STRUCTURE OF HYPERBOLIC UNITARY GROUPS
I. ELEMENTARY SUBGROUPS

Anthony Bak, Nikolai Vavilov

Fakultät für Mathematik
Universität Bielefeld
33615 Bielefeld, Deutschland

and

Department of Mathematics and Mechanics
University of Sanct-Petersburg
Petrodvorets, 198904, Russia

Abstract. This is the first in the series of papers dedicated to the structure of hyperbolic unitary groups over form rings and their subgroups. In this part we recall foundations of the theory and study elementary subgroup of the hyperbolic unitary group over a ring. In particular, using a variant of Suslin's patching method, we prove standard commutator formulae for relative elementary subgroups.

1991 *Mathematics Subject Classification.* 20G35, 20H35.

Key words and phrases. Unitary group, Unitary transvections, Elementary subgroup.

The second-named author gratefully acknowledges the support of the *Alexander von Humboldt-Stiftung* and the *SFB 343* and the hospitality of the *Universität Bielefeld* during the preparation of the present article.

Introduction

In this series of papers we intend to systematically study subgroups of hyperbolic unitary groups $\mathrm{U}(2n, R, \Lambda)$ over a form ring (R, Λ), corresponding to a symmetry λ (see [B1] – [B3], [Ba3], [HO'M]). Our primary objectives are two-fold. On one hand we plan to update the foundational results from [B1], [Ba3], removing, whenever possible, stability conditions, or replacing them by weaker commutativity or finiteness conditions. The excellent recent exposition [HO'M] concentrates on the Steinberg groups, structure of unitary groups over division rings and isomorphism theory and does not contain proofs of structure results in full generality. We fill this gap. In partivular, we carry over to unitary groups over form rings the main structure theorems over general rings and their generalizations obtained for the usual classical groups during the last 20 years: [A3], [A4], [B4], [B5], [BV3], [CK], [Go1] – [Go5], [GM1], [GM2], [Hb1], [Kh1] – [Kh3], [Ko1], [Ko3], [L1], [L2], [LL], [Md3], [Su2], [SK], [Ta1] – [Ta3], [Tu], [V2] – [V8], [Va1], [Va3] – [Va6], [VPS], [Wi][1]. On the other hand we generalize to unitary groups over form rings some more specialized structure results, related to the description of various classes of subgroups. In particular, we carry over results of [BV1] – [BV3], [Va1], [Va4], [Go4] (see [Va6], [Va7], [Za] for many further references).

In the first half of this first part we fix the general framework for the rest of the series. The second half studies elementary subgroups $\mathrm{EU}(2n, R, \Lambda)$ of hyperbolic unitary groups $\mathrm{U}(2n, R, \Lambda)$. We study also relative elementary groups $\mathrm{EU}(2n, I, \Gamma)$ corresponding to form ideals (I, Γ) in (R, Λ). In particular we prove the following result (recall that a ring R is called *almost commutative* if it is finitely generated as a module over its centre).

Theorem. *Let (R, Λ) be an almost commutative form ring, $n \geq 3$. Then for any form ideal (I, Γ) the corresponding elementary subgroup $\mathrm{EU}(2n, I, \Gamma)$ is normal in the hyperbolic unitary group $\mathrm{U}(2n, R, \Lambda)$. Moreover,*

$$\mathrm{EU}(2n, I, \Gamma) = [\mathrm{EU}(2n, R, \Lambda), \mathrm{CU}(2n, I, \Gamma)],$$

where $\mathrm{CU}(2n, I, \Gamma)$ is the full congruence subgroup of level (I, Γ).

[1] We refer to [Di], [HO'M], [Md2], [Md3] for a description of the earlier stages of development. For the linear groups the old era ended with [Ba1], [BMS], [Ba2]. The thesis [B1] was not published and was not easily available, especially in Russia and China. This did a lot of harm. In fact, many works were appearing *up to late 80-ies*, proving structure theorems for classical groups over such rings, as, say, zero-dimensional ones. We do not cite these publications in our bibliography, see, for example, references in [HO'M] and also in [CK], [V5], [V6], [Va6].

In [BV] we give another proof of this result based on a variant of "localization and patching". There we show also, that this result remains true for $n \geq 2$ under the additional assumption $R\Lambda + \Lambda R = R$. However in this paper we give a direct global proof developing ideas of A.A.Suslin [Su2], [Tu], [Kh1], [Kh3]. Of course, here we also must use "patching", but in the ring R itself. The proof requires very little information about the local case, in fact one has to assume only some transitivity properties of the elementary unitary group. This gives a lot of room for generalizations in the style of [V2] or [Kh1] – [Kh3]. As we shall see in § 7 the proof embraces wider classes of rings than just the almost commutative ones. Another advantage is that unlike the methods related to passage to localizations, the global approach gives explicit formulae for decomposition of a transvection into elementaries (incidentally, the formulae from the sections §§ 4 – 7 play an important role in the subsequent parts of this paper).

This result was known in many special cases. First, in the stable range, i.e. when n is large enough compared to some sort of dimension of the ring R, this result was established by A.Bak [B1] – [B3], compare also [Ba3] (for the general linear group $\mathrm{GL}(n, R)$ analogous result was discovered by H.Bass [Ba1], [Ba2], for some hyperbolic classical groups by L.N.Vaserstein [V1], we refer to the surveys [Va5], [Va6] for a systematic bibliography).

A.A.Suslin [Su2], [Tu] has proven that the elementary subgroup $\mathrm{E}(n, R)$ is normal in the general linear group $\mathrm{GL}(n, R)$ whenever $n \geq 3$ and R is almost commutative (see also [BV3], [GMi2], [HO'M], [Kh1] – [Kh3], [Md], [Sv], [SK], [V2], [Va5], [Va6] for various proofs of this result and its generalizations to wider classes of rings). For commutative case the original proof of A.A.Suslin [Su2] is based on solution of linear equations over rings, see [SK] for a very elegant reformulation in terms of anti-symmetric matrices. A particularly simple direct proof of Suslin's result based on decomposition of unipotents was found in 1987 by A.V.Stepanov (see [Sv], [VPS] and the expositions in [Va5], [Va6]).

Another approach is based on various local-global principles. Thus, already the proof of normality for almost commutative rings required "patching" [Tu][2]. Another version of this method was proposed by L.N.Vaserstein

[2]This result was first published in a paper of M.S.Tulenbaev, but as explicitly stated there, it is due to A.A.Suslin. The paper [Kh1] improves the result slightly (pushing it to rings *algebraic* over centre), it also corrects some misprints from [Tu] and organizes the proof in a better way. However many ring theoretical arguments are omitted in [Kh1] as well. To the best of our knowledge the most systematic exposition of these ideas is contained in the thesis of S.G.Khlebutin [Kh3].

[V2] and was christened "localization and patching". It consists of throwing in polynomial variables, passage to localizations with respect to central multiplicative subsets and using corresponding local results. In general the interrelation of the patching method of A.A.Suslin and the localization and patching method of L.N.Vaserstein is exactly the same as the interrelation of solutions of Serre's problem by A.A.Suslin and D.Quillen, respectively [La][3].

The methods above operate essentially in terms of the centre of the ground ring, as do the published works by the Moscow school. As a result one has to impose some commutativity or finiteness condition. I.Z.Golubchik has successfully used *non-commutative* localizations (more specifically *Ore localizations*, in particular, results of [GM]) to describe normal subgroups [Go1] – [Go5]. He and A.V.Mikhalev told us that they had applied non-commutative localizations also to the normality of elementary subgroups, but we have not seen any written proofs.

Observe that some conditions on n and R are necessary here since the subgroup $\mathrm{E}(2,R)$ is not necessarily normal in $\mathrm{GL}(2,R)$ even for some *very* good rings R, like, say, Dedekind rings (compare [Co1], [Su1], [Sw]). On the other hand using *universal* localizations V.N.Gerasimov [Ge] (see also [Va6]) has shown that for any given n there exist some very nasty rings R for which $\mathrm{E}(n,R)$ is not normal in $\mathrm{GL}(n,R)$. Of course, these rings furnish also counter-examples to the corresponding results for the unitary groups as well. These rings are *very* far from being commutative. In fact some of their factor-rings are not weekly finite (recall that a ring R is called weekly finite, if any one-sided invertible square matrix over R is two-sided invertible, see [Co2], [BV3], [Va7] for a discussion of this condition and its role in the theory of linear groups). Observe that by the same reasons as for the group $\mathrm{GL}(2,R)$ normality fails for $\mathrm{SO}(4,R)$ (see [Ko2]).

Almost immediately the result was generalized by A.A.Suslin and V.I.Kopeiko [SK], [Ko1] from linear groups to the hyperbolic symplectic and orthogonal groups $\mathrm{Sp}(2n,R)$, $n \geq 2$, and $\mathrm{SO}(2n,R)$, $n \geq 3$, over a commutative ring R (in our terminology these groups correspond to the cases when the involution on R is trivial and, respectively, $\lambda = -1$ and $\Lambda = R$, or $\lambda = 1$

[3]For historical accuracy one may note that the first application of the Quillen's method to linear groups is also due to A.A.Suslin. It appeared in his remarkable paper [Su2] dedicated to the K_1 analogue of Serre's problem. Many authors seem to ignore that already the proof of the "Quillen's Theorem" of Suslin (see § 3 of [Su2], especially Lemma 3.3 and proofs of Lemmas 3.4 and 3.7) contained *all* ingredients of localization and patching. This method was successfully used in late 70-ies by Suslin's students V.I.Kopeiko and M.S.Tulenbaev and later by E.Abe, D.Costa and others.

and $\Lambda = \Lambda_{\min}$). The result for the symplectic group has been partially reproven in [Ta1]. Compare also [GM1], [HO'M], [Ko3], [L1], [L2], [V5], [V6], [Va1], [Va4], [Va5], where one can find other proofs of these results and some generalizations to not necessarily hyperbolic groups.

The results for the split classical groups fall also under the umbrella of Chevalley groups. Normality result for the absolute elementary subgroups in simple Chevalley groups of rank at least 2 over an arbitrary commutative ring R was first proven by G.Taddei [Ta2], [Ta3], who used "localization and patching", compare also [V4], [A3], [A4]. Another approach to the proof of this result, based on Stepanov's idea of decomposition of unipotents, was proposed in [VPS], [Va5].

The most general published result in terms of rings and forms considered is due to I.Z.Golubchik and A.V.Mikhalev [GM1] (compare also [Ko3], [HO'M]). They use a version of Suslin's patching method[4]. They phrase their results in terms of arbitrary rings with involution (not just the ring $\mathrm{M}(n, R)$, as we do), and assume that Witt index is at least 2 (in their setting it is expressed in terms of behaviour of idempotents under involution). In our setting their main result applies to the groups with $\lambda = -1$, $\Lambda = \Lambda_{\max}$ and $n \geq 2$. Actually their restrictions on λ and Λ are explained by the fact that in their proof they consider only elements of *long* root type. As noted above, for $\lambda \neq -1$ this result simply does not hold for Witt index 2. Analogous results were announced by L.N.Vaserstein [V7], but the proof never appeared.

The second commutator formula in the theorem above was discovered at the stable level by H.Bass and A.Bak. For classical groups over commutative rings it was obtained independently by L.N.Vaserstein, Z.I.Borevich, N.A.Vavilov and Li Fuan [BV2], [BV3], [L1], [L2], [V2], [V5], [V6], [Va1], [Va4]. After the work of M.R.Stein [St] it is a common understanding that the questions about relative groups may be reduced to that for the absolute ones by appropriate change of rings. The first applications of this idea to the normality of relative elementary subgroups were given by J.Milnor [Mi] and by A.A.Suslin and V.I.Kopeiko [SK] (compare also [Hb1], [HO'M], [V3], [Va5], [Va6]).

In §§ 1, 2 we reproduce fundamental definitions and notation related to the unitary groups over form rings and their elementary subgroups. Most of this material comes from [B1] – [B3], [Ba3], [Hb1], [HO'M], but is updated and adapted to our needs. In §§ 3, 4 we discuss form ideals and the corresponding

[4]The proof is not easy to follow either. In 4.3 existence of certain elements λ, s and t is claimed, but these elements are never displayed. The formulae at the end of proof of 4.3 suggest that $1+\lambda = (1+\lambda_2)(1+\overline{\lambda}_1)(1+\lambda_1)$, but then it is not clear why the condition in the next paragraph is satisfied.

relative groups. In particular we prove the usual results about the generation of the relative elementary groups and reduce the relative case of the theorem above to the absolute one. In § 5 we introduce and study a notion of ESD-transvections, slightly more general, than the usual one. In § 6 we prove certain Whitehead-type lemmas, which guarantee that an ESD-transvection lies in $\mathrm{EU}(2n, R, \Lambda)$ if it is defined by columns with zero elements. Finally in § 7 we use Suslin's patching method to prove the theorem above in the absolute case.

The paper is essentially self-contained since we prove all the subsidiary results we need. We have to do so, since in this paper we consider our group $\mathrm{U}(2n, R, \Lambda)$ with respect to a Witt base $e_1, \dots, e_n, e_{-n}, \dots, e_{-1}$ (where e_i is orthogonal to all e_j except e_{-i}). In all previously published works, where hyperbolic unitary groups over form rings were considered, the base was always numbered either as $e_1, e_{-1}, \dots, e_n, e_{-n}$, or as $e_1, \dots, e_n, e_{-1}, \dots, e_{-n}$. Our numbering in this paper is inspired by [Ca] and is much more natural in many respects, since, for example, in this numbering the standard Borel subgroup is represented by upper triangular matrices. We pay a price for this convenience though – namely we have to restate many of the foundational results. This paper presents a common background for the subsequent papers by the authors dedicated to the study of the normal structure and subgroups of $\mathrm{U}(2n, R, \Lambda)$ and related Steinberg groups.

§ 1 Hyperbolic unitary groups

In this section we recall definition of the hyperbolic unitary groups, etc.

1^0. Λ-quadratic forms. Let R be a (not necessarily commutative) associative ring with 1. For natural m, n we denote by $\mathrm{M}(m, n, R)$ the additive group of $m \times n$ matrices with entries in R. In particular $\mathrm{M}(n, R) = \mathrm{M}(n, n, R)$ is the ring of matrices of degree n over R. For a matrix $x \in \mathrm{M}(m, n, R)$ we denote by x_{ij}, $1 \leq i \leq m$, $1 \leq j \leq n$, its entry in the position (i, j). Let $e = e_n$ be the identity matrix and e_{ij}, $1 \leq i, j \leq n$, be a standard matrix unit, i.e. the matrix which has 1 in the position (i, j) and zeros elsewhere. For $x \in \mathrm{M}(m, n, R)$ we denote by x^t the naive transpose of x, i.e. the matrix $x \in \mathrm{M}(n, m, R)$ which has x_{ij} in the position (j, i). When the ring R is not commutative, this transposition does not have good properties and will *always* be used in combination with involution of the ground ring (see [Ba2] for a correct definition of transpose).

Let $\alpha \mapsto \overline{\alpha}$ be an involution on R, i.e. an anti-automorphism of order two. In other words, for any $\alpha, \beta \in R$ one has $\overline{\alpha + \beta} = \overline{\alpha} + \overline{\beta}$, $\overline{\alpha\beta} = \overline{\beta}\,\overline{\alpha}$ and $\overline{\overline{\alpha}} = \alpha$.

Fix an element $\lambda \in \mathrm{Cent}(R)$ such that $\lambda\overline{\lambda} = 1$. Set

$$\Lambda_{\min} = \{\alpha - \lambda\overline{\alpha}, \alpha \in R\}, \qquad \Lambda_{\max} = \{\alpha \mid \alpha \in R, \alpha = -\lambda\overline{\alpha}\}.$$

A *form parameter* Λ is an additive subgroup of R such that

1) $\Lambda_{\min} \subseteq \Lambda \subseteq \Lambda_{\max}$,
2) $\alpha\Lambda\overline{\alpha} \subseteq \Lambda$ for all $\alpha \in R$.

The pair (R, Λ) is called a *form ring*. Sometimes when the choice of Λ is clear from the context, we use the shortcut R for (R, Λ) and call R a form ring. For example, in many cases $\Lambda_{\min} = \Lambda_{\max}$, so that there is a unique choice of Λ for a given involution and λ. This is so, for instance, when there exists a central element $\varepsilon \in \mathrm{Cent}(R)$ such that $\varepsilon + \overline{\varepsilon} \in R^*$ (in particular, when $2 \in R^*$).

