Algebra Colloquium **7**:2 (2000) 159–196

Algebra Colloquium

Structure of Hyperbolic Unitary Groups I: Elementary Subgroups

Anthony Bak
Fakultät für Mathematik, Universität Bielefeld
33615 Bielefeld, Deutschland
E-mail: bak@mathematik.uni-bielefeld.de

Nikolai Vavilov
Department of Mathematics and Mechanics
University of Sanct-Petersburg, Petrodvorets 198904, Russia
E-mail: vavilov@pdmi.ras.ru

Received 23 February 1999
Revised 25 June 1999

Communicated by Fu-an Li

Abstract. This is the first in a 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 the elementary subgroup of the hyperbolic unitary group over a form ring. In particular, using a variant of Suslin's patching method, we prove standard commutator formulae for relative elementary subgroups.

2000 Mathematics Subject Classification: 20G35, 20H25

Keywords: unitary group, unitary transvections, elementary subgroup

1 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 [6–8, 15, 36]). Our primary objectives are two-fold. On the one hand, we plan to update the foundational results from [6, 15], removing whenever possible stability conditions or replacing them by weaker commutativity or finiteness conditions. The excellent recent exposition [36] concentrates on Steinberg groups, the structure of unitary groups over division rings and isomorphism theory, but does not contain proofs of structure results in full generality. We fill this gap. In particular, we carry over to unitary groups over form rings the main structure theorems of the past 30 years for general linear groups over associative rings and their extensions to the usual classical groups [3, 4, 9, 10, 19, 23, 26–30, 32–34, 37–39, 42, 44, 46–48, 52, 57, 60, 62–65, 67–73, 76, 78–83, 85, 86]. Furthermore, we generalize to unitary groups over form rings more specialized structure results relating to the description of various classes of subgroups. In particular, we carry over results of [17–19, 29, 76, 79] (see [81, 83, 87] for many further references).

In the first half of this article, we fix the general framework for the rest of the series. In the second half, we study elementary subgroups $\mathrm{EU}(2n, R, \Lambda)$ of hyperbolic unitary groups $\mathrm{U}(2n, R, \Lambda)$, as well as 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 center).

Theorem 1.1. *Let* (R, Λ) *be an almost commutative form ring and* $n \geq 3$*. 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)*.*

The concept of a form ideal and its application to describing and analyzing normal subgroups of unitary groups appears for the first time in [6]. Here, the conclusion of the theorem above, as well as the sandwich classification of normal subgroups of unitary groups, which will appear in the second paper of this series, was obtained under the condition that R is almost commutative and $n \geq \max(3, \dim(\mathrm{Cent}(R)) + 2)$. Unfortunately, [6] was never published and is not easily available, especially in Russia and China. This did a lot of harm. In fact, many works appearing *up till the late eighties* were proving structure theorems for classical groups over rings covered in [6] such as zero-dimensional ones. We do not cite these publications in our bibliography (see, for example, references in [36] and also in [23, 70, 71, 83]).

In [12], we give another proof of the theorem above, based on a variant of "localization and patching". There, we also show that the theorem remains true for $n \geq 2$ under the additional assumption $R\Lambda + \Lambda R = R$. However, in the current paper, we give a direct global proof developing ideas of Suslin (cf. [37, 39, 57, 65]). 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 assumes only some transitivity properties of the elementary unitary group. This provides lots of room for generalizations in the style of [67] or [37–39]. As we will see in Sec. 8, the proof embraces wider classes of rings than just the almost commutative ones. Another advantage is that, unlike localization methods, the global approach gives explicit formulae for decomposing a transvection into elementaries (incidentally, the formulae from Secs. 5–8 play an important role in subsequent parts of this paper).

Theorem 1.1 was known in many special cases. First, as mentioned above, i.e., when n is large relative to some sort of dimension on the ring R, the result was established by Bak [6, 7] (compare also [15]). For the general linear group $\mathrm{GL}(n, R)$, the analogous result was discovered by Bass [13, 14], and for some hyperbolic classical groups, the analogous result was obtained by Vaserstein [66]. Refer to [80, 83] for a systematic bibliography.

Suslin [57] and [65] proved 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 [19, 33, 36–39, 51, 55, 60, 67, 80, 83] for various proofs of this result and its generalizations to wider classes of rings). For the commutative case, the original proof of Suslin [57] is based on solving linear equations over rings (see [60] for a very elegant reformulation in terms of anti-symmetric matrices). A particularly simple direct proof of Suslin's result based on decomposing unipotents was found in 1987 by Stepanov (see [55, 85] and the expositions in [56, 80, 83]).

The proof in [65] of the normality of $\mathrm{E}(n, R)$ requires "patching". The author states explicitly that the result and proof are due to Suslin. The paper [37] improves the result slightly (pushing it to rings *algebraic* over their centers). It also corrects some misprints from [65] and organizes the proof in a better way. However, many ring theoretical arguments are omitted in [37] as well. To the best of our knowledge, the most systematic exposition of these ideas is contained in [39]. Another version of this method was proposed by Vaserstein [67] and was christened "localization and patching". It consists of throwing in polynomial variables, passing to localizations with respect to central multiplicative subsets, and using corresponding local results. In general, the relationship of the patching method of Suslin to the localization and patching method of Vaserstein is the same as the relationship of the solution of Serre's problem by Suslin to that by Quillen (see [45]).

For historical accuracy, note that the first application of Quillen's method to linear groups is also due to Suslin. It appeared in his remarkable paper [57] dedicated to the K_1 analog of Serre's problem. Many authors seem to ignore that Suslin's proof of "Quillen's Theorem" (see Sec. 3 in [57], especially, Lemma 3.3 and proofs of Lemmas 3.4 and 3.7) already contained *all* the ingredients of localization and patching. This method was successfully used in the late seventies by Suslin's students Kopeiko and Tulenbaev, and later by Abe, Costa, and others.

The methods above operate essentially in terms of the center of the ground ring, as do the published works of the Moscow School. In order for this to work, one has to impose some commutativity or finiteness conditions on the rings considered. On the other hand, Golubchik has successfully used *non-commutative* localizations (more specifically, *Ore localizations*, in particular in [31]) to describe normal subgroups [26–30]. He and Mikhalev told us that they had applied non-commutative localization to establish 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 Dedekind rings (compare [21, 58, 61]). On the other hand, using *universal* localizations, Gerasimov [25] (see also [83]) 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 also furnish counterexamples to the corresponding results for unitary groups. 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 weakly finite if any one-sided invertible square matrix over R is two-sided invertible; see [19, 22, 83] for a discussion of this condition and its role in the theory of linear groups). Observe that normality fails in $\mathrm{SO}(4, R)$ (see [43]) by the same reasons as for the group $\mathrm{GL}(2, R)$.

Suslin's result on the normality of $\mathrm{E}(n, R)$ in $\mathrm{GL}(n, R)$ was generalized almost immediately by Suslin and Kopeiko [42, 60] from linear groups to hyperbolic symplectic groups $\mathrm{Sp}(2n, R)$ with $n \geq 2$ and orthogonal groups $\mathrm{SO}(2n, R)$ with $n \geq 3$ 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 $\lambda = 1$ and $\Lambda = \Lambda_{\min}$). The result for the symplectic group has been partially reproven in [62]. (Compare also [32, 36, 44, 46, 47, 56, 70, 71, 76, 79, 80], where one can find other proofs of these results and some generalizations to groups that are not necessarily hyperbolic.)

The results for split classical groups fall under the umbrella of Chevalley groups. The normality of absolute elementary subgroups of simple Chevalley groups of rank at least 2 over an arbitrary commutative ring R was first proven by Taddei [63, 64] who used "localization and patching" (compare also [3, 4, 69]). Another approach to prove this result, based on Stepanov's idea of decomposition of unipotents, was proposed in [56, 80, 82, 84, 85].

The most general published result for automorphism groups of forms is due to Golubchik and Mikhalev [32] (compare also [36, 44]). They use a version of Suslin's patching method. The proof is not easy to follow. The existence of certain elements λ, s, t is claimed, but these elements are never displayed. The formulae suggest that $1 + \lambda = (1 + \lambda_2)(1 + \overline{\lambda}_1)(1 + \lambda_1)$, but it is not clear why the condition in the next paragraph is satisfied. Their results are phrased in terms of arbitrary rings with involution (not just the ring $\mathrm{M}(n, R)$), and assume the Witt index is at least 2 (in their setting, it is expressed in terms of the behavior of idempotents under the involution). In our setting, their main result applies to the groups where $\lambda = -1$, $\Lambda = \Lambda_{\max}$ 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 Vaserstein [72], but a proof never appeared.

