# On groups of similitudes in associative rings

Evgenii L. Bashkirov

*Abstract.* Let $R$ be an associative ring with 1 and $R^\times$ the multiplicative group of invertible elements of $R$. In the paper, subgroups of $R^\times$ which may be regarded as analogues of the similitude group of a non-degenerate sesquilinear reflexive form and of the isometry group of such a form are defined in an abstract way. The main result states that a unipotent abstractly defined similitude must belong to the corresponding abstractly defined isometry group.

*Keywords:* associative rings, unipotent elements

*Classification:* 16U60, 20H25

First we give the definition of the main ring-theoretical notion we use and study in the present short paper.

**Definition.** Let $R$ be an associative ring with an identity element 1. An element $g \in R$ is called *unipotent* if $(g-1)^k = 0$ for some integer $k \geq 1$.

Unipotent elements of full matrix rings over fields, that is, matrices all eigenvalues of which equal 1 are meaningful for the theory of linear groups because the most important and interesting classical linear groups are generated by such elements. In particular, in papers devoted to linear groups over associative division rings, assertions about the impossibility for unipotent elements in the similitude group of a sesquilinear reflexive form to have the similitude multiplicator different from 1 appear remarkably often. In other words, these assertions state that every unipotent element in the similitude group of a certain sesquilinear reflexive form must belong to the isometry group of the form itself. Sometimes such statements are proved, mainly for division rings of particular kinds ([11, 1.18]) and also for special unipotent elements ([10, 4.2.7]), but largely they are taken for granted being regarded as manifest (see, for instance, [1], [2], [3], [4]). At the same time, a rather favorable frequency of appearing of such results in articles on classical linear groups allows us to say about their importance for the study of these groups and simultaneously suggests an idea of the existence of their general background. Indeed, in this note, we establish a result, which provides the desired background within the framework of the ring theory, and thereby implies the above mentioned assertions on similitude groups of reflexive sesquilinear forms as its specific cases. Therefore, the results obtained may prove helpful for better understanding of the nature of similitudes relative to sesquilinear forms. It is worth mentioning that our approach is elementary in nature and the prerequisite to the present paper is a knowledge of the content of Chapter 1 in Dieudonné's book [5]. We shall also base the proof of our main result on elementary calculations using well known properties of binomial coefficients.

So, to begin with, denote by $R$ an associative ring with an identity element 1 and let $R^\times$ be the multiplicative group of all invertible elements of $R$. Let $J$ be an antiautomorphism of $R$, that is, an automorphism of the additive group of $R$ such that $J(ab) = J(b)J(a)$ for all $a, b \in R$. Note that the mapping $J$ may be identity if $R$ is commutative. Now fix an element $r \in R^\times$ and let $C$ be a subsemigroup of the multiplicative semigroup of $R$ (that is, $ab \in C$ whenever $a, b \in C$) with the property $1 \in C$. Suppose $C$ consists of elements lying in the center of $R$. Then define $H(r, C, J)$ to be the collection of all $g \in R$ for which there exists $c \in C$ (depending on $g$) such that

$$J(g)rg = cr. \tag{1}$$

We refer to the elements of $H(r, C, J)$ as similitudes or, more precisely, as similitudes in the ring $R$. Since $r$ is invertible, $c$ figuring in (1) is uniquely defined by $g$. Therefore, $c$ is denoted by $c_g$ and this $c_g$ is called the multiplicator of the similitude $g$. One can directly check that $H(r, C, J)$ is a subsemigroup of the multiplicative semigroup of $R$. We call $H(r, C, J)$ the similitude semigroup (in the ring $R$). For any $g, h \in H(r, C, J)$, the equation $c_{gh} = c_g c_h$ is true because $c_g$, $c_h$ lie in the center of $R$, and this shows that the mapping $g \to c_g$ is a homomorphism of the semigroup $H(r, C, J)$ into the semigroup $C$. Now assume that $g$ is a similitude in $H(r, C, J)$ such that $g$ is invertible in $R$. Then (1) implies that $c_g \in R^\times$, $g^{-1} \in H(r, C, J)$ and $c_{g^{-1}} = c_g^{-1}$. Thus the set $H(r, C, J) \cap R^\times := S(r, C, J)$ is a subgroup of $R^\times$ and we can obtain a homomorphism of this group into the group $C \cap R^\times \leq R^\times$ by restricting the mapping $g \to c_g$ to $S(r, C, J)$. The image of $S(r, C, J)$ under this homomorphism is denoted by $M(S(r, C, J))$ and we designate its kernel by $G(r, C, J)$. Accordingly, $G(r, C, J)$ is a normal subgroup of $S(r, C, J)$ and

$$G(r, C, J) = \{g \in R^\times \mid J(g)rg = r\}.$$

We refer to $S(r, C, J)$ as a similitude group (in the ring $R$). This definition of a similitude group both supports and extends the traditional definition which is given in [5, Chapter 1, §9]. The group $G(r, C, J)$ may be thought of as a group comparable to the corresponding classical linear one. The quotient of $S(r, C, J)$ by $G(r, C, J)$ is isomorphic to the abelian group $M(S(r, C, J))$, and so $G(r, C, J)$ is a subgroup of the commutator subgroup of $S(r, C, J)$.