Consider a free right R-module $V \cong R^{2n}$ of rank $2n$. Fix a base $e_1, \dots, e_{2n}$ of the module V. We may think of elements $v \in V$ as columns with components in R. In particular, e_i is the column whose i-th coordinate is 1, while all other coordinates are zeros. Following [Ca] we will usually number the base as follows: $e_1, \dots, e_n, e_{-n}, \dots, e_{-1}$. According to this choice of base we write $v = (v_1, \dots, v_n, v_{-n}, \dots, v_{-1})^t$, where $v_i \in R$.

Denote by $p = p_n$ the matrix in $\mathrm{M}(n, R)$ which has 1's along the second (skew) diagonal and zeros elsewhere. In other words, p has the following matrix form:

$$\begin{pmatrix} 0 & \dots & 1 \\ \vdots & \iddots & \vdots \\ 1 & \dots & 0 \end{pmatrix}.$$

Now we consider the *sesquilinear form* f on V which has (with respect to the fixed base $e_1, \dots, e_{-1}$) the Gram matrix

$$\begin{pmatrix} 0 & p \\ 0 & 0 \end{pmatrix}.$$

In other words

$$f(u, v) = \overline{u}^t \begin{pmatrix} 0 & p \\ 0 & 0 \end{pmatrix} v = \overline{u}_1 v_{-1} + \dots + \overline{u}_n v_{-n}.$$

Now this form defines two other forms: an *even λ-hermitian form* $h = f + \lambda\overline{f}$, where $\overline{f}(u, v) = \overline{f(v, u)}$, and a *$\Lambda$-quadratic form* $q : V \to R/\Lambda$ by $q(v) = f(u, u) \mod \Lambda$. In other words, h is the form with the Gram matrix

$$\begin{pmatrix} 0 & p \\ \lambda p & 0 \end{pmatrix},$$

or, what is the same,

$$h(u, v) = f(u, v) + \lambda\overline{f(v, u)} = \overline{u}_1 v_{-1} + \dots + \overline{u}_n v_{-n} + \lambda\overline{u}_{-n} v_n + \dots + \lambda\overline{u}_{-1} v_1.$$

This form is in fact λ-hermitian, i.e. for any $u, v \in V$ one has

$$h(u,v)=\lambda\overline{h(v,u)}.$$

This fact will be used in the subsequent calculations without any reference. We will often omit h in the expression $h(u,v)$ and write simply (u,v) for the inner product of u, v with respect to the form h.

In turn, q is defined as follows:

$$q(u)=\overline{u}_1u_{-1}+\ldots+\overline{u}_nu_{-n} \mod \Lambda.$$

We refer to the module V equipped with the λ-hermitian form h and the Λ-quadratic form q as the hyperbolic Λ-quadratic module of rank $2n$ over R (with respect to the symmetry λ and the form parameter Λ).

The following easy fact immediately follows from the definition of h and q. It is crucial for our calculations in § 3.

Lemma 1.1. *For any $u, v \in V$ one has*

$$q(u+v)-q(u)-q(v)=h(u,v)+\Lambda.$$

Proof. In fact,

$$\begin{aligned} f(u+v,u+v)-f(u,u)-f(v,v)&=f(u,v)+f(v,u)=\\ &=\big(f(u,v)+\lambda\overline{f(v,u)}\big)+\big(f(v,u)-\lambda\overline{f(v,u)}\big), \end{aligned}$$

where the first summand is equal to $h(u,v)$, whereas the second one belongs to Λ.

2^0. Hyperbolic unitary groups. As usual we denote by $\mathrm{GL}(n,R)$ the group of all two-sided invertible matrices of degree n with entries from R. For a matrix $g \in GL(n,R)$ we denote by g^{-1} its inverse. For any matrix $g \in \mathrm{M}(n,R)$ we denote by g^* the "hermitian transpose" of g, i.e. the matrix which has $\overline{g}_{ji}$ in the position (i,j). Clearly the map $g \mapsto g^*$ is an anti-automorphism of $\mathrm{GL}(n,R)$, i.e. $(xy)^*=y^*x^*$ for any $x, y \in \mathrm{GL}(n,R)$. We consider the set of Λ-antihermitian matrices $\mathrm{AH}(n,R,\Lambda)$ which is defined as follows:

$$\mathrm{AH}(n,R,\Lambda)=\{a\in \mathrm{M}(n,R)\mid a=-\lambda a^* \text{ and } a_{ii}\in\Lambda \text{ for all } i=1,\ldots,n\}.$$

Now we define our principal object of study. Let $\mathrm{U}(2n,R,\Lambda)$ be the group consisting of all elements from $\mathrm{GL}(V)\cong \mathrm{GL}(2n,R)$ which preserve the λ-hermitian form h and the Λ-quadratic form q. In other words $g \in \mathrm{GL}(2n,R)$

belongs to $\mathrm{U}(2n, R, \Lambda)$ if and only if $h(gu, gv) = h(u, v)$ and $q(gu) = q(u)$ for all $u, v \in V$. The following result is a version of the Lemma 3.1 from [Ba3]. Write a matrix g of degree $2n$ in the block form with respect to the partition (n, n):

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

where a, b, c, d are matrices of degree n over R.

Lemma 1.2. *A necessary and sufficient condition for a matrix $g \in \mathrm{M}(2n, R)$ to belong to $\mathrm{U}(2n, R, \Lambda)$ is that*

1) $g^{-1} = \begin{pmatrix} pd^*p & \overline{\lambda}pb^*p \\ \lambda pc^*p & pa^*p \end{pmatrix}$,

2) $a^*pc,\ b^*pd \in \mathrm{AH}(n, R, \Lambda)$.

Proof. To belong to $\mathrm{U}(2n, R, \Lambda)$ a matrix g should preserve the forms h and q. the condition that g preserves h means that $h(gu, gv) = h(u, v)$. Since this holds for arbitrary $u, v \in V$, one should have $g^*hg = h$, or, in other words,

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^* \begin{pmatrix} 0 & p \\ \lambda p & 0 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 0 & p \\ \lambda p & 0 \end{pmatrix}.$$

This means that

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \begin{pmatrix} 0 & \overline{\lambda}p \\ p & 0 \end{pmatrix} \begin{pmatrix} a^* & c^* \\ b^* & d^* \end{pmatrix} \begin{pmatrix} 0 & p \\ \lambda p & 0 \end{pmatrix}.$$

Multiplying the factors on the right hand side we get condition in 1).

Now suppose that g already stabilizes h. What does it mean that g preserves q? Clearly g preserves q if and only if the Λ-quadratic form associated with the sesquilinear form $f(gu, gv) - f(u, v)$ is zero, or, what is the same,

$$z = \begin{pmatrix} a & b \\ c & d \end{pmatrix}^* \begin{pmatrix} 0 & p \\ 0 & 0 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} - \begin{pmatrix} 0 & p \\ 0 & 0 \end{pmatrix} =$$

$$= \begin{pmatrix} a^*pc & a^*pd - p \\ b^*pc & b^*pd \end{pmatrix} \in \mathrm{AH}(2n, R, \Lambda).$$

Now the condition in 1) implies that this last matrix is $\Lambda_{\max}$-antihermitian, i.e. belongs to $\mathrm{AH}(2n, R, \Lambda_{\max})$. Indeed, the equality

$$\begin{pmatrix} pd^*p & \overline{\lambda}pb^*p \\ \lambda pc^*p & pa^*p \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} e & 0 \\ 0 & e \end{pmatrix}$$

implies that

$$d^*pa + \overline{\lambda}b^*pc = p,$$

$$d^*pb + \overline{\lambda}b^*pd = 0,$$

$$\lambda c^*pa + a^*pc = 0,$$

$$\lambda c^*pb + a^*pd = p.$$

The second and third of these last equalities amount precisely to the fact that a^*pc and b^*pd belong to $\mathrm{AH}(n, R, \Lambda_{\max})$, while the last one of them says that $a^*pd - p = -\lambda\overline{b^*pc}$. In other words $z \in \mathrm{AH}(2n, R, \Lambda_{\max})$.

To belong to $\mathrm{AH}(2n, R, \Lambda)$ the matrix z has to satisfy additional condition that $z_{ii} \in \Lambda$ for all $i = 1, \dots, -1$. In view of the preceding paragraph this condition amounts precisely to saying that a^*pc, $b^*pd \in \mathrm{AH}(n, R, \Lambda)$.

The proof shows that the second condition in the lemma may be replaced by the following condition: the diagonal coefficients of the matrices a^*pc and b^*pd lie in Λ. In fact the above lemma may be stated in an equivalent but slightly different form.

Lemma 1.3. *A necessary and sufficient condition for a matrix $g \in \mathrm{M}(2n, R)$ to belong to the $\mathrm{U}(2n, R, \Lambda)$ is that*

1) $g'_{ij} = \lambda^{(\varepsilon(j)-\varepsilon(i))/2}\overline{g}_{-j,-i}$ *for all $i, j = 1, \dots, -1$,*
2) $\sum \overline{g}_{ij}g_{-ij} \in \Lambda$, $1 \leq i \leq n$, *for all $j = 1, \dots, -1$.*

Clearly in view of 1) condition 2) may be replaced by an analogous condition imposed on rows rather than on columns.

3^0. Polarity map. Sometimes it is convenient to express conditions from Lemmas 1.2 and 1.3 in terms of columns.

A vector $u \in V$ is called *isotropic*, if $q(u) = 0$, or, in other words, if $f(u, u) \in \Lambda$. Obviously, for an isotropic vector one has $h(u, u) = 0$. Vectors $u, v \in V$ are called *orthogonal* if they are orthogonal with respect to h, i.e. if $(u, v) = 0$. Since the form h is λ-hermitian, it is reflexive and thus the orthogonality relation is symmetric.

The definition of the unitary group implies that any column u of a matrix $g \in \mathrm{U}(2n, R, \Lambda)$ is isotropic. Let u and v be the i-th and the j-th columns of a matrix $g \in \mathrm{U}(2n, R, \Lambda)$ respectively, where $i \neq -j$. Then u and v are orthogonal.

Lemma 1.4. *If $v = (v_1, \dots, v_{-1})^t$, where $v_i \in R$, is the i-th column of a matrix g from $\mathrm{U}(2n, R, \Lambda)$, then the $(-i)$-th row $\widetilde{v}$ of the inverse matrix g^{-1}*

is expressed as follows:

$$\widetilde{v} = \begin{cases} (\lambda\overline{v}_{-1}, \dots, \lambda\overline{v}_{-n}, \overline{v}_n, \dots, \overline{v}_1), & \text{if } i = 1, \dots, n, \\ (\overline{v}_{-1}, \dots, \overline{v}_{-n}, \overline{\lambda}\overline{v}_n, \dots, \overline{\lambda}\overline{v}_1), & \text{if } i = -n, \dots, -1. \end{cases}$$

Proof. Calculate explicitly the $(-i)$-th row of the matrix on the right hand side of the formula in 1) of the previous lemma.

The formulae for $\widetilde{v}$ in this lemma differ only by a scalar factor (the second one of them is obtained from the first one by multiplication by $\overline{\lambda}$). Thus we can define the 'polarity' map from R^{2n} to the free left module of rank $2n$. Following [Co2] we denote this module by ${}^{2n}R$. We can identify ${}^{2n}R$ with the module consisting of all rows of length $2n$ with components from R. Then the polarity map $\widetilde{\ }: R^{2n} \longrightarrow {}^{2n}R$ is defined as follows: for a column $u = (u_1, \dots, u_n, u_{-n} \dots, u_{-1})^t \in R^{2n}$ we define the row $\widetilde{u} \in {}^{2n}R$ as

$$\widetilde{u} = (\lambda\overline{u}_{-1}, \dots, \lambda\overline{u}_{-n}, \overline{u}_n, \dots, \overline{u}_1).$$

In other words if $u = (x, y)^t$, where x and y are columns of height n, then $\widetilde{u} = (\lambda\overline{y}^t p, \overline{x}^t p)$. Clearly one has $h(u, v) = \widetilde{u}v$.

It is clear that the polarity map is *involutory linear*, i.e.

$$\widetilde{u+v} = \widetilde{u} + \widetilde{v}, \qquad \widetilde{u\xi} = \overline{\xi}\widetilde{u},$$

for all $u, v \in V$, $\xi \in R$. This property is often use in the sequel without explicit reference. Now we can reformulate Lemma 1.3 in yet another form. Let e^i be the dual base of the module ${}^{2n}R$. We may think of e^i as the row of length $2n$ whose i-th coordinate is 1, while all other coordinates are zeros.

Lemma 1.5. *For any vector $v \in V$ and any $g \in U(2n, R, \Lambda)$ one has $\widetilde{gv} = \widetilde{u}g^{-1}$.*

Proof. Since the polarity map is involutory linear, it is sufficient to prove the lemma only for the base vectors $v = e_i$. By definition $\widetilde{\ }$ maps e_i to e^{-i} if $i = 1, \dots, n$ and to λe^{-i} if $i = -n, \dots, -1$. Now ge_i is precisely the i-th column of g, while $\widetilde{e^{-i}}g^{-1}$ is the $(-i)$-th row of g^{-1} multiplied by λ if $i = -n, \dots, -1$. It remains to apply the previous lemma.

§ 2 Elementary hyperbolic unitary group

In this section we recall definition of the elementary unitary group.

1^0. Elementary unitary transvections. We consider the two following types of transformations in $\mathrm{U}(2n, R, \Lambda)$ which we call *elementary unitary* *transvections*. Denote the set $\{1,\dots,n,-n,\dots,-1\}$ of indices by Ω. Then $\Omega=\Omega^+\cup\Omega^-$, where $\Omega^+=\{1,\dots,n\}$ and $\Omega^-=\{-n,\dots,-1\}$. For an element $i\in\Omega$ we denote by $\varepsilon(i)$ the sign of Ω, i.e.

$$\varepsilon(i)=\begin{cases}1, & \text{if } i\in\Omega^+,\\ -1, & \text{if } i\in\Omega^-.\end{cases}$$

The transvections $T_{ij}(\xi)$ correspond to the pairs $i,j\in\Omega$ such that $i\neq j$. If moreover $i\neq -j$, the for any $\xi\in R$ we set

$$T_{ij}(\xi)=e+\xi e_{ij}-\lambda^{(\varepsilon(j)-\varepsilon(i))/2}\overline{\xi}e_{-j,-i}.$$

We will refer to these elements as the "elementary short root elements".

On the other side for $j=-i$ and $\alpha\in\lambda^{-(\varepsilon(i)+1)/2}\Lambda$ we set

$$T_{i,-i}(\alpha)=e+\alpha e_{i,-i}.$$

We will refer to these elements as the "elementary long root elements". Note that $\overline{\Lambda}=\overline{\lambda}\Lambda$. In fact, for any element $\alpha\in\Lambda$ one has $\overline{\alpha}=-\overline{\lambda}\alpha$ and thus $\overline{\Lambda}$ coinicdes with the set of products $\overline{\lambda}\alpha$, $\alpha\in\Lambda$. This means that in the definition above $\alpha\in\overline{\Lambda}$ when $i\in\Omega^+$ and $\alpha\in\Lambda$ when $i\in\Omega^-$.

A straightforward calculation shows that these elements actually do belong to $\mathrm{U}(2n,R,\Lambda)$. Now we describe these matrices explicitly, depending on the signs of i and j.

First of all, if the signs of i and j coincide, then the power of λ which appears in the definition of $T_{ij}(\xi)$ is 1. Thus, the corresponding transvections have the shape

$$e+\xi e_{ij}-\overline{\xi}e_{-j,-i}.$$

They are in the image of the hyperbolic embedding of the general linear group $\mathrm{GL}(n,R)$ in the unitary group $\mathrm{U}(2n,R,\Lambda)$. More precisely for any $a\in\mathrm{GL}(n,R)$ we set

$$H(a)=\begin{pmatrix}a & 0\\ 0 & p(a^*)^{-1}p\end{pmatrix}.$$

A straightforward calculation shows that $H(a)\in\mathrm{U}(2n,R,\Lambda)$. In these notation the transvection $T_{ij}(\xi)$, $i,j\in\Omega^+$, above is just the image $H(t_{ij}(\xi))$ of the ordinary linear transvection $t_{ij}(\xi)=e+\xi e_{ij}$ under the hyperbolic embedding. It is clear that $T_{ij}(\xi)=T_{-j,-i}(-\overline{\xi})$ when $i,j\in\Omega^-$ gives the same set of transvections.