The commutator formula in the theorem above was discovered at the stable level by Bass and Bak. For classical groups over commutative rings, it was obtained independently by Vaserstein, Borevich, Vavilov and Li (see [18, 19, 46, 47, 67, 70, 71, 76, 79]). After the work of Stein [54], it is a common understanding that questions about relative groups may be reduced to those for absolute groups by an appropriate change of rings. The first applications of this idea to demonstrating the normality of relative elementary subgroups were given by Milnor [53], and Suslin and Kopeiko [60] (compare also [34, 36, 68, 80, 83]).

The rest of this paper is organized as follows. In Secs. 2 and 3, we reproduce fundamental definitions and notation concerning unitary groups over form rings and their elementary subgroups. Most of the material comes from [6–8, 15, 34, 36], but is updated and adapted to our needs. In Secs. 4 and 5, we discuss form ideals and the corresponding relative groups. In particular, we prove the usual results about the generation of relative elementary groups and reduce the relative case of the theorem above to the absolute one. In Sec. 6, we introduce and study a notion of ESD-transvections, which is slightly more general than the usual one. In Sec. 7, 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 containing zero elements. Finally, in Sec. 8, 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. In fact, we have to do so since we define the group $\mathrm{U}(2n, R, \Lambda)$ with respect to the ordered basis $e_1, \dots, e_n, e_{-n}, \dots, e_{-1}$, where e_i is orthogonal to each e_j except e_{-i}, and the inner product of e_{-i} and e_i is 1. In all previously published works, where hyperbolic unitary groups over form rings were considered, either the ordered basis $e_1, e_{-1}, \dots, e_n, e_{-n}$ or the ordered basis $e_1, \dots, e_n, e_{-1}, \dots, e_{-n}$ is used. The ordering in this paper is inspired by [20] and is more natural in many respects than the alternatives since, for example, under this ordering, 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 subsequent papers by the authors dedicated to the study of the normal structure of $\mathrm{U}(2n, R, \Lambda)$ and its Steinberg group.

2 Hyperbolic Unitary Groups

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

1^0. Λ-quadratic forms. This notion was invented by Bak and first appeared in [6]. Let R be a (not necessarily commutative) associative ring with 1. For natural numbers m, n, we denote by $\mathrm{M}(m, n, R)$ the additive group of $m \times n$ matrices with entries in R. In particular, $\mathrm{M}(n, R) = \mathrm{M}(n, n, R)$ 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} its entry in the position (i, j). Let $e = e_n$ be the identity matrix and e_{ij} 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 an involution on R (see [14] 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 $\overline{\overline{\alpha}} = \alpha$.

The ingredient which distinguishes Λ-quadratic forms from ordinary quadratic forms is the notion of a form parameter Λ, which we define next.

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

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

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 ith coordinate is 1, while all other coordinates are zeros. Following [20], 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.

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 Λ-*quadratic form* $q : V \to R/\Lambda$ by $q(u) = f(u,u) \mod \Lambda$. In other words, h is the form with the Gram matrix $\begin{pmatrix} 0 & p \\ \lambda p & 0 \end{pmatrix}$, or

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

This form is in fact λ-hermitian, i.e., $h(u,v) = \lambda\overline{h(v,u)}$ for any $u, v \in V$. This fact will be used in 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:

$$q(u) = \overline{u}_1 u_{-1} + \cdots + \overline{u}_n u_{-n} \mod \Lambda.$$

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 definitions of h and q. It is crucial for our calculations in Sec. 4.

Lemma 2.1. *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. These groups were discovered by Bak and appeared first in [6]. 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-anti-hermitian matrices $\mathrm{AH}(n,R,\Lambda)$ which is defined as

$$\mathrm{AH}(n,R,\Lambda) = \{a \in \mathrm{M}(n,R) \mid a = -\lambda a^*,\ a_{ii} \in \Lambda\ \forall\ i = 1,\dots,n\}\,.$$

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 provides a matrix description of the elements of $\mathrm{U}(2n,R,\Lambda)$ which is due to Bak [6]. Write a matrix g of degree $2n$ in the block form $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ with respect to the partition (n,n), where a, b, c, d are matrices of degree n over R.

Lemma 2.2. *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 obtain condition (1).

Suppose g already stabilizes h. 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

$$\begin{aligned} z &= \begin{pmatrix} a & b \\ c & d \end{pmatrix}^* \begin{pmatrix} 0 & p \\ 0 & 0 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} - \begin{pmatrix} 0 & p \\ 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} a^*pc & a^*pd - p \\ b^*pc & b^*pd \end{pmatrix} \in \mathrm{AH}(2n, R, \Lambda). \end{aligned}$$

Now condition (1) implies that this last matrix is Λ_{max}-anti-hermitian, i.e., belongs to $\mathrm{AH}(2n, R, \Lambda_{\mathrm{max}})$. 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

$$\begin{aligned} d^*pa + \overline{\lambda} b^*pc &= p, \\ d^*pb + \overline{\lambda} b^*pd &= 0, \\ \lambda c^*pa + a^*pc &= 0, \\ \lambda c^*pb + a^*pd &= p. \end{aligned}$$

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

To belong to $\mathrm{AH}(2n, R, \Lambda)$, the matrix z has to satisfy the 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 Lemma 2.2 may be replaced by the condition that the diagonal coefficients of the matrices a^*pc and b^*pd

lie in Λ. In fact, Lemma 2.2 may be stated in an equivalent but slightly different form.

Lemma 2.3. *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) $g'_{ij} = \lambda^{(\varepsilon(j)-\varepsilon(i))/2}\overline{g}_{-j,-i}$ *for all* $i, j = 1, \dots, -1$,

(2) $\sum_{1\leq i\leq n} \overline{g}_{ij} g_{-ij} \in \Lambda$ *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.

3⁰. Polarity map. Sometimes it is convenient to express conditions from Lemmas 2.2 and 2.3 in terms of columns.

A vector $u \in V$ is called *isotropic* if $q(u) = 0$ or, in other words, $f(u, u) \in \Lambda$. 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., $(u, v) = 0$. Since the form h is λ-hermitian, 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 ith and jth columns of $g \in \mathrm{U}(2n, R, \Lambda)$, respectively, where $i \neq -j$. Then u and v are orthogonal.

Lemma 2.4. *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*

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

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 Lemma 2.4 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 [22], 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 $\widetilde{\ }\colon R^{2n} \to {}^{2n}R$ 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 $\widetilde{u} = (\lambda\overline{u}_{-1}, \dots, \lambda\overline{u}_{-n}, \overline{u}_n, \dots, \overline{u}_1)$. Clearly, one has $h(u, v) = \widetilde{u}v$.

It is clear that the polarity map is *involutory linear*, i.e., $\widetilde{u + v} = \widetilde{u} + \widetilde{v}$ and $\widetilde{u\xi} = \overline{\xi}\widetilde{u}$ for all $u, v \in V$ and $\xi \in R$. This property is often used in the sequel without explicit reference. Now we can reformulate Lemma 2.3 in 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 ith coordinate is 1, while all other coordinates are zeros.

Lemma 2.5. *For any* $v \in V$ *and* $g \in \mathrm{U}(2n, R, \Lambda)$, *one has* $\widetilde{gv} = \widetilde{v}g^{-1}$.

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 ith 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. □

3 Elementary Hyperbolic Unitary Group

In this section, we recall the 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., $\varepsilon(i) = 1$ if $i \in \Omega^+$ and $\varepsilon(i) = -1$ if $i \in \Omega^-$.

The transvections $T_{ij}(\xi)$ correspond to the pairs $i, j \in \Omega$ such that $i \neq j$. Moreover, if $i \neq -j$, 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 hand, 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 $H(a) \in \mathrm{U}(2n, R, \Lambda)$. With this notation, the transvection $T_{ij}(\xi)$ where $i, j \in \Omega^+$ 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.