Now we make a simple observation which turns out to be very useful for our further study. Define $T$ to be the mapping from $R$ into $R$ such that $T(g) = r^{-1}J(g)r$

for each $g \in R$. It is straightforward to verify that $T$ is an antiautomorphism of $R$ and

$$H(r,C,J) = H(1,C,T), \quad S(r,C,J) = S(1,C,T), \quad G(r,C,J) = G(1,C,T).$$

Thus without loss of generality we may restrict ourselves to the case $r = 1$. We write $S(C,J)$ and $G(C,J)$ instead of $S(1,C,T)$ and $G(1,C,J)$ respectively. Moreover, since $G(C,J)$ does not depend on $C$, we may also write $G(J)$ instead of $G(C,J)$.

Next we want to supply the readers with some examples, which can be studied in a truly elementary way, to illustrate the notions introduced. With this end in view, let us assume that $J$ leaves $C$ invariant as a whole, that is, $J(c) \in C$ whenever $c \in C$. In this case, $S(C,J)$ contains enough much elements, namely, every $c \in C$ is a similitude in $S(C,J)$, the multiplicator of $c$ being $J(c)c$. In general, if $g \in G(J)$ and $z \in C$, then $zg = gz$ is a similitude in $S(C,J)$ such that the multiplicator of $z$ is equal to $J(z)z$. Conversely, if $s$ is a similitude in $S(C,J)$ and the multiplicator of $s$ has the form $J(z)z$ with $z \in C$, then $sz^{-1} \in G(J)$.

Now we are ready to formulate and prove our main result.

**Proposition 1.** *Suppose the ring $R$ is an algebra over its commutative subring $A$. Assume $A$ contains no nilpotent elements and $M(S(C,J)) \leq A^{\times}$. If $g \in H(C,J)$ and $(g-1)^k = 0$ for some integer $k \geq 1$, then $g \in G(J)$.*

Proof: As is well known, in every associative ring each nilpotent element being added to the identity always gives an invertible element. So $g$ is invertible and we have actually that $g \in S(C,J)$. If $k = 1$, then there is nothing to prove. Suppose $k \geq 2$ and let $c$ be the multiplicator of $g$, that is $J(g)g = c$. It follows that $g$ and $J(g)$ are permutable, and so we have

$$J(g)^k g^k = c^k. \tag{2}$$

Further, the condition $(g-1)^k = 0$ implies that

$$g^k = \sum_{m=1}^{k} \binom{k}{m} g^{k-m} (-1)^{m-1}, \tag{3}$$

where $\binom{k}{m}$ denotes, as usual, the binomial coefficient $\binom{k}{m} = \frac{k!}{m!(k-m)!}$ modulo the characteristic of $R$. On substituting (3) into (2) we obtain

$$\sum_{m=1}^{k} \sum_{l=1}^{k} \binom{k}{m} \binom{k}{l} J(g)^{k-m} g^{k-l} (-1)^{m+l-2} = c^k. \tag{4}$$

But $J(g)^{k-m}g^{k-m} = c^{k-m}$ whence it follows that $J(g)^{k-m} = c^{k-m}g^{-k+m}$. Therefore, $J(g)^{k-m}g^{k-l} = c^{k-m}g^{m-l}$, and hence

$$\sum_{m=1}^{k}\sum_{l=1}^{k}\binom{k}{m}\binom{k}{l}c^{k-m}g^{m-l}(-1)^{m+l-2} - c^k = 0.$$

Multiplying both sides of the last equation by $g^{k-1}$ yields

$$\sum_{m=1}^{k}\sum_{l=1}^{k}\binom{k}{m}\binom{k}{l}c^{k-m}g^{k+m-l-1}(-1)^{m+l-2} - c^k g^{k-1} = 0. \tag{5}$$

Next let us write $g$ as the sum $1+h$ with $h \in R$. On substituting this expression into (5) we find that

$$\sum_{m=1}^{k}\sum_{l=1}^{k}\binom{k}{m}\binom{k}{l}c^{k-m}(1+h)^{k+m-l-1}(-1)^{m+l-2} - c^k(1+h)^{k-1} = 0. \tag{6}$$