Let, next, $i\in\Omega^+$ and $j\in\Omega^-$. If $i\neq -j$ then the corresponding transvection has the shape

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

Clearly $T_{-j,-i}(\xi) = T_{ij}(-\overline{\lambda}\,\overline{\xi})$. These transvections may be considered together with the transvections $T_{i,-i}(\alpha)$, $i \in \Omega^+$, as follows. In this case $T_{i,-i}(\alpha)$ may be viewed as the usual linear transvection $t_{i,-i}(\alpha) = e + \alpha e_{i,-i}$, where α runs over $\overline{\Lambda}$. The transvections of the two types above come from the unipotent embedding of the (additive) group $\mathrm{AH}(n, R, \overline{\Lambda})$ of $\overline{\Lambda}$-antihermitian matrices into $\mathrm{U}(2n, R, \Lambda)$. This embedding is defined as follows. For $b \in \mathrm{AH}(n, R, \overline{\Lambda})$ set

$$X^+(b) = \begin{pmatrix} e & bp \\ 0 & e \end{pmatrix}.$$

A direct calculation shows that $X^+(b) \in \mathrm{U}(2n, R, \Lambda)$. Clearly both types of the transvections above are in the image of X^+. Namely, they are the images of the matrices $\xi e_{i,-j} - \overline{\lambda}\,\overline{\xi} e_{-j,i}$, $\xi \in R$, and αe_{ii}, $\alpha \in \overline{\Lambda}$, respectively.

Let, finally, $i \in \Omega^-$ and $j \in \Omega^+$. If $i \neq -j$ then the corresponding transvection has the shape

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

Clearly $T_{-j,-i}(\xi) = T_{ij}(-\lambda\overline{\xi})$. These transvections may be considered together with the transvections $T_{i,-i}(\alpha)$, $i \in \Omega^-$, as follows. In this case $T_{i,-i}(\alpha)$ is the usual linear transvection $t_{i,-i}(\alpha) = e + \alpha e_{i,-i}$, where α runs over Λ. The transvections of the two types above come from the unipotent embedding of the (additive) group $\mathrm{AH}(n, R, \Lambda)$ of Λ-antihermitian matrices into $\mathrm{U}(2n, R, \Lambda)$. This embedding is defined as follows. For $c \in \mathrm{AH}(n, R, \Lambda)$ set

$$X^-(c) = \begin{pmatrix} e & 0 \\ pc & e \end{pmatrix}.$$

A direct calculation shows that $X^-(b) \in \mathrm{U}(2n, R, \Lambda)$. Clearly both types of the transvections above are in the image of X^-. Namely, they are the images of the matrices $\xi e_{-i,j} - \lambda\overline{\xi} e_{j,-i}$, $\xi \in R$, and $\alpha e_{-i,-i}$, $\alpha \in \Lambda$, respectively.

2^0. Elementary relations. In this subsection we list the obvious relations among the elementary unitary transvections.

It immediately follows from the definition, that

(R1) $$T_{ij}(\xi) = T_{-j,-i}(\lambda^{(\varepsilon(j)-\varepsilon(i))/2}\overline{\xi}).$$

The maps $T_{ij} : R^+ \to \mathrm{U}(2n, R, \Lambda)$ are additive, i.e.

(R2) $$T_{ij}(\xi)T_{ij}(\zeta) = T_{ij}(\xi + \zeta).$$

Recall that for any two elements x, y of a group G one denotes by $[x, y] = xyx^{-1}y^{-1}$ their commutator. For $h \neq j, -i$, $k \neq i, -j$, one has

(R3) $$[T_{ij}(\xi), T_{hk}(\zeta)] = e.$$

Finally we have three remaining elementary relations, corresponding to the case of two short roots whose sum is a short root, two short roots, whose sum is a long root, and short and long roots, whose sum is a root respectively.

First we consider the case of two short roots whose sum is a short root. Let $i, h \neq \pm j$, $i \neq \pm h$. Then

$$[T_{ij}(\xi), T_{jh}(\zeta)] = T_{ih}(\xi\zeta). \tag{R4}$$

Applying relation (R1) to one or both transvections on the right hand side of this relation, we can get three equivalent forms of this relation:

$$[T_{ij}(\xi), T_{h,-j}(\zeta)] = T_{i,-h}(-\lambda^{(\varepsilon(-j)-\varepsilon(h))/2}(\xi\overline{\zeta}),$$
$$[T_{-j,i}(\xi), T_{jh}(\zeta)] = T_{-i,h}(-\lambda^{(\varepsilon(i)-\varepsilon(-j))/2}\overline{\xi}\zeta),$$
$$[T_{ji}(\xi), T_{hj}(\zeta)] = T_{hi}(-\zeta\xi).$$

The last of these relations is also obtained from (R4) by applying the following obvious formula: $[x, y]^{-1} = [y, x]$. In what follows we refer to these relation also as (R4)

Now we consider the case of two short roots whose sum is a long root. Let $i \neq \pm j$. Then

$$[T_{ij}(\xi), T_{j,-i}(\zeta)] = T_{i,-i}(\xi\zeta - \lambda^{-\varepsilon(i)}\overline{\zeta}\,\overline{\xi}). \tag{R5}$$

As before we will use the following versions of this relation, also referring to them as (R5):

$$[T_{ij}(\xi), T_{i,-j}(\zeta)] = T_{i,-i}(\lambda^{(\varepsilon(-j)-\varepsilon(i))/2}\xi\overline{\zeta} - \lambda^{(\varepsilon(j)-\varepsilon(i))/2}\zeta\overline{\xi}),$$
$$[T_{-j,i}(\xi), T_{ji}(\zeta)] = T_{-i,i}(-\lambda^{(\varepsilon(i)-\varepsilon(-j))/2}\overline{\xi}\zeta + \lambda^{(\varepsilon(i)-\varepsilon(j))/2}\overline{\zeta}\xi),$$
$$[T_{ji}(\xi), T_{-i,j}(\zeta)] = T_{-i,i}(-\zeta\xi + \lambda^{\varepsilon(i)}\overline{\xi}\,\overline{\zeta}).$$

Finally, we turn to the case of a long and a short root. If $i \neq \pm j$, then

$$[T_{i,-i}(\alpha), T_{-i,j}(\xi)] = T_{ij}(\alpha\xi)T_{-j,j}(-\lambda^{(\varepsilon(j)-\varepsilon(-i))/2}\overline{\xi}\alpha\xi). \tag{R6}$$

This equation may be given also the following alternative forms, to which we also refer as (R6):

$$[T_{i,-i}(\alpha), T_{-j,i}(\xi)] = T_{ij}(-\lambda^{(\varepsilon(i)-\varepsilon(-j))/2}\alpha\overline{\xi})T_{-j,j}(\lambda^{(\varepsilon(i)-\varepsilon(-j))/2}\xi\alpha\overline{\xi}),$$
$$[T_{ji}(\xi), T_{i,-i}(\alpha)] = T_{j,-i}(\xi\alpha)T_{j,-j}(\lambda^{(\varepsilon(i)-\varepsilon(j))/2}\xi\alpha\overline{\xi}).$$

The relations (R1) – (R6) define the unitary Steinberg group. Thus the calculations where only these relations are used hold in fact already for the Steinberg group and not just for the elementary unitary group.

3^0. Elementary unitary group. In this subsection we introduce the subgroup of $\mathrm{U}(2n,R,\Lambda)$ which will be the main object of our study in the rest of the paper. Namely, the subgroup generated by all elementary unitary transvections $T_{ij}(\xi)$, $i\neq\pm j$, $\xi\in R$, and $T_{i,-i}(\alpha)$, $\alpha\in\Lambda$, is called the *elementary unitary group* and is denoted by $\mathrm{EU}(2n,R,\Lambda)$.

Lemma 2.1. *For any $b\in\mathrm{AH}(n,R,\overline{\Lambda})$ and any $c\in\mathrm{AH}(n,R,\Lambda)$ one has $X^+(b),X^-(c)\in\mathrm{EU}(2n,R,\Lambda)$.*

Proof. Indeed, one has

$$X^+(b)=\prod T_{i,-j}(b_{ij}),\qquad X^-(c)=\prod T_{-i,j}(c_{ij}),$$

where the first product is taken by $i,j\in\Omega^+$ such that $i\leq j$, while the second one is taken by $i,j\in\Omega^+$ such that $i\geq j$.

As usual $\mathrm{E}(n,R)$ denotes the elementary subgroup of the general linear group $\mathrm{GL}(n,R)$, i.e. the subgroup generated by all elementary transvections $t_{ij}(\xi)$, $1\leq i\neq j\leq n$, $\xi\in R$. Then by definition the image $H(g)$ of any matrix $g\in\mathrm{E}(n,R)$ under hyperbolic embedding lies in $\mathrm{EU}(2n,R,\Lambda)$. Actually, unless $n=2$ it lies already in the subgroup of $\mathrm{EU}(2n,R,\Lambda)$, generated by the images of X^+ and X^-. The following result is essentially proposition 5.1 of [Ba3].

Lemma 2.2. *Suppose that either $n\neq 2$, or $R=\Lambda R+R\Lambda$. Then*

$$\mathrm{EU}(2n,R,\Lambda)=\big\langle X^+(b),b\in\mathrm{AH}(n,R,\overline{\Lambda});X^-(c),c\in\mathrm{AH}(n,R,\Lambda)\big\rangle.$$

Proof. Denote the right hand side of this formula by G. We have to prove that $T_{ij}(\xi)$ belongs to G for any $i,j\in\Omega^+$, $i\neq j$, and any $\xi\in R$. If $n=1$, there is nothing to prove, thus we may assume that $n\geq 2$. If $n\geq 3$, we may choose $h\in\Omega^+$, $h\neq i,j$. Then relation (R4) implies that $T_{ij}(\xi)=[T_{i,-h}(\xi),T_{-h,j}(1)]$, where the factors on the right hand side belong to G. Assume finally that $n=2$. Then relation (R6) implies that

$$T_{ij}(\xi\alpha)=[T_{i,-j}(\xi),T_{-j,j}(\alpha)]T_{i,-i}(-\lambda^{(\varepsilon(-j)-\varepsilon(i))/2}\xi\alpha\overline{\xi})\in G,$$

By the same relation

$$T_{ij}(\alpha\xi)=[T_{-i,j}(\xi),T_{i,-i}(\alpha)]T_{-j,j}(-\lambda^{(\varepsilon(j)-\varepsilon(-i))/2}\overline{\xi}\alpha\xi)\in G.$$

These two inclusions show that $T_{ij}(\Lambda R+R\Lambda)\subseteq G$.

By definition $\mathrm{EU}(2n,R,\Lambda)$ is generated by both the long and the short root unipotents. However under some assumptions on Λ only the short root unipotents *or* the long root unipotents would suffice.

Lemma 2.3. *Suppose $\Lambda = \Lambda_{\min}$ and $n \geq 2$. Then the group* $\mathrm{EU}(2n, R, \Lambda)$ *is generated by elementary short root unipotents.*

Proof. Let $T_{i,-i}(\alpha)$ be an elementary long root unipotent. Since $\Lambda = \Lambda_{\min}$, there exists $\xi \in R$ such that $\alpha = \xi - \lambda^{-\varepsilon(i)}\overline{\xi}$. Pick up an index $j \neq \pm i$. Then by the relation (R5) one has

$$T_{i,-i}(\alpha) = [T_{ij}(\xi), T_{-j,i}(1)].$$

Of course the group $\mathrm{EU}(2n, R, \Lambda)$ for $n \geq 2$ is never generated by *elementary* long root unipotents. But under certain assumptions it might be generated by their conjugates. Recall that a conjugate of an elementary long/short root element is called a *long/short root element* or a *long/short root unipotent*.

Lemma 2.4. *Assume that $\Lambda R + R\Lambda = R$. Then the group* $\mathrm{EU}(2n, R, \Lambda)$ *is generated by long root unipotents.*

Proof. Take any $i \neq \pm j$, $\xi \in R$ and $\alpha \in \Lambda$. Then one can rewrite formulae appearing in the proof of Lemma 2.2 as follows:

$$T_{ij}(\xi\alpha) = \big(T_{i,-j}(\xi)T_{-j,j}(\alpha)T_{i,-j}(-\xi)\big)T_{-j,j}(-\alpha)T_{i,-i}(-\lambda^{(\varepsilon(-j)-\varepsilon(i))/2}\xi\alpha\overline{\xi}),$$

$$T_{ij}(\alpha\xi) = \big(T_{-i,j}(\xi)T_{i,-i}(\alpha)T_{-i,j}(-\xi)\big)T_{i,-i}(-\alpha)T_{j,-j}(-\lambda^{(\varepsilon(j)-\varepsilon(-i))/2}\overline{\xi}\alpha\xi),$$

where all factors on the right hand side are long root elements. This proves the lemma.

§ 3 Congruence groups

In this section we recall definitions of the form ideals and corresponding relative groups, see [B1] – [B3], [Ba3], which play crucial role in the description of the normal subgroups in unitary groups.

1^0. Form ideals. Let (R, Λ) be a form ring. Let I be a (two-sided) ideal in R invariant with respect to the involution, i.e. such that $\overline{I} = I$. Set $\Gamma_{\max} = I \cap \Lambda$ and

$$\Gamma_{\min} = \left\{\xi - \lambda\overline{\xi} \mid \xi \in I\right\} + \left\{\xi\alpha\overline{\xi} \mid \xi \in I,\ \alpha \in \Lambda\right\}.$$

By definition $\Gamma_{\min}$ and $\Gamma_{\max}$ depend both on the *absolute* form parameter Λ *and* an ideal I in R. The form parameter Λ is fixed and will not be accounted in the notation. Sometimes it is necessary to stress the dependence of $\Gamma_{\min}$ and $\Gamma_{\max}$ on I. In such cases we write $\Gamma_{\min}(I)$ and $\Gamma_{\max}(I)$.

A *relative form parameter* Γ in (R,Λ) *of level* I is an additive subgroup of I such that

1) $\Gamma_{\min}\subseteq\Gamma\subseteq\Gamma_{\max}$,
2) $\alpha\Gamma\overline{\alpha}\subseteq\Gamma$ for all $\alpha\in R$.

Again when we deal with various ideals we write $\Gamma(I)$ to indicate that Γ is a form parameter of level I.

A *form ideal* in (R,Λ) is a pair (I,Γ), where I is an involution invariant ideal in R and Γ is a relative form parameter of level I.

The form ideals play the same role in the category of form rings as (two-sided) ideals in the category of rings. This notion of a form ideal was introduced by A.Bak [B1] – [B3]. It is crucial for the description of normal subgroups of $\mathrm{U}(2n,R,\Lambda)$, $n\geq 3$. It was used by H.Bass [Ba3] under the name of a *unitary ideal*. For commutative rings with trivial involution notion of *special submodule* essentially equivalent to the notion of a relative form parameter was introduced independently by E.Abe [A1],[A3], [A4], [AS]. In the book [HO'M] of A.Hahn and O.T.O'Meara they are called simply *ideals* of form rings. L.N.Vaserstein [V7] christened form ideals 'quasi-ideals', which does not seem a good choice, since this term is overcharged already. Finally in a very important recent paper [CK] D.Costa and G.Keller introduce the notion of a *radix* and discuss the interrelations of form ideals, Jordan ideals of $\mathrm{M}(n,R)$ with transpose as the involution (R is commutative) and radices. It turns out that for $n\geq 3$ form ideals, Jordan ideals and radices coincide and correspond to Abe's special submodules. However for $n=2$ not all radices are form ideals. This is precisely why it is much more difficult to describe the normal subgroups of $\mathrm{U}(4,R,\Lambda)$ than the normal subgroups of $\mathrm{U}(2n,R,\Lambda)$, $n\geq 3$.

2^0. Doubling a form ring. To treat the relative groups corresponding to the form ideals we have to recall some notation related to Stein's relativization [St].

Let first I be any ideal of a ring R. Then we can define the *double* $R\times_I R$ of a ring R *along an ideal* $I\leq R$ by the Cartesian square

$$\begin{array}{ccc} R\times_I R & \xrightarrow{\pi_1} & R \\ {\scriptstyle\pi_2}\big\downarrow & & \big\downarrow{\scriptstyle\pi} \\ R & \xrightarrow[\pi]{} & R/I \end{array}$$

In a more down to earth language $R\times_I R$ consists of all pairs $(a,b)\in R\times R$ such that $a\equiv b\pmod I$,

$$R\times_I R=\big\{(a,b)\in R\times R\mid a-b\in I\big\},$$

with the component-wise operations and $\pi_1(a,b)=a$, $\pi_2(a,b)=b$. Clearly $\operatorname{Ker}\pi_1=(0,I)$ and $\operatorname{Ker}\pi_2=(I,0)$. The diagonal embedding $\delta: R\to R\times_I R$ given by $\delta(a)=(a,a)$ splits both π_1 and π_2.

