolics of type A_{l-1} , the case of E_6 can be reduced to two parabolics of type D_5 , and the case E_7 can be reduced to two parabolics of type E_6 . We plan to come up with detailed proofs in the sequel of the present paper.

On the other hand, for all multiply laced systems, apart from B_l , as well as for E_8 , there are some complications already at the level of Lemma 1 and Lemma 2. It appears that to get good bounds for C_l even in the absolute case one should not proceed directly from [17], Variation 7, but use a shorter decomposition developed by Andrei Lavrenov in [12] instead. We believe that there is a similar short decomposition of unipotents for F_4 , based on the fact that it is the twisted form of E_6 , but such a decomposition seems to be quite a bit fancier, and we intend to address it separately.

Another challenge would be, of course, to generalize Theorem 1 to birelative commutator subgroups in the style of [27, 28].

This iteration of the present paper resulted from a discussion with Pavel Gvozdevsky on his work on the bounded elementary generation of relative subgroups in the number case, see [7], where [a special case of] our Theorem 1 serves as the base of induction.

REFERENCES

1. A. Bak, *Non-abelian K-theory: The nilpotent class of K_1 and general stability*. — *K-Theory* **4** (1991), 363–397.
2. A. Bak, N. A. Vavilov, *Structure of hyperbolic unitary groups. I. Elementary subgroups*. — *Algebra Colloq.* **7**, No. 2 (2000), 159–196.
3. H. Bass, *K-theory and stable algebra*. — *Publ. Math. Inst. Hautes Études Sci.* No. 22 (1964), 5–60.
4. H. Bass, J. Milnor, J.-P. Serre, *Solution of the congruence subgroup problem for SL_n ($n \geq 3$) and Sp_{2n} ($n \geq 2$)*. — *Publ. Math. Inst. Hautes Études Sci.* **33** (1967), 59–137.
5. 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.
6. A. J. Hahn, O. T. O’Meara, *The classical groups and K-theory*, Springer, Berlin et al., 1989.
7. P. B. Gvozdevsky *Width of $SL(n, \mathcal{O}_S, I)$* , [arXiv:2206.11101v1](https://arxiv.org/abs/2206.11101v1) [math GR] 22 June 2022, 16p.
8. R. Hazrat, A. Stepanov, N. Vavilov, Zuhong Zhang, *Commutator width in Chevalley groups*. — *Note di Matematica* **33**, No. 1 (2013), 139–170.
9. R. Hazrat, N. Vavilov, *Bak’s work on K-theory of rings (with an appendix by Max Karoubi)*. — *J. K-Theory* **4**, No. 1 (2009), 1–65.
10. R. Hazrat, N. Vavilov, Zuhong Zhang, *The commutators of classical groups*. — *J. Math. Sci.* **222**, No. 4 (2017), 466–515.
11. B. Kunyavskii, E. Plotkin, N. Vavilov, *Bounded generation and commutator width of Chevalley groups: function case*, 2022. <https://arxiv.org/abs/2204.10951>, 55p.
12. A. Lavrenov, *Another presentation for symplectic Steinberg groups*. — *J. Pure Appl. Algebra* **219**, No. 9 (2015), 3755–3780.
13. S. Sinchuk, A. Smolensky, *Decompositions of congruence subgroups of Chevalley groups*. — *Internat. J. Algebra Comput.* **28**, No. 6 (2018), 935–958.
14. A. Sivatski, A. Stepanov, *On the word length of commutators in $GL_n(R)$* . — *K-Theory* **17** (1999), 295–302.
15. M. R. Stein, *Generators, relations and coverings of Chevalley groups over commutative rings*. — *Amer. J. Math.* **93** (1971), No. 4, 965–1004.
16. A. Stepanov, *Structure of Chevalley groups over rings via universal localization*. — *J. Algebra* **450** (2016), 522–548.
17. A. Stepanov, N. Vavilov, *Decomposition of transvections: a theme with variations*. — *K-Theory* **19**, No. 2 (2000), 109–153.

18. A. A. Suslin, *The structure of the special linear group over polynomial rings.* — Math. USSR Izv. **11** (1977), No. 2, 235–253.
19. A. A. Suslin, V. I. Kopeiko, *Quadratic modules and orthogonal groups over polynomial rings.* — J. Sov. Math. **20**, No. 6 (1982), 2665–2691.
20. J. Tits, In: *Systèmes générateurs de groupes de congruences.* — C. R. Acad. Sci. Paris, Sér A **283** (1976), 693–695.
21. 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.
22. L. N. Vaserstein, *Bounded reduction of invertible matrices over polynomial rings by addition operations*, unpublished, 2006, <http://www.personal.psu.edu/lxv1/pm2.pdf>, 12p.
23. L. N. Vaserstein, A. A. Suslin, *Serre's problem on projective modules over polynomial rings, and algebraic K-theory.* — Math. USSR Izv. **10** (1978), 937–1001.
24. N. Vavilov, *A third look at weight diagrams.* — Rend. Sem. Mat. Univ. Padova **104** (2000), 201–250.
25. N. A. Vavilov, A. V. Stepanov, *Standard commutator formula.* — Vestnik St. Petersburg State Univ., Ser. 1, **41**, No. 1 (2008), 5–8.
26. N. A. Vavilov, A. V. Stepanov, *Linear groups over general rings I. Generalities.* — J. Math. Sci. **188**, No. 5 (2013), 490–550.
27. N. Vavilov, Z. Zhang, *Multiple commutators of elementary subgroups: end of the line.* — Linear Algebra and Applications **599** (2020), 1–17.
28. N. Vavilov, Z. Zhang, *Commutators of elementary subgroups: curiouser and curiouser.* — Transformation Groups, (2021), 1–18. <https://doi.org/10.1007/s00031-021-09662-z>

St. Petersburg State University
E-mail: mihail.buryakov@gmail.com
E-mail: nikolai-vavilov@yandex.ru

Поступило 28 октября 2022 г.