DECOMPOSITION OF TRANSVECTIONS: A THEME WITH VARIATIONS

ALEXEI STEPANOV $^\flat$, NIKOLAI VAVILOV $^\sharp$

$^\flat$ Chair of Mathematics II
Sankt-Petersburg Electrotechnical University
197376, Sankt-Petersburg, Russia
and
$^\sharp$ Department of Mathematics and Mechanics
Sankt-Petersburg State University
198904, Petrodvorets, Russia

INTRODUCTION

The study of linear groups over fields goes back to the mid 19-th century. A landmark in their history is the book of J.Dieudonné [Di], where the theory was extended to skew-fields. However the theory of linear groups over rings – apart from some arithmetical and zero-dimensional ones – lay dormant until mid 60-ies, when the real revolution of generality was initiated by the seminal paper of H.Bass [Ba1] (see also [Ba2], [BMS]). In particular H.Bass introduced a new notion of dimension of rings, the so called stable rank, and discovered that with some slight changes the principal structure results of the theory of linear groups over skew-fields remain valid for groups whose degree is large with respect to the stable rank. J.Milnor initiated also the study of Steinberg group in this context [Mi]. This theory was extended to other classical groups by A.Bak [B1] – [B3], [Ba3]. Very important contributions

---

Key words and phrases. General linear group, elementary subgroup, congruence subgroups, normal structure, Steinberg group, centrality of K_2, transvections, intermediate subgroups, almost commutative rings, linear dependence, orthogonal group, symplectic group, Chevalley groups.

The authors gratefully acknowledge support of the *Alexander von Humboldt-Stiftung*, the *SFB 343 an der Universität Bielefeld*, the *INTAS 93–436EXT*, the *Cariplo Foundation for Scientific Research* through the *Landau Network Cooperation Centre* and the *RFBR contract 98-01-00825*. Part of the final text was written while the second-named author visited the *Isaac Newton Institute for Mathematical Sciences*.

in this direction are due to R.K.Dennis, W.van der Kallen, M.R.Stein, A.A.Suslin and L.N.Vaserstein, among others.

In another decade the next major breakthrough was triggered by the works of J.S.Wilson [Wi], I.Z.Golubchik [Go1] and A.A.Suslin [Su1] who discovered that for commutative rings the above results remain valid starting from degree 3 independently of dimension of the ground ring. New proofs of these results and their generalizations were obtained by A.A.Suslin – V.I.Kopeiko [Tu], [SK], [Ko], L.N.Vaserstein [V1], Z.I.Borewicz – N.A.Vavilov [BV], G.Taddei [Ta1], and many others. W.van der Kallen [vdK] and M.S.Tulenbaev [Tu] obtained analogous results for Steinberg groups (starting from degree 4). I.Z.Golubchik, A.V.Mikhalev, S.G.Khlebutin, L.N.Vaserstein and others (see, for example, [Go3], [Go5], [GM2]) generalized these results to wide classes of non-commutative rings, including, in particular, von Neumann regular rings, PI-rings and other important classes. One may find an account of this theory in a remarkable book of A.J.Hahn and O.T.O'Meara [HO].

To be more specific, let us introduce some notation (see §§ 1–3, 7 for definitions of these and other related notions). Let R be an associative ring with 1, $G = \mathrm{GL}(n, R)$ be the general linear group of degree n over R, $\mathrm{E}(n, R)$ be the elementary subgroup of G, $\mathrm{St}(n, R)$ be the corresponding Steinberg group.

The work of H.Bass and J.Milnor attracted major attention to the following questions, among others:

(1) Normality of $\mathrm{E}(n, R)$ in $\mathrm{GL}(n, R)$,

(2) Classification of subgroups in $\mathrm{GL}(n, R)$, normalized by $\mathrm{E}(n, R)$,

(3) Centrality of the extension $\mathrm{St}(n, R) \longrightarrow \mathrm{E}(n, R)$.

There are *standard* answers to these question, which mimic the behaviour of these groups at the stable level [Ba1], [Ba2], [Mi]. For example, if R is commutative, and $n \geq 3$, then $\mathrm{E}(n, R)$ is normal in $\mathrm{GL}(n, R)$ (Suslin theorem), the standard description of subgroups in $\mathrm{GL}(n, R)$, normalized by $\mathrm{E}(n, R)$, holds (Wilson – Golubchik theorem) and the extension $\mathrm{St}(n, R) \longrightarrow \mathrm{E}(n, R)$ is central if $n \geq 4$ (van der Kallen theorem). The example of V.N.Gerasimov [Ge] shows that these results do not in general hold for arbitrary rings. In fact in his example the elementary subgroup $\mathrm{E}(n, R)$ is a free factor of $\mathrm{GL}(n, R)$ and thus is as far from being normal, as one can possibly imagine. Below we formulate the central problem of the theory of linear groups over rings.

Problem. *Determine the exact class (classes?) of rings when the above questions admit the standard answers.*

The conjecture formulated in [BoV2], [BoV3] turned out to be a failure, and was disproved (in one direction) by two independent counter-examples due to A.V.Stepanov [Sv2] and S.G.Khlebutin [Kh3]. In the present paper we formulate a necessary and sufficient condition for the elementary subgroup to be normal in the general linear group (see also [Sv5]), but it would be a non-trivial task to describe the resulting class of rings in purely ring-theoretical terms.

In the sequel we always assume that $n \geq 3$. It is a common knowledge that behaviour of linear groups of degree 2 is pathological. In fact, after the initial successes (where one should especially quote the seminal work of P.M.Cohn [Co1]) for decades the theory for $n = 2$ entered a retarded stage, characterized by numerous attempts to force the standard answers to hold. Of course here we do not even mention extremely interesting recent work of many authors which specifically studies GL_2 in the cases when standard answers *fail* to hold, like the very elegant papers of A.Mason or the $(2,3)$-generation project which brought truly remarkable fruits in the last few years in the works of M.C.Tamburini, J.S.Wilson, G.Malle, L.Di Martino and the second-named author, M.Liebeck and A.Shalev, and many others, resulting in a solution of an outstanding Hurwitz groups problem.

But the recent years witnessed an unexpected sea change, due to the work of D.Costa and G.Keller [CK1] – [CK3] which gave a new vitality to the whole subject of linear groups over rings, by returning GL_2 to the general context of GL_n, $n \geq 2$. Informally, their philosophy may be rendered by the following paraphrase of J.Conrad: there is nothing abnormal about the linear groups of degree 2, they are simply *extra-normal*. In fact, Costa and Keller introduced a different notion of standardness, based on a class of subgroups much wider than that of relative elementary subgroups. We are convinced that a suitable truly non-commutative version of the Costa and Keller approach must have far-reaching consequences in the theory of linear groups of higher degrees and would be the key to a complete solution of the above Problem. But even for GL_3 we are not yet quite up to it, so in the present paper we do not pursue this line any further and concentrate exclusively on the *standard* results.

There are now several methods to prove these results. Let us concentrate at the easiest one of them, normality of the elementary subgroup. One has to prove that for any elementary transvection $t_{ij}(\xi) = e + \xi e_{ij}$, $i \neq j$, $\xi \in R$, and any matrix $a \in G$ the conjugate $x = a t_{ij}(\xi) a^{-1}$ again belongs to $\mathrm{E}(n, R)$. The leading idea is to reduce to groups of smaller degree and then to apply something like Whitehead lemma. Observe that the above conjugate has the form $x = e + u\xi v$, where u is the i-th column of a and v is the j-th row of a^{-1}. At the stable level this reduction results from the surjective stability of K_1-functors. The original proof of A.A.Suslin [Su1] was based on the possibility to express any solution of a single linear equation as a sum of solutions with zero components. In fact this proof proves somewhat more, than simply normality of the elementary subgroup, namely that any transvection of the form $e + u\xi v$ belongs to $\mathrm{E}(n, R)$ if the column u is unimodular. Compare an extremely elegant rephrasing of this proof in terms of antisymmetric matrices in [SK], which also gives a slightly stronger result.

Several proofs, like Suslin's proof for almost commutative rings [Tu], Vaserstein's proof [V1] and Golubchik – Mikhalev's proof [GM2] used versions of localization method. For example the idea of Vaserstein's 'localization and patching' method (derived from Quillen – Suslin's solution of Serre's problem, see [La]) was to decompose $t_{ij}(\xi)$ as a product of transvections $t_{ij}(\zeta_h\xi)$ with respect to an appropriate partition $1=\sum\zeta_h$ in the centre of the ring R and then to apply the corresponding local results. To struggle with the zero divisors in R one had to throw in independent variables, like in the original Quillen's lemma. In [B5], [BV1] Bak proposed yet another way to overcome the difficulties due to the presence of zero divisors. One can find a survey and a comparison of these methods in [Va5], [Va9], [BV2].

I.Z.Golubchik and A.V.Mikhalev initiated the use of more powerful non-commutative Ore localizations. A striking application of these ideas is a result of I.Z.Golubchik [Go3], [Go5] on the normal structure of $\mathrm{GL}(n,R)$ over rings of finite localization dimension [GMa] all of whose primitive factor-rings are Ore. In terms of rings it is the most general result known so far. A remarkable feature of his proof, which dates not later than 1981, is that the standard commutator formulae – in particular, normality of $\mathrm{E}(n,R)$ – were not known in this generality, so that he could not invoke the usual induction procedures and had to extract a transvection in several steps (compare also [Go4], where the same trick is applied in a slightly different context). Unfortunately these outstanding results are still not fully published in a form, accessible to a Western reader and their true significance is largely underestimated.

In the present paper we propose yet another approach to the proof of these and other related results, an approach which is not ring theoretic or K-theoretic, but geometric in essence. The Theme and the Variations below all say that under certain conditions the elementary subgroup may be generated by transvections whose conjugates enjoy prescribed properties. The very general idea of this approach is to decompose a transvection $t_{ij}(\xi)$ as a product of *some* transvections x_h such that ax_ha^{-1} satisfies the assumptions of a reduction lemma (an appropriate analogue of Whitehead lemma). This is, of course, automatically the case when the transvections x, such that axa^{-1} falls in the above assumptions, generate the whole elementary group $\mathrm{E}(n,R)$. This last statement may be rephrased as a certain 'linear dependence' condition in the ring R, being a simultaneous generalization of conditions like commutativity, stable rank and Ore condition. This linear dependence condition is sufficient – sometimes even *necessary* and sufficient, compare § 8 below and [Sv5] – for the normality of $\mathrm{E}(n,R)$, standard description of normal subgroups, and further analogous problems.

For a given ring R there usually is a large freedom in constructing the above decomposition $t_{ij}(\xi)=\prod x_h$. For example in many cases the transvections x_h may be found already in the conjugates of the upper unitriangular group $\mathrm{U}(n,R)$ by permutation matrices. When R is commutative the proofs are especially elegant. In fact in this case undergraduate linear algebra tells us that the above ring theoretical conditions are obviously satisfied (the "Theme" in § 5). This observation from [Sv1] – 'Stepanov's proof of Suslin's theorem' (see [VPS], [Va5]) – was the starting point of our approach. A remarkable feature of this proof is that it provides explicit formulae expressing any conjugate of an elementary transvection as a product of factors from proper parabolic subgroups. These formulae are *geometric* in the sense that (unlike all proofs using localizations) the number of factors and their shape do not depend on the ground ring. This number was used in [SSv] to give an upper bound for the length of commutators in $\mathrm{GL}(n, R)$ where R is commutative and $n \geq 3$.

In the present paper we rewrite in the same spirit also the proof of Wilson – Golubchik theorem. The resulting proof is so elementary and conceptually clear (and even in the case of fields so much shorter than the proofs of Jordan – Dickson theorem contained in the standard textbooks) that it is almost incomprehensible why the normal structure of linear groups over commutative rings was not understood long before 1973. We discuss also other variations of this idea. For example, we show how to modify the decomposition to make it applicable to the study of Steinberg groups (van der Kallen – Tulenbaev theorem), linear groups over non-commutative rings (among other things we give a necessary and sufficient condition for the normality of elementary subgroup and a proof of Vaserstein theorem based on our method, see [Sv5] for further non-commutative applications of our ideas) and subgroups normalized by a part of elementary matrices.

As for the last topic, we give a new proof of the description of subgroups normalized by a group $E(\nu)$ of elementary block diagonal matrices. This result is essentially due to Borewicz – Vavilov [BoV2], where subgroups containing $E(\nu)$ are described. That one can pass to subgroups normalized by $E(\nu)$ here was noticed by Golubchik [Go4]. Later Stepanov [Sv4] gave a general method to obtain a description of subgroups normalized by a given subgroup from the description of subgroups containing it and its normal subgroups. In fact our methods can be applied for many other classes of subgroups, see [Va6] for an overview. Also we show how to prove the normality of the elementary net subgroup in the net normalizer for a very general class of nets.

The present paper serves as an elementary introduction to two much broader projects. One of them carried forward by the second-named author and E.B.Plotkin is concerned with similar results for groups of Lie type over commutative rings. It is described in [VPS], [Va5], [VP], where one can find many further details and additional references. Theorem 4 of [Va7], is a further variation, which transcribes the theme for the 27-dimensional representation of the Chevalley group of type E_6 and the 56-dimensional representation of the group of type E_7. Recently in a joint work of the second-named author and E.Ya.Perelman [VPe] this program was realised also for the polyvector representation of GL_n, thus giving a generalisation and a better proof of our Variation 1. Another running project is a joint work of the authors with A.Bak (see [BV2], [BV3] and following) and is concerned with generalizations to classical groups over non-commutative rings. In both cases there are substantial technical difficulties, stemming from representation theory, or ring theory, respectively. The present paper explains the basic ideas underlying the proofs as elementary and with as few technicalities, as possible.

The Theme and the present proofs of Theorem 1 go back to [Sv1], but were never published as such in this elementary form (they first appeared in print in the setting of Chevalley groups, in [VPS] and [Va5]). The proof of Theorem 2 stems essentially from [BoV3], but the reader might notice that the present exposition with the explicit use of the Theme accounts for further considerable simplifications. We do not reproduce a complete proof of Theorem 3 – it seems almost impossible to improve upon the central piece of van der Kallen's brilliant exposition in [vdK]. But we do just enough to show that his proof works with our decomposition of transvections, instead of the one which comes from Suslin's lemma. Theorem 4 is due to Vaserstein [V1] and we show that it too gets a very natural proof based on our idea. Theorem 5 stems essentially from [BoV3]. But here again we can give better proofs (of stronger results), by the same reason as above, but also because of further simplifications, proposed in [Va1] and [Va6]. Finally, Theorem 6 is a takeover of a theme from [Sv3], which is yet another generalisation of Suslin's normality theorem as well as some results from [BoV3].

Results of § 13 and § 14 essentially stem from the works of the second-named author on subgroups of the classical groups [Va1], [Va2], [Va4], [Va5] (for the even orthogonal case see also [VPS]), but were not explicitly stated there. Here we introduce further important simplifications as compared with [Va5]. The most important of them is as follows. In the symplectic group an explicit formula expressing an elementary root unipotent as a product of symplectic transvection stabilizing a column is quite complicated. But in most applications one does not need such a formula. Usually it is quite sufficient to know that the orthogonal/symplectic transvections stabilizing a column of an orthogonal/symplectic matrix, generate the corresponding elementary subgroup.

This work contains several new results, specifically the non-commutative generalizations and most of the Variations were never explicitly stated before in this form; some other results (like the ones on intermediate subgroups) were never published in a form, accessible to a Western reader. But its *Schwerpunkt* is not there. Our goal is rather to provide an extremely elementary unified treatment of the known results giving substantially shorter proofs than the published ones, and to show that they depend on the geometry of the natural module rather than on the properties of the ring. This is why we decided to make the paper entirely self-contained and included in §§ 1–4, 10 all the necessary definitions as well as the proofs of all auxiliary results we use (the exceptions are made for some group theoretical lemmas and some purely technical facts necessary only to prove some results of §§ 9 and 11, where we give an explicit reference).

We believe that a thorough understanding of how the proofs work in this simplest case should facilitate further generalizations be it in the context of classical groups over associative rings, in the context of (exceptional) Lie type groups over commutative rings, or in further contexts mentioned in § 15.

§ 1. General linear group

Let G be any group. For two elements $x, y \in G$ we denote by $[x, y] = xyx^{-1}y^{-1}$

their commutator, by $x^y = y^{-1}xy$ and ${}^y x = yxy^{-1}$ the conjugates of x by y and y^{-1} respectively. The double commutator $[[x,y],z]$ is denoted by $[x,y,z]$. Occasionally we use the Hall commutator identity in the form

$$[x, y^{-1}, z^{-1}]^x \, [z, x^{-1}, y^{-1}]^z \, [y, z^{-1}, x^{-1}]^y = e.$$

We write $H \leq G$ to denote that H is a subgroup of G (as opposed to the notation $X \subseteq G$ for arbitrary subsets) and $H \trianglelefteq G$ to denote that H is a normal subgroup of G. For a subset $X \subseteq G$ we denote by $\langle X \rangle \leq G$ the subgroup generated of G by X and by $\langle X \rangle^H$ the smallest subgroup which contains X and is normalized by H. Thus, for example, $\langle X \rangle^G$ is the normal subgroup of G generated by X. For two subgroups $F, H \leq G$ we denote by $[F, H]$ the corresponding relative commutator subgroup, generated by all commutators $[f, h]$, $f \in F$, $h \in H$.

Let R be an associative ring with 1. By an *ideal* I of a ring R we always understand a *two-sided ideal*. To denote that I is an ideal of R we write $I \trianglelefteq R$. As usually we denote by R^* the *multiplicative group* of a ring R. It consists of all *two-sided invertible* elements $x \in R$ for which there exists $y \in R$ such that $xy = 1 = yx$. We do not impose the following condition at the outset, but actually for the most part the rings we consider will be *Dedekind finite* (alias, *1-weakly finite*) i.e. such that invertibility is equivalent to one sided invertibility, or, what is the same, $xy = 1$ implies $yx = 1$.

For two natural numbers m, n we denote by $\mathrm{M}(m,n,R)$ the additive group of $m \times n$ matrices with entries in R. In particular, $\mathrm{M}(n,R) = \mathrm{M}(n,n,R)$ be the full matrix ring of degree n over R. The multiplicative group $\mathrm{M}(n,R)^*$ is denoted by $\mathrm{GL}(n,R)$ and is called the *general linear group of degree* n *over* R. Let $\mathrm{Cent}(R)$ be the centre of a ring R. Then the centre $\mathrm{C}(n,R)$ of the group $\mathrm{GL}(n,R)$ consists of the *scalar transformations* λe, $\lambda \in \mathrm{Cent}(R)^*$.

As usual we denote by e the identity matrix of degree n (specified by context) and by e_{ij} a *standard matrix unit*, i.e. the matrix which has 1 in the position (i,j) and zeros elsewhere. For a matrix $a \in \mathrm{M}(m,n,R)$ we denote by a_{ij} its matrix entry in the position (i,j), so that $a = (a_{ij}) = \sum a_{ij}e_{ij}$, $1 \leq i \leq m$, $1 \leq j \leq n$. The j-th column (a_{ij}), $1 \leq i \leq m$ of the matrix a is denoted by a_{*j}, whereas its i-th row (a_{ij}), $1 \leq j \leq n$ is denoted by a_{i*}. In the case when $a \in \mathrm{GL}(n,R)$ the inverse of a is denoted by $a^{-1} = (a'_{ij})$. With the above conventions notation a'_{*j} and a'_{i*} becomes self-explanatory. For a matrix $a \in \mathrm{M}(m,n,R)$ we denote by $a^t \in \mathrm{M}(n,m,R)$ its *formal transpose*, i.e. the matrix defined by $(a^t)_{ji} = a_{ij}$ (notice that this operation does not have good properties, the correctly defined transpose of a belongs to $\mathrm{M}(n,m,R^o)$, where R^o is the ring opposite to R).