One can define embeddings $\iota_1,\iota_2: I\to R\times_I R$ by $\iota_1(c)=(c,0)$ and $\iota_2(c)=(0,c)$ respectively. Clearly the image of ι_2 coincides with the kernel of π_1. In other words, the sequence

$$1\to I\xrightarrow{\iota_2} R\times_I R\xrightarrow{\pi_1} R\to 1$$

is exact. Moreover, it is split by δ. We can rewrite the above in a slightly different form.

For a pair (R,I) one can define semidirect product $R\ltimes I$ of R and I as the set of pairs (a,c), $a\in R$, $c\in I$, with the component-wise addition and the multiplication given by the following formula: $(a,c)(b,d)=(ab,ad+cb+cd)$.

Lemma 3.1. *The ring $R\times_I R$ is isomorphic to the semidirect product $R\ltimes I$ of $\delta(R)\cong R$ and $\operatorname{Ker}\pi_1\cong I$.*

Proof. Define a map from $R\times_I R$ to $R\ltimes I$ by $(a,b)\mapsto(a,b-a)$. The definition of multiplication in $R\ltimes I$ implies that this is a homomorphism. The inverse homomorphism is defined by $(a,c)\mapsto(a,a+c)$.

Of course if we write the Cartesian squares above in terms of semidirect products, rather then doubles, we must define π_1, π_2 and δ by $\pi_1(a,c)=a$, $\pi_2(a,c)=a+c$ and $\delta(a)=(a,0)$.

In the sequel we identify I with $\operatorname{Im}\iota_2=\operatorname{Ker}\pi_1$. Usually the relative questions for the ideal I in R can be reduced to the absolute questions for the ring $R\times_I R$.

Let now (I,Γ) be a form ideal in a form ring (R,Λ). Then we can define the double of Λ along Γ by exactly the same formula, as above:

$$\Lambda\times_\Gamma\Lambda=\big\{(a,c)\in\Lambda\times\Lambda\mid a-c\in\Gamma\big\}.$$

It is easy to observe that $\Lambda\times_\Gamma\Lambda$ is a form parameter in $R\times_I R$ (see [HO'M], Lemma 5.2.15).

Lemma 3.2. *$(R\ltimes I,\Lambda\ltimes\Gamma)$ is a form ring with respect to the component-wise involution and $\lambda=(\lambda,0)$.*

Another form ring which can be associated with this form ideal is the factor ring $(R/I,\Lambda/\Gamma_{\max})$ (see [HO'M], Lemma 5.2.12). Then we have a *commutative* square of form rings:

$$\begin{array}{ccc}(R\times_I R,\Lambda\times_\Gamma\Lambda) & \xrightarrow{\pi_1} & (R,\Lambda)\\ {\scriptstyle\pi_2}\big\downarrow & & \big\downarrow{\scriptstyle\pi}\\ (R,\Lambda) & \xrightarrow[\pi]{} & (R/I,\Lambda/\Gamma_{\max})\end{array}$$

analogous to the Cartesian square above. This commutative square is actually *Cartesian* when $\Gamma = \Gamma_{\max}$. Since the functor U_{2n} from rings to groups commutes with limits the commutative/Cartesian squares of form rings above lead to commutative/Cartesian squares of groups U_{2n}. This will be amply used in the rest of this section and the next one.

3⁰. Principal congruence subgroups. Let (I,Γ) be a form ideal of (R,Λ). The *principal congruence subgroup* $\mathrm{U}(2n,I,\Gamma)$ of level (I,Γ) in $\mathrm{U}(2n,R,\Lambda)$ consists of those

$$g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{U}(2n,R,\Lambda),$$

which are congruent to $e = e_{2n}$ modulo I and preserve $f(u,u)$ modulo Γ:

$$f(gu,gu) \in f(u,u) + \Gamma, \qquad u \in V.$$

One can give the following characterization of $\mathrm{U}(2n,I,\Gamma)$ analogous to Lemma 1.2.

Lemma 3.3. *Let (I,Γ) be a form ideal in (R,Λ). A necessary and sufficient condition for a matrix $g \in \mathrm{U}(2n,R,\Lambda)$ to belong to $\mathrm{U}(2n,I,\Gamma)$ is that*

1) $g \equiv e \pmod{I}$,
2) $a^*pc, b^*pd \in \mathrm{AH}(n,R,\Gamma)$.

Proof. Replace Λ with Γ in the second part of the proof of Lemma 1.2.

Clearly $\iota_2 : I \to R \ltimes I$, $\xi \mapsto (0,\xi)$, defines an embedding of the corresponding groups. In fact it is easy to see (see [HO'M], Lemma 5.2.17) that we get the following split exact sequence of groups:

$$1 \to \mathrm{U}(n,I,\Gamma) \xrightarrow{\iota_2} \mathrm{U}(n,R\times_I R,\Lambda\times_\Gamma \Lambda) \underset{\delta}{\overset{\pi_1}{\rightleftarrows}} \mathrm{U}(n,R,\Lambda) \to 1.$$

This is precissely the Stein's definition of a relative group.

By definition $\mathrm{U}(2n,I,\Gamma)$ is a subgroup of $\mathrm{U}(2n,R,\Lambda)$. In fact it is a *normal* subgroup of $\mathrm{U}(2n,R,\Lambda)$. This was first shown by A.Bak in [B2], Theorem 4.1.4 by direct computation. However there is a more elegant approach based on a variant of Stein's relativization. Below we reproduce the argument (compare also [OM'H], § 5.2 and [Hb1], Lemma 2.6).

Lemma 3.4. *For any form ideal (I,Γ) in (R,Λ) the corresponding principal congruence group $\mathrm{U}(2n,I,\Gamma)$ is a normal subgroup of $\mathrm{U}(2n,R,\Lambda)$.*

Proof. By the previous subsection we have a commutative square of groups:

$$\begin{array}{ccc}
\mathrm{U}(2n, R\times_I R, \Lambda\times_\Gamma \Lambda) & \xrightarrow{\pi_1} & \mathrm{U}(2n,R,\Lambda) \\
{\scriptstyle\pi_2}\big\downarrow & & \big\downarrow{\scriptstyle\pi} \\
\mathrm{U}(2n,R,\Lambda) & \xrightarrow[\pi]{} & \mathrm{U}(2n,R/I,\Lambda/\Gamma_{\max})
\end{array}$$

Clearly, the kernel of π_1 is a normal subgroup of $\mathrm{U}(2n, R\times_I R, \Lambda\times_\Gamma \Lambda)$. Being split π_2 is surjective. This implies that the image of $\operatorname{Ker}\pi_1$ under π_2 is a normal subgroup of $\mathrm{U}(2n,R,\Lambda)$. But $\operatorname{Ker}\pi_1=\mathrm{U}(2n,\iota_2(I),i_2(\Gamma))$ and its image under π_2 coincides with $\mathrm{U}(2n,I,\Gamma)$.

4^0. Full congruence subgroups. Let first G be any group, F and H be subgroups of G. One can define

$$C_G(F,H)=\{g\in G \mid [g,f]\in H \text{ for all } f\in F\}\,.$$

In general $C_G(F,H)$ is just a subset of G. However if H is a *normal* subgroup of G, then the standard equalities for the commutators imply that $C_G(F,H)$ is a subgroup of G. We will use the notation $C_G(F,H)$ *only* when $H\trianglelefteq G$. If, moreover, F is also a normal subgroup of G, then $C_G(F,H)$ is a normal subgroup of G. The normal subgroup $C_G(G,H)$ will be denoted by $C_G(H)$, or, when the group G is clear from the context, simply by $C(H)$.

Return to the case when $G=\mathrm{U}(2n,R,\Lambda)$. Let further (I,Γ) be a form ideal in (R,Λ). We can define a normal subgroup $\mathrm{CU}(2n,I,\Gamma)$ of G as $C\big(\mathrm{U}(2n,I,\Gamma)\big)$. In other words,

$$\mathrm{CU}(2n,I,\Gamma)=\{g\in \mathrm{U}(2n,R,\Lambda) \mid [g,\mathrm{U}(2n,R,\Lambda)]\subseteq \mathrm{U}(2n,I,\Gamma)\}\,.$$

This group is called the *full congruence subgroup* in G of level (I,Γ). Although it is not reflected in the notation, this group depends not only on (I,Γ), but also on (R,Λ). This definition of the full congruence group follows [Ba2]. Often this group (denoted also by $\mathrm{U}'(2n,I,\Gamma)$ or by $\mathrm{U}^{\sim}(2n,I,\Gamma)$) is defined in a slightly different way. For example, [HO'M] define the full congruence subgroup setting

$$\mathrm{U}^{\sim}(2n,I,\Gamma)=\{g\in \mathrm{U}(2n,R,\Lambda) \mid [g,\mathrm{EU}(2n,R,\Lambda)]\subseteq \mathrm{EU}(2n,I,\Gamma)\}\,.$$

But in interesting situations these groups coincide, as we shall see in the next part of this work.

For the general linear group $\mathrm{GL}(n,R)$ the full congruence subgroup of level I is the full preimage of the centre of the group $\mathrm{GL}(n,R/I)$ under reduction homomorphism modulo I. This is also true in our case, when the relative form parameter is *maximal*. Namely $\mathrm{CU}(2n,I,\Gamma_{\max})$ is the full preimage of the centre of the group $U(2n,R/I,\Lambda/\Gamma_{\max})$ under the reduction homomorphism $U(2n,R,\Lambda)\longrightarrow U(2n,R/I,\Lambda/\Gamma_{\max})$.

§ 4 Relative elementary groups

In this section we define the relative elementary groups. We prove a standard result about their generators and show that the standard commutator formulae follow from the normality of the absolute elementary subgroup.

1^0. Relative elementary subgroups. An elementary unitary transvection $T_{ij}(\xi)$, $i \neq j$ is called *elementary of level* (I, Γ) if $\xi \in I$ and, moreover, $\xi \in \Gamma$ if $i = -j$. Denote the subgroup generated by all (I, Γ)-elementary transvections by $\mathrm{FU}(2n, I, \Gamma)$:

$$\mathrm{FU}(2n, I, \Gamma) = \big\langle T_{ij}(\xi) \mid i \neq j,\ \xi \in I \text{ with } \xi \in \Gamma \text{ if } i = -j \big\rangle.$$

By definition $\mathrm{FU}(2n, I, \Gamma)$ is a subgroup of the absolute elementary subgroup $\mathrm{EU}(2n, R, \Lambda)$. However the subgroup $\mathrm{FU}(2n, I, \Gamma)$ is *very* seldom normal in $\mathrm{EU}(2n, R, \Lambda)$. The *elementary subgroup* $\mathrm{EU}(2n, I, \Gamma)$ *of level* (I, Γ) is defined as the *normal closure* of $\mathrm{FU}(2n, I, \Gamma)$ in $\mathrm{EU}(2n, R, \Lambda)$:

$$\mathrm{EU}(2n, I, \Gamma) = \mathrm{FU}(2n, I, \Gamma)^{\mathrm{EU}(2n,R,\Lambda)}.$$

Although it is not reflected in the notation, the group $\mathrm{EU}(2n, I, \Gamma)$ depends also on R and Λ. Since the principal congruence subgroup $\mathrm{U}(2n, I, \Gamma)$ is normal in $\mathrm{U}(2n, R, \Lambda)$, and contains all elementary transvections of level (I, Γ), it follows that $\mathrm{EU}(2n, I, \Gamma) \leq \mathrm{U}(2n, I, \Gamma)$.

The following result is a relative analogue of Lemma 2.2.

Lemma 4.1. *Suppose that either $n \neq 2$, or $I = \Lambda I + I\Lambda$. Then*

$$\mathrm{EU}(2n, I, \Gamma) = \big\langle X^+(b), b \in \mathrm{AH}(n, I, \overline{\Gamma}); X^-(c), c \in \mathrm{AH}(n, I, \Gamma)\big\rangle^{\mathrm{EU}(2n,R,\Lambda)}.$$

Proof. For $n \neq 2$ the proof of Lemma 2.2 carries over without any changes. For $n = 2$ one takes $\xi \in I$, $\alpha \in \Lambda$ and observes that $\xi\alpha\overline{\xi} \in \Gamma_{\min} \leq \Gamma$. Now the same argument as in Lemma 2.2 shows that $T_{ij}(\Lambda I + I\Lambda)$ is contained in the right hand side for any $i, j \in \Omega^+$, $i \neq j$.

The following result does not hold for $n = 2$ without some additional assumptions on the ring R.

Lemma 4.2. *Suppose that either $n \neq 2$, or $I = \Lambda I + I\Lambda$. Then*

$$\mathrm{EU}(2n, I, \Gamma) = [\mathrm{EU}(2n, I, \Gamma), \mathrm{EU}(2n, R, \Lambda)].$$

Proof. By definition the right hand side is contained in the left hand side. Indeed, setting $\xi = 1$ in (R4) we get that $T_{ij}(\zeta)$, $i \neq \pm j$, $\zeta \in I$, belongs to the right hand side. Now (R6) (again with $\xi = 1$) implies that

$$T_{j,-j}(\alpha) = [T_{i,-i}(-\lambda^{(\varepsilon(-j)-\varepsilon(-i))/2}\alpha), T_{-i,-j}(1)]T_{ij}(\lambda^{(\varepsilon(-j)-\varepsilon(-i))/2}\alpha)$$

belongs to the right hand side for $\alpha \in \Gamma$ or $\alpha \in \overline{\Gamma}$, depending on the sign of j. Thus all the generators of $\mathrm{EU}(2n, I, \Gamma)$ as a normal subgroup of $\mathrm{EU}(2n, R, \Lambda)$ are contained in the right hand side. But it is itself normal in $\mathrm{EU}(2n, R, \Lambda)$.

The following obvious fact will be often used without reference.

Lemma 4.3. *Suppose that* $\mathrm{EU}(2n, I_1, \Gamma_1) \leq \mathrm{U}(2n, I_2, \Gamma_2)$. *Then* $I_1 \leq I_2$ *and* $\Gamma_1 \leq \Gamma_2$.

2^0. Generation of relative elementary subgroups. As we know from the preceding subsection, it is not true in general that $\mathrm{EU}(2n, I, \Gamma)$ is generated by *elementary* transvections of level (I, Γ). In fact, fix $i \neq j$ and consider matrices

$$Z_{ij}(\xi, \zeta) = T_{ji}(\zeta) T_{ij}(\xi) T_{ji}(-\zeta),$$

where $\xi \in I$, $\zeta \in R$ if $i \neq -j$ and $\xi \in \Gamma$, $\zeta \in \Lambda$ otherwise[5]. These matrices do not in general belong to $\mathrm{FU}(2n, I, \Gamma)$. However this is essentially the unique counter-example. The following result is a unitary version of a result by A.A.Suslin and L.N.Vaserstein [VS] (see also [V2], [B5], [Va6]).

Proposition 1. *Let* $n \geq 3$. *The for any form ideal* (I, Γ) *the corresponding relative elementary group* $\mathrm{EU}(2n, I, \Gamma)$ *is generated by all matrices* $Z_{ij}(\xi, \zeta) = T_{ji}(\zeta) T_{ij}(\xi) T_{ji}(-\zeta)$, *where either* $\xi \in I$, $\zeta \in R$ *and* $i \neq \pm j$, *or* $\xi \in \Gamma$, $\zeta \in \Lambda$, $i = -j$.