Next, let $i \in \Omega^+$ 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 $T_{i,-i}(\alpha)$ for $i \in \Omega^+$ as follows. The transvection $T_{i,-i}(\alpha)$ 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 both types above come from the unipotent embedding of the (additive) group $\mathrm{AH}(n, R, \overline{\Lambda})$ of $\overline{\Lambda}$-anti-hermitian 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 $\xi e_{i,-j} - \overline{\lambda}\,\overline{\xi} e_{-j,i}$ for $\xi \in R$ and αe_{ii} for $\alpha \in \overline{\Lambda}$, respectively.

Finally, let $i \in \Omega^-$ 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 $T_{i,-i}(\alpha)$ for $i \in \Omega^-$ as follows. The transvection $T_{i,-i}(\alpha)$ is the usual linear transvection $t_{i,-i}(\alpha) = e + \alpha e_{i,-i}$, where α runs over Λ. The transvections of both types above come from the unipotent embedding of the (additive) group $\mathrm{AH}(n, R, \Lambda)$ of Λ-anti-hermitian 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 $X^-(c) \in \mathrm{U}(2n, R, \Lambda)$. Clearly, both types of the transvections above are in the image of X^-. Namely, they are the images of the matrices $\xi e_{-i,j} - \lambda\overline{\xi} e_{j,-i}$ for $\xi \in R$ and $\alpha e_{-i,-i}$ for $\alpha \in \Lambda$, 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

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

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

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

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

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

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 a short and a long roots whose sum is a root, respectively.

First, we consider the case of two short roots whose sum is a short root. Let $i, h \neq \pm j$ and $i \neq \pm h$. 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 obtain three equivalent forms of this relation:

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

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

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

Finally, we turn to the case of a long root 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 the following alternative forms, to which we also refer as (R6):

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

The relations (R1)–(R6) define the unitary Steinberg group. In fact, the calculations, where only these relations are used, already hold for the Steinberg group and not just for the elementary unitary group (compare with (3.16) and (3.17) of [8]).

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*, denoted by $\mathrm{EU}(2n, R, \Lambda)$. It was first constructed in [6].

Lemma 3.1. *For any* $b \in \mathrm{AH}(n, R, \overline{\Lambda})$ *and* $c \in \mathrm{AH}(n, R, \Lambda)$*, one has* $X^+(b)$, $X^-(c) \in \mathrm{EU}(2n, R, \Lambda)$.

Proof. Indeed, $X^+(b) = \prod T_{i,-j}(b_{ij})$ and $X^-(c) = \prod T_{-i,j}(c_{ij})$, where the first product is taken over all $i, j \in \Omega^+$ such that $i \leq j$, while the second one is taken over all $i, j \in \Omega^+$ such that $i \geq j$. □

As usual, $\mathrm{E}(n, R)$ denotes the elementary subgroup of $\mathrm{GL}(n, R)$, i.e., the subgroup generated by all elementary transvections $t_{ij}(\xi)$ for $1 \leq i \neq j \leq n$ and $\xi \in R$. By definition, the image $H(g)$ of any $g \in \mathrm{E}(n, R)$ under the hyperbolic embedding lies in $\mathrm{EU}(2n, R, \Lambda)$. Actually, unless $n = 2$, it already lies 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 [15].

Lemma 3.2. *Suppose either* $n \neq 2$ *or* $R = \Lambda R + R\Lambda$*. Then*

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

Proof. Denote the right-hand side of this formula by G. We have to prove that $T_{ij}(\xi)$ belongs to G for any $i, j \in \Omega^+$ with $i \neq j$ and $\xi \in R$. If $n = 1$, there is nothing to prove. Thus, we may assume $n \geq 2$. If $n \geq 3$, we may choose $h \in \Omega^+$ with $h \neq i, j$. 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. When $n = 2$, (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 short root unipotents. However, under some assumptions on Λ, only the short root unipotents *or* the long root unipotents would suffice.

Lemma 3.3. *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 $j \neq \pm i$. By (R5), one has $T_{i,-i}(\alpha) = [T_{ij}(\xi), T_{-j,i}(1)]$. □

Of course, $\mathrm{EU}(2n, R, \Lambda)$ 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 (resp., short) root element is called a *long* (resp., *short*) *root element* or a *long* (resp., *short*) *root unipotent*.

Lemma 3.4. *Assume* $\Lambda R + R\Lambda = R$. *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 the formulae appearing in the proof of Lemma 3.2 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. □

4 Congruence Groups

In this section, we recall the definitions of form ideals and their corresponding relative groups (see [6–8, 15]), which play a crucial role in describing the normal subgroups of unitary groups.

1⁰. Form ideals. Let (R, Λ) be a form ring. Let I be a (two-sided) ideal in R, which is invariant with respect to the involution, i.e., $\overline{I} = I$. Set $\Gamma_{\max} = I \cap \Lambda$ and $\Gamma_{\min} = \big\{\xi - \lambda\overline{\xi} \mid \xi \in I\big\} + \big\{\xi\alpha\overline{\xi} \mid \xi \in I,\ \alpha \in \Lambda\big\}$. By definition, $\Gamma_{\min}$ and $\Gamma_{\max}$ depend both on the *absolute* form parameter Λ and the 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.

Form ideals play the same role in the category of form rings as (two-sided) ideals in the category of rings. This notion was introduced by Bak [6–8]. It is crucial for the description of normal subgroups of $\mathrm{U}(2n, R, \Lambda)$, $n \geq 3$. It was used by Bass [15] under the name of a *unitary ideal*. For commutative rings with trivial involution, the notion of *special submodule* is essentially equivalent to the notion of relative form parameter and was introduced independently by Abe [1, 3–5]. In the book of Hahn and O'Meara [36], they are simply called *ideals* of form rings. Vaserstein [72] 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 [23], Costa and 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 $\mathrm{U}(2n, R, \Lambda)$ for $n \geq 3$.

2^0. Doubling a form ring. To treat the relative groups corresponding to form ideals, we have to recall some notation related to Stein's relativization [54].

First, let I be any ideal of a ring R. Then we can define the *double* $R \times_I R$ of R *along* I by the Cartesian square:

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

In a more down to earth language, $R \times_I R$ consists of all pairs $(a, b) \in R \times R$ such that $a \equiv b \pmod{I}$, i.e., $R \times_I R = \big\{(a, b) \in R \times R \mid a - b \in I\big\}$ with the component-wise operations of addition and multiplication 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 the *semidirect product* $R \ltimes I$ of R and I as the set of pairs (a, c) for $a \in R$ and $c \in I$ with component-wise addition and multiplication given by $(a, c)(b, d) = (ab, ad + cb + cd)$.

Lemma 4.1. *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 than 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, relative questions for the ideal I in R can be reduced to absolute ones for the ring $R \times_I R$.

Now let (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:

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

It is easy to observe that $\Lambda \times_\Gamma \Lambda$ is a form parameter in $R \times_I R$ (see [36, Lemma 5.2.15]).

Lemma 4.2. *$(R \ltimes I, \Lambda \ltimes \Gamma)$ is a form ring with respect to the componentwise 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 [36, Lemma 5.2.12]). Then we have a *commutative* square of form rings:

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

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 (resp., Cartesian) square of form rings above leads to the commutative (resp., Cartesian) square of groups U_{2n}. This will be amply used in the rest of this section and in Sec. 5.

3^0. Principal congruence subgroups. These groups were introduced in [6]. 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:

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

One can give the following characterization of $\mathrm{U}(2n, I, \Gamma)$ analogous to Lemma 2.2.

Lemma 4.3. *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 2.2. □

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 obtain the following split exact sequence of groups (see [36, Lemma 5.2.17]):

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

This is precisely 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 Bak in Theorem 4.1.4 of [7] by a direct computation. However, there is a more elegant approach based on a variant of Stein's relativization. Below, we reproduce the argument (compare also [36, Sec. 5.2] and [34, Lemma 2.6]).

Lemma 4.4. *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:

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

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

4^0. Full congruence subgroups. First, let G be any group, and F and H subgroups of G. One can define $C_G(F, H) = \{g \in G \mid [g, f] \in H\ \forall f \in F\}$. In general, $C_G(F, H)$ is only a subset of G. However, if H is a *normal* subgroup of G, then the standard equalities for commutators imply that $C_G(F, H)$ is a subgroup of G. We will use the notation $C_G(F, H)$ *only* when $H \trianglelefteq G$. Moreover, if 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 $G = \mathrm{U}(2n, R, \Lambda)$. Further, let (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 is that given in [6]. Later, this group (also denoted by $\mathrm{U}'(2n, I, \Gamma)$ or $\mathrm{U}^{\sim}(2n, I, \Gamma)$) has been defined slightly differently. For example, in [36],

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

But in interesting situations, these groups coincide as we will 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 center of the group $\mathrm{GL}(n,R/I)$ under the 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 center of the group $U(2n,R/I,\Lambda/\Gamma_{\max})$ under the reduction homomorphism $U(2n,R,\Lambda)\to U(2n,R/I,\Lambda/\Gamma_{\max})$.