Let R^n be a free right R-module consisting of all columns u of height n with components from R. Then $\mathrm{GL}(n,R)$ may be identified with the automorphism group of R^n. Namely $a \in \mathrm{GL}(n,R)$ acts on R^n by left multiplication: $u \mapsto au$, for $u \in R^n$. Following P.M.Cohn we denote the free left R-module consisting of all rows of length n with components from R by ${}^nR$. Then $\mathrm{GL}(n,R)$ acts on ${}^nR$

on the right: $v \mapsto va$ for $a \in \mathrm{GL}(n, R)$ and $v \in {}^nR$. There is a natural pairing $R^n \times {}^nR \to R$ furnished by multiplication.

The standard bases of R^n and ${}^nR$ are denoted by $e_1, \dots, e_n$ and $f_1, \dots, f_n$ respectively. In other words e_i is the i-th column e_{*i} of the identity matrix e of order n (specified by the context), while f_i is the i-th row e_{i*} of e. Clearly for a column $u \in R^n$ one has $u_i = f_i u$ and, similarly, $v_i = ve_i$ for a row $v \in {}^nR$. In general, for a matrix $a = (a_{ij}) \in \mathrm{M}(m, n, R)$ one has $a_{*j} = ae_j$, $a_{i*} = f_i a$ and $a_{ij} = f_i a e_j$. These formulae will be occasionally used in the sequel to facilitate calculation of matrix entries.

A column $u = (u_1, \dots, u_n)^t \in R^n$ is called *unimodular* if it is *left invertible* as a matrix, i.e. if there exists a row $v = (v_1, \dots, v_n) \in {}^nR$ such that $vu = v_1u_1 + \dots + v_nu_n = 1$. This amounts to saying that the components of u generate the unit *left* ideal, $Ru_1 + \dots + Ru_t = R$. Similarly, a row $v = (v_1, \dots, v_n) \in {}^nR$ is called *unimodular* if it is *right invertible* as a matrix, or, what is the same, if its components generate the unit *right* ideal, $v_1R + \dots + v_nR = R$. Columns and rows of an invertible matrix are unimodular (in particular, e_i's and f_i's are), but the converse is not true in general: not necessarily a unimodular column/row can be included into an invertible matrix.

The set of all unimodular columns of height n will be denoted by $U(R^n)$ and the set of all unimodular rows of length n will be denoted by $U({}^nR)$. In the sequel we shall quite often use columns and rows with zero components. The set of columns of height n, respectively the set of rows of length n, which have at least one zero component will be denoted by $N(R^n)$, respectively by $N({}^nR)$.

Now we recall definitions of the congruence subgroups modulo an ideal. Let I be an ideal in R. There is a canonical projection $\rho_I : R \to R/I$, sending any element $\lambda \in R$ to the element $\overline{\lambda} = \lambda \bmod I = \lambda + I$. It defines the corresponding reduction homomorphism

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

("*reduction modulo I*") such that the image of a matrix $a = (a_{ij})$ equals $\overline{a} = (\overline{a}_{ij})$.

The kernel of the homomorphism ρ_I is denoted by $\mathrm{GL}(n, R, I)$ and is called the *principal congruence subgroup in* $\mathrm{GL}(n, R)$ *of level* I. It consists of all matrices congruent to e modulo I. There is another natural normal subgroup associated to an ideal. Namely, consider the centre $\mathrm{C}(n, R/I)$ of $\mathrm{GL}(n, R/I)$. Then the inverse image of $\mathrm{C}(n, R/I)$ under the reduction homomorphism ρ_I is denoted by $\mathrm{C}(n, R, I)$ and is called the *full congruence subgroup of level* I. It consists of all matrices congruent to a scalar matrix modulo I. For a proper ideal $I \trianglelefteq R$ the groups $\mathrm{GL}(n, R, I)$, $\mathrm{C}(n, R, I)$ will be called *relative*, as opposed to the *absolute* group $\mathrm{GL}(n, R) = \mathrm{GL}(n, R, R) = \mathrm{C}(n, R, R)$. Notice that $\mathrm{C}(n, R) = \mathrm{C}(n, R, 0)$ is the absolute centre.

Occasionally we consider the group $\mathrm{GL}(n, R)$ embedded in $\mathrm{GL}(n+1, R)$. Unless otherwise specified the embedding is given by the usual rule

$$a \mapsto a \oplus 1 = \begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix}, \tag{1}$$

where $\oplus$ denotes the direct sum of matrices. Iterating this embedding one gets the embedding $\mathrm{GL}(n,R) \leq \mathrm{GL}(m,R)$ for any $m \geq n$.

§ 2. Elementary subgroups

An *elementary transvection* $t_{ij}(\xi)$ is a matrix of the form $t_{ij}(\xi) = e+\xi e_{ij}$, $\xi \in R$, $1 \leq i \neq j \leq n$. Elementary transvections are subject to the following well-known relations. The first one of them is additivity in ξ:

$$t_{ij}(\xi)t_{ij}(\eta) = t_{ij}(\xi+\eta) \tag{2}$$

and the following is a special case of the Chevalley commutator formula

$$[t_{ij}(\xi), t_{hk}(\eta)] = \begin{cases} e, & \text{if } j \neq h,\ i \neq k, \\ t_{ik}(\xi\eta), & \text{if } j = h,\ i \neq k, \\ t_{hj}(-\eta\xi), & \text{if } j \neq h,\ i = k. \end{cases} . \tag{3}$$

Notice, that the case of two *opposite* transvections $t_{ij}(\xi)$ and $t_{ji}(\zeta)$ is excluded here.

The subgroup $\mathrm{E}(n,R)$ of the general linear group $\mathrm{GL}(n,R)$, generated by all the elementary transvections

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

is called the *elementary subgroup* of the general linear group. In the traditional usage the expression "elementary matrices" is reserved to the elementary transvections $t_{ij}(\xi)$ and the elementary pseudo-reflections $d_i(\varepsilon) = e + (\varepsilon - 1)e_{ii}$, $\varepsilon \in R^*$. But we will follow the usual K-theory slang and refer to all elements of $\mathrm{E}(n,R)$ as *elementary matrices*.

Consider the elementary subgroup $\mathrm{E}(n,I)$ generated by the elementary transvections of level I:

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

Obviously, this group does not depend on the choice of an ambient ring R, but only on I itself. In general the group $\mathrm{E}(n,I)$ has no chances to be normal in $\mathrm{GL}(n,R)$. To be normal in $\mathrm{GL}(n,R)$ a group has at least to be normalized by $\mathrm{E}(n,R)$.

The normal subgroup of the elementary group $\mathrm{E}(n,R)$ generated by the elementary transvections of level I is called the *elementary subgroup of level I* and is denoted by $\mathrm{E}(n,R,I)$. In other words, $\mathrm{E}(n,R,I)$ is the normal closure of $\mathrm{E}(n,I)$ in $\mathrm{E}(n,R)$:

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

For a proper ideal $I \trianglelefteq R$ we refer to the elements of $\mathrm{E}(n,R,I)$ as the elementary matrices *relative* to I. It easily follows from (3) that $\mathrm{E}(n,R,I)$ is generated by the matrices $t_{ij}(\zeta)t_{ji}(\xi)t_{ij}(-\zeta)$ (see [VS], [V1] or [Va8] for details).

The following elementary corollary of (3) is constantly used in the sequel.

Lemma 1. *Let $n \geq 3$. Then for any ideal $I \trianglelefteq R$ one has*

$$[\mathrm{E}(n,R),\mathrm{E}(n,R,I)] = \mathrm{E}(n,R,I).$$

In particular, the absolute elementary subgroup $\mathrm{E}(n,R)$ is perfect,

$$[\mathrm{E}(n,R),\mathrm{E}(n,R)] = \mathrm{E}(n,R)$$

.

We need also the following technical fact.

Lemma 2. *Let $n \geq 3$. If for some $a \in \mathrm{GL}(n,R)$ one has $[\mathrm{E}(n,R),a] \leq \mathrm{C}(n,R)$, then $a \in \mathrm{C}(n,R)$.*

Proof. Clearly, a matrix which commutes with $\mathrm{E}(n,R)$ is central, i.e. belongs to $\mathrm{C}(n,R)$ (it is enough to look at the commutators $[t_{ij}(1),a]$, for $1 \leq i \neq j \leq n$). By assumption, $[\mathrm{E}(n,R),a,\mathrm{E}(n,R)] = [a,\mathrm{E}(n,R),\mathrm{E}(n,R)] = 1$. But then by the preceding lemma and the Hall identity one has

$$[\mathrm{E}(n,R),g] = [\mathrm{E}(n,R),\mathrm{E}(n,R),g] \leq [\mathrm{E}(n,R),g,\mathrm{E}(n,R)][g,\mathrm{E}(n,R),\mathrm{E}(n,R)] = 1.$$

It turns out that in many important situations (at the stable level, over a commutative ring, etc.) the group $\mathrm{E}(n,R,I)$ is normal in $\mathrm{GL}(n,R)$ for any ideal $I \trianglelefteq R$, or, in other words,

$$[\mathrm{GL}(n,R),\mathrm{E}(n,R,I)] = \mathrm{E}(n,R,I). \tag{4}$$

We refer to this fact and the formula

$$[\mathrm{E}(n,R),\mathrm{C}(n,R,I)] = \mathrm{E}(n,R,I) \tag{5}$$

as the *standard commutator formulae.*

Usually one says that the *standard description* of subgroups in $\mathrm{GL}(n,R)$ normalized by $\mathrm{E}(n,R)$ holds, if for every such subgroup H there exists a unique ideal $I \trianglelefteq R$ such that

$$\mathrm{E}(n,R,I) \leq H \leq \mathrm{C}(n,R,I). \tag{6}$$

Colloquially one often refers to this also as the *standard description of normal subgroups* in $\mathrm{GL}(n,R)$. In the sequel we adopt the following more restrictive definition. We use the expression *standard description* only in the presence of the standard commutator formulae, i.e. when *both* (4) and (5) hold for all ideals $I \trianglelefteq R$ and (6) holds for all subgroups, normalized by $\mathrm{E}(n,R)$.

§ 3. Transvections

First, let $R = K$ be a skew-field. In this case a *transvection* is simply a conjugate of an elementary transvection $t_{ij}(\xi)$, $\xi \in K$, $1 \leq i, j \leq n$. The class of transvections admits an intrinsic description. Namely, recall that $a \in \mathrm{GL}(n, K)$ is called *unipotent* if $a - e$ is nilpotent and that the *residue* of a equals by definition the rank of $a - e$, $\mathrm{res}(a) = \mathrm{rk}(a - e)$. Now a transvection may be characterized as a unipotent matrix of residue 1. Over rings this notion ramifies into a variety of different concepts, according to what is called a matrix of rank 1 (see [Va5], [VPe] and [Va9] for comparison of some of these definitions).

Now, let R be any associative ring. A *transvection* is a matrix of the form $t_{uv}(\xi) = e + u\xi v$, where $u \in R^n$, $v \in {}^nR$ are a column and a row such that $vu = 0$, and $\xi \in R$. The parameter ξ above is superfluous, since it may be concealed in u or v as follows:

$$t_{uv}(\xi) = t_{\xi u,v}(1) = t_{u,\xi v}(1). \tag{7}$$

Indeed, many authors use the notation $t_{uv} = t_{uv}(1)$ (geometrically u describes the *centre* of t_{uv}, whereas the hyperplane orthogonal to v is its *axis*). But sometimes, especially when handling the one-parameter subgroups of transvections with a given centre and axis, it is more convenient to explicitly document the dependence on ξ. Denote by $\mathrm{T}(n, R)$ the *transvection subgroup* of $\mathrm{GL}(n, R)$, generated by all transvections $t_{uv}(\xi)$ (see “Winnie-the-Pooh comments” in [Va5] for the definition of other related groups). For an ideal $I \trianglelefteq R$ one can define the corresponding relative group $\mathrm{T}(n, R, I)$ generated by the transvections $t_{uv}(\xi)$, where $\xi \in I$.

An elementary transvection $t_{ij}(\xi)$ is a transvection (set $u = f_i$, $v = e_j$). This shows that $\mathrm{E}(n, R, I) \leq \mathrm{T}(n, R, I)$ but the converse is not true in general. More generally, for any column $u \in R^n$ and any row $v \in {}^nR$ the matrices

$$t_{*j}(u) = e + (u - u_j e_j) f_j = \prod t_{ij}(u_i), \quad i \neq j,$$

$$t_{i*}(v) = e + e_i(v - v_i f_i) = \prod t_{ij}(v_j), \quad j \neq i,$$

are transvections. Moreover these transvections belong to $\mathrm{E}(n, R, I)$ for $u_i, v_j \in I$ $(i \neq j)$. These elements provide the simplest examples of transvections not conjugate to the elementary ones. Indeed, take a row $v = (0, v_2, \dots, v_n)$ such that the right ideal generated by $v_2, \dots, v_n$ cannot be generated by less than $n - 1$ elements. Then the transvection $t_{1*}(v) = \prod t_{1j}(v_j)$, $j \neq 1$, cannot be possibly conjugate to an elementary transvection. The elements $t_{*j}(u)$ and $t_{i*}(v)$ are a special case of what is called “elementary fake root unipotents” in [Va5], [Va7], [VP].

Below we collect the first properties of transvections, which will be used in the sequel without explicit reference. First of all, a conjugate of a transvection is a transvection. More precisely, let $a \in \mathrm{GL}(n,R)$, $u \in R^n$, $v \in {}^nR$, $vu = 0$, $\xi \in R$. Then

$$at_{uv}(\xi)a^{-1} = t_{au,va^{-1}}(\xi). \tag{8}$$

The following additivity property generalizes relation (2). Let $u, w \in R^n$ be two columns, and $v \in {}^nR$ be a row such that $vu = vw = 0$ and $\xi \in R$. Then

$$t_{u+w,v}(\xi) = t_{uv}(\xi)t_{wv}(\xi). \tag{9}$$

Finally, one has a commutator formula, generalizing (3). Namely, let $u, x \in R^n$ be two columns, and $v, y \in {}^nR$ be two rows such that $vu = yx = yu = 0$, $\xi, \zeta \in R$. Then

$$[t_{uv}(\xi), t_{xy}(\zeta)] = t_{uy}(\xi vx\zeta). \tag{10}$$

The following result is crucial in the reduction to smaller ranks. It is a variant of the Whitehead lemma due to Bass [Ba1], later Vaserstein further generalized it (see [Su1], [VS], [V1], [Va5] for details). It asserts that $\mathrm{T}(n,R,I) \leq \mathrm{E}(n+1,R,I)$ with respect to the embedding (1).

Lemma 3. *Let I be an ideal of R, $\xi \in I$. Further, let, $u \in R^n$ and $v \in {}^nR$ be a column and a row such that $vu = 0$ and $u \in N(R^n)$ or $v \in N({}^nR)$. Then the transvection $e + u\xi v$ belongs to $\mathrm{E}(n,R,I)$.*

Proof. Let $v_j = 0$ for some j. Representing u as $u = u_je_j + (u - u_je_j)$ and using (9) and (10) one has

$$\begin{aligned} e + u\xi v &= (e + u_je_j\xi v)(e + (u - u_je_j)\xi v) = \\ &= (e + u_je_j\xi v)[e + (u - u_je_j)f_j, e + e_j\xi v]. \end{aligned} \tag{11}$$

Now $e + u_je_j\xi v$ and $e + e_j\xi v$ belong to $\mathrm{E}(n,R,I)$, whereas $e + (u - u_je_j)f_j \in \mathrm{E}(n,R)$. But $\mathrm{E}(n,R,I)$ being normal in $\mathrm{E}(n,R)$ this implies that $e + u\xi v \in \mathrm{E}(n,R,I)$. The proof for the case, when $u_i = 0$ for some i is similar.

The following corollary singles out the main idea behind the proof of normality of the elementary subgroup at the stable level due to Bass [Ba1], [Ba2]. Actually, Bass uses surjective stability for the K_1-functor, or, in other words, the decomposition $\mathrm{GL}(n,R) = \mathrm{GL}(n-1,R)\,\mathrm{E}(n,R)$, but the reason why it works is exceedingly simple.

Corollary. *Assume that for any unimodular column $u \in U(R^n)$ there exists an elementary matrix $a \in \mathrm{E}(n,R)$ such that $au \in N(R^n)$. Then $\mathrm{E}(n,R,I) \trianglelefteq \mathrm{GL}(n,R)$.*

Proof. It suffices to show that $gt_{ij}(\xi)g^{-1} \in \mathrm{E}(n,R,I)$ for all $\xi \in I$, $1 \leq i \neq j \leq n$, $g \in \mathrm{GL}(n,R)$. Let g_{*i} be the i-th column of g. By assumption there exists $a \in \mathrm{E}(n,R)$ such that $u = ag_{*i} \in N(R^n)$. Let, say, $u_h = (ag)_{hi} = 0$. Set $v = g'_{j*}a^{-1}$. The transvection $(ag)t_{ij}(\xi)(ag)^{-1} = e + u\xi v$ falls into the conditions of Lemma 3 and thus belongs to $\mathrm{E}(n,R,I)$. But then

$$gt_{ij}(\xi)g^{-1} = a^{-1}(ag)t_{ij}(\xi)(ag)^{-1}a \in \mathrm{E}(n,R,I)$$

by the very definition of the latter group.

§ 4. Reduction to smaller degree

In this section we collect some well-known technical results. They reduce the problem of the standard description of normal subgroups of $\mathrm{GL}(n,R)$ to the possibility to extract a non-central matrix with a zero entry from a subgroup normalized by $\mathrm{E}(n,R)$. For the most part the following lemmas are to be found already in [Ba1], see also [V1] or [BoV3]. They are included to make our proof of Theorem 2 completely self-contained. Recall that we always assume $n \geq 3$. First we show how to lift an elementary transvection modulo an ideal.

Lemma 4. *Let H be a subgroup of the general linear group $\mathrm{GL}(n,R)$ normalized by the elementary subgroup $\mathrm{E}(n,R)$ and let I be set of all $\xi \in R$ such that $t_{12}(\xi)$ belongs to H. Then I is an ideal of R and $\mathrm{E}(n,R,I) \leq H$.*

Proof. It follows from (2) that I is an additive subgroup. Let $\xi \in I$ and $\zeta \in R$. Applying (3) twice we get

$$t_{12}(\zeta\xi) = \big[t_{13}(1), [t_{31}(\zeta), t_{12}(\xi)]\big], \qquad t_{12}(\xi\zeta) = \big[[t_{12}(\xi), t_{32}(\zeta)], t_{23}(1)\big] \quad \in H,$$

so that I is indeed an ideal. To finish the proof it remains to observe that $\mathrm{E}(n,R,I)$ coincides with the normal closure of $\{t_{12}(\xi) \mid \xi \in I\}$ in $\mathrm{E}(n,R)$.