Proof. Since $Z_{ij}(\xi, 0) = T_{ij}(\xi)$ is a usual elementary unitary transvection, the subgroup generated by all $Z_{ij}(\xi, \zeta)$ contains $\mathrm{FU}(2n, I, \Gamma)$. By definition $\mathrm{EU}(2n, I, \Gamma)$ is generated by matrices ${}^{x}T_{ij}(\xi) = x T_{ij}(\xi) x^{-1}$, where $x \in \mathrm{EU}(2n, R, \Lambda)$, $\xi \in I$ and $i \neq j$. We shall proceed by induction on the length t of a shortest expression of x as a product of elementary matrices. When $t \leq 1$ there is nothing to prove. If $t \geq 2$, we can write x in the form $T_{hk}(\zeta) y$, where $y \in \mathrm{EU}(2n, R, \Lambda)$, $\zeta \in R$ and $h \neq k$. Now we may apply the formula ${}^{ab}c = {}^{a}[b, c]\,{}^{a}c$ valid for any three elements a, b, c of a group. Thus

$${}^{T_{hk}(\zeta) y} T_{ij}(\xi) = {}^{T_{hk}(\zeta)}[y, T_{ij}(\xi)]\, {}^{T_{hk}(\zeta)} T_{ij}(\xi).$$

If $(h, k) \neq (j, i)$ we can apply (R3) – (R6) and the definition of a relative form parameter to conclude that $z = {}^{T_{hk}(\zeta)} T_{ij}(\xi)$ belongs to $\mathrm{FU}(2n, I, \Gamma)$, and if $(h, k) = (j, i)$, then $z = Z_{ij}(\xi, \zeta)$. On the other hand, since y is shorter, than x, the commutator $[y, T_{ij}(\xi)]$ is a product of factors of the form $Z_{lm}(\omega, \vartheta)$, $\omega \in I$, $\vartheta \in R$, $l \neq m$. Now for $w = {}^{T_{hk}(\zeta)} Z_{lm}(\omega, \vartheta)$ we have three options:

---

[5]Starting from this point we usually do not destinguish between the cases $i \neq -j$ and $i = -j$. We simply write $\xi \in I$ or $\zeta \in R$ assuming that automatically $\xi \in \Gamma$ and $\zeta \in \Lambda$ when $i = -j$.

If $(h,k)\neq(l,m),(m,l)$, then $w\in Z_{lm}(\omega,\vartheta)\,\mathrm{FU}(n,I,\Gamma)$ by (R3) – (R6).

If $(h,k)=(m,l)$, then $w=Z_{1m}(\omega,\vartheta+\zeta)$.

Finally, let $(h,k)=(l,m)$. If $l\neq -m$, we can argue exactly as in the case of the general linear group. Namely, take an index $p\neq\pm h,\pm k$ and express $T_{lm}(\omega)$ as $T_{lm}(\omega)=[T_{lp}(1),T_{pm}(\omega)]$. Then

$$\begin{aligned}{}^{T_{lm}(\zeta)}Z_{lm}(\omega,\vartheta) &= {}^{T_{lm}(\zeta)T_{ml}(\vartheta)}T_{lm}(\omega) = \\ &= {}^{T_{lm}(\zeta)T_{ml}(\vartheta)}[T_{lp}(1),T_{mp}(\omega)] = \\ &= [{}^{T_{lm}(\zeta)T_{ml}(\vartheta)}T_{lp}(1),{}^{T_{lm}(\zeta)T_{ml}(\vartheta)}T_{pm}(\omega)] = \\ &= [T_{mp}(\vartheta)T_{lp}(1+\zeta\vartheta),T_{pl}(-\omega\vartheta)T_{pm}(\omega(1+\vartheta\zeta))]\end{aligned}$$

Now decomposing the last commutator into four factors using the formula $[ab,cd]={}^{a}[b,c]\cdot{}^{ac}[b,d]\cdot[a,c]\cdot{}^{c}[a,d]$ we see that all these factors are products of transvections from $\mathrm{FU}(2n,I,\Gamma)$ and factors of the form $Z_{ij}(\xi,\zeta)$ belonging to the $\mathrm{EU}(2n,I,\Gamma)$ (see [B5] or [Va6] for a detailed calculation).

The proof for the case $l=-m$ is similar and even easier. Pick up a $p\neq\pm l$ and express $T_{l,-l}(\omega)$ as in Lemma 4.2:

$$T_{l,-l}(\omega)=[T_{p,-p}(-\mu\omega),T_{-p,-l}(1)]T_{pl}(\mu\omega),$$

where $\mu=\lambda^{(\varepsilon(-l)-\varepsilon(-p))/2}$. Then

$$\begin{aligned}{}^{T_{l,-l}(\zeta)}Z_{l,-l}(\omega,\vartheta) &= {}^{T_{l,-l}(\zeta)T_{-l,l}(\vartheta)}T_{l,-l}(\omega) = \\ &= {}^{T_{l,-l}(\zeta)T_{-l,l}(\vartheta)}[T_{p,-p}(-\mu\omega),T_{-p,-l}(1)]T_{pl}(\mu\omega) =\end{aligned}$$

This expression is a product of a matrix from $\mathrm{FU}(2n,I,\Gamma)$ and the commutator

$$\begin{aligned}&[{}^{T_{l,-l}(\zeta)T_{-l,l}(\vartheta)}T_{p,-p}(-\mu\omega),{}^{T_{l,-l}(\zeta)T_{-l,l}(\vartheta)}T_{-p,-l}(1)] = \\ &\qquad = [T_{p,-p}(-\mu\omega),T_{-p,-l}(1+\vartheta\zeta)T_{-p,l}(-\vartheta))]\end{aligned}.$$

Decomposing this commutator according to the formula $[a,bc]=[a,b]{}^{b}[a,c]$ one gets a product of a matrix from $\mathrm{FU}(2n,I,\Gamma)$ with the commutator

$${}^{T_{-p,-l}(1+\vartheta\zeta)}[T_{p,-p}(-\mu\omega),T_{-p,l}(-\vartheta))]$$

which in turn is a product of a matrix from $\mathrm{FU}(2n,I,\Gamma)$ and a matrix of the form $Z_{pl}(\xi,\zeta)$. This finishes the proof.

3^0. Relativization. In this subsection we show how relative results follow from the results for the absolute case. In particular the theorem in the introduction follows from the special case $(I,\Gamma)=(R,\Lambda)$, i.e. from the normality of the absolute elementary subgroup. This idea is due to M.R.Stein [St] and was applied to the normality of the elementary subgroups by J.Milnor [Mi], A.A.Suslin and V.I.Kopeiko [SK] and others.

Everything said above about the congruence groups modulo a form ideal applies also to the relative elementary groups. In particular, we have the following split exact sequence of groups (see [HO'M], Lemma 5.3.22):

$$1 \to \mathrm{EU}(n,I,\Gamma) \xrightarrow{\iota_2} \mathrm{EU}(n,R\times_I R,\Lambda\times_\Gamma\Lambda) \underset{\delta}{\overset{\pi_1}{\rightleftarrows}} \mathrm{EU}(n,R,\Lambda) \to 1.$$

As above we may identify $\iota_2(\mathrm{EU}(2n,I,\Gamma))$ with $\mathrm{EU}(2n,\iota_2(I),\iota_2(\Gamma))$. The following easy observation (see [Mi], Lemma 4.2 for the linear case) reduces relative questions about the elementary subgroups to the absolute ones.

Lemma 4.4. *Let (I,Γ) be a form ideal of a form ring (R,Λ). Then one has*

$$\iota_2(\mathrm{U}(2n,I,\Gamma))\cap \mathrm{EU}(2n,R\times_I R,\Lambda\times_\Gamma\Lambda)=\iota_2(\mathrm{EU}(2n,I,\Gamma)).$$

Proof. Clearly the right hand side is contained in the left hand side. Conversely, let (a,b) be an element from

$$\iota_2(\mathrm{U}(2n,I,\Gamma))\cap \mathrm{EU}(2n,R\times_I R,\Lambda\times_\Gamma\Lambda)=$$
$$\mathrm{U}(2n,\iota_2(I),\iota_2(\Gamma))\cap \mathrm{EU}(2n,R\times_I R,\Lambda\times_\Gamma\Lambda).$$

Then $(a,b)\in \mathrm{U}(2n,\iota_2(I),\iota_2(\Gamma))$ implies that $a=e$ and $b\in \mathrm{U}(2n,I,\Gamma)$. On the other hand, since $(e,b)\in \mathrm{EU}(2n,R\times_I R,\Lambda\times_\Gamma\Lambda)$, it can be expressed in the form

$$(e,b)=(c_1,c_1d_1)\dots(c_t,c_td_t)$$

where $c_i\in \mathrm{EU}(2n,R,\Lambda)$ and $d_i\in \mathrm{EU}(2n,I,\Gamma)$ are some elementary unitary transvections. Clearly, $c_1\dots c_t=e$. Now we may rewrite b in the form

$$b=(c_1d_1c_1^{-1})(c_1c_2d_2c_2^{-1}c_1^{-1})\dots(c_1\dots c_td_tc_t^{-1}\dots c_1^{-1})\in \mathrm{EU}(2n,I,\Gamma),$$

and thus, clearly, $(e,b)\in \mathrm{U}(2n,\iota_2(I),\iota_2(\Gamma))$.

Corollary. *Normality of the absolute elementary subgroup implies all other statements of the Theorem above.*

Proof. First we prove that all relative elementary subgroups are normal. Indeed, let $a \in \mathrm{U}(2n, R, \Lambda)$, $\beta \in \mathrm{EU}(2n, I, \Gamma)$. Then since $\mathrm{EU}(2n, R \times_I R, \Lambda \times_\Gamma \Lambda)$ is normal in $\mathrm{U}(2n, R \times_I R, \Lambda \times_\Gamma \Lambda)$, one has

$$(e, aba^{-1}) = (a, a)(e, b)(a, a)^{-1} \in \mathrm{EU}(2n, R \times_I R, \Lambda \times_\Gamma \Lambda).$$

On the other hand clearly $aba^{-1} \in \mathrm{U}(2n, I, \Gamma)$. Then by the above lemma $aba^{-1} \in \mathrm{EU}(2n, I, \Gamma)$.

Now we prove the second commutator formula. Let $a \in \mathrm{EU}(2n, R, \Lambda)$ and $b \in \mathrm{CU}(2n, I, \Gamma)$. Then normality of $\mathrm{EU}(2n, R \times_I R, \Lambda \times_\Gamma \Lambda)$ in $\mathrm{U}(2n, R \times_I R, \Lambda \times_\Gamma \Lambda)$ implies that

$$(e, [a, b]) = [(a, a), (e, b)] \in \mathrm{EU}(2n, R \times_I R, \Lambda \times_\Gamma \Lambda).$$

On the other hand $[a, b] \in \mathrm{U}(2n, I, \Gamma)$ by the definition of $\mathrm{CU}(2n, I, \Gamma)$. Again by the above lemma this implies that $[a, b] \in \mathrm{EU}(2n, I, \Gamma)$.

§ 5 Eichler–Siegel–Dickson transformations

In this section we introduce the elements which play the crucial role in what follows.

1^0. ESD-transvections. Let u, v be two orthogonal vectors in V such that u is isotropic. Let further $\xi \in R$ and $\alpha \in \Lambda$. We introduce the following matrix:

$$T_{uv}(\xi, \alpha) = e + u\xi\widetilde{v} - v\overline{\lambda}\,\overline{\xi}\widetilde{u} - u\overline{\lambda}\xi\big(f(v, v) + \alpha\big)\overline{\xi}\widetilde{u}.$$

As we shall see now, this matrix always belongs to the unitary group.

Lemma 5.1. *For any orthogonal vectors u, v such that u is isotropic and any $\xi \in R$, $\alpha \in \Lambda$ one has $T_{uv}(\xi, \alpha) \in \mathrm{U}(2n, R, \Lambda)$.*

Proof. Denote $T_{uv}(\xi, \alpha)$ by g. We have to prove that g preserves both h and q. Start with h. Let x, y be arbitrary vectors from V. Recall that we write simply (x, y) instead of $h(x, y)$. Then clearly

$$(gx, gy) - (x, y) = (gx - x, gy - y) + (gx - x, y) + (x, gy - y).$$

Here

$$gx - x = u\xi(v, x) - v\overline{\lambda}\,\overline{\xi}(u, x) - u\overline{\lambda}\xi\big(f(v, v) + \alpha\big)\overline{\xi}(u, x),$$

and the same holds for y. Since the vectors u and v are orthogonal and u is isotropic, one has

$$(gx - x, gy - y) = (-v\lambda\overline{\xi}(u, x), -v\lambda\overline{\xi}(u, y)) = \overline{(u, x)}\xi(v, v)\overline{\xi}(u, y).$$

One the other side,

$$(gx-x,y)=\overline{(v,x)}\overline{\xi}(u,y)-\overline{(u,x)}\lambda\xi(v,y)-\overline{(u,x)}\lambda\xi(\overline{f(v,v)}+\overline{\alpha})\overline{\xi}(u,y),$$

$$(x,gy-y)=(x,u)\xi(v,y)-(x,v)\overline{\lambda}\,\overline{\xi}(u,y)-(x,u)\overline{\lambda}\xi(f(v,v)+\alpha)\overline{\xi}(u,y).$$

Since h is λ-hermitian, the first summand of $(gx-x,y)$ cancels with the second summand of $(x,gy-y)$ and vice versa. Thus, finally,

$$\begin{aligned}(gx-x,y)+(x,gy-y)&=-\overline{(u,x)}\lambda\xi\overline{\alpha}\overline{\xi}(u,y)-(x,u)\overline{\lambda}\xi\alpha\overline{\xi}(u,y)=\\&=-\overline{(u,x)}\xi\Big(\lambda\big(\overline{f(v,v)}+\overline{\alpha}\big)+f(v,v)+\alpha\Big)\overline{\xi}(u,y)\end{aligned}.$$

Since $\alpha\in\Lambda$, this sum cancels with $f(gx-x,gy-y)$ and the preservation of h is established.

Now we pass to the Λ-quadratic form q. Take an arbitrary vector $x\in V$. Lemma 1.1 asserts that

$$q(gx)-q(x)=q(gx-x)+(x,gx-x)+\Lambda.$$

Thus it remains only to show that $f(gx-x,gx-x)+(x,gx-x)\in\Lambda$. The first summand equals to

$$f(gx-x,gx-x)=\overline{(u,x)}\xi f(v,v)\overline{\xi}(u,x),$$

while the second summand equals to

$$(x,gx-x)=(x,u)\xi(v,x)-(x,v)\overline{\lambda}\,\overline{\xi}(u,x)-(x,u)\overline{\lambda}\xi(f(v,v)+\alpha)\overline{\xi}(u,x).$$

Finally,

$$f(gx-x,gx-x)+(x,gx-x)=(x,u)\xi(v,x)-(x,v)\overline{\lambda}\,\overline{\xi}(u,x)-(x,u)\overline{\lambda}\xi\alpha\overline{\xi}(u,x).$$

The last summand in this expression has the form $\zeta\alpha\overline{\zeta}$ for $\zeta=(x,u)\xi$ and thus belongs to Λ by the definition (recall that $\alpha\in\Lambda$). On the other side

$$(x,u)\xi(v,x)-(x,v)\overline{\lambda}\,\overline{\xi}(u,x)=(x,u)\xi(v,x)-\lambda\overline{(v,x)}\overline{\xi}\overline{(x,u)}\in\Lambda,$$

again by definition of Λ. Thus the form q is also preserved by g, which finishes proof of the lemma.

2^0. Basic properties of ESD-transvections. In this subsection we state some obvious properties of the elements $T_{uv}(\xi,\alpha)$. Definition of ESD-transvections and Lemma 1.4 immediately imply that a conjugate of an ESD-transvection by an element of $\mathrm{U}(2n,R,\Lambda)$ is again an ESD-transvection:

Lemma 5.2. *For any $g \in \mathrm{U}(2n, R, \Lambda)$ one has*

$$gT_{uv}(\xi,\alpha)g^{-1} = T_{gu,gv}(\xi, \alpha + f(v,v) - f(gv,gv)).$$

Notice that the expression $\alpha + f(v,v) - f(gv,gv)$ on the right hand side belongs to Λ by the definition of $\mathrm{U}(2n, R, \Lambda)$.

Lemma 5.3. *For any $\xi, \zeta \in R$ one has*

$$T_{uv}(\xi\zeta, \alpha) = T_{u\xi,v}(\zeta, \alpha) = T_{u,v\overline{\zeta}}(\xi, \zeta\alpha\overline{\zeta}).$$

In particular this lemma shows that the parameter ξ in the definition of the ESD-transvections is optional and can be concealed either in u or in v, for example $T_{uv}(\xi,\alpha) = T_{u\xi,v}(1,\alpha)$. However we think that it's not advisable to do so, especially when one wants to speak about one-parameter subgroups of ESD-transvections, etc. In some situations it is convenient to assume that u or v enjoy some special properties, for example, are unimodular. This can be done only if we explicitly keep the parameter ξ.

The following lemma establishes additivity property of $T_{uv}(\xi,\alpha)$ in v when u is fixed.