5 Relative Elementary Groups

In this section, we define the relative elementary groups of Bak [6]. They are the key to establishing the sandwich classification theorem for $\mathrm{EU}(2n,A,\Lambda)$-normal subgroups of $\mathrm{U}(2n,A,\Lambda)$, which will appear in the second paper of this series. Below, 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 $T_{ij}(\xi)$ is called *elementary of level* (I,Γ) if $\xi\in I$, and moreover, $\xi\in\lambda^{-(\varepsilon(i)+1)/2}\Gamma$ if $i=-j$. Denote the subgroup generated by all (I,Γ)-elementary transvections by

$$\mathrm{FU}(2n,I,\Gamma)=\left\langle T_{ij}(\xi)\mid \xi\in I \text{ with } \xi\in\lambda^{-(\varepsilon(i)+1)/2}\Gamma \text{ if } i=-j\right\rangle.$$

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 $\mathrm{EU}(2n,I,\Gamma)\le\mathrm{U}(2n,I,\Gamma)$.

The following result is a relative analog of Lemma 3.2.

Lemma 5.1. *Suppose either $n\neq 2$ or $I=\Lambda I+I\Lambda$. Then*

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

Proof. For $n\neq 2$, the proof of Lemma 3.2 carries over without any changes. For $n=2$, one takes $\xi\in I$, $\alpha\in\Lambda$, and observes that $\xi\alpha\overline{\xi}\in\Gamma_{\min}\le\Gamma$. Now the same argument as in Lemma 3.2 shows that $T_{ij}(\Lambda I+I\Lambda)$ is contained in the right-hand side for any $i,j\in\Omega^+$ with $i\neq j$. □

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

Lemma 5.2. *Suppose 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 obtain that $T_{ij}(\zeta)$ for $i \neq \pm j$ and $\zeta \in I$ 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 5.3. *Suppose $\mathrm{EU}(2n, I_1, \Gamma_1) \leq \mathrm{U}(2n, I_2, \Gamma_2)$. 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 $\xi \in \lambda^{-(\varepsilon(i)+1)/2}\Gamma$, $\zeta \in \lambda^{-(\varepsilon(i)+1)/2}\Lambda$ otherwise. (Starting from this point, we usually do not distinguish between the cases $i \neq -j$ and $i = -j$. We simply write $\xi \in I$ or $\zeta \in R$ assuming $\xi \in \Gamma$ and $\zeta \in \Lambda$ automatically when $i = -j$.) In general, these matrices do not belong to $\mathrm{FU}(2n, I, \Gamma)$. However, this is essentially the unique counterexample. The following result is a unitary version of a result by Suslin and Vaserstein [75] (see also [10, 67, 83]).

Proposition 5.1. *Let $n \geq 3$. For any form ideal (I, Γ), the corresponding relative elementary group $\mathrm{EU}(2n, I, \Gamma)$ is generated by all matrices $Z_{ij}(\xi, \zeta)$, where either $\xi \in I$, $\zeta \in R$ and $i \neq \pm j$, or $\xi \in \lambda^{-(\varepsilon(i)+1)/2}\Gamma$, $\zeta \in \lambda^{-(\varepsilon(i)+1)/2}\Lambda$ and $i = -j$.*

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 ${}^{x}T_{ij}(\xi) = xT_{ij}(\xi)x^{-1}$, where $x \in \mathrm{EU}(2n, R, \Lambda)$, $\xi \in I$ and $i \neq j$. We will proceed by induction on the length t of a shortest expression of x as a product of elementary matrices. When

$t = 0$, there is nothing to prove. Suppose $t = 1$. 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 $z = Z_{ij}(\xi, \zeta)$. 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 ${}^{ab}c = {}^{a}[b, c] \cdot {}^{a}c$ valid for any three elements a, b, c of a group. Thus,

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

On the other hand, since y is shorter than x, the commutator $[y, T_{ij}(\xi)]$ is a product of factors of the form $Z_{lm}(\omega, \vartheta)$, where $\omega \in I$, $\vartheta \in R$, $l \neq m$. Now for $w = {}^{T_{hk}(\zeta)}Z_{lm}(\omega, \vartheta)$, we have three options:

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 $w = Z_{lm}(\omega, \vartheta + \zeta)$.

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 $\mathrm{EU}(2n, I, \Gamma)$ (see [10] or [83] for a detailed calculation).

The proof for the case $l = -m$ is similar and even easier. Pick up $p \neq \pm l$ and express $T_{l,-l}(\omega) = [T_{p,-p}(-\mu\omega),\ T_{-p,-l}(1)]T_{pl}(\mu\omega)$ as in Lemma 5.2, 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

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

Decomposing this commutator according to the formula $[a, bc] = [a, b]\cdot{}^{b}[a, c]$, one obtains 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 completes 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 Stein [54] and was applied to establish the normality of relative elementary subgroups by Milnor [53], Suslin and Kopeiko [60], 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 [36, 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 [53, Lemma 4.2] for the linear case) reduces questions about relative elementary subgroups to ones about absolute elementary subgroups.

Lemma 5.4. *Let (I,Γ) be a form ideal of a form ring (R,Λ). Then*

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

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

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

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 $(e,b)=(c_1,c_1d_1)\cdots(c_t,c_td_t)$, where $c_i\in\mathrm{EU}(2n,R,\Lambda)$ and $d_i\in\mathrm{EU}(2n,I,\Gamma)$ are some elementary unitary transvections. Clearly, $c_1\cdots c_t=e$. Now we may rewrite b in the form

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

Thus, $(e,b)\in\mathrm{U}(2n,\iota_2(I),\iota_2(\Gamma))$. □

Corollary 5.1. *The normality of the absolute elementary subgroup implies all other statements of Theorem 1.1.*

Proof. First, we prove that all relative elementary subgroups are normal. Indeed, let $a\in\mathrm{U}(2n,R,\Lambda)$, $\beta\in\mathrm{EU}(2n,I,\Gamma)$. 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, $aba^{-1}\in\mathrm{U}(2n,I,\Gamma)$. Thus, by the above lemma, $aba^{-1}\in\mathrm{EU}(2n,I,\Gamma)$. Now we prove the second commutator formula. Let $a\in$

$\mathrm{EU}(2n, R, \Lambda)$, $b \in \mathrm{CU}(2n, I, \Gamma)$. The 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 $[a, b] \in \mathrm{EU}(2n, I, \Gamma)$. □

6 Eichler–Siegel–Dickson Transformations

In this section, we introduce transformations which play a crucial role in what follows.

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

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

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

Lemma 6.1. *For any orthogonal vectors u, v with u 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. 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 $(gx - x, gy - y) = (-v\lambda\overline{\xi}(u, x), -v\lambda\overline{\xi}(u, y)) = \overline{(u, x)}\xi(v, v)\overline{\xi}(u, y)$. On the other hand,

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

Since h is λ-hermitian, the first summand of $(gx - x,\, y)$ cancels with the second summand of $(x, gy - y)$ and vice versa. 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 2.1 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 $\overline{(u,x)}\xi f(v,v)\overline{\xi}(u,x)$, while the second summand equals $(x,u)\xi(v,x) - (x,v)\overline{\lambda}\,\overline{\xi}(u,x) - (x,u)\overline{\lambda}\xi(f(v,v)+\alpha)\overline{\xi}(u,x)$. 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 definition (recall that $\alpha \in \Lambda$). On the other hand,

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

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

2^0. Basic properties of ESD-transvections. In this subsection, we state some obvious properties of the elements $T_{uv}(\xi,\alpha)$. The definition of ESD-transvections and Lemma 2.5 immediately imply that a conjugate of an ESD-transvection by an element of $\mathrm{U}(2n,R,\Lambda)$ is again an ESD-transvection.