Lemma 5. *Let R be a ring such that for all ideals $I \triangleleft R$ the following two conditions are satisfied:*

1. *Any non-central subgroup of $\mathrm{GL}(n,R/I)$, normalized by $\mathrm{E}(n,R/I)$ contains a non-trivial elementary transvection;*
2. *The commutator formula $[\mathrm{E}(n,R), \mathrm{GL}(n,R,I)] = \mathrm{E}(n,R,I)$ holds.*

Then the normal structure of $\mathrm{GL}(n,R)$ is standard.

Proof. Let I be the ideal associated with the group in Lemma 4. Consider the image $\overline{H}$ of H in $\mathrm{GL}(n,R/I)$ under the reduction homomorphism modulo I. Clearly $\overline{H}$ is normalized by $\mathrm{E}(n,R/I)$ and if it is non-central, then by assumption there exists a non-trivial elementary transvection $t_{ij}(\overline{\xi}) \in \overline{H}$, where $\xi \notin I$. This means that $t_{ij}(\xi) = bc$, where $b \in H$, $c \in \mathrm{GL}(n,R,I)$, and thus for any $a \in \mathrm{E}(n,R)$ one has $[a, t_{ij}(\xi)] = [a, bc] = [a,b]b[a,c]b^{-1}$. But $[a,b] \in H$ since H is normalized by $\mathrm{E}(n,R)$ and $[a,c] \in \mathrm{E}(n,R,I) \in H$ by the standard commutator formula so that again $b[a,c]b^{-1} \in H$. This means that $[a, t_{ij}(\xi)] \in H$ for any elementary matrix, and, in particular, $t_{hj}(\xi) = [t_{hi}(1), t_{ij}(\xi)] \in H$, implying that $\xi \in I$, which is absurd. This contradiction shows that $\overline{H}$ is central in $\mathrm{GL}(n,R/I)$ and thus $H \leq \mathrm{C}(n,R,I)$.

In turn to prove that H contains an elementary transvection it is enough to prove much simpler facts.

Lemma 6. *Let H be a subgroup in $\mathrm{GL}(n,R)$, normalized by the elementary subgroup $\mathrm{E}(n,R)$. If H contains a non-central matrix g such that $a_{*j} = e_j\lambda$ for some $\lambda \in R^*$, then H contains a non-trivial elementary transvection.*

Proof. Clearly, we may suppose that $j = 1$, or, in other words, that H contains a non-central matrix of the form

$$a = \begin{pmatrix} \lambda & u \\ 0 & c \end{pmatrix},$$

where $\lambda \in R^*$, $u \in {}^{n-1}R$ and $c \in \mathrm{GL}(n-1, R)$. First we show that H contains a matrix of the form

$$b = \begin{pmatrix} \lambda & u \\ 0 & \lambda e \end{pmatrix},$$

where either $u \neq 0$, or λ is non-central. This is obvious if $c = \lambda e$ (recall that a is non-central). Let $c \neq \lambda e$. Take a row $f_k \in {}^{n-1}R$, $k \in \{1, \dots, n-i\}$ and consider the following commutator:

$$\left[\begin{pmatrix} 1 & f_k \\ 0 & e \end{pmatrix}, \begin{pmatrix} \lambda & u \\ 0 & c \end{pmatrix}\right] = \begin{pmatrix} 1 & f_k(c - \lambda e)c^{-1} \\ 0 & e \end{pmatrix} \in H$$

Since $c - \lambda e \neq 0$, one can choose k in such a way that $f_k(c - \lambda e)c^{-1} \neq 0$. Thus H always contains a non-central matrix b.

Now if λ is non-central, then there exists $\alpha \in R$ such that $\lambda\alpha \neq \alpha\lambda$ and thus $t_{12}(\lambda\alpha\lambda^{-1} - \alpha) = [b, t_{12}(\alpha)] \in H$ is a non-trivial elementary transvection. On the other hand if λ is central, but $u \neq 0$, then $t_{ij}(\lambda^{-1}u_i) = [b, t_{ij}(1)] \in H$, where $i \neq 1$ is chosen so that $u_i \neq 0$ and $j \neq 1, i$.

Remark. For a Dedekind finite ring it is sufficient to require that a column of a be proportional to the corresponding column of the identity matrix, $a_{*j} = e_j\lambda$. Then automatically the coefficient λ will be invertible. But in general this is not the case, since the inverse of an upper triangular matrix need not to be upper triangular, as the following example shows. Take $\lambda, \mu \in R$ such that $\mu\lambda = 1$ but $\lambda\mu \neq 1$. Then

$$\begin{pmatrix} \lambda & 1 \\ 0 & -\mu \end{pmatrix}^{-1} = \begin{pmatrix} \mu & 1 \\ 1 - \lambda\mu & -\lambda \end{pmatrix}.$$

In general if $a_{*j} = e_j\lambda$ for an invertible matrix a, one can conclude only that λ is left invertible. Of course an analogue of this lemma holds if any row of the matrix a is proportional to the corresponding row of the identity matrix. In general, here one can conclude from $a_{i*} = \lambda f_i$ only that λ is right invertible.

Lemma 7. *Let H be a subgroup in $\mathrm{GL}(n, R)$, normalized by the elementary subgroup $\mathrm{E}(n, R)$. If H contains a non-central matrix a commuting with an elementary transvection $t_{ij}(1)$, $1 \leq i \neq j \leq n$, then H contains a non-trivial elementary transvection.*

Proof. If a commutes with $t_{ij}(1)$, then $a_{hj} + a_{hi} = a_{hj}$ for all $h \neq i$ and thus the i-th column of a equals λe_i. To apply the preceding lemma it remains to show that

λ is invertible. Invertibility of a shows that λ is *left* invertible. But now the same argument as above applied to the j-th row of a shows that $a_{ih} + a_{jh} = a_{ih}$ for all $h \neq j$ and, finally, $a_{ij} + a_{ii} = a_{ij} + a_{jj}$. Thus the j-th row of a equals λf_j, which shows that λ is also right invertible and thus the preceding lemma implies that H contains an elementary transvection.

Lemma 8. *Let H be a subgroup in $\mathrm{GL}(n, R)$, $n \geq 3$, normalized by the elementary subgroup $\mathrm{E}(n, R)$. If H contains a non-central matrix a with a zero entry, then H contains a non-trivial elementary transvection.*

Proof. Conjugating by a monomial matrix from $\mathrm{E}(n, R)$ we may suppose that either $a_{11} = 0$, or $a_{n1} = 0$. Consider $b = [a, t_{12}(1)] \in H$. By Lemma 7 we may suppose that $b \neq e$. If $a_{n1} = 0$, then the last row of b is the same as in the identity matrix. Since $b \neq e$, it is non-central and we can apply the Lemma 6. On the other hand, if $a_{11} = 0$, then the first row of b coincides with the first row of $t_{12}(-1)$, i.e. is equal to $(1, -1, 0, \dots, 0)$. Since $b \neq e$, it is non-central, and since $n \geq 3$ we have $b_{1n} = 0$. Thus we can apply to b the previous case.

§ 5 Theme: decomposition of transvections

In this section we present the idea in its simplest form.

Theme. *Let R be a commutative ring. Then for any $a \in \mathrm{GL}(n, R)$, $n \geq 3$, the elementary subgroup $\mathrm{E}(n, R)$ is generated by the transvections $e + e_i v$, where $1 \leq i \leq n$ and the row v is such that v_i=0 and va^{-1} has at least one zero component.*

Proof. In other words, we have to prove that there exists a set of rows $Y = Y(a) \subseteq {}^nR$ such that

i) $Y \subseteq N({}^nR)$, or, in other words, each row $v \in X$ has at least one zero component, say $v_i = 0$ for some $i = i(v)$;

ii) $Ya^{-1} \subseteq N({}^nR)$, or, in other words, each row va^{-1}, $v \in Y$ has at least one zero element;

iii) transvections $t_{i*}(v) = e + e_i v$, $i = i(v)$, generate $\mathrm{E}(n, R)$ as v ranges over Y.

Take pair-wise distinct indices $i, j, h \in \{1, \dots, n\}$ (this is where one uses the assumption $n \geq 3$), an arbitrary index $k \in \{1, \dots, n\}$, an arbitrary element $\xi \in R$, and set
$$v = a_{kh}\xi a'_{hk} f_j - a_{kh}\xi a'_{jk} f_h.$$

Clearly such a v satisfies condition (i) above, since $v_l = 0$ for all $l \neq j, h$, and in particular, $v_i = 0$. To check condition (ii) observe that commutativity of R implies that
$$(va^{-1})_k = a_{kh}\xi a'_{hk} a'_{jk} - a_{kh}\xi a'_{jk} a'_{hk} = 0.$$

It remains to verify that transvections $t_{i*}(v) = e + e_i v$, $v \in Y$, span the elementary subgroup $\mathrm{E}(n, R)$. It is sufficient to express any elementary transvection $t_{ij}(\xi)$ as a product of $t_{i*}(v)$'s for the above rows v. With this end note that

$$t_{i*}(v) = t_{ij}(a_{kh}\xi a'_{hk})t_{ih}(-a_{kh}\xi a'_{jk})$$

and that commutativity of R implies that

$$t_{ij}(\xi) = \prod_{k=1}^{n} t_{ij}(a_{kh}\xi a'_{hk})t_{ih}(-a_{kh}\xi a'_{jk}),$$

This finishes the proof.

Remark. In fact, the above proof shows somewhat more, namely, that for any ideal $I \triangleleft R$ the *relative* elementary subgroup $\mathrm{E}(n, R, I)$ is generated by the transvections $e + e_i\xi v$, where the rows v are as above, while $\xi \in I$. Alternatively, one can take the same transvections $e + e_i v$, as above, but restrict oneself to the rows v with entries in I. But in most cases this generalization is not necessary, since there are standard ways to reduce problems about relative elementary subgroups to those about absolute elementary subgroups, possibly over a different ring.

Now we are in a position to prove the main structure theorems in the general linear group over a commutative ring. The following result combines Suslin normality theorem [Su1] and the second standard commutator formula, proven in this case by Vaserstein [V1] and Borewicz – Vavilov [BoV3].

Theorem 1. *Let R be a commutative ring, I be an ideal of R and $n \geq 3$. Then the standard commutator formulae hold. In other words,* $\mathrm{E}(n, R, I)$ *is normal in* $\mathrm{GL}(n, R)$ *and* $[\mathrm{E}(n, R), \mathrm{C}(n, R, I)] = \mathrm{E}(n, R, I)$.

Proof. First we prove normality of the elementary subgroup in the absolute case. We have to check that for any matrix $a \in \mathrm{GL}(n, R)$ one has $a\,\mathrm{E}(n, R)a^{-1} \leq \mathrm{E}(n, R)$. Let $Y = Y(a)$ be the set of rows constructed in the Theme. Since va^{-1} has a zero entry,

$$a(e + e_i v)a^{-1} = e + (ae_i)(va^{-1}) \in \mathrm{E}(n, R)$$

by Lemma 3. Since the transvections $e + e_i v$, $i = i(v)$, generate $\mathrm{E}(n, R)$ as v ranges over Y, it follows that $a\,\mathrm{E}(n, R)a^{-1} \leq \mathrm{E}(n, R)$.

Next we prove that for any ideal $I \leq R$ one has

$$[\mathrm{E}(n, R), \mathrm{GL}(n, R, I)] = \mathrm{E}(n, R, I). \tag{12}$$

Indeed, take a matrix $a \in \mathrm{GL}(n, R, I)$ and let, as above, $Y = Y(a)$ be the set of rows constructed in the Theme. Since the transvections $e + e_i v$, $i = i(v)$, generate $\mathrm{E}(n, R)$ as v ranges over Y, it suffices to prove that the commutator $c = [e + e_i v, a]$ belongs to $\mathrm{E}(n, R, I)$. Let, say, $(va^{-1})_j = 0$. Then by Lemma 3 one has

$$\begin{aligned} c &= (e + e_i v)\big(e - (ae_i)(va^{-1})\big) = \\ &= (e + e_i v)(e - a_{ji}e_j(va^{-1}))[e + (ae_i - a_{ji}e_j)f_j, e - e_j va^{-1}]. \end{aligned}$$

All these factors belong to $\mathrm{E}(n,R)$ and now we shall consider this product modulo the normal subgroup $\mathrm{E}(n,R,I)$. Since $a \in \mathrm{GL}(n,R,I)$, we have $va^{-1}-v \in {}^nI$ and $ae_i - e_i \in I^n$.

First, let $j=i$. In this case $ae_i - a_{ji}e_j \in {}^nI$ and thus the commutator belongs to $\mathrm{E}(n,R,I)$. On the other hand

$$e - a_{ji}e_j\xi va^{-1} \equiv e - e_i v \,(\operatorname{mod} \mathrm{E}(n,R,I)).$$

Now, let $j \neq i$. Then $a_{ji} \in I$ so that the second factor in the above expression for c belongs to $\mathrm{E}(n,R,I)$, while the commutator is congruent to

$$[e + e_i f_j, e - e_j\xi v] \equiv e - e_i v \,(\operatorname{mod} \mathrm{E}(n,R,I)).$$

Thus in both cases $c \in \mathrm{E}(n,R,I)$.

Now we are in a position to finish the proof of the theorem. Indeed by (12) the relative elementary subgroup $\mathrm{E}(n,R,I)$ is normal in $\mathrm{GL}(n,R)$ being the mutual commutator subgroup of two normal subgroups $\mathrm{E}(n,R)$ and $\mathrm{GL}(n,R,I)$. The second standard commutator formula follows from (12), normality of $\mathrm{E}(n,R,I)$, and Lemma 1 by a standard group theoretic argument (see, for example, [Ba2], Chapter V, Lemma 5.1).

Remark. Actually, a peculiarity of the commutative case is that the construction of the rows in the Theme leads to an *explicit* expression of $at_{ij}(\xi)a^{-1}$ as a product of elementary transvections. Namely,

$$at_{ij}(\xi)a^{-1} = \prod at_{ij}(a_{kh}\xi a'_{hk})t_{ih}(-a_{kh}\xi a'_{jk})a^{-1}, \quad 1 \le k \le n,$$

where each factor on the right hand side satisfies conditions of Lemma 3 and thus may be further decomposed according to (11). There are n factors and it follows from (11) that each of them is a product of at most $4(n-1)$ elementary transvections. This gives a formula to express $at_{ij}(\xi)a^{-1}$ as a product of at most $4n(n-1)$ elementary transvections which is *polynomial* in a_{hk} and ξ. This is not necessarily the most economical expression, as far as the constant 4 is concerned, but even over a field one cannot improve the order of growth. For example the leading term of the expression resulting from Bruhat (or Birkhoff) decomposition equals $n(n-1)$ (the number of roots of $\mathrm{SL}(n,R)$, there are further summands depending on fundamental roots).

The following theorem was discovered by Wilson [Wi] and Golubchik [Go1]. Later several further proofs of this result and its broad generalizations were proposed by various authors, as described in the introduction.

Theorem 2. *Let R be a commutative ring and $n \geq 3$. Then for any subgroup H in $\mathrm{GL}(n,R)$ normalized by $\mathrm{E}(n,R)$ there exists a unique ideal $I \leq R$ such that*

$$\mathrm{E}(n,R,I) \le H \le \mathrm{C}(n,R,I).$$

Proof. The property of being commutative is inherited by all factors of R. It follows from Theorem 1 and Lemma 4 that it suffices to prove that if H is non-central then it contains a non-trivial elementary transvection. In turn, Lemma 7 implies that for this it suffices to find in H a non-central matrix with a zero entry. Indeed, let $a \in H$ be a non-central matrix. Let $Y = Y(a)$ be the set of rows constructed in the Theme.

First, let $c = [e + e_i v, a] \in H$ be non-central for some $v \in Y(a)$, $i = i(v)$. Let $va^{-1}e_j = (va^{-1})_j = 0$. We claim that $c_{hj} = 0$ for all $h \neq i, j$ (such an h exists since $n \geq 3$). In fact, since $f_h e_i = f_h e_j = 0$, one has

$$c_{hj} = f_h(e + e_i v)(e - (ae_i)(va^{-1}))e_j = f_h a e_i v a^{-1} e_j = 0,$$

in which case c is the desired non-central matrix with a zero entry.

Thus we may assume that for all transvections $e + e_i v$, $v \in Y(a)$, $i = i(v)$, the commutator $[e + e_i v, a]$ belongs to $\mathrm{C}(n, R)$. Since the above transvections generate $\mathrm{E}(n, R)$, this implies that $[\mathrm{E}(n, R), a] \leq \mathrm{C}(n, R)$. Then by Lemma 2 a itself is central, contrary to the assumption. This contradiction shows that H always contains a non-central matrix with a zero entry and we are done.

§ 6. First variation: minors

The Theme of the preceding section is just the simplest case of a decomposition of an elementary transvection $t_{ij}(\xi)$ into a product of transvections satisfying conditions of Whitehead lemma. Usually there is an enormous freedom in the choice of such a decomposition, and here we start with these variations. The following generalization is useful in the analysis of the Steinberg group. It was also used in the proofs of similar results (especially the analogues of Theorems 1 and 2) for other classical groups, specifically for symplectic groups in [Va1], [Va5] and, more recently, for unitary groups in [BV2, BV3]. In the following lemma we strengthen the conclusion of the Theme by imposing that the rows vg^{-1} have *two* zero components. This has a price: the lower bound on n must be increased by 1.

Variation 1. *Let R be a commutative ring. Then for any $a \in \mathrm{GL}(n, R)$, $n \geq 4$, the elementary subgroup $\mathrm{E}(n, R)$ is generated by the transvections $e + e_i v$, where $1 \leq i \leq n$ and the row v is such that v_i=0 and va^{-1} has at least two zero components.*

Proof. In other words, we have to prove that there exists a set of rows $Y = Y(a) \subseteq {}^nR$ satisfying the same conditions (i) and (iii) as in the proof of Theme but with condition (ii) replaces by the following stronger condition

ii)$'$ Each row va^{-1}, $v \in Y$, has at least two zero components.

In the construction of Y we will use the following notation. For a matrix $a \in \mathrm{GL}(n, R)$ and four indices $r \neq s$, $p \neq q$, we consider the corresponding minors $\Delta_{sr}^{pq} = \Delta_{sr}^{pq}(a)$ of a of order two, standing on the intersection of the r-th and the s-th rows with the p-th and the q-th columns. We will need also to consider the minors of order two of the inverse matrix, and will use the notation $\widetilde{\Delta}_{rs}^{pq} = \widetilde{\Delta}_{rs}^{pq}(a) = \Delta_{rs}^{pq}(a^{-1})$. In other words,

$$\Delta_{rs}^{pq} = \det \begin{pmatrix} a_{rp} & a_{rq} \\ a_{sp} & a_{sq} \end{pmatrix} \qquad \text{and} \qquad \widetilde{\Delta}_{rs}^{pq} = \det \begin{pmatrix} a'_{rp} & a'_{rq} \\ a'_{sp} & a'_{sq} \end{pmatrix} .$$

Take pair-wise distinct indices $i, j, h, k \in N = \{1, \dots, n\}$ (this is where one uses the assumption $n \geq 4$), two arbitrary distinct indices $p, q \in N$, an arbitrary element $\xi \in R$, and set

$$v = \Delta_{pq}^{hk}\xi(\widetilde{\Delta}_{hk}^{pq} f_j - \widetilde{\Delta}_{jk}^{pq} f_h + \widetilde{\Delta}_{jh}^{pq} f_k).$$

Clearly such a v satisfies condition (i) above, since $v_l = 0$ for all $l \neq j, h, k$, and in particular, $v_i = 0$. To check condition (ii)$'$ consider the matrix

$$\begin{pmatrix} a'_{jr} & a'_{jp} & a'_{jq} \\ a'_{kr} & a'_{kp} & a'_{kq} \\ a'_{lr} & a'_{lp} & a'_{lq} \end{pmatrix},$$

where $r = p$ or $r = q$. The determinant of this matrix is zero and expanding it in the first column one sees that the components of the row va^{-1} in positions p and q are zeros.