Lemma 5.4. *For any two vectors v, w orthogonal to an isotropic vector u, any $\xi, \zeta \in R$ and any $\alpha, \beta \in \Lambda$ one has*

$$T_{uv}(\xi,\alpha)T_{uw}(\zeta,\beta) = T_{u,v\overline{\xi}+w\overline{\zeta}}(1, \xi\alpha\overline{\xi} + \zeta\beta\overline{\zeta} - \zeta f(w,v)\overline{\xi} + \lambda\xi\overline{f(w,v)}\overline{\zeta}).$$

The formulae expressing additivity in the first argument are more complicated. We state them in the next subsection for elements of short root type.

3^0. Elements of long and short root types. Now we introduce two special cases of the ESD-transvections. These are really the only types which appear in the analysis of the conjugates of the elementary unitary transvections. For an isotropic vector $u \in V$ and an element $\alpha \in \Lambda$ we denote the element $T_{u0}(1, -\alpha)$ by $T_u(\alpha)$ and call it an *element of long root type*. One has

$$T_u(\alpha) = e + u\overline{\lambda}\alpha\widetilde{u}.$$

For two orthogonal vectors $u, v \in V$ such that u is isotropic and an element $\xi \in R$ we denote the element $T_{uv}(\xi, 0)$ by $T^{\bullet}_{uv}(\xi)$. One has

$$T^{\bullet}_{uv}(\xi) = e + u\xi\widetilde{v} - v\overline{\lambda}\,\overline{\xi}\widetilde{u} - u\overline{\lambda}\xi f(v,v)\overline{\xi}\widetilde{u}.$$

If v is also isotropic, we denote the element $T_{uv}(\xi, -f(v,v))$ by $T_{uv}(\xi)$. Clearly

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

We call the elements $T^{\bullet}_{uv}(\xi)$ and $T_{uv}(\xi)$ *elements of short root type*. Clearly the elementary unipotents introduced in the previous section may be interpreted as elements of root type. More precisely,

Lemma 5.5. *For any $i \neq \pm j$ and any $\xi \in R$, $\alpha \in \Lambda$, one has*

$$T_{i,-i}(\alpha) = T_{e_i}(\lambda^{(\varepsilon(i)+1)/2}\alpha),$$

$$T_{ij}(\xi) = T_{e_i,e_{-j}}(\lambda^{-(\varepsilon(j)+1)/2}\xi) = T^{\bullet}_{e_i,e_{-j}}(\lambda^{-(\varepsilon(j)+1)/2}\xi).$$

In particular Lemma 5.2 implies that a conjugate of an elementary unitary transvection is an ESD-transvection. More precisely, a conjugate of an element $T_{i,-i}(\alpha)$ is an element of long root type, while a conjugate of $T_{ij}(\xi)$, $i \neq \pm j$, is an element of short root type.

Remark. The reader may ask, why $T_{uv}(\xi,\alpha)$, depending on two parameters? Well, for extra-short roots. In $\mathrm{U}(2n,R,\Lambda)$ there are no extra-short roots, but there are extra-short root subgroups. Indeed, look at the group $\mathrm{U}(2n+1,R)$. It is well-known that the corresponding root system is not reduced, and has roots of *three* different lengths: long, short and *extra-short*, the last ones being halves of the long ones. The extra-short root subgroups are not abelian, see, for example, [Ca], [A2]. The extra-short roots disappear when one passes to $\mathrm{U}(2n,R)$ – but the corresponding root subgroups survive.

Actually any ESD-transvection may be presented as a product of *commuting* short and long root type elements. This essentially reduces the study of the ESD-transvection to these two cases.

Lemma 5.6. *One has*

$$T_{uv}(\xi,\alpha) = T^{\bullet}_{uv}(\xi)T_u(-\xi\alpha\overline{\xi}).$$

If v is also isotropic, then

$$T_{uv}(\xi,\alpha) = T_{uv}(\xi)T_u(-\xi(f(v,v)+\alpha)\overline{\xi}).$$

4^0. Basic properties of elements of root type. Some of the formulae for ESD-transvections simplify considerably when restricted to the elements of root type. For example for elements of long root type Lemma 4.4 becomes.

Lemma 5.7. *Let u be an isotropic vector. Then for any $\alpha,\beta \in \Lambda$, $\xi \in R$ one has*

$$T_u(\alpha)T_u(\beta) = T_u(\alpha+\beta), \qquad T_{u\xi} = T_u(\xi\alpha\overline{\xi}).$$

Analogous additivity formula holds for elements of long root type.

Lemma 5.8. *Let u,v be orthogonal isotropic vectors. Then for any $\xi,\zeta \in R$ one has*

$$T_{uv}(\xi)T_{uv}(\zeta) = T_{uv}(\xi+\zeta).$$

The following additivity property of elements of long root type (compare the proof of Lemma 2.4) plays crucial role in the second part of the work.

Lemma 5.9. *Let u, v be two ortogonal isotropic vectors. Then for any $\alpha \in \Lambda$ one has*

$$T_u(\alpha)T_v(\alpha) = T_{u+v}(\alpha)T_{u,v}(-\overline{\lambda}\alpha).$$

The following lemma (compare Lemma 1.7 of [SK]) establishes additivity property of the elements of short root type $T_{uv}(\xi)$ in u when v is fixed.

Lemma 5.10. *Let u, v be ortogonal isotropic vectors which are both orthogonal to w. Then for any $\xi, \zeta \in R$ one has*

$$T^\bullet_{uw}(\xi)T^\bullet_{vw}(\zeta) = T^\bullet_{u\xi+v\zeta,w}(1)T_{u\xi,v\zeta}(\overline{f(w,w)}).$$

The following lemma (compare Formula (h) on page 214 of [HO'M], although the involution seems to be missing there) expresses symmetry of $T_{uv}(\xi)$ with respect to u and v when *both* of them are isotropic.

Lemma 5.11. *Assume that both u and v are isotropic. Then $T_{uv}(\xi) = T_{v,-u}(\overline{\lambda}\,\overline{\xi})$.*

§ 6 Whitehead type Lemmas

In this section we prove that a unitary transvection $T_{u,v}(\xi,\alpha)$ belongs to the elementary subgroup $\mathrm{EU}(2n,R,\Lambda)$ if u or v (or both) have zero components.

1^0. Heisenberg group. In this subsection we give explicit shape of elements from the unipotent radical of the standard parabolic subgroup P_1 which we need in subsequent subsections.

Lemma 6.1. *Let $v = (v_2, \dots, v_n, v_{-n}, \dots, v_{-2})^t$ be any vector of length $2l-2$. Then matrices*

$$Y^+_\bullet(v) = \begin{pmatrix} 1 & -\overline{\lambda}\widetilde{v} & -\overline{\lambda}f(v,v) \\ 0 & e & v \\ 0 & 0 & 1 \end{pmatrix}, \qquad Y^-_\bullet(v) = \begin{pmatrix} 1 & 0 & 0 \\ v & e & 0 \\ -f(v,v) & -\widetilde{v} & 1 \end{pmatrix},$$

belong to $\mathrm{EU}(2n,R,\Lambda)$. *Moreover, one has*

$$Y^+_\bullet(v)^{-1} = \begin{pmatrix} 1 & \overline{\lambda}\widetilde{v} & -\overline{f(v,v)} \\ 0 & e & -v \\ 0 & 0 & 1 \end{pmatrix}, \qquad Y^-_\bullet(v)^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -v & e & 0 \\ -\lambda\overline{f(v,v)} & \widetilde{v} & 1 \end{pmatrix}.$$

Proof. Direct calculation shows that

$$Y^+_\bullet(v) = \prod T_{i,-1}(v_i), \qquad Y^-_\bullet(v) = \prod T_{i1}(v_i),$$

where both products are taken over $i = 2, \dots, -2$ in the natural order. It remains to observe that all factors on the right hand side belong to $\mathrm{EU}(2n,R,\Lambda)$. The formulae for the inverse matrices are verified by a straightforward calculation.

Lemma 6.2. *Let $v = (v_2, \dots, v_n, v_{-n}, \dots, v_{-2})^t$ be an isotropic vector of length $2l-2$. Then matrices*

$$Y^+(v) = \begin{pmatrix} 1 & -\overline{\lambda}\widetilde{v} & 0 \\ 0 & e & v \\ 0 & 0 & 1 \end{pmatrix}, \qquad Y^-(v) = \begin{pmatrix} 1 & 0 & 0 \\ v & e & 0 \\ 0 & -\widetilde{v} & 1 \end{pmatrix},$$

belong to $\mathrm{EU}(2n, R, \Lambda)$. *Moreover,* $Y^+(v)^{-1} = Y^+(-v)$ *and* $Y^-(v)^{-1} = Y^-(-v)$.

Proof. Since v is isotropic, one has $f(v,v) \in \Lambda$. Clearly

$$Y^+(v) = T_{1,-1}(\overline{\lambda} f(v,v)) Y^+_\bullet(v), \qquad Y^-(v) = T_{-1,1}(f(v,v)) Y^-_\bullet(v),$$

where the factors on the right hand sides belong to $\mathrm{EU}(2n, R, \Lambda)$. The last statement of the lemma follows from the fact that since u is isotropic, one has $\widetilde{u}u = (u,u) = 0$.

The following two lemmas are straightforward. They will be used in the sequel of this paper.

Lemma 6.3. *For any $u, v \in R^{2n-2}$ one has*

$$[Y^+_\bullet(u), Y^+_\bullet(v)] = T_{1,-1}(\overline{\lambda}\widetilde{v}u - \overline{\lambda}\widetilde{u}v),$$
$$[Y^-_\bullet(u), Y^-_\bullet(v)] = T_{-1,1}(\widetilde{v}u - \widetilde{u}v).$$

In fact the matrices $Y^+_\bullet(v), \dots, Y^-(v)$ introduced above are ESD-transvections, corresponding to the case when the first of the vectors u, v is a standard base vector, namely:

$$Y^+_\bullet(v) = T^\bullet_{e_1,v}(-\overline{\lambda}), \qquad Y^-_\bullet(v) = T^\bullet_{e_{-1},v}(-1),$$

and, respectively,

$$Y^+(v) = T_{e_1,v}(-\overline{\lambda}), \qquad Y^-(v) = T_{e_{-1},v}(-1).$$

Now we start proving that an ESD-transvection is elementary if u and/or v has enough zeros in one sense or other.

2^0. Whitehead-Vaserstein lemma for $\mathbf{P_1}$: long root type. In this subsection we consider the elements of the form $T_u(\alpha) = T_{u0}(1, -\alpha)$, where u is an isotropic vector and $\alpha \in \Lambda$. Recall that $T_u(\alpha) = e + u\overline{\lambda}\alpha\widetilde{u}$.

Lemma 6.4. *Suppose that $u_i = u_{-i} = 0$ for some $i \in I$. Then for all $\alpha \in \Lambda$ one has $T_u(\alpha) \in \mathrm{EU}(2n, R, \Lambda)$.*

Proof. Since the group $\mathrm{EU}(2n, R, \Lambda)$ is normalized by all permutation matrices $H(\pi)$, where $\pi \in \mathrm{S}_n$, we may without loss of generality assume that $i = 1$. Let v be the column of height $2n-2$ which is obtained from u by dropping the coordinates with indices ± 1. then v is isotropic. A direct calculation using the fact that $\widetilde{v}v = 0$, shows that

$$T_u(\alpha) = T_{-1,1}(\alpha)Y^{+}(v)T_{-1,1}(-\alpha)Y^{-}(v\alpha)Y^{+}(-v),$$

where all factors on the right hand side belong to $\mathrm{EU}(2n, R, \Lambda)$ by the previous lemma.

Lemma 6.5. *Suppose $u_i = 0$ for some $i \in I$. Then $T_u(\alpha) \in \mathrm{EU}(2n, R, \Lambda)$ for all $\alpha \in \Lambda$.*

Proof. By the same reason as in the previous lemma and since $\mathrm{EU}(2n, R, \Lambda)$ is normalized by the matrix

$$\begin{pmatrix} 0 & p \\ \lambda p & 0 \end{pmatrix},$$

we may without loss of generality assume that $u_{-1} = 0$. Present u in the form $u = v + e_1 u_1$, where v is the column obtained from u by changing its 1-st coordinate to 0 and $u_1 = u_1$. Since $u_{-1} = 0$, the vector v is isotropic. Then

$$T_u(\alpha) = T_v(\alpha)Y^{+}(v\overline{\lambda}\alpha\overline{u}_1)T_{1,-1}(\overline{\lambda}u_1\alpha\overline{u}_1),$$

where the two first factors belong to $\mathrm{EU}(2n, R, \Lambda)$ by the two preceding lemmas and the third one by the definition of Λ.

3^0. Whitehead-Vaserstein lemma for $\mathbf{P_1}$: general case. In this subsection we consider the elements of the form $T^{\bullet}_{uv}(\xi) = T_{uv}(\xi, 0)$, where u, v are orthogonal vectors, u is isotropic and $\xi \in R$. Notice that for the time being *we do not assume v to be isotropic*. Recall that $T^{\bullet}_{uv}(\xi) = e + u\xi\widetilde{v} - v\overline{\lambda}\,\overline{\xi}\widetilde{u} - u\overline{\lambda}\xi f(v, v)\overline{\xi}\widetilde{u}$.

Lemma 6.6. *Suppose that $u_i = u_{-i} = v_i = v_{-i} = 0$ for some $i \in I$. Then $T^{\bullet}_{uv}(\xi) \in \mathrm{EU}(2n, R, \Lambda)$ for all $\xi \in R$.*

Proof. As in the previous lemma we may without loss of generality assume that $u_1 = u_{-1} = v_1 = v_{-1} = 0$. Let x and y be the columns of height $2n-2$ which are obtained from u and v respectively by dropping the coordinates with indices ± 1. Then x, y are orthogonal, and x is isotropic, in particular,

$\widetilde{x}x = \widetilde{x}y = \widetilde{y}x = 0$. On the other side, $f(y,y) = f(v,v)$, and thus $\widetilde{y}y = f(v,v) + \lambda\overline{f(v,v)}$. Now a direct calculation using these equalities shows that

$$T^{\bullet}_{uv}(\xi) = Y^{-}(x\lambda\xi)Y^{+}_{\bullet}(-y)Y^{-}(-x\lambda\xi)Y^{+}_{\bullet}(y - x\lambda\xi\overline{f(v,v)}),$$

where the right hand side belongs to $\mathrm{EU}(2n,R,\Lambda)$ by Lemmas 6.1 and 6.2.

Lemma 6.7. *Suppose that $u_i = v_i = 0$ for some $i \in I$. Then $T^{\bullet}_{uv}(\xi) \in \mathrm{EU}(2n,R,\Lambda)$ for all $\xi \in R$.*

Proof. As in the previous lemmas we may without loss of generality assume that $u_1 = v_1 = 0$. Present u and v in the form $u = x + e_1u_1$, $v = y + e_1v_1$, where x and y are the columns obtained from u and v respectively by changing their 1-st coordinates to 0. Since $u_{-1} = v_{-1} = 0$, the vectors x and y are orthogonal and x is isotropic. Then

$$T^{\bullet}_{uv}(\xi) = T^{\bullet}_{xy}(\xi)Y^{+}_{\bullet}(x\xi\overline{v}_1 - y\overline{\lambda}\,\overline{\xi}\,\overline{u}_1)T_{1,-1}(u_1\xi\overline{v}_1 - \overline{\lambda}v_1\overline{\xi}\overline{u}_1),$$

where the first two factors belong to $\mathrm{EU}(2n,R,\Lambda)$ by the Lemmas 6.5 and 6.1 respectively and the last one by the definition of Λ.

Assume now that the vector v is also isotropic. Then the following version of the preceding lemma holds.

Lemma 6.8. *Suppose that $u_i = v_i = 0$ for some $i \in I$. Then $T_{uv}(\xi) \in \mathrm{EU}(2n,R,\Lambda)$ for all $\xi \in R$.*

Proof. Recall that

$$T_{uv}((\xi) = T^{\bullet}_{uv}(\xi)T_u(\xi f(v,v)\overline{\xi}).$$

But the factors on the right hand side belong to $\mathrm{EU}(2n,R,\Lambda)$ by Lemmas 6.6 and 6.4 respectively.

Of course this could be proven directly, for example, the formula in the proof of Lemma 6.5 may be simplified to the following commutator relation:

$$T_{uv}(\xi) = [Y^{-}(x\lambda\xi), Y^{+}(-y)].$$

Now we can prove the 'Whitehead-Vaserstein Lemma' which says in particular that the Eichler subgroup $\mathrm{TU}(2n-2,R,\Lambda)$ is contained in the elementary group $\mathrm{EU}(2n,R,\Lambda)$ of larger degree.