Lemma 6.2. *For any $g \in \mathrm{U}(2n,R,\Lambda)$, one has*

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

Note 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 6.3. *For any $\xi,\zeta \in R$, $T_{uv}(\xi\zeta,\alpha) = T_{u\xi,v}(\zeta,\alpha) = T_{u,v\overline{\zeta}}(\xi,\zeta\alpha\overline{\zeta})$.*

In particular, this lemma shows that the parameter ξ in the definition of 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 is 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 u and v enjoy some special properties, for example, they are unimodular. This can be done only if we explicitly keep the parameter ξ.

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

Lemma 6.4. *For any two vectors v,w orthogonal to an isotropic vector u, for any $\xi,\zeta \in R$ and $\alpha,\beta \in \Lambda$, 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 types of ESD-transvections. These are really the only ones which appear in the analysis of conjugates of 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.

Lemma 6.5. *For any $i \neq \pm j$, $\xi \in R$, and $\alpha \in \Lambda$, 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 6.2 implies that a conjugate of an elementary unitary transvection is an ESD-transvection. More precisely, a conjugate of $T_{i,-i}(\alpha)$ is an element of long root type, while a conjugate of $T_{ij}(\xi)$ for $i \neq \pm j$ is an element of short root type.

Remark. The reader may ask why $T_{uv}(\xi, \alpha)$ depends on two parameters. It is because of 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, [2, 20]). 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 6.6. *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.

Lemma 6.7. *Let u be an isotropic vector. Then for any $\alpha, \beta \in \Lambda$ and $\xi \in R$, one has $T_u(\alpha)T_u(\beta) = T_u(\alpha + \beta)$ and $T_{u\xi} = T_u(\xi\alpha\overline{\xi})$.*

The analogous additivity formula holds for elements of short root type.

Lemma 6.8. *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 with the proof of Lemma 3.4) plays a crucial role in the second part of the work.

Lemma 6.9. *Let u, v be orthogonal 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 with [60, Lemma 1.7]) establishes the additivity property of elements of short root type $T_{uv}(\xi)$ in u when v is fixed.

Lemma 6.10. *Let u, v be orthogonal 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 expresses symmetry of $T_{uv}(\xi)$ with respect to u and v when *both* of them are isotropic (compare with Formula (h) on page 214 of [36], although the involution seems to be missing there).

Lemma 6.11. *Assume both u and v are isotropic. Then $T_{uv}(\xi) = T_{v,-u}(\overline{\lambda\xi})$.*

7 Whitehead Type Lemmas

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

1^0. Heisenberg group. In this subsection, we describe the explicit shape of elements from the unipotent radical of the standard parabolic subgroup P_1. This will be used in subsequent subsections.

Lemma 7.1. *Let $v = (v_2, \dots, v_n, v_{-n}, \dots, v_{-2})^t$ be any vector of length $2n - 2$. Then the 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,*

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

Proof. A 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 7.2. *Let* $v = (v_2, \dots, v_n, v_{-n}, \dots, v_{-2})^t$ *be an isotropic vector of length* $2n-2$*. Then the 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 $\widetilde{u}u = (u,u) = 0$ since u is isotropic. □

The following lemma is straightforward. It will be used in the sequel of this paper.

Lemma 7.3. *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 another.

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

Lemma 7.4. *Suppose $u_i = u_{-i} = 0$ for some $i \in I$. Then for all $\alpha \in \Lambda$, one has $T_u(\alpha) \in \mathrm{EU}(2n, R, \Lambda)$.*

Proof. Since $\mathrm{EU}(2n, R, \Lambda)$ is normalized by all permutation matrices $H(\pi)$, where $\pi \in \mathrm{S}_n$, without loss of generality, we may assume $i = 1$. 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 7.5. *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 Lemma 7.4, since $\mathrm{EU}(2n, R, \Lambda)$ is normalized by the matrix $\begin{pmatrix} 0 & p \\ \lambda p & 0 \end{pmatrix}$, without loss of generality, we may assume $u_{-1} = 0$. Present u in the form $u = v + e_1 u_1$, where v is the column obtained from u by changing its first coordinate to 0. 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 Lemmas 7.3 and 7.4, 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 $\xi \in R$. Note that we do not assume v to be isotropic.

Lemma 7.6. *Suppose $u_i = u_{-i} = v_i = v_{-i} = 0$ for some $i \in I$. Then $T^\bullet_{uv}(\xi) \in \mathrm{EU}(2n, R, \Lambda)$ for all $\xi \in R$.*

Proof. As in Lemma 7.5, without loss of generality, we may assume $u_1 = u_{-1} = v_1 = v_{-1} = 0$. 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 hand, $f(y, y) = f(v, v)$, 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 7.1 and 7.2. □

Lemma 7.7. *Suppose* $u_i = v_i = 0$ *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, without loss of generality, we may assume $u_1 = v_1 = 0$. Present u and v in the form $u = x + e_1 u_1$ and $v = y + e_1 v_1$, where x and y are the columns obtained from u and v, respectively, by changing their first coordinates to 0. Since $u_{-1} = v_{-1} = 0$, the vectors x and y are orthogonal and x is isotropic. Thus,

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

where the first two factors belong to $\mathrm{EU}(2n, R, \Lambda)$ by Lemmas 7.5 and 7.1, respectively, and the last one by the definition of Λ. □

Now assume the vector v is also isotropic. Then the following version of the preceding lemma holds.

Lemma 7.8. *Suppose* $u_i = v_i = 0$ *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 7.7 and 7.5. □

Of course, this could be proven directly because, for example, the formula in the proof of Lemma 7.6 may be simplified to the 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 7.1. *Let* u, v *be any orthogonal vectors with* u *isotropic. Assume* $u_i = v_i = 0$ *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 Sec. 6, 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 7.8 and 7.5, respectively. □

4^0. Kopeiko–Taddei lemma. Here, we show that the addition formulae for ESD-transvections from Sec. 6 (see Lemmas 6.4 and 6.9) imply that $T_{uv}(\xi, \alpha)$ belongs to the elementary group if u has two zeros in symmetric positions. As is known from [42, 62], this already suffices to prove the normality in the classical symplectic case (when $\lambda = -1$ and the involution is trivial).

Lemma 7.9. *Suppose* $u_i = u_{-i} = 0$ *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. Without loss of generality, we may assume $i = 1$. By Lemmas 6.6 and 7.4, it is 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 in this case, one easily sees that $T^{\bullet}_{uv}(\xi) = Y^{+}(\overline{\lambda}x\xi\,\overline{v}_1)Y^{-}(x\xi\,\overline{v}_{-1})$. □

Lemma 7.10. *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 $x = e_1v_1 + e_{-1}v_{-1}$ and $z = v - x$. By Lemmas 6.4 and 6.3, one has $T_{uv}(\xi, \alpha) = T_{ux}(\xi, \beta)T_{uz}(\xi, \gamma)$ for appropriate $\beta, \gamma \in \Lambda$. Now the second factor on the right-hand side belongs to $\mathrm{EU}(2n, R, \Lambda)$ by Lemma 7.6 and the first one by Lemma 7.9. □

8 Normality of the Elementary Subgroup: Suslin's Approach

In this section, we prove the normality of $\mathrm{EU}(2n, R, \Lambda)$ in $\mathrm{U}(2n, R, \Lambda)$ for the case where R is almost commutative and $n \geq 3$. We do it following the Suslin approach in [57, 60, 65]. For *long* root elements in the case $\lambda = -1$ and $\Lambda = \Lambda_{\min}$, this approach was also used in [31]. For *short* root elements in the *symplectic* group, it was used in [42, 62]. Here, we show how it works in the general case.

1^0. Unitary Suslin lemma. The original Suslin's proof [57] of the normality of $\mathrm{E}(n, R)$ in $\mathrm{GL}(n, R)$ $(n \geq 3)$ over a *commutative* ring R was based on the following observation, which is sometimes referred to as Suslin's lemma (compare [12, 36, 51]). Recall that a row $u = (u_1, \dots, u_n) \in {}^nR$ is called *unimodular* if $u_1R + \cdots + u_nR = R$, 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 analog 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 8.1. *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 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 u 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 $w \in R^{2n}$ 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 over 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 $i \in I^+$ and $j \in I^-$. □

Now we state a non-commutative version of this lemma (compare [37, 65]). Let $R_0 = \mathrm{Cent}(R)$ be the center of the ring R. For an element $\xi \in R$, we denote by $\mathcal{O}(\xi)$ the ideal in R_0 of all *central multiples* of ξ, i.e., of all elements from R_0 which have the form $\xi\zeta = \zeta\xi$ for some $\zeta \in R$. Note that we do not assume below that u is unimodular. The price we pay is that we express only certain multiples of the vector v as a linear combination of vectors orthogonal to u with at most 2 non-zero coordinates.