It remains to verify that transvections $t_{i*}(v) = e + e_i v$, $v \in Y$, span the elementary subgroup $\mathrm{E}(n, R)$. It is sufficient to express any elementary transvection $t_{ij}(\xi)$ as a product of $t_{i*}(v)$'s for above rows v. With this end note that

$$t_{i*}(v) = t_{ij}(\Delta_{pq}^{hk}\xi\widetilde{\Delta}_{hk}^{pq}) t_{ih}(-\Delta_{pq}^{hk}\xi\widetilde{\Delta}_{jk}^{pq}) t_{ik}(\Delta_{pq}^{hk}\xi\widetilde{\Delta}_{jh}^{pq}).$$

and that by Binet-Cauchy theorem

$$\sum \widetilde{\Delta}_{rs}^{pq}\Delta_{pq}^{rs} = 1, \qquad 1 \leq p < q \leq n,$$

whereas

$$\sum \widetilde{\Delta}_{lm}^{pq}\Delta_{pq}^{rs} = 0, \qquad 1 \leq p < q \leq n,$$

if $\{l, m\} \neq \{r, s\}$. Now commutativity of R implies that

$$t_{ij}(\xi) = \prod t_{ij}(\Delta_{pq}^{hk}\xi\widetilde{\Delta}_{hk}^{pq}) t_{ih}(-\Delta_{pq}^{hk}\xi\widetilde{\Delta}_{jk}^{pq}) t_{ik}(\Delta_{pq}^{hk}\xi\widetilde{\Delta}_{jh}^{pq}), \quad 1 \leq p < q \leq n.$$

This finishes the proof.

By considering minors of order m instead of those of order 2 is immediate to generalize this result as follows: let R be commutative and $m \leq n-2$, then for any matrix $a \in \mathrm{GL}(n, R)$ the elementary subgroup $\mathrm{E}(n, R)$ is generated by the transvections $e + e_i v$, where $1 \leq i \leq n$ and the row v is such that v_i=0 and va^{-1} has at least m zero components. We stated the above lemma only for minors of order two to avoid awkward notation and since it is the most important case. But in the symplectic case which we outline in § 14 we use a similar result for higher order minors, so now we briefly sketch the construction to fix the notation.

Let $I, J \subseteq N = \{1, \dots, n\}$. If $|I| = |J| = m$ we denote by $\Delta_I^J = \Delta_I^J(a)$ the minor of order m of the matrix a which stands at the intersection of rows with indices from I and columns with indices from J. One sets also $\widetilde{\Delta}_I^J = \Delta_I^J(a^{-1})$. In other words,

$$\Delta_I^J = \det(a_{ij})_{i\in I, j\in J}, \qquad \text{and} \qquad \widetilde{\Delta}_I^J = \det(a'_{ij})_{i\in I, j\in J}.$$

Now we fix an invertible matrix $a \in \mathrm{GL}(n, R)$ and an m-element subset $J \subseteq N$, $|J| = m$, and consider the columns (a'_{*j}), $j \in J$, of the inverse matrix a^{-1} with indices from J. It is a routine matter to generalize Variation 1 to this situation and we explain only how to stabilize the columns. For this choose an $(m+1)$-element[1] subset $L \subseteq N$. For an index h denote by $\rho(L, h)$ the number of indices $k \in L$ strictly smaller, than h. The following result is an immediate corollary of the Laplace theorem.

Lemma 9. *The row $w = (w_h)$ defined by*

$$w_h = \begin{cases} (-1)^{\rho(L,h)} \widetilde{\Delta}^J_{L\setminus\{h\}}, & h \in L, \\ 0, & \text{otherwise}, \end{cases}$$

is orthogonal to all columns (a'_{*j}), $j \in J$.

Proof. Since zeros do not influence a product, we may without loss of generality restrict ourselves to the case $L = N$. With this convention the product of w by (a'_{*j}) equals $\sum(-1)^{h-1}\widetilde{\Delta}^J_{L\setminus\{h\}} a'_{hj}$. This is the expansion in the last column of the determinant obtained from the matrix $b = (a'_{hj})_{h\in L, j\in J}$, by appending the j-th column of b. Having two equal columns this determinant equals 0.

At this point it is completely clear, how to decompose an elementary transvection $t_{ij}(\xi)$ into factors preserving at least m columns of a^{-1}, where $m \leq n-2$. We fix an $(m+1)$-element subset $L \subseteq N$ containing j but not i and set

$$t_{ij}(\xi) = \prod \Big(\prod t_{ih}(\Delta_J^{L\setminus\{j\}} \xi (-1)^{\rho(L,j)+\rho(L,h)} \widetilde{\Delta}^J_{L\setminus\{h\}}) \Big),$$

where the internal product is taken over all $h \in L$, whereas the external product is taken over all m-element subsets $J \subseteq N$. Then by the last Lemma the J-th factor stabilizes all columns (a'_{*j}), $j \in J$, whereas the Binet-Cauchy theorem applied to

[1]Later we have to choose such a subset which, moreover, does not contain a fixed index i, this is exactly why the decomposition itself works when $n \geq m+2$. But for Lemma 9 it is enough to assume that $n \geq m+1$

the columns of a with indices from $L\setminus\{j\}$ and the rows of a^{-1} with indices from $L\setminus\{h\}$ guarantees that the appropriate sum equals 0, unless $h=j$, in which case it equals 1.

Remark. One can rewrite the decomposition of $t_{ij}(\xi)$ above using the fact that $\widetilde{\Delta}_J^K=\frac{(-1)^{\Sigma(J\cup K)}}{\det a}\cdot\Delta_{N\setminus K}^{N\setminus J}$ where $N=\{1,\dots,n\}$ and $\Sigma(L)=\sum\limits_{l\in L} l$. For $m=n-2$ this observation delivers the following decomposition

$$t_{ij}(\xi)=\prod_{1\le r<s\le n}\Big(\prod_{h\neq i} t_{ih}\big((a'_{ir}a'_{js}-a'_{is}a'_{jr})\xi(a_{ri}a_{sh}-a_{si}a_{rh})\big)\Big).$$

Conjugating the last formula by the matrix a^{-1} we get

$$e+a_{*i}\xi a'_{j*}=\prod_{1\le r<s\le n}\big(e+a_{*i}(a'_{ir}a'_{js}-a'_{is}a'_{jr})\xi(a_{ri}f_s-a_{si}f_r)\big),$$

which is exactly the same as the decomposition given by Suslin's lemma.

In [VPe] the second-named author and E.Ya.Perelman take over Variation 1 and get a similar but more general result with a much better proof in terms of the polyvector representation of GL_n. An analogue of Variation 1 is proven there for any column of Grassman coordinates, subject to Plücker relations. not just for minors of an invertible matrix.

§ 7. Centrality of K_2

Now we pass to the first application of results of § 6, another one comes in § 14. Recall that the Steinberg group $\mathrm{St}(n,R)$ is a group defined by generators and relations as follows. The generators are $x_{ij}(\xi)$, $1\le i\neq j\le n$, $\xi\in R$, whereas the relations are analogues of (2) and (3) with $t_{ij}(\xi)$ replaced by $x_{ij}(\xi)$. Since the elementary transvections $t_{ij}(\xi)$ satisfy these relations, there is the natural epimorphism $\pi:\mathrm{St}(n,R)\to\mathrm{E}(n,R)$ sending $x_{ij}(\xi)$ to $t_{ij}(\xi)$. The kernel of the epimorphism π is denoted by $\mathrm{K}_2(n,R)$ and is called the (unstable) K_2-functor.

The work of Milnor [Mi1] posed the problem whether $\mathrm{K}_2(n,R)$ is contained in the centre of the Steinberg group $\mathrm{St}(n,R)$. The most important situations, when one knows that centrality of $\mathrm{K}_2(n,R)$ holds, are as follows:

(1) At the stable level, say if $n\ge\operatorname{sr}R+2$ (it follows from Dennis – Vaserstein theorem about the surjective stability of K_2);
(2) When R is a ring with weak Euclidean algorithm, for example a free associative algebra over a field (Silvester – Cohn);
(3) When $n\ge 4$ and R is almost commutative (van der Kallen [vdK] and Tulenbaev [Tu]).

Now we show that in the proof of van der Kallen's theorem one can use Variation 1 instead of Suslin's lemma.

Theorem 3. *Let R be a commutative ring and $n \geq 4$. Then $K_2(n, R)$ is central in* $\mathrm{St}(n, R)$.

Sketch of proof. Of course the proof closely follows that in van der Kallen [vdK] and uses many lemmas therefrom, so we give only an outline. The idea of [vdK] is to construct another presentation of the Steinberg group, symmetric with respect to an arbitrary change of base. We consider a modified Steinberg group $\widetilde{\mathrm{St}}(n, R)$ defined by the following presentation. This group is generated by symbols $x_{uv}(\xi)$ where u is a column of an invertible matrix $a \in \mathrm{GL}(n, R)$, v is a row of length n such that $vu = 0$, $\xi \in R$, subject to the following relations:

- **(R1)** $x_{uv}(\xi) = x_{u,\xi v}(1)$,
- **(R2)** $x_{uv}(\xi)x_{uw}(\xi) = x_{u,v+w}(\xi)$,
- **(R3)** $x_{uv}(1)x_{yz}(\xi)x_{uv}(1)^{-1} = x_{y+u(vy),z-(zu)v}(\xi)$.

In particular, these relations imply that $x_{u0}(\xi) = x_{uv}(0) = e$ and that $x_{uv}(1)^{-1} = x_{uv}(-1) = x_{u,-v}(1)$. We usually write simply x_{uv} instead of $x_{uv}(1)$.

Now centrality of $\mathrm{K}_2(n, R)$ in $\mathrm{St}(n, R)$ would automatically follow from the existence of an isomorphism $\mathrm{St}(n, R) \longrightarrow \widetilde{\mathrm{St}}(n, R)$ mapping $x_{ij}(\xi)$ to $x_{e_i,f_j}(\xi)$. This amounts to defining symbols $x_{uv}(\xi)$ satisfying the above relations R1 – R3 in terms of the elementary generators $x_{ij}(\xi)$. As compared with [vdK] in our case such a construction is somewhat simplified by the fact that we work only with columns of invertible matrices instead of arbitrary unimodular columns. Indeed, since any row orthogonal to the i-th column a_{*i} of a may be expressed as a linear combination of the rows a'_{j*}, $i \neq j$, it follows from R1 and R2 that $\widetilde{\mathrm{St}}(n, R)$ is generated by the elements $x_{a_{*i},a'_{j*}}(\xi)$, where $a \in \mathrm{GL}(n, R)$, $1 \leq i \neq j \leq n$, $\xi \in R$. This immediately implies that the group $\widetilde{\mathrm{St}}(n, R)$ is perfect whenever $n \geq 3$. Indeed, taking $h \neq i, j$, we get from R3 and R2 that $x_{a_{*i},a'_{j*}}(\xi) = [x_{a_{*i},a'_{h*}}(\xi), x_{a_{*h},a'_{j*}}(1)]$.

Now in case where $u_i = 0$ or $v_i = 0$ for some i the symbol $x_{uv}(\xi)$ may be defined exactly as in [vdK], on the base of Formula (11). Say, if u and v are *arbitrary* column and row such that $vu = 0$ and $v_j = 0$, we set

$$x_{uv}(\xi) = \prod x_{jh}(u_j\xi v_h)\bigl[\prod x_{hj}(u_h), \prod x_{jh}(\xi v_h)\bigr], \tag{13}$$

where all products are taken in $h \neq j$. In the case $u_i = 0$ the symbol $x_{uv}(\xi)$ can be defined similarly, and 3.9, 3.10 of [vdK] show that these definitions are indeed consistent. Now at the first sight it appears that since we are interested only in constructing $x_{a_{*i},a'_{j*}}(\xi)$, we could proceed directly from the Theme and define it by the following formula

$$x_{a_{*i},a'_{j*}}(\xi) = \prod x_{a_{*i},a'_{hk}a'_{j*}+a'_{jk}a'_{h*}}(a_{kh}\xi),$$

where the product on the right hand side is taken over all $k \in N = \{1, \dots, n\}$ and all factors are already defined by 13. However it is not immediately obvious why this definition is consistent[2] and we choose a much cheaper way and proceed from Variation 1. Namely, for $n \geq 4$ we take *four* pair-wise distinct indices i, j, h, k and we define $x_{a_{*i},a'_{j*}}(\xi)$ as follows:

$$x_{a_{*i},a'_{j*}}(\xi) = \prod x_{a_{*i},\tilde{\Delta}^{pq}_{hk}a'_{j*}-\tilde{\Delta}^{pq}_{jk}a'_{h*}+\tilde{\Delta}^{pq}_{jh}a'_{k*}}(\Delta^{hk}_{pq}\xi),$$

where the product on the right hand side is taken over all $1 \leq p < q \leq n$ and all factors are defined by (13). Since now the rows in these factors have at least *two* zeros, the remaining part of van der Kallen's proof (3.18 of [vdK], etc.) can be repeated word by word with our definition of $x_{uv}(\xi)$, instead of his.

§ 8. Second variation: linear dependence

The moral of § 5 is that to prove normality of $\mathrm{E}(n, R)$ it is enough to find for any $a \in \mathrm{GL}(n, R)$ *some* system of generators $\Omega(a)$ of $\mathrm{E}(n, R)$ such that for any $x \in \Omega(a)$ one has $axa^{-1} \in \mathrm{E}(n, R)$. In the Theme and Variation 1 we considered as such generators certain transvections of the form $e+uv = e+e_iv$. But obviously one can generalize this in the two following directions. First, nobody forces us to set $u = e_i$, we can consider any set of transvections $e + uv$, such that $a(e + uv)a^{-1} \in \mathrm{E}(n, R)$. Second, transvections of the form $e + e_iv$ were automatically inside the elementary subgroup. In general, we do not have to require from the outset that our new transvections $e+uv$ belong to $\mathrm{E}(n, R)$. It is sufficient that they generate a subgroup, which contains $\mathrm{E}(n, R)$.

There is usually an enormous freedom in construction of such generating sets of transvections. Below we give an obvious example. Let G be a group, H be its subgroup. Usually one defines normality of H by the condition $aHa^{-1} \leq H$ for all $a \in G$. But obviously, one could require instead that $H \leq a^{-1}Ha$. This leads to the following amusing reformulation of normality. Unlike the rest of the paper here we assume only that $n \geq 2$.

Variation 2. *Let R be any associative ring with 1 and $n \geq 2$. The elementary subgroup $\mathrm{E}(n, R)$ is normal in $\mathrm{GL}(n, R)$ if and only if for all matrices $a \in \mathrm{GL}(n, R)$ the group generated by transvections $e + a'_{*i}\xi a_{j*}$, $\xi \in R$, $1 \leq i \neq j \leq n$, contains $\mathrm{E}(n, R)$.*

Proof. Clearly, $\mathrm{E}(n, R)$ is normal in $\mathrm{GL}(n, R)$ if and only if for any matrix $a \in \mathrm{GL}(n, R)$ and any elementary transvection $t_{ij}(\xi)$, $\xi \in R$, $1 \leq i \neq j \leq n$, the conjugate $x = a^{-1}t_{ij}(\xi)a$ belongs to $\mathrm{E}(n, R)$. But by (8) the matrix x is itself a transvection $x = e + a'_{*i}\xi a_{j*}$ and thus $a^{-1}\,\mathrm{E}(n, R)a \leq \mathrm{E}(n, R)$ is equivalent to claiming that all these transvections belong to $\mathrm{E}(n, R)$. Interchanging a and

[2] On the other hand there are no published examples of a commutative ring R such that that $K_2(3, R)$ is not central in $\mathrm{St}(3, R)$, so that a priori it is not excluded that this definition may be proven to be consistent for all $n \geq 3$ with some extra effort.

a^{-1} we see that actually in this case these transvections *generate* $\mathrm{E}(n,R)$. For a $a \in \mathrm{GL}(n,R)$ let $H(a)$ denote the subgroup generated by the transvections $x = e + a'_{*i}\xi a_{j*}$. Then normality of $\mathrm{E}(n,R)$ implies that $H(a) \geq \mathrm{E}(n,R)$. Conversely, one has $aH(a)a^{-1} \leq \mathrm{E}(n,R)$ and thus $H(a) \geq \mathrm{E}(n,R)$ implies that $a\,\mathrm{E}(n,R)a^{-1} \leq \mathrm{E}(n,R)$.

Using Variation 2 we easily obtain a necessary and sufficient condition for the normality of the elementary subgroup.

Proposition 1. *Let R be any associative ring with 1 and $n \geq 2$. For the elementary subgroup $\mathrm{E}(n,R)$ to be normal in $\mathrm{GL}(n,R)$ it is necessary and sufficient that the inclusion*

$$\big\langle e + uv \mid u \in R^n,\ v \in {}^nR,\ vu = 0,\ va^{-1} \in N({}^nR)\big\rangle \geq \mathrm{E}(n,R)$$

holds for all $a \in \mathrm{GL}(n,R)$.

Proof. Clearly this condition is sufficient. In fact, by Lemma 3 the transvection $a(e+uv)a^{-1} = e + (au)(va^{-1})$ belongs to $\mathrm{E}(n,R)$ since the row va^{-1} contains a zero component. To prove necessity observe that if $\mathrm{E}(n,R)$ is normal in $\mathrm{GL}(n,R)$, then by Variation 2 already the transvections $e + a'_{*i}\xi a_{j*}$, $\xi \in R$, $1 \leq i \neq j \leq n$, generate $\mathrm{E}(n,R)$. But these transvections are obviously of required type, since setting $u = a'_{*i}\xi$ and $v = a_{i*}$ we get a row $va^{-1} = f_i$ with $n-1$ zeros.

Of course, the condition in this proposition is not an intrinsic ring theoretical characterization of the class of rings, where $\mathrm{E}(n,R)$ happens to be normal, but simply a fancier reformulation of normality. But this new condition is easier to check in many concrete situations. For example the Theme asserts exactly that this condition is obviously satisfied when R is commutative and $n \geq 3$.