Proposition 2. *Let u, v be any orthogonal vectors such that u is isotropic. Assume that $u_i = v_i = 0$ for some i. Then $T_{uv}(\xi,\alpha) \in \mathrm{EU}(2n,R,\Lambda)$ for all $\xi \in R$ and $\alpha \in \Lambda$.*

Proof. From the previous section we know that

$$T_{uv}(\xi,\alpha) = T_{uv}(\xi,0)T_{u0}(\xi,\alpha).$$

But the factors on the right hand side belong to $\mathrm{EU}(2n, R, \Lambda)$ by Lemmas 6.7 and 6.4 respectively.

4^0. Kopeiko-Taddei Lemma. Here we show that the addition formulae for ESD-transformations from § 5 (see Lemmas 5.4 and 5.9) imply that $T_{uv}(\xi, \alpha)$ belongs to the elementary group if u has two zeros in symmetric positions. As is known from [Ko1], [Ta1], this already suffices to prove normality in the classical symplectic case (when $\lambda = -1$ and the involution is trivial).

Lemma 6.9. *Suppose that $u_i = u_{-i} = 0$ and $v = e_i v_i + e_{-i} v_{-i}$. Then $T_{uv}(\xi, \alpha) \in \mathrm{EU}(2n, R, \Lambda)$ for all $\xi \in R$ and $\alpha \in \Lambda$.*

Proof. As always we may without loss of generality assume that $i = 1$. By Lemmas 5.6 and 6.4 it's enough to prove that $T_{uv}(\xi, \alpha) \in \mathrm{EU}(2n, R, \Lambda)$ for *some* $\alpha \in \Lambda$, for example, for $\alpha = 0$. But for this case one easily sees that

$$T_{uv}(\xi) = Y^+(\overline{\lambda} x \xi \overline{v}_1) Y^-(x \xi \overline{v}_{-1}).$$

Lemma 6.10. *Suppose $u_i = u_{-i} = 0$ for some i, and let v be any vector orthogonal to u. Then $T_{uv}(\xi, \alpha) \in \mathrm{EU}(2n, R, \Lambda)$ for all $\xi \in R$ and $\alpha \in \Lambda$.*

Proof. Express v in the form $v = x + z$, where $x = e_1 v_1 + e_{-1} v_{-1}$ and $y = v - x$. Then by Lemmas 11 and 12 one has

$$T_{uv}(\xi, \alpha) = T_{ux}(\xi, \beta) T_{uy}(\xi, \gamma),$$

for appropriate $\beta, \gamma \in \Lambda$. Now the first factor on the right hand side belongs to $\mathrm{EU}(2n, R, \Lambda)$ by Lemma 6.6 and the second one by the previous Lemma.

§ 7. Normality of elementary subgroup: Suslin's approach

In this section we prove normality of $\mathrm{EU}(2n, R, \Lambda)$ in $\mathrm{U}(2n, R, \Lambda)$ for the case when R is almost commutative and $n \geq 3$. We do it following the Suslin's approach from [Su], [Tu], [SK]. For *long* root elements in the case $\lambda = -1$, $\Lambda = \Lambda_{\min}$ this approach was used also in [GM]. For *short* root elements in the *symplectic* group it was used in [K], [Ta1]. Here we show how it works in the general case.

1^0. Unitary Suslin's Lemma. The original Suslin's proof [Su2] of normality of the elementary subgroup $\mathrm{E}(n, R)$ in the general linear group $\mathrm{GL}(n, R)$, $n \geq 3$, over a *commutative* ring R was based on the following observation, which is sometimes referred to as a Suslin's lemma (compare [Md], [BV], [HO'M]). Recall that a row $u = (u_1, \dots, u_n) \in {}^nR$ is called *unimodular*, if $u_1R+\ldots+u_nR=R$, i.e. the (right) ideal generated by the components of u coincides with R. It is equivalent to saying that there exists a column $w\in R^n$ such that $uw=1$. Suslin's lemma asserts that if $u\in {}^nR$ is a unimodular row of length $n\geq 2$ then any solution $v\in R^n$ of the homogeneous linear equation $uv=0$ is a linear combination of solutions $u_je_i-u_ie_j$, $i\neq j$, which have at most two non-zero coordinates.

First we state a unitary analogue of this Lemma for the commutative case. Note that unimodularity of a column u is equivalent to the unimodularity of the corresponding row $\widetilde{u}$. It is clear that if u is unimodular, there exists a column w such that $h(u,w)=1$. We are interested in decomposing columns orthogonal to u into sums of columns having enough zero coordinates.

Lemma 7.1. *Suppose $u\in R^{2n}$ is a unimodular column of height $n\geq 2$ over a commutative ring R. Then any vector $v\in R^{2n}$ orthogonal to u, i.e. such that $h(u,v)=0$, may be expressed as a linear combination of vectors $u_{ij}=\lambda^{(\varepsilon(i)-1)/2}\overline{u}_{-j}e_i-\lambda^{(\varepsilon(j)-1)/2}\overline{u}_{-i}e_j$, $i\neq j$, which are all orthogonal to v and have at most two non-zero coordinates.*

Proof. It is clear that that all the vectors u_{ij} are orthogonal to u. Now the lemma is proven by essentially the same formula as the original Suslin's lemma. Namely, fix a $w\in R^{2n}$ such that $h(u,w)=1$. Then a direct calculation using the definition of h and the equalities $h(u,v)=0$, $h(u,w)=1$, shows that

$$v=\sum(w_jv_i-w_iv_j)u_{ij},$$

where the sum is taken by all $i<j$ in our sense, i.e. by such pairs (i,j) that either the signs of i and j coincide and $i<j$ in the usual sense, or else $i\in I^+$, $j\in I^-$.

Now we state a non-commutative version of this lemma (compare [Tu], [Kh]). Let $R_0=\mathrm{Cent}(R)$ be the centre of the ring R. For an element $\xi\in R$ we denote by $\mathcal{O}(\xi)$ the ideal in R_0, generated by all *central multiples* of ξ, i.e. by all elements from R_0 which have the form $\xi\zeta=\zeta\xi$ for some $\zeta\in R$. Notice that we do not assume any more u to be unimodular: the price we pay is that now we express as a linear combination not the vector v itself, but some of its multiples.

Lemma 7.2. *Suppose $u\in R^{2n}$ is a column of height $n\geq 2$ and $\theta=\overline{u}_i\zeta=\zeta\overline{u}_i$, $\zeta\in R$, be a central multiple of $\overline{u}_i$ for a fixed $i=1,\ldots,-1$. Then for any vector $v\in R^{2n}$ orthogonal to u its multiple $v\theta$ may be expressed as $v\theta=\sum w_{j,-i}$, $j\neq i$, where each of the vectors $w_{j,-i}$ is orthogonal to u and has at most two nonzero coordinates.*

Proof. Set

$$w_{j,-i}=v_j\theta e_j-\lambda^{(1-\varepsilon(i))/2}\lambda^{(\varepsilon(j)+1)/2}\zeta\overline{u}_{-j}v_je_{-i},$$

Every $w_{j,-i}$ is orthogonal to u and if one denotes $w = \sum w_{j,-i}$, $j \neq -i$, then, clearly, $w_j = v_j\theta$ for all $j \neq -i$. Finally,

$$w_{-i} = \lambda^{(1-\varepsilon(i))/2} \sum_{j\neq -i} \lambda^{(\varepsilon(j)+1))/2} \zeta \overline{u}_{-j} v_j = \zeta \overline{u}_i v_{-i} = v_{-i}\theta.$$

The decompositions of root type elements in the following subsections are based on this lemma. Thus in the next subsection we apply it to an *isotropic* vector $v = u$. In this case we want our summands to be isotropic as well. Clearly, $w_{i,-i}$ is not isotropic. But actually we do not need to decompose $v\theta$ into summands which have only two non-zero elements, it is usually enough to obtain *one* zero element. Therefore we are usually done simply by presenting $v\theta$ in the form $v\theta = w_{j,-i} + (v\theta - w_{j,-i})$ for some $j \neq \pm i$.

2^0. Decomposition of long root unipotents. In this subsection we decompose a long root unipotent $T_u(\alpha)$ under the assumption that $\alpha \in \theta\Lambda\overline{\theta}$ for some $\theta \in \mathcal{O}(\overline{u}_i)$, $i = 1, \dots, -1$.

Proposition 3. *Let $u \in R^{2n}$ be an isotropic column, $\alpha \in \Lambda$. Assume that $n \geq 3$ and $\alpha \in \theta\Lambda\overline{\theta}$ for $\theta \in \mathcal{O}(\overline{u}_i)$, $i = 1, \dots, -1$. Then $T_u(\alpha) \in \mathrm{EU}(2n, R, \Lambda)$.*

Proof. Without loss of generality we may assume that $\alpha = \theta\beta\overline{\theta}$, where $\theta \in \mathcal{O}(u_{-1})$, $\beta \in \Lambda$. Then by Lemma 5.7 one has $T_u(\theta\beta\overline{\theta}) = T_{u\theta}(\beta)$. On the other hand, the column $u\theta$ may be decomposed as follows $u\theta = v + w$, where v and w are orthogonal isotropic vectors of the form:

$$v = (\zeta, u_2\theta, 0, \dots, 0, 0)^t, \qquad w = (u_1\theta - \zeta, 0, u_3\theta, \dots, u_{-2}\theta, u_{-1}\theta)^t.$$

Indeed, let $\xi \in R$ be such that $\overline{u}_{-1}\xi = \xi\overline{u}_{-1} \in R_0$. Then one may take $\zeta = -\xi\overline{u}_2 u_{-2}$. A direct check using the fact that u is isotropic and centrality of ζ shows that v and w are indeed orthogonal and isotropic. Now by Lemma 5.10 one has

$$T_{u\theta}(\beta) = T_v(\beta)T_w(\beta)T_{vw}(\beta),$$

where all there factors on the right hand side are elementary: the first two by the Whitehead–Vaserstein lemma (say, our Lemma 6.4), the last one by the Kopeiko–Taddei lemma (Lemma 6.9).

3^0. Decomposition of short root unipotents. In this subsection we decompose a short root unipotent $T_{uv}(\xi)$ under the assumption that $\xi \in \overline{\mathcal{O}(\overline{u}_i)}\mathcal{O}(\overline{u}_j)R$ for some $j \neq \pm i$.

Proposition 4. *Let $u \in R^{2n}$ be an isotropic vector, $v \in R^{2n}$ be any vector, orthogonal to u, and $\xi \in R$. Assume that $n \geq 3$ and $\xi \in \overline{\mathcal{O}(\overline{u}_i)}\mathcal{O}(\overline{u}_j)R$ for some $i \neq \pm j$. Then $T_{uv}(\xi) \in \mathrm{EU}(2n, R, \Lambda)$.*

Proof. Without loss of generality we may assume that $\xi = \overline{\eta}\theta\rho$, where $\eta \in \mathcal{O}(\overline{u}_{-1})$, $\theta \in \mathcal{O}(\overline{u}_{-2})$ and $\rho \in R$. By Lemma 5.3 one has $T_{uv}(\eta\theta\rho) = T_{u,v\eta}(\theta\rho)$. On the other hand by Lemma 7.2 the column $v\eta$ may be decomposed into the sum of columns $w_{j,-i}$ all of which are orthogonal to u and have at most two non-zero coordinates. By Lemmas 5.4 and 5.6 one has

$$T_{u,v\eta}(\theta\rho) = T_{u,w_{i,-i}}(\theta\rho)T_u(*)\prod T_{u,w_{j,-i}}(\theta\rho),$$

where the product is taken over all $j \neq \pm i$. The second factor on the right hand side is elementary by the previous section, every factor in the product is elementary by the Kopeiko-Taddei lemma (since all $w_{j,-i}$, $j \neq \pm i$, are isotropic). It remains only to consider the first factor.

We cannot apply Kopeiko–Taddei lemma to the element $T_{u,w_{i,-i}}(\theta\rho)$ since $w_{i,-i}$ is not isotropic. But we can repeat arguments from the previous section to apply the Whitehead–Vaserstein lemma. Namely, by Lemma 5.3 one has $T_{u,w_{i,-i}}(\theta\rho) = T_{u\theta,w_{i,-i}}(\rho)$. Now, precisely as in the proof of the preceding proposition, we can decompose the column $u\theta$ as $u\theta = w + z$, where w and z are orthogonal isotropic vectors of the form:

$$w = (0, \zeta, u_3\theta, 0, \dots, 0, 0)^t, \qquad z = (u_1\theta, u_2\theta - \zeta, 0, u_4\theta, \dots, u_{-2}\theta, u_{-1}\theta)^t.$$

Moreover, since u and w are orthogonal to $w_{i,-i}$, z also is. Now by Lemma 5.10 one has

$$T_{u\theta,w_{i,-i}}(\rho) = T_{w,w_{i,-i}}(\rho)T_{z,w_{i,-i}}(\rho)T_{w,z}(*).$$

Here the first two factors on the right hand side are elementary by the Whitehead–Vaserstein lemma ($w_{-2} = z_3 = 0$ together with the corresponding elements in $w_{i,-i}$) and the last one by the Kopeiko–Taddei lemma (w is isotropic and $w_1 = w_{-1} = 0$). This finishes the proof of the proposition.

4^0. Partitions of 1. In this subsection we show that when we replace a unimodular column u by the columns cu, $c \in \mathrm{EU}(2n, R, \Lambda)$, the elements appearing in the propositions 3 and 4 actually generate the whole ring R_0. Our argument reproduces – with an obvious complication due to the presence of an involution – the original argument of A.A.Suslin (see [Tu], Lemma 1, [Kh1], Lemma 3, or [Kh3], pp. 18–20).

The only difference with the proof for the linear case consists in the following. Maximal ideals of the ring R_0 are not necessarily invariant with respect to the involution. As a result it is possible that θ does not belong to a maximal ideal $\mathfrak{m} \in \mathrm{Max}(R_0)$, but $\overline{\theta}$ does. It is technically more convenient to pass to a subring A of R_0, where maximal ideals enjoy this property and to prove a formally stronger fact, that already the products of the multiples of $(cu)_i$ lying in A generate the whole ring A as c ranges over $\mathrm{EU}(2n, R, \Lambda)$.

Let A be the subring, generated by the norms of elements from R_0, i.e. by all $\theta\overline{\theta}$, where $\theta \in R_0$. Clearly A contains also traces of elements from R_0. Indeed, $\theta + \overline{\theta} = (1+\theta)\overline{(1+\theta)} - \theta\overline{\theta} - 1 \in A$. Clearly there are at most two maximal ideals $\mathfrak{m}_1, \mathfrak{m}_2 \in \mathrm{Max}(R_0)$, $\overline{\mathfrak{m}_1} = \mathfrak{m}_2$, lying over a maximal ideal $\mathfrak{m} \in \mathrm{Max}(A)$. For a maximal ideal $\mathfrak{m} \in \mathrm{Max}(A)$ we denote by $A_{\mathfrak{m}}$, $R_{\mathfrak{m}}$, etc. localizations of A, R, etc. with respect to the multiplicative set $S_{\mathfrak{m}} = A \setminus \mathfrak{m}$: $A_{\mathfrak{m}} = {S_{\mathfrak{m}}}^{-1}A$, $R_{\mathfrak{m}} = {S_{\mathfrak{m}}}^{-1}R$, etc. As usual we denote by $\mathrm{Rad}(R)$ the Jacobson radical of a ring R.

Lemma 7.3. *Let (R, Λ) be an almost commutative form ring. Then for every maximal ideal $\mathfrak{m} \in \mathrm{Max}(A)$ the ring $Q_{\mathfrak{m}} = R_{\mathfrak{m}}/\operatorname{Rad}(R_{\mathfrak{m}})$ is classically semisimple and the canonical morphism $\phi = \phi_{\mathfrak{m}} : R \longrightarrow Q_{\mathfrak{m}}$ is surjective.*

Proof. First observe that $R_{\mathfrak{m}}$ is algebraic over $A_{\mathfrak{m}}$. Indeed, being a root of the equation $x^2 - (\theta + \overline{\theta})x + \theta\overline{\theta} = 0$ with the coeffitients in A, every elements $\theta \in R_0$ is algebraic over A. Since R is a finitely generated module over R_0, it is also algebraic over A and our claim follows.