Lemma 8.2. *Suppose $u \in R^{2n}$ is a column of height $n \geq 2$ and $\theta = \overline{u}_i\zeta = \zeta\,\overline{u}_i$ $(\zeta \in R)$ is a central multiple of $\overline{u}_i$ for a fixed i. Then for any vector $v \in R^{2n}$ orthogonal to u, its multiple $v\theta$ may be expressed as a sum $v\theta = \sum_{j\neq i} w_{j,-i}$, where each of the vectors $w_{j,-i}$ is orthogonal to u and has at most two non-zero 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 lets $w = \sum_{j\neq -i} w_{j,-i}$, then clearly, $w_j = v_j\theta$ for all $j \neq -i$. Finally,

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

The decompositions of root type elements in the following subsections are based on this lemma. In the next subsection, we will apply the lemma 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 that each summand has *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 8.1. *Let $u \in R^{2n}$ be an isotropic column and $\alpha \in \Lambda$. Assume $n \geq 3$ 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 $\alpha = \theta\beta\overline{\theta}$, where $\theta \in \mathcal{O}(\overline{u}_{-1})$ and $\beta \in \Lambda$. By Lemma 6.7, one has $T_u(\theta\beta\overline{\theta}\,) = T_{u\theta}(\beta)$. On the other hand, the column $u\theta$ may be decomposed as $u\theta = v+w$, 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 $\theta = \overline{u}_{-1}\xi = \xi\,\overline{u}_{-1} \in R_0$. Then one may take $\zeta = -\xi\,\overline{u}_2u_{-2}$. A direct check using the fact that u is isotropic and ζ central shows that v and w are indeed orthogonal and isotropic. By Lemma 6.10, one has $T_{u\theta}(\beta) = T_v(\beta)T_w(\beta)T_{vw}(\beta)$, where all factors on the right-hand side are elementary: the first two by the Whitehead–Vaserstein lemma (Lemma 7.5) and the last one by the Kopeiko–Taddei lemma (Lemma 7.10). □

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 8.2. *Let $u \in R^{2n}$ be an isotropic vector, $v \in R^{2n}$ any vector orthogonal to u, and $\xi \in R$. Assume $n \geq 3$ 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 $\xi = \overline{\eta}\theta\rho$, where $\eta \in \mathcal{O}(\overline{u}_{-1})$, $\theta \in \mathcal{O}(\overline{u}_{-2})$, and $\rho \in R$. By Lemma 6.3, one has $T_{uv}(\eta\theta\rho) = T_{u,v\eta}(\theta\rho)$. On the other hand, by Lemma 8.2, the column $v\eta$ may be decomposed into a sum of columns $w_{j,-i}$, all of which are orthogonal to u and have at most two non-zero coordinates. By Lemmas 6.4 and 6.6, one has

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

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 the 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 6.3, one has $T_{u,w_{i,-i}}(\theta\rho) = T_{u\theta,w_{i,-i}}(\rho)$. As in the proof of Proposition 8.1, we can decompose the column $u\theta$ as $u\theta = w + z$, where w and z are orthogonal isotropic vectors of the form

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

Moreover, since u and w are orthogonal to $w_{i,-i}$, so is z. By Lemma 6.10, 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 ($w_{-3} = z_3 = 0$ 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 completes the proof. □

4^0. Partitions of 1. In this subsection, it is shown that, when a single unimodular column u is replaced by all columns cu, where c ranges over

$\mathrm{EU}(2n, R, \Lambda)$, then the elements appearing in Propositions 8.1 and 8.2 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 Suslin (see [65, Lemma 1], [37, Lemma 3], or [39, pp. 18–20]).

The only divergence from the proof of the linear case is that 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 $\mathfrak{m} \in \mathrm{Max}(R_0)$, whereas $\overline{\theta}$ does. It is technically more convenient to pass to a subring A of R_0, in which the involution is trivial, and to prove a formally stronger fact that the products of 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 also contains the 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 given maximal ideal $\mathfrak{m} \in \mathrm{Max}(A)$. For a maximal ideal $\mathfrak{m} \in \mathrm{Max}(A)$, we denote by $A_{\mathfrak{m}} = {S_{\mathfrak{m}}}^{-1}A$, $R_{\mathfrak{m}} = {S_{\mathfrak{m}}}^{-1}R$, etc. the localizations of A, R, etc. with respect to the multiplicative set $S_{\mathfrak{m}} = A \backslash \mathfrak{m}$. As usual, we denote by $\mathrm{Rad}(R)$ the Jacobson radical of a ring R.

Lemma 8.3. *Let (R, Λ) be an almost commutative form ring. Then for any maximal ideal $\mathfrak{m} \in \mathrm{Max}(A)$, the ring $Q_{\mathfrak{m}} = R_{\mathfrak{m}} / \mathrm{Rad}(R_{\mathfrak{m}})$ is classically semisimple and the canonical morphism $\phi = \phi_{\mathfrak{m}} : R \to Q_{\mathfrak{m}}$ 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 element $\theta \in R_0$ 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 $\mathrm{Rad}(A_{\mathfrak{m}}) = \mathfrak{m}A_{\mathfrak{m}}$, Nakayama's lemma implies that $\mathfrak{m}$ is contained in every maximal ideal of $A_{\mathfrak{m}}(\xi)$, and thus, $\mu \in \mathrm{Rad}(A_{\mathfrak{m}}(\xi))$. Hence, $1 + \mu\xi$ is invertible in $A_{\mathfrak{m}}(\xi)$.

Now it is easy to prove that ϕ is surjective. In fact, let $\xi/\theta \in R_{\mathfrak{m}}$, where $\xi \in R$ and $\theta \in S_{\mathfrak{m}}$. 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 obtain $\xi/\theta = \eta\theta + \mu\xi/\theta \in \eta\theta + \mathrm{Rad}(R_{\mathfrak{m}})$.

Finally, observe that $Q_{\mathfrak{m}} = R_{\mathfrak{m}} / \mathrm{Rad}(R_{\mathfrak{m}})$ is a finitely generated module over $(R_0)_{\mathfrak{m}} / \mathrm{Rad}((R_0)_{\mathfrak{m}})$. The latter 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 $R_{\mathfrak{m}} / \mathrm{Rad}(R_{\mathfrak{m}})$ is both semisimple and artinian. □

We set $\Gamma_{\mathfrak{m}} = \phi_{\mathfrak{m}}(\Lambda)$. Then the homomorphism $\phi_{\mathfrak{m}} : (R, \Lambda) \to (Q_{\mathfrak{m}}, \Gamma_{\mathfrak{m}})$ of form rings is surjective. Thus, we obtain 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 [6, 36, 49, 66]). In particular, for all $n \geq 3$, the group acts transitively on the set of all isotropic unimodular vectors (see [36, Theorem 9.1.3] or [49, Theorem 8.1] for much stronger results).

Lemma 8.4. *Let $u \in R^{2n}$ be a unimodular isotropic column over an almost commutative ring R. Then the ideal in R_0 generated by all $\theta\overline{\theta}$, where $\theta \in \mathcal{O}((cu)_{-1})$ and 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 $Q_{\mathfrak{m}}$. 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 $\xi \in R$ and $\theta \in S_{\mathfrak{m}}$. Thus, $(cu)_{-1}\xi = \xi(cu)_{-1} = \theta \in S_{\mathfrak{m}} \cap \mathcal{O}((cu)_{-1})$. Clearly, $\theta\overline{\theta} = \theta^2 \in S_{\mathfrak{m}} \cap \mathcal{O}((cu)_{-1})$. 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 8.5. *Let $u \in R^{2n}$ be a unimodular isotropic column over an almost commutative ring R. Then the ideal in R_0 generated by all $\overline{\eta}\theta$, 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 Lemma 8.4. 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 $((cu)_{-1})^{-1} = \xi/\eta$ and $((cu)_{-2})^{-1} = \zeta/\theta$, where $\xi, \zeta \in R$ and $\eta, \theta \in S_{\mathfrak{m}}$. 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, $\overline{\eta}\theta = \eta\theta \in S_{\mathfrak{m}} \cap \mathcal{O}((cu)_{-1})\mathcal{O}((cu)_{-2})$. Thus, for every $\mathfrak{m} \in \mathrm{Max}(A)$, there exists $c \in \mathrm{EU}(2n, R, \Lambda)$ such that $\mathcal{O}((cu)_{-1})\mathcal{O}((cu)_{-2}) \nsubseteq \mathfrak{m}$. □

Note that we actually used very little of the ring $Q_{\mathfrak{m}}$. Nothing changes in the proof of the last two lemmas if only $\mathrm{EU}(2n, Q_{\mathfrak{m}}, \Gamma_{\mathfrak{m}})$ acts transitively on the set of isotropic unimodular columns. This is the case, for example, when $n \geq \mathrm{asr}(Q_{\mathfrak{m}})+2$ (see [49, Theorem 8.1]) or even better when $n \geq \Lambda S(Q_{\mathfrak{m}})+1$ (see Sec. 3 and the proof of Lemma 4.1 in [11]). 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 [37–39] or when R is von Neumann regular, etc.