The above proposition could be easily formulated also in the relative case, but usually there is no need for this. In fact, there is a standard procedure to reduce questions about relative elementary subgroups to the absolute case, called relativization. The idea, due to Stein [St] (and used later in the works of Milnor, Bak, Suslin, Vaserstein, among others), is to use the double $R \times_I R$ of the ring R along the ideal I (also called the smash product of the ring R and the ideal I). By definition $R \times_I R$ is the subring of the direct product $R \times R$, consisting of all pairs (λ, μ) such that $\lambda - \mu \in I$. Then the groups $\mathrm{GL}(n, R \times_I R)$ and $\mathrm{E}(n, R \times_I R)$ can be identified with semi-direct products $\mathrm{GL}(n,R) \leftthreetimes \mathrm{GL}(n,R,I)$ and $\mathrm{E}(n,R) \leftthreetimes \mathrm{E}(n,R,I)$ respectively. The following result is a version of relativization (see Proposition 3.1 of [Sv5] for a proof).

Lemma 10. *For the elementary subgroup $\mathrm{E}(n, R \times_I R)$ to be normal in $\mathrm{GL}(n, R \times_I R)$ it is necessary and sufficient that $\mathrm{E}(n,R)$ is normal in $\mathrm{GL}(n,R)$ and the standard commutator formula $[\mathrm{GL}(n,R,I), \mathrm{E}(n,R)] = \mathrm{E}(n,R,I)$ holds. In particular, normality of $\mathrm{E}(n, R \times_I R)$ in $\mathrm{GL}(n, R \times_I R)$ implies normality of $\mathrm{E}(n,R,I)$ in $\mathrm{GL}(n,R)$.*

For non-commutative rings it is often difficult to check the condition of Proposition 1 directly. Luckily an apparently much weaker condition suffices to prove normality of the elementary subgroup. Namely, above we applied Lemma 3 to the transvection $d = a(e + uv)a^{-1}$ itself. But as we have already seen in the corollary of Lemma 3, it already suffices that a conjugate of d by an elementary matrix falls into conditions of Lemma 3. Indeed, let $b \in \mathrm{E}(n, R)$ be such that Lemma 3 can be applied to bdb^{-1}. Then $bdb^{-1} \in \mathrm{E}(n, R)$ and thus also $d = b^{-1}(bdb^{-1})b \in \mathrm{E}(n, R)$.

§ 9. Third variation: description of normal subgroups

A result on the normal structure of $\mathrm{GL}(n, R)$ which uses the idea of Variation 2 and the last observation was proven in [Sv5]. Here we explain the main idea of that result and apply it to get a new proof of the theorem of L.N.Vaserstein on the normal subgroups of $\mathrm{GL}(n, R)$ from [V1].

Note that in the proof of Theorem 2 we used the condition that the columns from $N(R^n)$ annihilating a row of a given matrix generate R^n. Here we show that this condition is much more restrictive – terribly much more restrictive – than what one really needs to describe normal subgroups. In fact it suffices to be able to find for each matrix just one, almost arbitrary, column from $N(R^n)$ orthogonal to a row of this matrix.

To formulate the exact condition on the above column we need some notation. For a matrix $c \in \mathrm{GL}(n, R)$ we denote by $I(c)$ its *level*, i.e. the smallest ideal in R such that $c \in \mathrm{C}(n, R, I)$. In other words, $I(c)$ is the ideal generated by the elements c_{ij} and $c_{ii}\xi - \xi c_{jj}$ for all $i \neq j$ and $\xi \in R$. By $I^{\perp}(c)$ we denote the intersection of the right and left annihilators of $I(c)$. Roughly speaking, $I^{\perp}(c)$ is orthogonal to the part of the ring R which we are interested in. Note that $I(a) = I(a^{-1})$ and that for any matrix $b \in \langle a \rangle^{\mathrm{E}(n,R)}$ we have $I(b) \leq I(a)$, and hence $I^{\perp}(b) \geq I^{\perp}(a)$. Observe also that $I^{\perp}(a)$ is a proper ideal if and only if a is non-central.

Variation 3. *Let H be a subgroup of $\mathrm{GL}(n, R)$ normalized by $\mathrm{E}(n, R)$, $a \in H$, and $u \in N(R^n)$. If $au \in N(R^n)$ and $u \notin \big(I^{\perp}(a)\big)^n$ then H contains a non-trivial elementary transvection.*

Proof. Let $u_i = 0$ and $(au)_j = 0$. Take an arbitrary element $\lambda \in R$ and consider the matrix

$$c = at_{*i}(u)a^{-1}t_{*i}(-u) = (e + aua'_{i*})t_{*i}(-u) \in H$$

Calculation shows that $c_{jj} = 1$ and $c_{jk} = 0$ for any $k \neq i, j$. Hence, by Lemma 8 either H contains a non-trivial elementary transvection and we are done or $c \in \mathrm{C}(n, R)$. In the latter case we obviously have $c = e$. Thus, $aua'_{i*} = uf_i$. In particular, $aua'_{il} = 0$, and hence $ua'_{il} = 0$, for any $l \neq i$. Since $u \notin \big(I^{\perp}(a)\big)^n$, by Lemma 4.3 of [Sv5], H contains a non-trivial elementary transvection.

Variation 3 can be considered as a broad generalization of Lemma 8. The next result generalizes Lemma 7 in the same style.

Lemma 11. *Let H be a subgroup of $\mathrm{GL}(n,R)$ normalized by $\mathrm{E}(n,R)$, $a \in H$, $u \in N(R^n)$, and $v \in {}^nR$. Suppose that $vu = 0$ and $uv \notin \mathrm{M}\big(n, I^{\perp}(a)\big)$. If a commutes with the transvection $e + uv$ then H contains a non-trivial elementary transvection.*

Proof. It follows that $auv = uva$, and, since $u_j = 0$ for some index j, we have $a_{j*}uv = 0$. Thus, for an index k we have $auv_k \in N(R^n)$ and $uv_k \notin \big(I^{\perp}(a)\big)^n$, and it remains to apply Variation 3 to get the result.

For a technical reason which will soon become clear we need to consider a bit bigger ideal than $I^{\perp}(a)$. Namely, denote by $J(a)$ the set of all $\xi \in R$ such that $\xi I(a)^m = I(a)^m \xi = 0$ for some positive integer m. Obviously, $I^{\perp}(a) \leq J(a)$ and if $J(a) = R$ then $a \in \mathrm{C}(n, R, \operatorname{Rad} R)$ where $\operatorname{Rad} R$ is the Jacobson radical of R. The main advantage of considering $J(a)$ instead of $I^{\perp}(a)$ is that $J(a)$ has the main property of a radical, whereas $I^{\perp}(a)$ does not. More precisely, let $\overline{a}$ be the image of a in $\mathrm{GL}\big(n, R/J(a)\big)$ under the reduction homomorphism. Then $J(\overline{a})$ (and, a fortiori, $I^{\perp}(\overline{a})$) is zero.

The next lemma shows that an elementary transvection can be lifted modulo $J(a)$.

Lemma 12. *Let H be the smallest subgroup of $\mathrm{GL}(n,R)$ containing an invertible matrix a and normalized by $\mathrm{E}(n,R)$. Suppose that there is a non-central matrix $b \in H$ such that its image in $\mathrm{GL}\big(n, R/J(a)\big)$ under the reduction homomorphism has a zero entry in a non-diagonal place. Then H contains a non-trivial elementary transvection.*

Proof. As we have already noticed, $I(b) \leq I(a)$, and hence $J(a) \leq J(b)$. Thus, for some indices $i \neq j$ we have $b_{ij} \in J(b)$. In particular, $(b_{ij}R)^m = 0$ for some m which implies that b_{ij} belongs to the Jacobson radical of R. But then by Lemma 4.1 of [Sv5] H contains a non-trivial elementary transvection.

Now we are ready to prove Vaserstein's theorem [V1].

Theorem 4. *Let A be a ring with the center R. Suppose that for any maximal ideal M of R there is a multiplicative subset $S \subseteq R \setminus M$ such that $\operatorname{sr}(S^{-1}A) \leq n-1$. Then for any $H \leq \mathrm{GL}(n,A)$ normalized by $\mathrm{E}(n,R)$ there exists a unique ideal I such that*

$$\mathrm{E}(n,R,I) \leq H \leq \mathrm{C}(n,R,I)$$

Proof. By Statement (6) of [V1] and Lemma 5 we have to prove that for each ideal I any non-central subgroup H of $\mathrm{GL}(n, A/I)$ normalized by $\mathrm{E}(n, A/I)$ contains a non-trivial elementary transvection. Let a be a non-central matrix from H. Clearly we may assume that H is the smallest subgroup containing a and normalized by $\mathrm{E}(n, A/I)$. Denote by B the quotient ring $(A/I)/J(a)$ and by b the image of a in $\mathrm{GL}(n,B)$ under the reduction homomorphism. If b is central then we are done by Lemma 12. If b is non-central then by the same argument it remains to prove that the image $\overline{H}$ of H in $\mathrm{GL}(n,B)$ contains a non-trivial elementary transvection.

Choose a maximal ideal M in R containing $R \cap \operatorname{Ker}(A \to B)$. Then there is a multiplicative subset $S \subseteq R \setminus M$ such that $\operatorname{sr}(S^{-1}A) \le n-1$. If T is the image of S in B then $T^{-1}B$ is a quotient of $S^{-1}A$, and therefore $\operatorname{sr}(T^{-1}B) \le n-1$. By definition of stable rank there are $\lambda_1, \ldots, \lambda_n \in B$ and $\alpha \in T$ such that the row $(b_{11} + b_{1n}\frac{\lambda_1}{\alpha}, \ldots, b_{1,n-1} + b_{1n}\frac{\lambda_{n-1}}{\alpha})$ is unimodular in $T^{-1}B$. Hence, there are $\mu_1, \ldots, \mu_{n-1} \in T^{-1}B$ not all zeroes and such that $\sum\limits_{k=1}^{n-1} \big(b_{1k} + b_{1n}\frac{\lambda_k}{\alpha}\big)\mu_k = 0$ (if a unimodular row of length $m > 1$ is linearly independent then it defines an isomorphism $(T^{-1}B)^m \cong (T^{-1}B)^1$ which contradicts the finiteness of $\operatorname{sr}(T^{-1}B)$). Multiplying (if necessary) the elements μ_k by an element from T we get in the ring B the formula

$$\sum_{k=1}^{n-1} b_{1k}\alpha\eta_k + b_{1n}\sum_{k=1}^{n-1} \lambda_k\eta_k = 0,$$

where $\eta_k \in B$ are such that their images in $T^{-1}B$ are not all zeroes.

Let $u = b(\alpha\eta_1, \ldots, \alpha\eta_{n-1}, \sum\limits_{k=1}^{n-1} \lambda_k\eta_k)^t$ and $v = (\lambda_1, \ldots, \lambda_{n-1}, -\alpha)b^{-1}$. Clearly, $u_1 = 0$ and $vu = 0$. Note that $uv \neq 0$ because the first $n-1$ entries of its last column are $-\alpha^2\eta_1, \ldots, -\alpha^2\eta_{n-1}$ and they do not vanish in $T^{-1}B$. Therefore, if b commutes with the transvection $e + uv$ then we are done by Lemma 11. Otherwise, consider the non-central matrix

$$c = (e + uv)b^{-1}(e - uv)b = (e + uv)(e - b^{-1}uvb) \in \overline{H}.$$

Since the first row of $e + uv$ is f_1, the first row of c coincides with the first row of $e - b^{-1}uvb$ which equals

$$c_{1*} = (1 - \alpha\eta_1\lambda_1, -\alpha\eta_1\lambda_2, \ldots, -\alpha\eta_1\lambda_{n-1}, \alpha^2\eta_1).$$

But then $c_{1*}(0, \ldots, 0, \alpha, \lambda_{n-1})^t = 0$ and we can apply Variation 3.

§ 10. Net subgroups

A square array $\sigma = (\sigma_{ij})$, $1 \le i, j \le n$, consisting of n^2 additive subgroups $\sigma_{ij} \subseteq R$ is called a *net over R of degree n*, if $\sigma_{ih}\sigma_{hj} \subseteq \sigma_{ij}$ for all i, j, h. A net σ is called *unital* if $1 \in \sigma_{ii}$ for all i. A net σ is called a D-*net* if $\sigma_{ii} = R$ for all i. In this case the definition of a net implies that all σ_{ij} are two-sided ideals in R.

For a given net σ we denote by $\mathrm{M}(\sigma)$ the collection of all matrices $a = (a_{ij}) \in \mathrm{M}(n, R)$ such that $a_{ij} \in \sigma_{ij}$ for all i, j. Clearly, $\mathrm{M}(\sigma)$ is a subring of $\mathrm{M}(n, R)$. The identity matrix is contained in $\mathrm{M}(\sigma)$ if and only if σ is a unital net.

The collection $e + \mathrm{M}(\sigma)$, consisting of the matrices $e + a$, where $a \in \mathrm{M}(\sigma)$, is a multiplicative system. The following concept was introduced by Z.I.Borewicz in [Bo1], [Bo2] (see [BoV1] – [BoV3] or [Va6] for further references). The largest subgroup of the general linear group $G = \mathrm{GL}(n, R)$ contained in the multiplicative system $e+\mathrm{M}(\sigma)$ is called the net subgroup of $\mathrm{GL}(n,R)$, corresponding to the net σ and is denoted by $G(\sigma)$.

We introduce a partial order on the set of all nets σ over R of fixed degree n. Namely, for given nets σ, τ we write $\sigma\le\tau$ if $\sigma_{ij}\subseteq\tau_{ij}$ for all i,j. The conditions $\sigma\le\tau$ and $G(\sigma)\le G(\tau)$ are equivalent.

One can associate with a net σ some further subgroups of $G=\mathrm{GL}(n,R)$. The most important of them is the *elementary net subgroup* $E(\sigma)$ defined as the group generated by all elementary transvections of level σ:

$$E(\sigma)=\big\langle t_{ij}(\xi)\,|\,\xi\in\sigma_{ij},1\le i\ne j\le n\big\rangle.$$

Now, let ν be an equivalence relation on the index set $I=\{1,\dots,n\}$. We write simply $i\sim j$ to denote that i is equivalent to j with respect to ν. Let $I_1,\dots,I_t$ be the equivalence classes of ν on I. We will denote by $h(\nu)$ the smallest of the orders $n_1=|I_1|,\dots,n_t=|I_t|$ of these classes.

We associate to the equivalence relation ν on I a D-net $[\nu]$ defined by $[\nu]_{ij}=R$ whenever $i\sim j$ and $[\nu_{ij}]=0$ otherwise. The corresponding net subgroup $\mathrm{G}([\nu])$ is also denoted by $D(\nu)$. If each of the equivalence classes $I_1,\dots,I_t$ consists of consecutive integers, then ν is completely specified by the partition $(n_1,\dots,n_t)$, $n=|\nu|=n_1+\ldots+n_t$. In this case $D(\nu)=D(n_1,\dots,n_t)$ is the usual group of block diagonal matrices with diagonal blocks of sizes $n_1,\dots,n_t$ respectively.

For an arbitrary equivalence relation ν the group $D(\nu)$ is conjugate by a permutation matrix to the group $D(n_1,\dots,n_t)$ of ordinary block diagonal matrices. On the basis of this fact we will call $D(\nu)$ the group of "block diagonal matrices" of type ν even in the general case.

The elementary net subgroup $E([\nu])$, corresponding to the net $[\nu]$ will be denoted by $E(\nu)$ and will be called the elementary block diagonal group of type ν. In other words, $E(\nu)$ is generated by all elementary transvections $t_{ij}(\xi)$, where $\xi\in R$, $1\le i\ne j\le n$, $i\sim j$.

In a sequel we will consider only equivalence relations ν such that $h(\nu)\ge 3$ (this corresponds to the restriction $n\ge 3$ of the preceding sections). Let ν be an equivalence relation on I. If $\sigma_{ij}=\sigma_{pq}$ provided $i\sim p$ and $j\sim q$ then σ is called a *ν-net*. Clearly, if $\sigma\ge[\nu]$ then σ is a ν-net.

Another very important subgroup corresponding to a net σ is the normalizer $N_G(\mathrm{G}(\sigma))$ of the net subgroup $\mathrm{G}(\sigma)$ in $G=\mathrm{GL}(n,R)$, which will be denoted by $\mathrm{N}(\sigma)$. Let $a\in G$ be a matrix in G. It is clear that if for all quadruples of indices i,j,r,s we have

$$a_{ir}\sigma_{rs}a'_{sj}\subseteq\sigma_{ij} \tag{14}$$

then the matrix a normalizes the subgroup $G(\sigma)$, i.e. $a\in N(\sigma)$. It is natural to ask, to what extent the above conditions are necessary. For the class of nets we are interested in this question is answered by the following result (see [BoV3], § 3, and [Sv1]).

Lemma 13. *Let R be an arbitrary ring, ν be an equivalence relation on $I = \{1, \dots, n\}$ such that $h(\nu) \geq 3$, and σ be a unital ν-net of degree n. A matrix $a \in \mathrm{GL}(n, R)$ belongs to the net normalizer $\mathrm{N}(\sigma)$ if and only if inclusions (14) are valid for all indices i, j, r, s. Here the inclusions for $r = s$ follow from those for $r \neq s$.*

The last sentence of the lemma may be stated as follows.

Corollary. *Assume as above, then $N(\sigma) = \{a \in \mathrm{GL}(n, R) \mid aE(\sigma)a^{-1} \leq G(\sigma)\}$.*

In the next section we consider subgroups of $\mathrm{GL}(n, R)$ containing an elementary block diagonal group $E(\nu)$. The following lemma shows that with such a subgroup we can associate a net.

Lemma 14. *Let $H \geq E(\nu)$, where $h(\nu) \geq 3$. Then there exists a unique largest D-net $\sigma = (\sigma_{ij})$ over R such that $E(\sigma) \leq H$.*

Proof. Set $\sigma_{ij} = R$ if $i \sim j$ and $\sigma_{ij} = \{\xi \in R \mid t_{ij}(\xi) \in H\}$ otherwise. Then (3) immediately implies that all σ_{ij} are ideals and that $\sigma = (\sigma_{ij})$ is indeed a net (see [BoV3], § 4 for details. The maximality of σ is obvious.

The net constructed in this lemma will be called the net *associated* with the subgroup H.

§ 11. Fourth Variation: part of elementary matrices

In this section we consider a generalization of structure theorems to subgroups normalized by a part of elementary matrices. More precisely, let $H \leq \mathrm{E}(n, R)$ be a group of elementary matrices. We are interested in description of subgroups in $G = \mathrm{GL}(n, R)$ normalized by H. As noticed in [Sv4] in many cases this problem may be reduced to the following two special cases: 1) description of subgroups of G containing H, 2) description of normal subgroups of H.

Here we give all details only for subgroups containing a group of elementary block diagonal matrices $H = E(\nu)$ for a partition ν such that $h(\nu) \geq 3$. But in fact a similar approach works in many further situations, some of which are listed in [Va6]. Roughly speaking, H should have enough unipotents to reduce the problems to proper parabolic subgroups and to control what happens inside proper parabolic subgroups. There are many instances, when this is possible, and below we list some most immediate examples:

- elementary classical subgroups, like $H = \mathrm{EO}(n, R)$ inside $G = \mathrm{GL}(n, R)$, provided $n \geq 6$ (see the remark at the end of the next section);
- tensored elementary subgroups, like $H = \mathrm{E}(n_1, R) \otimes \ldots \otimes \mathrm{E}(n_t, R)$ inside $G = \mathrm{GL}(n_1 \cdot \ldots \cdot n_t, R)$, provided $n_1, \dots, n_t \geq 3$;
- elementary block triangular subgroup $H = \mathrm{EB}(\nu)$ inside $G = \mathrm{GL}(|\nu|, R)$, provided $h(\nu) \geq 2$.