Next we prove that $\mathfrak{m}$ is contained in $\mathrm{Rad}(R_{\mathfrak{m}})$. Indeed, take arbitrary $\mu \in \mathfrak{m}$ and $\xi \in R_{\mathfrak{m}}$. We have to prove that $1 + \mu\xi$ is invertible in $R_{\mathfrak{m}}$. Since $R_{\mathfrak{m}}$ is algebraic, the subalgebra $A_{\mathfrak{m}}(\xi) \subseteq R_{\mathfrak{m}}$ is finite dimensional. Since $\mathrm{Rad}(A_{\mathfrak{m}}) = \mathfrak{m}A_{\mathfrak{m}}$, the Nakayama's lemma implies that $\mathfrak{m}$ is contained in every maximal ideal of $A_{\mathfrak{m}}(\xi)$ and thus $\mu \in \mathrm{Rad}(A_{\mathfrak{m}}(\xi))$. Thus $1 + \mu\xi$ is invertible already in $A_{\mathfrak{m}}(\xi)$.

Now it is easy to prove that ϕ is surjective. In fact, let $\xi/\theta R_{\mathfrak{m}}$, where $\xi \in R$, $\theta \in S_{\mathfrak{m}}$. Since $\theta A + \mathfrak{m} = A$, there exist $\eta \in A$ and $\mu \in \mathfrak{m}$, such that $\theta\eta + \mu = 1$. Multiplying this equality by ξ/θ we get $\xi/\theta = \eta\theta + \mu\xi/\theta \in \eta\theta + \mathrm{Rad}(R_{\mathfrak{m}})$.

Finally observe that $Q_{\mathfrak{m}} = R_{\mathfrak{m}}/\operatorname{Rad}(R_{\mathfrak{m}})$ is a finitely generated module over $(R_0)_{\mathfrak{m}}/\operatorname{Rad}((R_0)_{\mathfrak{m}})$. This last ring is either a field $(R_0)_{\mathfrak{m}}/\mathfrak{m}(R_0)_{\mathfrak{m}}$, or a direct sum of two fields $(R_0)_{\mathfrak{m}}/\mathfrak{m}_1(R_0)_{\mathfrak{m}} \oplus (R_0)_{\mathfrak{m}}/\mathfrak{m}_2(R_0)_{\mathfrak{m}}$. Thus the ring $R_{\mathfrak{m}}/\operatorname{Rad}(R_{\mathfrak{m}})$ is both semisimple and artinian.

We set $\Gamma_{\mathfrak{m}} = \phi_{\mathfrak{m}}(\Lambda_{\mathfrak{m}})$. Then the homomorphism of form rings $\phi_{\mathfrak{m}} : (R, \Lambda) \longrightarrow (Q_{\mathfrak{m}}, \Gamma_{\mathfrak{m}})$ is surjective. Thus we get an epimorphism of the corresponding elementary unitary groups

$$\phi_{\mathfrak{m}} : \mathrm{EU}(2n, R, \Lambda) \longrightarrow \mathrm{EU}(2n, Q_{\mathfrak{m}}, \Gamma_{\mathfrak{m}}).$$

Since $Q_{\mathfrak{m}}$ is semisimple, the elementary unitary group $\mathrm{EU}(2n, Q_{\mathfrak{m}}, \Gamma_{\mathfrak{m}})$ has very strong transitivity properties, see [B1], [V1], [MKV], [HO'M]. In particular for all $n \geq 3$ the group acts transitively on the set of all isotropic unimodular vectors (see [HO'M], Theorem 9.1.3 or [MKV], Theorem 8.1 for much stronger results).

Lemma 7.4. *Let $u \in R^{2n}$ be a unimodular isotropic column over an almost commutative ring R. Then the ideal in R_0 generated by $\theta\overline{\theta}$, where $\theta \in \mathcal{O}((cu)_{-1})$, c runs over* $\mathrm{EU}(2n, R, \Lambda)$, *coincides with R_0.*

Proof. Let $\mathfrak{m} \in \mathrm{Max}(A)$ be any maximal ideal of A. Then the vector $\phi_\mathfrak{m}(u)$ is an isotropic unimodular vector over a classically semisimple ring $Q_\mathfrak{m}$. Then the transitivity of $\mathrm{EU}(2n, Q_\mathfrak{m}, \Gamma_\mathfrak{m})$ and the surjectivity of $\phi_\mathfrak{m}$ imply that there exists $c \in \mathrm{EU}(2n, R, \Lambda)$ such that $\phi_\mathfrak{m}((cu)_{-1}) \in {Q_\mathfrak{m}}^*$, or, in other words, $(cu)_{-1}$ is invertible in $R_\mathfrak{m}$. This means that $((cu)_{-1})^{-1} = \xi/\theta$, where $\xi \in R$, $\theta \in S_\mathfrak{m}$. Then $(cu)_{-1}\xi = \xi(cu)_{-1} = \theta \in S_\mathfrak{m} \cap \mathcal{O}((cu)_{-1})$. Clearly also $\theta\overline{\theta} = \theta^2 \in S_\mathfrak{m} \cap \mathcal{O}((cu)_{-1})$. Thus for every $\mathfrak{m} \in \mathrm{Max}(A)$ there exist $c \in \mathrm{EU}(2n, R, \Lambda)$ and $\theta \in \mathcal{O}((cu)_{-1})$ such that $\theta\overline{\theta} \notin \mathfrak{m}$.

Lemma 7.5. *Let $u \in R^{2n}$ be a unimodular isotropic column over an almost commutative ring R. Then the ideal in R_0 generated by $\overline{\eta}\theta$, where $\eta \in \mathcal{O}((cu)_{-1})$, $\theta \in \mathcal{O}((cu)_{-2})$, and c runs over* $\mathrm{EU}(2n, R, \Lambda)$, *coincides with R_0.*

Proof. We argue as in the preceding lemma. Let $\mathfrak{m} \in \mathrm{Max}(A)$ be any maximal ideal of A. Again the transitivity of $\mathrm{EU}(2n, Q_\mathfrak{m}, \Gamma_\mathfrak{m})$ and the surjectivity of $\phi_\mathfrak{m}$ imply that there exists $c \in \mathrm{EU}(2n, R, \Lambda)$ such that $\phi_\mathfrak{m}((cu)_{-1}), \phi_\mathfrak{m}((cu)_{-2}) \in {Q_\mathfrak{m}}^*$, or, in other words, $(cu)_{-1}$, $(cu)_{-2}$ are invertible in $R_\mathfrak{m}$. This means that $((cu)_{-1})^{-1} = \xi/\eta$, $((cu)_{-2})^{-1} = \zeta/\theta$, where $\xi, \zeta \in R$, $\eta, \theta \in S_\mathfrak{m}$. Then $(cu)_{-1}\xi = \xi(cu)_{-1} = \eta \in S_\mathfrak{m} \cap \mathcal{O}((cu)_{-1})$ and $(cu)_{-2}\zeta = \zeta(cu)_{-2} = \theta \in S_\mathfrak{m} \cap \mathcal{O}((cu)_{-2})$. Clearly $\overline{\eta}\theta = \eta\theta \in S_\mathfrak{m} \cap \overline{\mathcal{O}((cu)_{-1})}\mathcal{O}((cu)_{-2})$. Thus for every $\mathfrak{m} \in \mathrm{Max}(A)$ there exists $c \in \mathrm{EU}(2n, R, \Lambda)$ such that $\overline{\mathcal{O}((cu)_{-1})}\mathcal{O}((cu)_{-2}) \not\subseteq \mathfrak{m}$.

One can note that actually we used very little about the ring $Q_\mathfrak{m}$. Nothing changes in the proof of the last two lemmas if only $\mathrm{EU}(2n, Q_\mathfrak{m}, \Gamma_\mathfrak{m})$ is transitive on the set of isotropic unimodular columns. This is the case, for example, when $n \geq \mathrm{asr}(Q_\mathfrak{m}) + 2$, see [MKV], Theorem 8.1. One may impose various other ring theoretic conditions on R to guarantee the validity of the last two lemmas. For example, the proof works when R is algebraic over R_0 and satisfies some further finiteness conditions, as in [Kh1] – [Kh3], or when R is von Neumann regular, etc.

5^0. Proof of the theorem for the absolute case. Corollary from Lemma 4.4 shows that we have only to prove the theorem in the absolute case. In view of Lemmas 5.2 and 5.5 this amounts to proving that the elements $T_u(\alpha)$ and $T_{uv}(\xi)$, where u and v are the i-th and the j-th columns of a matrix $g \in \mathrm{U}(2n, R, \Lambda)$, $\alpha \in \Lambda$, $\xi \in R$, belong to $\mathrm{EU}(2n, R, \Lambda)$. In fact we will prove a stronger statement that these elements belong to $\mathrm{EU}(2n, R, \Lambda)$ whenever u is a *unimodular* isotropic column and v is orthogonal to u.

First we prove that unipotent elements of *long* root type $T_u(\alpha)$ belong to $\mathrm{EU}(2n,R,\Lambda)$ whenever u is unimodular. Indeed, if $\theta\in\mathcal{O}((cu)_{-1})$, where $c\in\mathrm{EU}(2n,R,\Lambda)$, then

$$T_u(\theta\alpha\overline{\theta})=c^{-1}T_{cu}(\theta\alpha\overline{\theta})c\in\mathrm{EU}(2n,R,\Lambda)$$

by Proposition 3. But since u is unimodular Lemma 7.4 shows that $\theta\overline{\theta}$ generate the unit ideal in R_0 as c ranges over $\mathrm{EU}(2n,R,\Lambda)$. Choose $c_1,\dots,c_t\in\mathrm{EU}(2n,R,\Lambda)$ such that there exists a partition of 1 of the form $\theta_1\overline{\theta}_1+\ldots+\theta_t\overline{\theta}_t=1$, $\theta_h\in\mathcal{O}((c_hu)_{-1})$. Then

$$T_u(\alpha)=\prod T_u(\theta_h\alpha\overline{\theta}_h)=\prod {c_h}^{-1}T_{c_hu}(\theta_h\alpha\overline{\theta}_h)c_h,$$

where all the factors on the right hand side belong to $\mathrm{EU}(2n,R,\Lambda)$.

Now we prove that unipotent elements of *short* root type $T_{uv}(\xi)$ belong to $\mathrm{EU}(2n,R,\Lambda)$ whenever u is unimodular. Indeed, if $\overline{\eta}\in\mathcal{O}((cu)_{-1})$, $\theta\in\mathcal{O}((cu)_{-2})$, where $c\in\mathrm{EU}(2n,R,\Lambda)$, then

$$T_{uv}(\overline{\eta}\theta\xi)=c^{-1}T_{cu,cv}(\overline{\eta}\theta\xi)c\in\mathrm{EU}(2n,R,\Lambda)$$

by Proposition 4. But since u is unimodular Lemma 7.5 shows that $\theta\overline{\eta}$ generate the unit ideal in R_0 as c ranges over $\mathrm{EU}(2n,R,\Lambda)$. Choose $c_1,\dots,c_t\in\mathrm{EU}(2n,R,\Lambda)$ such that there exists a partition of 1 of the form $\theta_1\overline{\eta}_1+\ldots+\theta_t\overline{\eta}_t=1$, $\overline{\eta}_h\in\mathcal{O}((c_hu)_{-1})$, $\theta_h\in\mathcal{O}((c_hu)_{-2})$. Then

$$T_{uv}(\xi)=\prod T_{uv}(\theta_h\xi\overline{\eta}_h)=\prod {c_h}^{-1}T_{c_hu,c_hv}(\theta_h\xi\overline{\eta}_h)c_h,$$

where all the factors on the right hand side belong to $\mathrm{EU}(2n,R,\Lambda)$.

This finishes the proof of the theorem 1 for the absolute case and thus, in view of § 4, for all cases.

Remark. In fact nothing changes in the proof for all other situations mentioned in the preceding subsection. In particular, we have proven the following result. Let (R,Λ) be an form ring. Assume that $n\geq 3$ is such that for all maximal ideals $\mathfrak{m}\in\mathrm{Max}(A)$ the group $\mathrm{EU}(2n,R_{\mathfrak{m}},\Lambda_{\mathfrak{m}})$ acts transitively on the set of all unimodular isotropic columns of height $2n$ over $R_{\mathfrak{m}}$. Then for any form ideal (I,Γ) the corresponding elementary subgroup $\mathrm{EU}(2n,I,\Gamma)$ is normal in $\mathrm{U}(2n,R,\Lambda)$ and

$$\mathrm{EU}(2n,I,\Gamma)=[\mathrm{EU}(2n,R,\Lambda),\mathrm{CU}(2n,I,\Gamma)].$$

As mentioned after the proof of Lemma 7.5, the condition of the corollary is satisfied, for example, when $n\geq\operatorname{asr}R_{\mathfrak{m}}+2$ for all $\mathfrak{m}\in\mathrm{Max}(A)$, see [MKV]. One could state many further generalizations like that, in the style of [V2], or [Kh1] – [Kh3].

Bibliography

[**A1**] Abe E., *Chevalley groups over local rings*, Tôhoku Math. J. **21** (1969), no. 3, 474–494.

[**A2**] Abe E., *Coverings of twisted Chevalley groups over commutative rings*, Sci. Repts. Tokyo Kyoiku Daigaku **13** (1977), 194–218.

[**A3**] Abe E., *Chevalley groups over commutative rings*, Proc. Conf. Radical Theory, (Sendai – 1988), Uchida Rokakuho Publ. Comp., Tokyo, 1989, pp. 1–23.

[**A4**] 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.

[**B6**] Bak A., *Computing sandwiches in sandwich classifications*, submitted.

[**BV**] Bak A., Vavilov N.A., *Normality for elementary subgroup functors*, Mathe Proc. Cambridge Phil. Soc. **118** (1995), no. No. 1, 35 - 47.

[**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.

[**BV1**] 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.

[**BV2**] Borewicz Z.I., Vavilov N.A., *On the subgroups of the general linear group over a commutative ring*, Sov. Math. Dokl. **26** (1982), no. 3, 679–681.

[**BV3**] 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.

[Ca] Carter R.W., *Simple groups of Lie type*, Wiley, London et al., 1972.

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

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

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

[Di] Dieudonné J., *La géométrie des groupes 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).

[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.

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

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

[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).

[KM1] Klein I.S., Mikhalev A.V., *Steinberg orthogonal group over a ring with involution*, Algebra and Logic **9** (1970), no. 2, 88–103.

[KM2] Klein I.S., Mikhalev A.V., *Unitary Steinberg group over a ring with involution*, Algebra and Logic **9** (1970), no. 5, 305–312.

[Ko1] Kopeiko V.I., *The stabilization of symplectic groups over a polynomial ring*, Math. U.S.S.R. Sbornik **34** (1978), 655–669.

[Ko2] Kopeiko V.I, *On a theorem of Suslin*, J. Sov. Math. **132** (1983), 119–121.

[Ko3] Kopeiko V.I., *Unitary and orthogonal groups over rings with involution*, Algebra and Discrete Math., Kalmytsk. Gos. Univ., Elista, 1985, pp. 3–14, (in Russian).

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

[MKV] Magurn B., van der Kallen W., Vaserstein L.N., *Absolute stable rank and Witt cancellation for noncommutative rings*, Invent. Math. **91** (1988), 525–543.

[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.

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

[Sv] Stepanov A.V., *Stability conditions in the theory of linear groups over rings*, Ph.D. Thesis. Leningrad State Univ., 1987.

[Su1] Suslin A.A., *On a theorem of Cohn*, J. Sov. Math. **17** (1981), no. 2.

[Su2] Suslin A.A., *On the structure of the general linear group over polynomial rings*, Soviet Math. Izv. **41** (1977), no. 2, 503–516..

[Su3] 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.

[Sw] Swan R.G., *Generators and relations for certain special linear groups*, Adv. Math. **6** (1971), 1–77.

[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 groupes élémentaire dans les groupes 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., *Stabilization of unitary and orthogonal groups over a ring*, Math. U.S.S.R. Sbornik **10** (1970), 307–326.

[V2] Vaserstein L.N., *On normal subgroups of* GL_n *over a ring*, Lecture Notes Math. **854** (1981), 456–465.

[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.

[VM] Vaserstein L.N., Mikhalev A.V., *The normal subgroups of the orthogonal group over a ring with involution*, Algebra and Logic **9** (1970), no. 6, 375–377.

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

[Va2] Vavilov N.A., *On the problem of normality of the elementary subgroup in a Chevalley group*, Algebraic and Discrete Systems, Ivanovo Univ., 1988, pp. 7–25, (in Russian).

[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., *Linear groups over general rings.* I, II, In preparation.

[Va7] 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.

[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: bak@mathematik.uni-bielefeld.de
vavilov@mathematik.uni-bielefeld.de