5^0. Proof of Theorem 1.1 in the absolute case. Corollary 5.1 shows that we only have to prove the theorem in the absolute case. In view of Lemmas 6.2 and 6.5, this amounts to proving that $T_u(\alpha)$ and $T_{uv}(\xi)$ belong to $\mathrm{EU}(2n, R, \Lambda)$, where u and v are, respectively, the ith and jth columns of a matrix $g \in \mathrm{U}(2n, R, \Lambda)$, $\alpha \in \Lambda$, and $\xi \in R$. In fact, we will prove the 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 8.1. But since u is unimodular, Lemma 8.4 shows that the elements $\theta\overline{\theta}$ 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 $\theta_1\overline{\theta}_1 + \dots + \theta_t\overline{\theta}_t = 1$, where $\theta_h \in \mathcal{O}((c_h u)_{-1})$. Then

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

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 $\overline{\eta} \in \mathcal{O}((cu)_{-1})$ and $\theta \in \mathcal{O}((cu)_{-2})$, 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 8.2. But since u is unimodular, Lemma 8.5 shows that the elements $\theta\overline{\eta}$ 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 $\theta_1\overline{\eta}_1 + \dots + \theta_t\overline{\eta}_t = 1$, where $\overline{\eta}_h \in \mathcal{O}((c_h u)_{-1})$ and $\theta_h \in \mathcal{O}((c_h u)_{-2})$. Then

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

where all the factors on the right-hand side belong to $\mathrm{EU}(2n, R, \Lambda)$. This completes the proof of the theorem for the absolute case, and thus, for all cases in view of Sec. 5. □

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 a form ring. Assume $n \geq 3$, and 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 8.5, the condition of Corollary 5.1 is satisfied, for example, when $n \geq \operatorname{asr} R_\mathfrak{m} + 2$ for all $\mathfrak{m} \in \mathrm{Max}(A)$ (see [49]) or even better, when $n \geq \Lambda S(R_{\mathfrak{m}})+1$ for all $\mathfrak{m} \in \mathrm{Max}(A)$ (see [11]). One could state many further generalizations like this in the style of [67] or [37–39].

Acknowledgements. The second 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 this paper.

References

[1] E. Abe, Chevalley groups over local rings, *Tôhoku Math. J.* **21** (3) (1969) 474–494.
[2] E. Abe, Coverings of twisted Chevalley groups over commutative rings, *Sci. Repts. Tokyo Kyoiku Daigaku* **13** (1977) 194–218.
[3] E. Abe, Chevalley groups over commutative rings, in: *Proc. Conf. Radical Theory* (Sendai, 1988), Uchida Rokakuho Publ. Comp., Tokyo, 1989, pp. 1–23.
[4] E. Abe, Normal subgroups of Chevalley groups over commutative rings, *Contemp. Math.* **83** (1989) 1–17.
[5] E. Abe, K. Suzuki, On normal subgroups of Chevalley groups over commutative rings, *Tôhoku Math. J.* **28** (1) (1976) 185–198.
[6] A. Bak, The stable structure of quadratic modules, Thesis, Columbia University, 1969.
[7] A. Bak, On modules with quadratic forms, in: *Algebraic K-Theory and Its Geometric Applications*, Lecture Notes in Mathematics, Vol. 108, Springer-Verlag, Berlin, 1969, pp. 55–66.
[8] A. Bak, *K-Theory of Forms*, Annals of Mathematics Studies, Vol. 98, Princeton University Press, Princeton, 1981.
[9] A. Bak, Subgroups of the general linear group normalized by relative elementary groups, in: *Algebraic K-Theory, Part II*, Lecture Notes in Mathematics, Vol. 967, Springer-Verlag, Berlin, 1982, pp. 1–22.
[10] A. Bak, Nonabelian K-theory: The nilpotent class of K_1 and general stability, *K-Theory* **4** (1991) 363–397.
[11] A. Bak, G. Tang, Stability for Hermitian K_1, *J. Pure Appl. Algebra* (to appear).
[12] A. Bak, N.A. Vavilov, Normality for elementary subgroup functors, *Math Proc. Cambridge Phil. Soc.* **118** (1) (1995) 35–47.
[13] H. Bass, K-theory and stable algebra, *Publ. Math. Inst. Hautes Et. Sci.* (22) (1964) 5–60.
[14] H. Bass, *Algebraic K-Theory*, Benjamin, New York-Amsterdam, 1968.
[15] H. Bass, Unitary algebraic K-theory, in: *Algebraic K-Theory, III: Hermitian K-Theory and Geometric Applications*, Lecture Notes in Mathematics, Vol. 343, Springer-Verlag, Berlin, 1973, pp. 57–265.
[16] H. Bass, J. Milnor, J.-P. Serre, Solution of the congruence subgroup problem for SL_n ($n \geq 3$) and Sp_{2n} ($n \geq 2$), *Publ. Math. Inst. Hautes Et. Sci.* (1967), No. 33, 59–137.
[17] Z.I. Borewicz, N.A. Vavilov, Subgroups of the general linear group over a semilocal ring containing the group of diagonal matrices, *Proc. Steklov Inst. Math.* (4) (1980) 41–54.

[18] Z.I. Borewicz, N.A. Vavilov, On the subgroups of the general linear group over a commutative ring, *Sov. Math. Dokl.* **26** (3) (1982) 679–681.

[19] Z.I. Borewicz, N.A. Vavilov, The distribution of subgroups in the general linear group over a commutative ring, *Proc. Steklov Inst. Math.* (3) (1985) 27–46.

[20] R.W. Carter, *Simple Groups of Lie Type*, Wiley, London et al., 1972.

[21] P. Cohn, On the structure of the GL_2 of a ring, *Publ. Math. Inst. Hautes Et. Sci.* (30) (1966) 365–413.

[22] P. Cohn, *Free Rings and Their Relations*, London Mathematical Society Monographs, No. 2, Academic Press, London, 1971.

[23] D. Costa, G. Keller, Radix redux: normal subgroups of symplectic group, *J. Reine Angew. Math.* **427** (1) (1992) 51–105.

[24] J. Dieudonné, *La Géométrie des Groupes Classiques*, 3rd ed., Springer-Verlag, Berlin, 1971.

[25] V.N. Gerasimov, The group of units of a free product of rings, *Math. USSR Sbornik* **62** (1) (1989) 41–63.

[26] I.Z. Golubchik, On the general linear group over an associative ring, *Uspehi Mat. Nauk* **28** (3) (1973) 179–180 (in Russian).

[27] I.Z. Golubchik, On the normal subgroups of orthogonal group over an associative ring with involution, *Uspehi Mat. Nauk* **30** (6) (1975) 165 (in Russian).

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

[29] I.Z. Golubchik, On the subgroups of the general linear group $\mathrm{GL}_n(R)$ over an associative ring R, *Russian Math. Surveys* **39** (1) (1984) 157–158.

[30] I.Z. Golubchik, On the normal subgroups of the linear and unitary groups over associative rings, in: *Spaces over Algebras and Some Problems in the Theory of Nets*, Ufa, 1985, pp. 122–142 (in Russian).

[31] I.Z. Golubchik, V.T. Markov, Localization dimension of PI-rings, *Trudy Sem. Petrovskogo* (6) (1981) 39–46 (in Russian).

[32] I.Z. Golubchik, A.V. Mikhalev, Elementary subgroup of a unitary group over a PI-ring, *Vestnik Mosk. Univ.* (Ser. 1 – Mat., Mekh.) (1) (1985) 30–36.