With slight variations this can be generalized also to some of the elementary groups parameterized by a certain subset of the ring R. For the usual relative subgroups $\mathrm{E}(n, R, I)$, $I \trianglelefteq R$, this is done [Wi], [B4], [V3], [LL], [Va3], [V8]. There are some further papers which consider description of subgroups in $\mathrm{GL}(n, R)$ normalized by $H = \mathrm{E}(n, S)$, where S is a subring of R, or even by $H = \mathrm{E}(\nu, S)$, where $h(\nu)$ is large enough, but these results depend on deeper ring-theoretical conditions and we do not discuss them here (see references in [Sv3], [Va6]).

The following result is a partial generalization of Theorem 2 with the elementary subgroup $\mathrm{E}(n, R)$ replaced by the group of elementary block diagonal matrices.

Theorem 5. *Let R be a commutative ring, $G = \mathrm{GL}(n, R)$ be the general linear group of degree n over R, ν be an equivalence relation on the set $I = \{1, \dots, n\}$ such that $h(\nu) \geq 3$. Further, let H be a subgroup of G containing the group $\mathrm{E}(\nu)$ of elementary block diagonal matrices over R of type ν. Then there exists a unique net $\sigma \geq [\nu]$ such that*

$$\mathrm{E}(\nu) \leq H \leq \mathrm{N}(\nu).$$

More precisely, σ will be the D-net associated with the subgroup H, as constructed in Lemma 17, i.e. the largest net σ such that $E(\sigma) \leq H$. Clearly $E(\sigma) \geq E(\nu)$ so that $\sigma \geq \nu$. The following analogue of Lemma 1 is an elementary corollary of formulae (3) (see [BoV3], Lemmas 9 and 11).

Lemma 15. *Let R be any ring and ν be an equivalence relation on the index set $I = \{1, \dots, n\}$ such that $h(\nu) \geq 3$. If σ is a net containing ν, then* $\mathrm{E}(\sigma) = [\mathrm{E}(\sigma), \mathrm{E}(\sigma)] = [\mathrm{E}(\sigma), \mathrm{E}(\nu)]$.

To prove the theorem we need also the following purely technical result similar in spirit to Lemma 6, but somewhat more precise (see [BoV3], Lemma 3). We fix the following setting: R denotes any ring (later it will be assumed to be commutative), ν is an equivalence relation on the index set $I = \{1, \dots, n\}$ such that $h(\nu) \geq 3$. Further, we consider a subgroup H of $G = \mathrm{GL}(n, R)$ containing $E(\nu)$ and denote by σ the D-net associated with H. Then, $E(\nu) \leq E(\sigma) \leq H$, so that, in particular, $\sigma \geq [\nu]$. Finally, $a = (a_{ij}) \in H$ denotes a matrix belonging to H.

Lemma 16. *Assume that $a_{*q} = e_q$. Then $a_{ij} \in \sigma_{ij}$ for all i and all $j \sim q$.*

In the following lemma it is essential that the ring R is commutative.

Lemma 17. *Let $b = a t_{rs}(\xi) a^{-1}$ for some $r \neq s$ and some $\xi \in R$. Then for all pair-wise distinct indices i, j, h one has*

$$b_{ij} = b_{ij} b'_{hh} - b_{ih} b'_{hj} = -b'_{ij} = -\widetilde{\Delta}^{jh}_{ih}.$$

Proof. Straightforward calculation.

The following modification of the Theme is immediate.

Variation 4. *Let R be a commutative ring. Then for any $g \in \mathrm{GL}(n, R)$ the group $\mathrm{E}(\nu)$ of elementary block diagonal matrices of type ν, $h(\nu) \geq 3$, is generated by the transvections $e + e_h v \in \mathrm{E}(\nu)$, where $1 \leq h \leq n$ and the row v is such that $v_h = 0$ and vg^{-1} has at least one zero component.*

Proof. Take in the proof of the Theme three *equivalent* pair-wise distinct indices.

Proof of Theorem 5. We have to prove that $H \leq \mathrm{N}(\sigma)$. By corollary of Lemma 16 it is enough to verify that $a\,\mathrm{E}(\sigma)a^{-1} \leq \mathrm{G}(\sigma)$ for any $a \in H$. The idea is to apply Variation 4 to the case when g is a transvection $b = at_{rs}(\xi)a^{-1}$, $a \in H$, $\xi \in \sigma_{rs}$. We want to check that $b \in \mathrm{G}(\sigma)$.

In fact it is already enough that to consider the following special type of transvections of the form $e + e_h v$ contained in the elementary block diagonal group $E(\nu)$. Look at the transvections $t = t_{hp}(a_{qr})t_{hq}(-a_{pr})$, where h, p, q ranges over triples of pair-wise distinct equivalent indices and $\zeta \in R$. Now a direct calculation shows that the entry of the transvection $z = btb^{-1}$ in position (i, j) equals

$$z_{ij} = \delta_{ij} + b_{ih}(a'_{qr}b'_{pj} - a'_{pr}b'_{qj}).$$

Then a straightforward calculation shows that $z_{ij} = \delta_{ij}$ for all i and all $j \neq p, q$ and $z_{ip} = \delta_{ip} + b_{ih}a_{qr}$. Now we can apply to z Lemma 16 and conclude that $z_{ip} \in \sigma_{ip} = \sigma_{ih}$, or, in other words, $b_{ih}a_{qr} \in \sigma_{ih}$ for $q \sim h$, $q \neq h$. But these last inclusions applied to *all* conjugates of elementary transvections in H already imply that $b \in G(\sigma)$, as an argument from [Va6] shows.

Indeed, these inclusions hold for *all* transvections of the form $b = at_{rs}(\xi)a^{-1}$, $a \in H$. Now consider the transvection $c = bt_{jp}(1)b^{-1} \in H$. The inclusions above applied to c instead of b show that $c_{ip}b_{jj} \in \sigma_{ip} = \sigma_{ij}$. But $c_{ip} = b_{ip}b'_{pp}$, where $b'_{pp} = 1 - a_{pr}\xi a'_{sp}$, so that $c_{ip} \equiv b_{ij} (\operatorname{mod} \sigma_{ij})$. Thus $b_{ij}b_{jj} \in \sigma_{ij}$. Applying this last inclusion to c we get $c_{ip}c_{pp} \in \sigma_{ip} = \sigma_{ij}$. But $c_{ip}b_{pp} \equiv b_{ij} (\operatorname{mod} \sigma_{ij})$. Thus, finally, $b_{ij} \in \sigma_{ij}$ for all i, j. Since this holds for all $r \neq s$ and $\xi \in \sigma_{rs}$ one has $aE(\sigma)a^{-1} \leq G(\sigma)$ for all $a \in H$, and this is exactly what remained to be proven.

§ 12. Fifth Variation: normality of $\mathrm{E}(\sigma)$ in $\mathrm{N}(\sigma)$.

Sometimes the standard description of intermediate subgroups, similar to the one in Theorem 5, holds in a much more general setting. Let, for example, Q be a subring of a ring R. Then, under certain conditions the description of subgroups of $\mathrm{GL}(n, R)$ containing $\mathrm{E}(\nu, Q) = \langle t_{ij}(\xi) \mid i \neq j,\ \xi \in Q \rangle$ is standard (see [Sv3] for the precise statement)[3]. One of the most natural situations when such a description holds is delivered by a Dedekind domain Q in its field of fractions R. The nets arising in this description are unital ν-nets instead of D-nets in Theorem 5.

[3]This result has never been published in a form accessible to a Western reader. But the technique of its proof is quite different from the methods of the present paper, so we do not reproduce this result here.

The following theorem from [Sv3] seems to naturally complement the above mentioned result. In the proof of Theorem 6 alongside Suslin's lemma we use yet another decomposition of transvections, the simplest one, of which you would think.

Theorem 6. *Let ν be an equivalence on $I = \{1, \dots, n\}$ with $h(\nu) \geq 3$. Further,let σ be a unital ν-net of degree n over a commutative ring R. Then $\mathrm{E}(\sigma)$ is normal in $\mathrm{N}(\sigma)$.*

Proof. Let $\xi \in \sigma_{ij}$ and $a \in \mathrm{N}(\sigma)$. We have to prove that $b = at_{ij}(\xi)a^{-1} \in \mathrm{E}(\sigma)$. Decomposing b by Suslin's lemma as at the end of § 6 we get

$$b = e + a_{*i}\xi a'_{j*} = \prod_{1 \leq r \neq s \leq n} \big(e + a_{*i}a'_{ir}a'_{js}\xi(a_{ri}f_s - a_{si}f_r)\big).$$

Unfortunately, any attempt to represent a factor of the product above as $e + uv$ such that u_k's and v_k's belong to appropriate additive subgroups of the net σ fails. Therefore, we shall further decompose each factor of the product.

Variation 5. *If $a \in \mathrm{GL}(n, R)$ and $i \in \{1, \dots, n\}$ then a transvection $e+uv$ can be decomposed as follows:*

$$e + uv = \prod_{k=1}^{n} \big(e + (ua'_{ik})(a_{ki}v)\big).$$

Using this obvious observation and commutativity of R we obtain the formula

$$b = \prod_{1 \leq r \neq s \leq n} \prod_{k=1}^{n} \big(e + (a_{*i}a'_{ik})(a_{ki}a'_{ir}a_{ri}\xi a'_{js}f_s - a_{ki}\xi a'_{js}a_{si}a'_{ir}f_r)\big).$$

Fix a triple (r, s, k) and consider the corresponding factor of the above product. Choose an index $h \sim k$, $h \neq r, s$ (this is always possible because $n \geq 3$). Let $u = a_{*i}a'_{ik}$ and $v = a_{ki}a'_{ir}a_{ri}\xi a'_{js}f_s - a_{ki}\xi a'_{js}a_{si}a'_{ir}f_r$, so that the factor under consideration equals $e+uv$. By Lemma 13 one has $u_l \in \sigma_{lk} = \sigma_{lh}$ and $v_l \in \sigma_{kl} = \sigma_{hl}$. Since $v_h = 0$ the matrix $e + uv$ can be decomposed in accordance with (11), i.e. $e + uv = (e + u_h e_h v)[e + (u - u_h e_h)f_h, e + e_h v]$. Now one easily sees that all the elementary transvections present in the above formula are already in $\mathrm{E}(\sigma)$. This completes the proof.

Variation 5 was also used by the authors to prove the normality of $\mathrm{E}(\sigma)$ in $\mathrm{N}(\sigma)$ if the ring R satisfies some stability condition instead of commutativity.

The fact that Theorem 6 holds for an arbitrary unital ν-net over an arbitrary commutative ring is rather surprising. As a matter of comparison one can observe that the standard distribution of intermediate subgroups between $\mathrm{E}(\nu, Q)$ and $\mathrm{GL}(n, R)$ seems to fail even when Q is a field and R is a polynomial ring over Q, or when $Q = k[x, y]$ and R is its field of fractions.

§ 13. Sixth Variation: orthogonal transvections

In this section and the next one we briefly sketch how to modify the method of the present paper to make it work for other classical groups and Chevalley groups (see [VPS], [Va5], [Va7], [VP], [BV2] for more details). For the sake of simplicity we assume $2 \in R^*$, although this is not at all necessary. A systematic exposition for all classical groups without this assumption can be found in the joint works of the second-named author with A.Bak, see [BV2] and its sequel.

We number the indices from 1 to $n = 2l$ as follows: $1, \dots, l, -l, \dots, -1$ and introduce on the free R-module $V = R^n$, a quadratic form by

$$Q(x_1, \dots, x_{-1}) = x_1 x_{-1} + \dots + x_l x_{-l}.$$

This form has the maximal Witt index and makes V into an *Artin space*. We furnish V with an inner product given by the bilinear form $(\ ,\ )$ obtained by polarising Q. We identify a vector $u \in V$ with its column of coordinates and an endomorphism of the module V with its matrix in the Witt base $e_1, \dots, e_{-1}$. Recall that a vector u is called *isotropic* if $(u, u) = 0$. A submodule U of V is called *totally isotropic* if every vector $u \in U$ is isotropic, or what is the same, if $(u, v) = 0$ for all $u, v \in U$.

Now $\mathrm{SO}(2l, R)$ is the group of even isometries of the bilinear product. This group may be described as follows. Define the matrix $F = F_n$ as $F = \mathrm{sdiag}(1, \dots, 1)$, in other words one has $F_{ij} = \delta_{i,-j}$. Now $\mathrm{SO}(2l, R)$ consists of those matrices $a \in \mathrm{GL}(2l, R)$, for which $\det(a) = 1$ and $aF_n a^t = F_n$. In terms of the matrix entries the orthogonality condition for a matrix a is expressed by the following equations $a'_{ij} = a_{-j,-i}$. To be a column of an orthogonal matrix a vector $x = (x_1, \dots, x_{-1})$ has to satisfy the quadratic equation $x_1 x_{-1} + \dots + x_l x_{-l} = 0$.

The *elementary root unipotents* have the form

$$T_{ij}(\xi) = T_{-j,-i}(-\xi) = e + \xi e_{ij} - \xi e_{-j,-i},$$

for $i \neq \pm j$. They generate the subgroup $\mathrm{EO}(2l, R)$ of $\mathrm{SO}(2l, R)$ which is called the *elementary orthogonal group*.

Let now u and v be two orthogonal isotropic vectors, and $\xi \in R$ be a scalar. Then the linear transformation

$$T_{uv}(\xi)x = x - \xi(u, x)v + \xi(v, x)u,$$

where $x \in V$, belongs to $G = \mathrm{SO}(n, R)$. The transformations $T_{uv}(\xi)$ are sometimes called *orthogonal transvections*. Many authors use the terms *Eichler-Siegel-Dickson transvections* or *ESD-transvections*. The ESD-transvections are precisely the elements of (long) root type in the natural representation of orthogonal groups.

The elementary orthogonal transvections are obtained in the following way. Take $u = e_i$, $v = e_{-j}$. Then a straightforward calculation shows that $T_{uv}(\xi) = T_{ij}(\xi)$. Return to the matrix notation. For a column $u \in R^n$ we denote by $\widetilde{u} \in {}^nR = (R^n)^*$ the row $(Fu)^t$, where $F = F_n$. This means that the i-th component $\widetilde{u}_i$ of the row $\widetilde{u}$ is equal to the $(-i)$-component u_{-i} of the column u if $i \neq 0$. Comparing this notation with the definition of the inner product on $V = R^n$ we see that $(u,v) = \widetilde{v}u$. In particular a vector $u \in V$ is isotropic if and only if $\widetilde{u}u = 0$. Now with this notation the matrix of $T_{uv}(\xi)$ looks as follows. Let $\xi \in R$ and $u, v \in R^n$ are such that $\widetilde{u}u = \widetilde{v}u = 0$. Then in the Witt base $e_1, \dots, e_{-1}$ the ESD-transvection $T_{uv}(\xi)$ is given by the matrix

$$T_{uv}(\xi) = e + u\xi\widetilde{v} - v\xi\widetilde{u}.$$

Now we construct certain special elements of root type which play the crucial role in our calculations. Let $u, v, w \in V$ generate a totally isotropic submodule. Then for any $\eta, \zeta, \vartheta \in R$ the product

$$T_{uv}(\eta)T_{uw}(\zeta)T_{vw}(\vartheta) = T_{xy}(1),$$

where $x = u + v + w$ and $y = ((\eta + \zeta - \vartheta)u + (\eta - \zeta + \vartheta)v + (-\eta + \zeta + \vartheta)w)/2$, is a root type element. It is not generally speaking a root unipotent i.e. it is not conjugate to an element of the form $x_\alpha(\xi)$ – certainly not if the ideal generated by η, ζ, ϑ is not principal. Still it satisfies all the polynomial equations satisfied by the root elements and in particular admits a Whitehead type formula which expresses its conjugate in a proper parabolic subgroup as a product of elementary factors.

One gets a very important special case of the above elements taking $u = e_i$, $v = e_j$, $w = e_h$, where all the six indices $\pm i$, $\pm j$, $\pm h$ are pair-wise distinct. Then the element $t = T_{i,-j}(\eta)T_{i,-h}(\zeta)T_{j,-h}(\vartheta)$ is a long root type element for arbitrary $\eta, \zeta, \vartheta \in R$. Take for example $i = 1$, $j = 2$, $h = 3$. Then the matrix of t looks as follows:

$$t = \begin{pmatrix} 1 & 0 & 0 & \zeta & \eta & 0 \\ & 1 & 0 & \vartheta & 0 & -\eta \\ & & 1 & 0 & -\vartheta & -\zeta \\ & & & 1 & 0 & 0 \\ & & & & 1 & 0 \\ & & & & & 1 \end{pmatrix}$$

For any column $v = (v_1, \dots, v_{-1})^t$ it is easy to pick up $\eta, \zeta, \vartheta \in R$ in such a way that t stabilizes v. For example $tv = v$ for $t = T_{i,-j}(v_{-h})T_{i,-h}(-v_{-j})T_{j,-h}(v_{-i})$.

Varying v over the columns of the matrix g^{-1} we get an analogue of the theme for the orthogonal group.

Variation 6. *Let R be a commutative ring such that $2 \in R^*$. Then for any $g \in \mathrm{GL}(2l, R)$, $l \geq 3$, the elementary orthogonal group $\mathrm{EO}(2l, R)$ is generated by the orthogonal transvections $T_{uv}(\xi)$ such that $u_r = v_r = 0$ for some index r and $(\widetilde{u}g^{-1})_s = (\widetilde{v}g^{-1})_s = 0$ for some index s.*

Proof. The group $\mathrm{EO}(n, R)$ is generated by the elementary orthogonal transvections $T_{i,-j}(\xi)$, where $i \neq \pm j$, $\xi \in R$. Since $l \geq 3$ one can choose an index h such that $\pm i, \pm j, \pm h$ are pair-wise distinct. One has

$$T_{i,-j}(\xi) = \prod T_{i,-j}(-g_{k,-h}\xi g'_{-h,k})T_{i,-h}(g_{k,-h}\xi g'_{-j,k})T_{j,-h}(-g_{k,-h}\xi g'_{-i,k}),$$

where $k \in \{1, \dots, -1\}$. The $2l$ factors on the right hand side enjoy the required properties: take $r = -i$ and $s = k$ (see [Va5], § 10 for details).

This lemma allows to rewrite the proofs of the main structure theorems for orthogonal groups in the same spirit as we did it in § 5 for the general linear group. But, in fact, we can do a lot more. First, in constructing the decomposition of an orthogonal transvections we never assumed the matrix g itself to be orthogonal. This means that we are able not just to rewrite the description of normal subgroups of $\mathrm{SO}(n, R)$ in this style. We are able to prove that any non-central subgroup of $\mathrm{GL}(n, R)$ which is normalized by $\mathrm{EO}(n, R)$ contains an element from a proper parabolic subgroup. This allows to establish a "standard" description for the subgroups of $\mathrm{GL}(n, R)$ normalized by $\mathrm{EO}(n, R)$ in terms of *pairs* of ideals of the ground ring. To illustrate the scope of this result, let us mention that it is a common generalization of the description of normal subgroups of $\mathrm{GL}(n, R)$, the description of normal subgroups of $\mathrm{SO}(n, R)$ and the classification of overgroups of $\mathrm{SO}(n, R)$ in $\mathrm{GL}(n, R)$.

Another important feature of the proof is that one does not use the whole stock of unipotents of the split orthogonal group. The proof relies solely on the fact that there is a "singular plane" in the geometry of long root subgroups. Thus we do not have to assume that our group is split, it suffices if it is "isotropic enough". This is true at least when the Witt index of the underlying module V is at least three. In particular, the above lemma applies also to the odd orthogonal groups, starting from degree 7 (rank 3). As we shall see in the next section, with a slightly more sophisticated decomposition, in most cases it is enough to require that the index is two. Actually, [V9] covers a still larger class of isotropic orthogonal groups, defined in terms of the local Witt indices. By the geometric nature of our proof we have to get enough unipotents globally.

§ 14. Seventh variation: symplectic transvections

In this section we briefly outline how to modify the method of the present paper to make it work for the symplectic group. Since the geometry of long root subgroups in the symplectic group is extremely poor, the proofs are perceptibly trickier here than in the linear and orthogonal cases, and require several new ideas and below we list the typical arising complications. First, since the group $\mathrm{Sp}(2l, R)$ is not generated by the short root unipotents (when $2 \notin R^*$) one has to decompose both the short and the long root elements. Second, it is a tricky business to put the factors in the decomposition of a short root unipotent into proper parabolic subgroups and thus minors of g^{-1}, rather than its individual elements, are used to construct the decomposition. Third, some factors in the decomposition of a long root unipotent do not sit in a proper parabolic subgroup and to guarantee reduction to smaller ranks one needs further Whitehead type lemmas. But the most dramatic complication of all is that one cannot write an easy formula to decompose a long root unipotent, and proves a qualitative result instead.

With the same numbering of the indices from 1 to $n = 2l$ as above we introduce on the free R-module $V = R^n$, $n = 2l$, a symplectic form by

$$(x, y) = (x_1 y_{-1} - x_{-1} y_1) + \ldots + (x_l y_{-l} + x_{-l} y_l),$$

for $x = (x_1, \ldots, x_{-1})$; $y = (y_1, \ldots, y_{-1}) \in V$. By definition of the form the standard base $e_1, \ldots, e_{-1}$ of V is a Witt base of V.

The *symplectic group* $\mathrm{Sp}(2l, R)$, is the isometry group of this form. Define the matrix $J = \mathrm{sdiag}(1, \ldots, 1, -1, \ldots, -1)$, where the series of 1's and (-1)'s consist of l members each. Then $\mathrm{Sp}(2l, R)$ consists of those matrices $a \in \mathrm{GL}(n, R)$ for which $aJ_{2l}a^t = J_{2l}$. A matrix $a = (a_{ij})$ is symplectic if and only if $a'_{ij} = \varepsilon_i \varepsilon_j a_{-j,-i}$, where ε_i is the sign of i, i.e. $\varepsilon_i = 1$, if $1 \le i \le l$, and $\varepsilon_i = -1$, if $-l \le i \le -1$.

The *elementary root unipotents* $T_{ij}(\xi)$, $i \ne j$, $\xi \in R$, have the form

$$T_{i,-i}(\xi) = e + \xi e_{i,-i}, \qquad i = -j,$$

these are the *long root unipotents*, and the form

$$T_{ij}(\xi) = T_{-j,-i}(-\varepsilon_i \varepsilon_j \xi) = e + \xi e_{ij} - \varepsilon_i \varepsilon_j \xi e_{-j,-i}, \qquad i \ne \pm j,$$

these are the *short root unipotents*. The subgroup of $\mathrm{Sp}(2l, R)$ generated by all elementary root unipotents $T_{ij}(\xi)$, $i \ne j$, $\xi \in R$, is called the *elementary symplectic group* and denoted by $\mathrm{Ep}(2l, R)$.

Let $u \in V$ be a vector and $\xi \in R$ be a scalar. We define a linear transformation $T_u(\xi)$ of the module V by the formula $T_u(\xi)x = x + (u, x)\xi u$. If $u, v \in V$ are two orthogonal vectors we define $T_{u,v}(\xi)$ by the formula

$$T_{uv}(\xi)x = x + \xi(u, x)v + \xi(v, x)u.$$

When $2 \in R^*$ one has $T_u(\xi) = T_{uu}(\xi/2)$. Clearly for any $u, v \in V$, $\xi \in R$ such that $(u, v) = 0$, the transformations $T_u(\xi)$ and $T_{uv}(\xi)$ belong to $G = \mathrm{Sp}(n, R)$. The transformations $T_u(\xi)$ are called *symplectic transvections*. They are just the usual (linear) transvections contained in the symplectic group. The transformations $T_{uv}(\xi)$ are often called *symplectic transvections* too, but more commonly *Eichler-Siegel-Dickson transvections* or *ESD-transvections*. The transformations $T_u(\xi)$ are precisely what we call elements of long root type in the natural representation of symplectic groups while $T_{uv}(\xi)$ for linearly independent u and v are the elements of short root type in this representation.

The conjugation by $g \in \mathrm{Sp}(n, R)$ has the usual effect on $T_u(\xi)$ and on $T_{uv}(\xi)$:

$$gT_u(\xi)g^{-1} = T_{gu}(\xi), \quad gT_{uv}(\xi)g^{-1} = T_{gu,gv}(\xi).$$

The elementary symplectic transvections are obtained in the following way. Take $u = e_i$, $v = e_{-j}$. Then $T_{uv}(\xi) = T_{ij}(\xi)$ and in particular $T_u(\xi) = T_{i,-i}(\xi)$.

For a column $u \in R^n$ we denote by $\widetilde{u} \in {}^nR = (R^n)^*$ the row $(Fu)^t$. In other words if $u = (u_1, \ldots, u_l, u_{-l}, \ldots, u_{-1})^t$ then $\widetilde{u} = (u_{-1}, \ldots, u_{-l}, -u_l, \ldots, -u_1)$.

Comparing this notation with the definition of the inner product on $V = R^n$ we see that $(u, v) = \widetilde{v}u$. With this notation the matrices of $T_u(\xi)$ and $T_{uv}(\xi)$ in the Witt base $e_1, \dots, e_{-1}$ look as follows.

$$T_u(\xi) = e + u\xi\widetilde{u}, \qquad T_{uv}(\xi) = e + u\xi\widetilde{v} + v\xi\widetilde{u}.$$

A symplectic version of Lemma 3 implies that $T_u(\xi) \in \mathrm{Ep}(2l, R)$ if u has at least one zero component. Let now u and v be two orthogonal isotropic vectors. Then clearly

$$T_{uv}(\xi) = T_{u+v}(\xi)T_u(-\xi)T_v(-xi),$$

so that the elementary subgroup $\mathrm{Ep}(2l, R)$ is generated by the long root unipotents. Unfortunately, to decompose a long root unipotent one needs also the short root ones. As we shall see, it is relatively easy to decompose a short root unipotent as a product of short root unipotents whose conjugates by a given matrix g belong to $\mathrm{Ep}(2l, R)$. Unfortunately, short root unipotents do not generate the whole elementary symplectic group, when 2 is not invertible in R.

For short root unipotents it is not enough to assume that $u_i = 0$ to conclude that $T_{uv}(\xi) \in \mathrm{Ep}(2l, R)$. In fact one should have at least two zeros in the columns u and v. The most immediate analogue of Lemma 3 is when T_{uv} is contained in a conjugate of the parabolic subgroup P_1, or, in other words, when $u_r = v_r = 0$. But there is also another Whitehead type lemma, which asserts that $T_{uv}(\xi) \in \mathrm{Ep}(2l, R)$ if $u_r = u_{-r} = 0$. This lemma was first proven [Ko] and [Ta1] (see also [BV2], § 6).

In [Va5] we used yet another analogue of Whitehead lemma, with respect to the parabolic subgroup P_l. Namely, it is easy to check that $T_{uv}(\xi) \in \mathrm{Ep}(2l, R)$ if the following condition holds: for every $s \in \{1, \dots, -1\}$ at least one of the following holds: $u_s = 0$ or $u_{-s} = 0$. In particular, if $T_{uv}(\xi)$ satisfies the following condition: there exists an l-element subset $J \subset \{1, \dots, -1\}$ such that $u_s = 0$ for all $s \in J$, then one has the following alternative: either there exists an r such that $u_r = u_{-r} = 0$, or for every s at least one of u_s, u_{-s} equals 0. In both cases $T_{uv}(\xi) \in \mathrm{Ep}(2l, R)$.

This shows that the following result implies structure theorems for symplectic group.

Variation 7. *Let R be a commutative ring. Then for any $g \in \mathrm{Sp}(2l, R)$, $l \geq 3$, the elementary symplectic group $\mathrm{Ep}(2l, R)$ is generated by:*

1) The short root elements $T_{uv}(\xi)$ such that $u_r = v_r = 0$ for some index r and there exists an l-element subset $J \subset \{1, \dots, -1\}$ such that $(\widetilde{u}g^{-1})_s = 0$ for all indices $s \in J$;

2) The long root elements $T_u(\xi)$ such that $u_r = 0$ for some index r and $gu_s = 0$ for some index s.

Sketch of proof. We do not check all technical details but outline the general scheme. However we believe that just enough details are given for anyone to convert it into an actual proof within a quarter of an hour.

Step 1. One expresses any elementary short root element as a product of short root elements satisfying 1) above. This is done as follows. Fix an element $T_{ij}(\xi)$, $\iota \neq \pm j$, $\xi \in R$. It is immediate to check that for arbitrary $\zeta_h \in R$, $h \neq \pm i$, and $\eta \in R$ the product

$$t = T_{i,-i}(2\eta) \prod T_{ih}(\zeta_h), \quad h \neq \pm i,$$

is an element of short root type. Now we may decompose t as a product of two *linear* transvections $t = t_1 t_2$, where

$$t_1 = t_{i,-i}(\eta) \prod t_{h,-i}(\varepsilon_h \zeta_h), \qquad t_2 = t_{i,-i}(\eta) \prod t_{i,h}(\zeta_h),$$

where both products are taken in $h \neq \pm i$. Now taking as $\eta = \zeta_{-i}$ and ζ_h multiples of certain minors of order l of the matrix g^{-1}, as we did in § 6, we may choose them in such a way that the left multiplication by t_2 stabilizes l columns of g^{-1} (if it comes to that, we could stabilize $2l-2$ columns, but to apply the P_l version of Whitehead lemma l is enough).

Say, if we want to stabilize the columns of g^{-1} with indices from $J \subset \{1, \dots, -1\}$, $|J| = l$, we could define η and ζ_h as follows. Let $L = \{j\} \cup \{h \mid \varepsilon_h \neq \varepsilon_j\}$. Then we set

$$t_J = T_{i,-i}(2\widetilde{\Delta}^J_{L\setminus\{j\}}\xi\Delta_J^{L\setminus\{-i\}}) \prod T_{ih}(\widetilde{\Delta}^J_{L\setminus\{j\}}\xi\Delta_J^{L\setminus\{h\}}),$$

where the product is taken in $h \in L\{-i\}$. The matrix g being symplectic this automatically implies that the rows of g with indices from $-J$ are stabilized by the right multiplication by t_1 (here we used the fact that the matrix g is symplectic). This means that for any choice of J we get a short root element t such that gtg^{-1} is indeed elementary. As in § 6 Binet–Cauchy theorem immediately implies that $T_{ij}(\xi)$ is a product of t_J as J ranges over all l-element subsets of $\{1, \dots, -1\}$.

This already proves the lemma when $2 \in R^*$ since in this case $T_{i,-i}(\xi)$ can be expressed as the commutator of two short root elements. The general case requires somewhat more.

Step 2. Now we want to show that for any $\xi \in R$ one can express as a product of elements of type 2) above an element of the form the form $x = T_{i,-i}(\xi)T_{i,-j}(*)T_{j,-j}(*)$. Take $u, v \in U$ such that $(u, v) = 0$. Then for any $\eta, \vartheta \in R$ the product

$$t = T_u(\eta^2)T_{uv}(\eta\vartheta)T_v(\vartheta^2)$$

is a long root type element. In fact a straightforward calculation shows that $t = T_{\eta u + \vartheta v}(1)$. One gets a very important special case of this formula taking $u = e_i$, $v = e_j$, where all the four indices $\pm i$, $\pm j$ are pair-wise distinct. Then the element $t = T_{i,-i}(\eta^2)T_{i,-j}(\eta\vartheta)T_{j,-j}(\vartheta^2)$ is a long root type element for arbitrary $\eta, \vartheta \in R$. Take for example $i = 1$, $j = 2$. For any column v one may pick up $\eta, \vartheta \in R$ in such a way that t stabilizes v. Now we search for x in the form

$$\prod T_{i,-i}(\eta_h \xi (g'_{-j,h})^2) T_{i,-j}(-\eta_h \xi g'_{-i,h} g'_{-j,h}) T_{j,-j}(\eta_h \xi (g'_{-i,h})^2),$$

where $\eta_h \in R$. The row $(g'_{-j,1}, \dots, g'_{-j,-1})$ is unimodular being a row of an invertible matrix. Thus also the row $\big((g'_{-j,1})^2, \dots, (g'_{-j,-1})^2\big)$ is unimodular, so that one can choose η_h in such a way that $\sum \eta_h (g'_{-j,h})^2 = 1$. This means indeed one can express x as a product of elements of type 2).

Step 3. Now combining the results of Steps 1 and 2 we see that for any $\xi \in R$ we can express as a product of elements of the forms 1) and 2) above an element of the form $x = T_{i,-i}(\xi) T_{j,-j}(*)$. Obviously, these elements together with the elementary short root unipotents generate the group $\mathrm{Ep}(2l, R)$. In fact, one has $[T_{ji}(1), T_{i,-i}(\xi) T_{j,-j}(*)] T_{j,-i}(\mp\xi) = T_{j,-j}(\pm\xi)$ so that long root elements can be expressed as product of elements of the forms 1) and 2).

This lemma immediately implies that $\mathrm{Ep}(2l, R)$ is normal in $\mathrm{Sp}(2l, R)$. Other results in the spirit of § 5, 7, 11 easily follow by introducing appropriate changes in the proofs.

§ 15. Further variations

Here we mention also some further variations, in particular, generalizations to exceptional groups and similar decompositions for elements which are not unipotent. These decompositions involving semisimple elements only work over some special classes of rings, but when they do one easily gets much stronger results than in the general case.

Variation 8: Groups of types E_6 and E_7. This proof works with the simply connected Chevalley groups $G = \mathrm{G}(\Phi, R)$ and $\Phi = \mathrm{E}_6, \mathrm{E}_7$ in the minimal representations, of degree 27 and 56 respectively, with highest weights $\overline{\omega}_1$ and $\overline{\omega}_7$. When we speak about the elements of G as matrices we always refer to these representations, the columns and rows being numbered by the corresponding weights (for microweight representations this is unambiguous, all weights having multiplicity 1). The proof of analogues of Theorems 1 and 2 for this cases based on an appropriate generalization of the Theme was first sketched in [VPS]. Almost all details for the case E_6 were supplied in [Va5], but we omitted there an explicit verification of the fact that the signs of the structure constants of the action are compatible with the equations defining the orbit of the highest weight vector, which at that point used explicit computer calculations. The definitive proof for both E_6 and E_7 is given in [Va7] where we use a recipe to calculate the signs of the action constants and the equations based on the weight diagrams instead.

In principle the proof there follows very closely those for A_l and D_l reproduced in § 5 and 13 above and is much more straightforward than that for C_l in § 14, but now another complication occurs. We cannot construct a unipotent stabilizing an *arbitrary* column and since we are interested only in the columns of the matrices from $G = G(\mathrm{E}_6, R)$ or $G = G(\mathrm{E}_7, R)$ we construct such a unipotent only for *singular* columns, i.e. essentially columns satisfying the equations which in the case of a field would define the closure of the orbit of the highest weight vector (all columns of matrices from G are singular). The orbit of the highest weight vector is an intersection of quadrics whose equations are very well known (from the standard monomial theory or, otherwise, as the first or second partial derivatives of the higher degree invariants, the cubic form for E_6 and the quartic form for E_7).

Then in both cases for any matrix $g \in \mathrm{G}(\Phi, R)$ the elementary subgroup $\mathrm{E}(\Phi, R)$ is generated by the unipotents of root type, having the form

$$z_\mu = x_{\beta_1}(\pm\xi g'_{\lambda_1,\mu}) \dots x_{\beta_m}(\pm\xi g'_{\lambda_m,\mu}),$$

where $\beta_1, \dots, \beta_m$ runs through maximal subsets of roots of Φ forming pair-wise angles $\pi/3$ (clearly, $m = 5$ for E_6 and $m = 7$ for E_7), μ runs through all weight of the corresponding representations and $\lambda_1, \dots, \lambda_m$ are certain weight depending on the choice of $\beta_1, \dots, \beta_m$. These elements correspond to a special type of 'singular subspaces' in the geometry of root subgroups in the corresponding groups. The most delicate part is the choice of signs so that z_μ preserves the μ-th column of g^{-1}. To prove that z_μ indeed does preserve the column we have to use the fact that this column is singular, i.e. satisfies certain quadratic equations. Once this is done, the rest is relatively easy, since in each case $gz_\mu g^{-1}$ sits in an appropriate conjugate of the maximal parabolic subgroup corresponding to the highest weight of our representation (i.e. P_1 for E_6 and P_7 for E_7). Now an appropriate analogue of Whitehead lemma finishes our job.

Variation 9: Other exceptional cases. The proof described above for E_6 and E_7 is a very simple one. Actually it is far too simple to carry over to other exceptional cases without some further technical complications. In all other exceptional cases one simply does not have a "singular subspace" in the geometry of root subgroups large enough to stabilize an arbitrary column, or even a singular column. So one has develop a D_l-proof rather then the A_l-proof described above (the proofs for A_l, D_l, E_6 and E_7 were an A_2-proof, an A_3-proof, an A_5-proof and an A_7-proof, respectively).

The proof for all exceptional groups was worked out by the second-named author and E.B.Plotkin (see [Va5] and [VP] for a detailed description of the whole project). It involves more complicated patterns in the geometry of root subgroups and requires a more delicate analysis of the equations (here, unlike the above cases, one needs equations already to check that a certain element is in fact an element of root type, for singular subspaces this was automatic, since there are no equations). For the adjoint representations of E_6, E_7 and E_8 one checks that products of elements which are in turn products of 8, 10 or 14 elementary unipotents (so these are a D_5-proof, a D_6-proof and a D_8-proof respectively).