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

[34] G. Habdank, A classification of subgroups of Λ-quadratic groups normalized by relative elementary subgroups, Dissertation, Universität Bielefeld, 1987.

[35] G. Habdank, Homology stability for Λ-quadratic groups, SFB Universität Bielefeld, No. 74, 1990 (preprint).

[36] A.J. Hahn, O.T. O'Meara, *The Classical Groups and K-Theory*, Springer-Verlag, Berlin, 1989.

[37] S.G. Khlebutin, Sufficient conditions for the normality of the subgroup of elementary matrices, *Russian Math. Surveys* **39** (3) (1984) 245–246.

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

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

[40] I.S. Klein, A.V. Mikhalev, Steinberg orthogonal group over a ring with involution, *Algebra and Logic* **9** (2) (1970) 88–103.

[41] I.S. Klein, A.V. Mikhalev, Unitary Steinberg group over a ring with involution, *Algebra and Logic* **9** (5) (1970) 305–312.
[42] V.I. Kopeiko, The stabilization of symplectic groups over a polynomial ring, *Math. USSR Sbornik* **34** (1978) 655–669.
[43] V.I Kopeiko, On a theorem of Suslin, *J. Sov. Math.* **132** (1983) 119–121.
[44] V.I. Kopeiko, Unitary and orthogonal groups over rings with involution, in: *Algebra and Discrete Mathematics*, Kalmytsk. Gos. University, Elista, 1985, pp. 3–14 (in Russian).
[45] T.Y. Lam, *Serre's Problem*, Lecture Notes in Mathematics, Vol. 635, Springer-Verlag, Berlin, 1978.
[46] Fu-an Li, The structure of symplectic groups over arbitrary commutative rings, *Acta Math. Sinica* (New Ser.) **3** (3) (1987) 247–255.
[47] Fu-an Li, The structure of orthogonal groups over arbitrary commutative rings, *Chin. Ann. Math.* **10B** (3) (1989) 341–350.
[48] Fu-an Li, Mulan Liu, Generalized sandwich theorem, *K-Theory* **1** (1987) 171–184.
[49] B. Magurn, W. van der Kallen, L.N. Vaserstein, Absolute stable rank and Witt cancellation for noncommutative rings, *Invent. Math.* **91** (1988) 525–543.
[50] B.R. McDonald, *Geometric Algebra over Local Rings*, Marcel Dekker, New York, 1976.
[51] B.R. McDonald, Automorphisms of $\mathrm{GL}_n(R)$, *Trans. Amer. Math. Soc.* **246** (1978) 155–171.
[52] B.R. McDonald, *Linear Algebra over Commutative Rings*, Marcel Dekker, New York, 1984.
[53] J. Milnor, *Introduction to Algebraic K-Theory*, Princeton University Press, Princeton, 1971.
[54] M.R. Stein, Relativizing functors on rings and algebraic K-theory, *J. Algebra* **19** (1) (1971) 140–152.
[55] A.V. Stepanov, Stability conditions in the theory of linear groups over rings, Ph.D. Thesis, Leningrad State University, 1987.
[56] A.V. Stepanov, N.A. Vavilov, Decomposition of transvections: a theme with variations, Univ. Bielefeld, No. 4, 1999, pp. 1–46 (preprint).
[57] A.A. Suslin, On the structure of the general linear group over polynomial rings, *Soviet Math. Izv.* **41** (2) (1977) 503–516.
[58] A.A. Suslin, On a theorem of Cohn, *J. Sov. Math.* **17** (2) (1981) 1801–1803.
[59] A.A. Suslin, Algebraic K-theory, *J. Sov. Math.* **28** (6) (1985) 870–923.
[60] A.A. Suslin, V.I. Kopeiko, Quadratic modules and orthogonal groups over polynomial rings, *J. Sov. Math.* **20** (6) (1982) 2665–2691.
[61] R.G. Swan, Generators and relations for certain special linear groups, *Adv. Math.* **6** (1971) 1–77.
[62] G. Taddei, Invariance du sous-groupe symplectique élémentaire dans le groupe symplectique sur un anneau, *C. R. Acad. Sci. Paris* (Sér. I) **295** (2) (1982) 47–50.
[63] G. Taddei, Schémas de Chevalley–Demazure, fonctions représentatives et théorème de normalité, Thèse, Univ. de Genève, 1985.
[64] G. Taddei, Normalité des groupes élémentaire dans les groupes de Chevalley sur un anneau, *Contemp. Math.* **55** (2) (1986) 693–710.

[65] M.S. Tulenbaev, Schur multiplier of the group of elementary matrices of finite order, *J. Sov. Math.* **17** (4) (1981) 2062–2067.
[66] L.N. Vaserstein, Stabilization of unitary and orthogonal groups over a ring, *Math. USSR Sbornik* **10** (1970) 307–326.
[67] L.N. Vaserstein, On normal subgroups of GL_n over a ring, in: *Algebraic K-Theory*, Lecture Notes in Mathematics, Vol. 854, Springer-Verlag, Berlin, 1981, pp. 456–465.
[68] L.N. Vaserstein, The subnormal structure of general linear groups, *Math. Proc. Cambridge Phil. Soc.* **99** (1986) 425–431.
[69] L.N. Vaserstein, On normal subgroups of Chevalley groups over commutative rings, *Tôhoku Math. J.* **38** (1986) 219–230.
[70] L.N. Vaserstein, Normal subgroups of orthogonal groups over commutative rings, *Amer. J. Math.* **110** (5) (1988) 955–973.
[71] L.N. Vaserstein, Normal subgroups of symplectic groups over rings, *K-Theory* **2** (5) (1989) 647–673.
[72] L.N. Vaserstein, Normal subgroups of classical groups over rings and gauge groups, *Contemp. Math.* **83** (1989) 451–459.
[73] L.N. Vaserstein, The subnormal structure of general linear groups over rings, *Math. Proc. Cambridge Phil. Soc.* **108** (2) (1990) 219–229.
[74] L.N. Vaserstein, A.V. Mikhalev, The normal subgroups of the orthogonal group over a ring with involution, *Algebra and Logic* **9** (6) (1970) 375–377.
[75] L.N. Vaserstein, A.A. Suslin, Serre's problem on projective modules over polynomial rings and algebraic K-theory, *Sov. Math. Izv.* **10** (1976) 937–1001.
[76] N.A. Vavilov, Structure of split classical groups over commutative rings, *Soviet Math. Doklady* **37** (1988) 550–553.
[77] N.A. Vavilov, On the problem of normality of the elementary subgroup in a Chevalley group, in: *Algebraic and Discrete Systems*, Ivanovo University, 1988, pp. 7–25 (in Russian).
[78] N.A. Vavilov, A note on the subnormal structure of general linear groups, *Math. Proc. Cambridge Phil. Soc.* **107** (2) (1990) 193–196.
[79] N.A. Vavilov, On subgroups of the split classical groups, *Proc. Math. Inst. Steklov* (4) (1991) 27–41.
[80] N.A. Vavilov, Structure of Chevalley groups over commutative rings, in: *Proc. Conf. Non-Associative Algebras, Related Topics*, Hiroshima, 1990, World Scientific, Singapore, 1991, pp. 219–335.
[81] N.A. Vavilov, Intermediate subgroups in Chevalley groups, in: *Proc. Conf. Groups of Lie Type and Their Geometries*, Como, 1993, Cambridge University Press, 1995, pp. 233–280.
[82] N.A. Vavilov, A third look at weight diagrams, Univ. Bielefeld, No. 110, 1998, pp. 1–41 (preprint).
[83] N.A. Vavilov, Linear groups over general rings, I, II (in preparation).
[84] N.A. Vavilov, E.B. Plotkin, Chevalley groups over commutative rings, I. Elementary calculations, *Acta Applicandae Math.* **45** (1) (1996) 73–113.
[85] N.A. Vavilov, E.B. Plotkin, A.V. Stepanov, Calculations in Chevalley groups over commutative rings, *Soviet Math. Doklady* **40** (1) (1990) 145–147.
[86] J.S. Wilson, The normal and subnormal structure of general linear groups, *Proc. Cambridge Phil. Soc.* **71** (1972) 163–177.
[87] A.E. Zalesskii, Linear groups, *J. Sov. Math.* **31** (3) (1985) 2974–3004.