The most difficult case is that of F_4 over a ring where 2 is not invertible, which combines all the complications typical for the adjoint exceptional cases with those we encountered for the symplectic group. In this case the entries of the long root type unipotents generating $\mathrm{E}(\mathrm{F}_4, R)$ and stabilizing the μ-th column of a matrix g^{-1} (as μ runs over all weights of the 26-dimensional representation) depend *quadratically* on the coordinates of the column, having the form

$$
\begin{aligned}
&x_{\beta_1}(\pm(g'_{\lambda_1\mu})^2)x_{\beta_2}(\pm(g'_{\lambda_2\mu})^2)x_{\beta_3}(\pm(g'_{\lambda_3\mu})^2)\cdot\\
&x_{\beta_4}(\pm g'_{\lambda_1\mu}g'_{\lambda_2\mu})x_{\beta_5}(\pm g'_{\lambda_1\mu}g'_{\lambda_3\mu})x_{\beta_6}(\pm g'_{\lambda_2\mu}g'_{\lambda_3\mu})\cdot\\
&x_{\beta_7}(\pm g'_{\lambda_1\mu}g'_{\nu_1\mu}\pm g'_{\lambda_2\mu}g'_{\nu_2\mu}\pm g'_{\lambda_3\mu}g'_{\nu_3\mu})x_{\beta_8}(\pm g'_{\lambda_1\mu}g'_{\rho_1\mu}\pm g'_{\lambda_2\mu}g'_{\rho_2\mu}\pm g'_{\lambda_3\mu}g'_{\rho_3\mu})
\end{aligned}
$$

for certain roots β_i and certain weights $\lambda_1, \dots, \rho_3$. It is quite an exercise to show that the unipotents of this shape actually generate the whole elementary subgroup $\mathrm{E}(\mathrm{F}_4, R)$. For G_2 in the 7-dimensional representation the corresponding formulae are *cubic* in the components of u.

Variation 10: Unitary groups. Let $G = \mathrm{U}(2n, R, \Lambda)$ be the hyperbolic unitary group over a form ring (R, Λ), corresponding to a symmetry λ, $\lambda\overline{\lambda} = 1$ (see [B1] – [B3] or [HO] for a definition of group rings and unitary groups). The involution on R is denoted by $\xi \mapsto \overline{\xi}$. In the papers [BV2], [BV3] A.Bak and the second-named author develop fundamentals of decomposition of unipotents in G over certain non-commutative rings satisfying some weaker commutativity condition. The following types of unipotents are used in [BV3] to generate the elementary subgroup $\mathrm{EU}(2n, R, \Lambda)$ in the unitary analogue of the Theme. 1) Long root type unipotents depending on $\xi, \zeta \in R$ and $\alpha \in \Lambda$, looking as follows:

$$
t = \begin{pmatrix} 1 & 0 & \overline{\xi}\alpha\zeta & -\overline{\xi}\alpha\xi \\ 0 & 1 & -\overline{\zeta}\alpha\zeta & \overline{\zeta}\alpha\xi \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.
$$

2) Short root type unipotents depending on $\xi, \zeta, \vartheta \in R$ looking as follows:

$$
t = \begin{pmatrix} 1 & 0 & 0 & \xi\overline{\zeta} & \xi\overline{\xi} & 0 \\ 0 & 1 & 0 & -\overline{\lambda}\xi\overline{\vartheta} & 0 & -\overline{\lambda}\xi\overline{\xi} \\ 0 & 0 & 1 & \zeta\overline{\vartheta} - \overline{\lambda}\vartheta\overline{\zeta} & \vartheta\overline{\xi} & -\overline{\lambda}\zeta\overline{\xi} \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.
$$

For a commutative or an almost commutative ring R it is easy to construct an element of such form in the stabilizer of any column $u \in V$. One easily recognizes the unipotents constructed in § 13 and § 14 as special cases of these new unipotents, but now non-commutativity and presence of involution cause additional technical complications in the proofs.

Variation 11: Non-unipotents. For some special classes of rings one can find non-unipotent elements stabilizing columns of a given matrix. Here we cite only one example, based on an idea from [Se]. Let R be a commutative domain satisfying the following two conditions: 1) all proper factor-rings of R are finite, 2) R has a unit θ of infinite order. We want to stabilize a column $u = (u_i) \in R^n$ by an element of the shape $z = e + (a-1)e_{rr} + be_{rs}$, where $r \neq s$, $a \in R^*$ and $b \in R$. Clearly $(zu)_i = u_i$ for all $i \neq r$, so we have only to choose a and b so that $(zu)_r = u_r$. Since the factor-ring R/u_sR is finite there exist a power θ^m such that $\theta^m \equiv 1 (\operatorname{mod} u_s)$. In other words, $\theta^m = 1 + \eta u_s$ for some $\eta \in R$. Setting $a = \theta^m$ we get $(zu)_r = au_r + bu_s = \theta^m u_r + bu_s = u_r + \eta u_s u_r + bu_s$. Now it remains to set $b = -\eta u_r$. The fact that there are plenty of elements z with this property indeed played central role in the study of the normal structure of $\mathrm{GL}(2,R)$ in [Se], as well as in some further problems related to the description of subgroups in $\mathrm{GL}(n,R)$. In particular, these elements can be used to describe subgroups of $\mathrm{SL}(n,R)$ containing the group of upper-triangular matrices $B(n,R)$, or the group of block diagonal matrices $D(\nu)$ with $h(\nu) \geq 2$. The second named author has obtained a generalization of this construction to other Chevalley groups, which will be published in a separate paper.

Bibliography

[**A1**] Abe E., *Chevalley groups over commutative rings*, Proc. Conf. Radical Theory, (Sendai – 1988), Uchida Rokakuho Publ. Comp., Tokyo, 1989, pp. 1–23.

[**A2**] Abe E., *Normal subgroups of Chevalley groups over commutative rings.*, Contemp. Math. **83** (1989), 1–17.

[**AS**] Abe E., Suzuki K., *On normal subgroups of Chevalley groups over commutative rings*, Tôhoku Math. J. **28** (1976), no. 1, 185–198.

[**B1**] Bak A., *The stable structure of quadratic modules*, Thesis Columbia Univ., 1969,.

[**B2**] Bak A., *On modules with quadratic forms*, Lecture Notes Math. **108** (1969), 55–66.

[**B3**] Bak A., *K-theory of forms*, Ann. of Math. Stud., vol. 98, Princeton Univ. Press, Princeton, N.J., 1981.

[**B4**] Bak A., *Subgroups of the general linear group normalized by relative elementary groups*, Lecture Notes Math. **967** (1982), 1–22.

[**B5**] Bak A., *Nonabelian* K*-theory: The nilpotent class of* K_1 *and general stability*, K–Theory **4** (1991), 363–397.

[**BV1**] Bak A., Vavilov N.A., *Normality for elementary subgroup functors*, Math. Proc. Cambridge Phil. Soc. **118** (1995), no. 35, 35–47.

[**BV2**] Bak A., Vavilov N.A., *Structure of hyperbolic unitary groups.* I., Preprint Univ. Bielefeld (1988), no. 112, 1–44.

[**BV3**] Bak A., Vavilov N.A., *Structure of hyperbolic unitary groups.* II. *Normal subgroups*, to appear.

[Ba1] Bass H., *K-theory and stable algebra*, Publ. Math. Inst. Hautes Et. Sci. (1964), no. 22, 5–60.

[Ba2] Bass H., *Algebraic K-theory*, Benjamin, N.Y. – Amsterdam, 1968.

[Ba3] Bass H., *Unitary algebraic K-theory*, Lecture Notes Math. **343** (1973), 57–265.

[BMS] Bass H., Milnor J., Serre J.–P., *Solution of the congruence subgroup problem for* SL_n $(n \geq 3)$ *and* Sp_{2n} $(n \geq 2)$, Publ. Math. Inst. Hautes Et. Sci. (1967), no. 33, 59–137.

[BoV1] Borewicz Z.I., Vavilov N.A., *Subgroups of the general linear group over a semilocal ring containing the group of diagonal matrices*, Proc. Steklov Inst. Math. (1980), no. 4, 41–54.

[BoV2] Borewicz Z.I., Vavilov N.A., *On the definition of a net subgroup*, J. Sov. Math. **30** (1985), 1810–1816.

[BoV3] Borewicz Z.I., Vavilov N.A., *The distribution of subgroups in the general linear group over a commutative ring*, Proc. Steklov. Inst. Math. (1985), no. 3, 27–46.

[Co1] Cohn P., *On the structure of the* GL_2 *of a ring*, Publ. Math. Inst. Hautes Et. Sci. (1966), no. 30, 365–413.

[Co2] Cohn P., *Free rings and their relations*, Academic Press, London, 1971.

[CK1] Costa D., Keller G., *The* $\mathrm{E}(2, A)$ *sections of* $\mathrm{SL}(2, R)$, Ann. Math. **134** (1991), 159–188.

[CK2] Costa D., Keller G., *Radix redux: normal subgroups of symplectic group*, J. reine angew. Math. **427** (1992), no. 1, 51–105.

[CK3] Costa D., Keller G., *On the normal subgroups of* $\mathrm{G}_2(A)$, Preprint Univ. Virginia (1996), 1–46.

[Di] Dieudonné J., *La géométrie des groups classiques, 3rd ed.*, Springer, Berlin et al., 1971.

[Ge] Gerasimov V.N., *The group of units of a free product of rings*, Math. U.S.S.R. Sbornik **62** (1989), no. 1, 41–63.

[Go1] Golubchik I.Z., *On the general linear group over an associative ring*, Uspehi Mat. Nauk **28** (1973), no. 3, 179–180, (in Russian).

[Go2] Golubchik I.Z., *On the normal subgroups of orthogonal group over an associative ring with involution*, Uspehi Mat. Nauk **30** (1975), no. 6, 165, (in Russian).

[Go3] Golubchik I.Z., *The normal subgroups of linear and unitary groups over rings*, Ph.D. Thesis, Moscow State Univ., 1981, (in Russian).

[Go4] Golubchik I.Z., *On the subgroups of the general linear group* $\mathrm{GL}_n(R)$ *over an associative ring* R, Russian Math. Surveys **39** (1984), no. 1, 157–158.

[Go5] Golubchik I.Z., *On the normal subgroups of the linear and unitary groups over associative rings*, Spaces over algebras and some problems in the theory of nets, Ufa, 1985, pp. 122–142, (in Russian).

[Go6] Golubchik I.Z., *On the general linear group over weak Noetherian associative algebras*, Fundamental and Appl. Math. **1** (1995), no. 3, 661–668.

[GM] Golubchik I.Z., Markov V.T., *Localization dimension of PI–rings*, Trudy Sem. Petrovskogo (1981), no. 6, 39–46, (in Russian).

[GM1] Golubchik I.Z., Mikhalev A.V., *Elementary subgroup of a unitary group over a PI-ring*, Vestnik Mosk. Univ. ser.1, – Mat., Mekh. (1985), no. 1, 30–36.

[GM2] Golubchik I.Z., Mikhalev A.V., *On the elementary group over a PI-ring*, Studies in Algebra, Tbilisi, 1985, pp. 20–24, (in Russian).

[Hb1] Habdank G., *A classification of subgroups of* Λ*-quadratic groups normalized by relative elementary subgroups*, Dissertation Universität Bielefeld, 1987, pp. 1–71.

[Hb2] Habdank G., *Homology stability for* Λ*-quadratic groups*, Preprint SFB Universität Bielefeld (1990), no. 74, 1–18.

[HO] Hahn A.J., O'Meara O.T., *The classical groups and K-theory*, Springer, Berlin et al., 1989.

[vdK] van der Kallen W., *Another presentation for the Steinberg group*, Proc. Nederl. Akad. Wetensch. Ser. A **80** (1977), 304–312.

[Kh1] Khlebutin S.G., *Sufficient conditions for the normality of the subgroup of elementary matrices*, Russian Math. Surveys **39** (1984), no. 3, 245–246.

[Kh2] Khlebutin S.G., *Some properties of the elementary subgroup*, Algebra, Logic and Number Theory, Moscow State Univ., 1986, pp. 86–90, (in Russian).

[Kh3] Khlebutin S.G., *Elementary subgroups of linear groups over rings*, Ph. D. Thesis, Moscow State Univ., 1987, (in Russian).

[Ko] Kopeiko V.I., *The stabilization of symplectic groups over a polynomial ring*, Math. U.S.S.R. Sbornik **34** (1978), 655–669.

[La] Lam T.Y., *Serre's problem*, Lecture Notes Math., vol. 635, 1978.

[L1] Li Fuan, *The structure of symplectic groups over arbitrary commutative rings*, Acta Math. Sinica **3** (1987), no. 3, 247–255.

[L2] Li Fuan, *The structure of orthogonal groups over arbitrary commutative rings*, Chin. Ann. Math. **10B** (1989), no. 3, 341–350.

[LL] Li Fuan, Liu Mulan, *Generalized sandwich theorem*, K-theory **1** (1987), 171–184.

[Md1] McDonald B.R., *Automorphisms of* $\mathrm{GL}_n(R)$, Trans. Amer. Math. Soc. **246** (1978), 155–171.

[Md2] McDonald B.R., *Geometric algebra over local rings*, Marcel Dekker, N.Y., 1976.

[Md3] McDonald B.R., *Linear algebra over commutative rings*, Marcel Dekker, N.Y., 1984.

[Mi] Milnor J., *Introduction to algebraic K-theory*, Princeton Univ. Press, Princeton, N.J., 1971.

[Se] Serre J.-P., *Le problḿe des groups de congruences pour* SL_2, Ann. Math. **92** (1970), 489–527.

[SSv] Sivatski A.S., Stepanov A.V., *On the word length of commutators in* $\mathrm{GL}(n,R)$, to appear, K-Theory (1999).

[St] Stein M.R., *Relativizing functors on rings and algebraic K-theory*, J. Algebra **19** (1971), no. 1, 140–152.

[Sv1] Stepanov A.V., *Stability conditions in the theory of linear groups over rings*, Ph.D. Thesis. Leningrad State Univ., 1987, pp. 1–112, (in Russian).

[**Sv2**] Stepanov A.V., *A ring of finite stable rank is not necessarily finite in the sense of Dedekind*, Soviet Math. Doklady **36** (1988), no. 2, 301–304.

[**Sv3**] Stepanov A.V., *Description of subgroups of the general linear group over a ring with the use of stability conditions.*, In: Rings and linear groups. (1988), Krasnodar, 82–91, (in Russian).

[**Sv4**] Stepanov A.V., *On the distribution of subgroups normalized by a given subgroup*, J. Sov. Math. **64** (1993), 769–776.

[**Sv5**] Stepanov A.V., *On the normal structure of the general linear group over a ring.*, Zapiski nauchnyh seminarov POMI **236** (1997), 166–182, (in Russian, English transl. to appear in J. Math. Sci.).

[**Su1**] Suslin A.A, *On the structure of the general linear group over polynomial rings*, Soviet Math. Izv. **41** (1977), no. 2, 503–516.

[**Su2**] Suslin A.A., *Algebraic K-theory*, J. Sov. Math. **28** (1985), no. 6, 870–923.

[**SK**] Suslin A.A., Kopeiko V.I., *Quadratic modules and orthogonal groups over polynomial rings*, J. Sov. Math. **20** (1982), no. 6, 2665–2691.

[**Ta1**] Taddei G., *Invariance du sous-groupe symplectique élémentaire dans le groupe symplectique sur un anneau*, C. R. Acad. Sci. Paris, Sér. I **295** (1982), no. 2, 47–50.

[**Ta2**] Taddei G., *Schémas de Chevalley–Demazure, fonctions représentatives et théorème de normalité*, Thèse, Univ. de Genève, 1985.

[**Ta3**] Taddei G., *Normalité des groups élémentaire dans les groups de Chevalley sur un anneau*, Contemp. Math. **55** (1986), no. 2, 693–710.

[**Tu**] Tulenbaev M.S., *Schur multiplier of the group of elementary matrices of finite order*, J. Sov. Math. **17** (1981), no. 4.

[**V1**] Vaserstein L.N, *On normal subgroups of* GL_n *over a ring*, Lecture Notes Math. **854** (1981), 456–465.

[**V2**] Vaserstein L.N., *Classical groups over rings*, Canad. Math. Soc. Conf. Proc. **4** (1984), 131–140.

[**V3**] Vaserstein L.N., *The subnormal structure of general linear groups*, Math. Proc. Cambridge Phil. Soc. **99** (1986), 425–431.

[**V4**] Vaserstein L.N, *On normal subgroups of Chevalley groups over commutative rings*, Tôhoku Math. J. **38** (1986), 219–230.

[**V5**] Vaserstein L.N., *Normal subgroups of orthogonal groups over commutative rings*, Amer. J. Math. **110** (1988), no. 5, 955–973.

[**V6**] Vaserstein L.N., *Normal subgroups of symplectic groups over rings*, K-theory **2** (1989), no. 5, 647–673.

[**V7**] Vaserstein L.N., *Normal subgroups of classical groups over rings and gauge groups*, Contemp. Math. **83** (1989), 451–459.

[**V8**] Vaserstein L.N., *The subnormal structure of general linear groups over rings*, Math. Proc. Cambridge Phil. Soc. **108** (1990), no. 2, 219–229.

[**VS**] Vaserstein L.N., Suslin A.A., *Serre's problem on projective modules over polynomial rings and algebraic K-theory.*, Sov. Math. Izv. **10** (1976), 937–1001.

[**Va1**] Vavilov N.A., *Subgroups of split classical groups – Dr. Sci. Thesis*, Leningrad State Univ., 1987, pp. 1–334, (in Russian).

[**Va2**] Vavilov N.A., *Structure of split classical groups over commutative rings*, Soviet Math. Doklady **37** (1988), 550–553.

[**Va3**] Vavilov N.A., *A note on the subnormal structure of general linear groups*, Math. Proc. Cambridge Phil. Soc. **107** (1990), no. 2, 193–196.

[**Va4**] Vavilov N.A., *On subgroups of the split classical groups*, Proc. Math. Inst. Steklov (1991), no. 4, 27–41.

[**Va5**] Vavilov N.A., *Structure of Chevalley groups over commutative rings*, Proc. Conf. Non–Associative Algebras, Related Topics, Hiroshima – 1990, World Scientific, Singapore et al., 1991, pp. 219–335.

[**Va6**] Vavilov N.A., *Intermediate subgroups in Chevalley groups*, Proc. Conf. Groups of Lie Type and their Geometries, Como – 1993, Cambridge Univ. Press, 1995, pp. 233–280.

[**Va7**] Vavilov N.A., *A third look at weight diagrams*, Preprint Univ. Bielfeld (1998), no. 110, 1–41.

[**Va8**] Vavilov N.A., *Geometry of minimal modules for the Chevalley groups of types* E_6 *and* E_7, In preparation.

[**Va9**] Vavilov N.A., *Linear groups over general rings.* I, II, In preparation.

[**VPe**] Vavilov N.A., Perelman E.Ya., *Transvections in polyvector representations* (to appear).

[**VP**] Vavilov N.A., Plotkin E.B., *Chevalley groups over commutative rings. I. Elementary calculations*, Acta Appl. Math. **45** (1996), 73–113.

[**VPS**] Vavilov N.A., Plotkin E.B., Stepanov A.V., *Calculations in Chevalley groups over commutative rings*, Soviet Math. Doklady **40** (1990), no. 1, 145–147.

[**Wi**] Wilson J.S., *The normal and subnormal structure of general linear groups*, Proc. Cambridge Phil. Soc. **71** (1972), 163–177.

[**Z**] Zalesskii A.E., *Linear groups*, J. Sov. Math. **31** (1985), no. 3.

E-MAIL: ALEXEI@STEPANOV.SPB.SU
VAVILOV@MATHEMATIK.UNI-BIELEFELD.DE