Recall now that $h$ is nilpotent, $h^k = 0$. Then (6) shows that $h$ is a root of a polynomial, say $q$, in one variable with coefficients in $A$ whose degree $s$ does not exceed $k$. Calculate the constant term of $q$. To do this we use the following well known elementary formula

$$\sum_{l=1}^{k}\binom{k}{l}(-1)^l = -1. \tag{7}$$

Employing (7) allows us to deduce from (6) that the constant term of $q$ is equal to

$$\begin{aligned}
&\sum_{m=1}^{k}\sum_{l=1}^{k}\binom{k}{m}\binom{k}{l}c^{k-m}(-1)^{m+l-2} - c^k \\
&= -\left(c^k + \sum_{m=1}^{k}\left(\sum_{l=1}^{k}\binom{k}{l}(-1)^l\right)\binom{k}{m}c^{k-m}(-1)^{m-1}\right) \\
&= -\left(c^k + \sum_{m=1}^{k}\binom{k}{m}(-1)^m c^{k-m}\right) = -(c-1)^k.
\end{aligned}$$

This and equation (6) are combined to yield that for some $a_1, a_2, \ldots, a_s \in A$ we have

$$(c-1)^k = a_1 h + a_2 h^2 + \cdots + a_s h^s. \tag{8}$$

The subring $A$ is contained in the center of $R$. Consequently, the element in the right hand side of (8) is nilpotent, since it is the sum of nilpotent elements which are pairwise permutable. Hence $(c-1)^k$ is nilpotent and so is $c-1$. But the last element belongs to the ring $A$ which possesses no nonzero nilpotent elements. Thus $c-1=0$ completing the proof of the proposition. $\square$

It should be pointed out that in Proposition 1 we cannot remove the assumption that the subring $A$ contains no nonzero nilpotent elements. In order to make sure that this is so, we can consider the following example. Let $k$ be an integer, $k \geq 2$, $P$ a field of characteristic different from 2, $R$ a commutative and associative $P$-algebra with a basis $1, v, v^2, \ldots, v^{k-1}$, where 1 is an identity element and $v^k = 0$. We think of $R$ as an algebra over $A = R$ and take the identity automorphism of $R$ as $J$. Then $G(R^\times) = \{\pm 1\}$ and $g = 1 + v$ is a similitude in $S(R^\times, J)$ with the multiplicator $(1+v)^2 = 1 + 2v + v^2 \neq 1$. But $(g-1)^k = v^k = 0$, and so $g$ is a unipotent element in $S(R^\times, J) \setminus G(R^\times)$.

At this point, we turn to some specific examples of similitudes in rings. These examples are actually the motivating ones of similitudes. They can be merged into the facts about similitudes given by Proposition 1 thereby well justifying terminology we have introduced.

Let $E$ be a right vector space of finite dimension $n \geq 2$ over an associative division ring $K$. Assume $K$ admits an involution, that is, an antiautomorphism $a \to \overline{a}$ $(a \in K)$ such that $\overline{\overline{a}} = a$. If $U$ is a matrix whose coefficients belong to $K$, then $\overline{U}$ denotes the matrix obtained from $U$ by applying the involution $^-$ to each element of $U$; in turn ${}^t\overline{U}$ denotes the transposed matrix of $\overline{U}$. Let $f$ be a nondegenerate reflexive sesquilinear form on $E \times E$ with respect to $^-$; thus $f$ is a biadditive mapping $f : E \times E \to K$ such that $f(xa, yb) = \overline{a} f(x,y) b$ for all $x, y \in E$ and $a, b \in K$ and the relations $f(x,y) = 0$ and $f(y,x) = 0$ are equivalent. We remind that a linear mapping $u : E \to E$ is called a similitude relative to $f$ if there exists $c \in K^\times$ such that $f(u(x), u(y)) = cf(x,y)$ for all $x, y \in E$. Choose a basis $\{e_1, \ldots, e_n\} = (e_i)$ of $E$ and let $D$ be the matrix of $f$ relative to $(e_i)$. A linear transformation $u$ of $E$ is a similitude relative to $f$ if and only if its matrix $U$ in $(e_i)$ satisfies the condition ${}^t\overline{U} D U = cU$. It is readily seen that $c$ belongs to the center of $K$ (see [5, Chapter 1, §9]). Since the mapping $J : U \to {}^t\overline{U}$ is an involution of the ring of all matrices of degree $n$ over $K$, the definition of similitudes in associative rings given in the present paper is quite comparable with the definition of similitudes relative to a reflexive sesquilinear form. Therefore Proposition 1 has the following corollary which was the main motivation to write the present work.

**Proposition 2.** *Let $n$ be an integer, $n \geq 2$, and let $f$ denote one of the following:*

(a) *a nondegenerate $^-$-skew-Hermitian form in $n$ variables over a division ring $K$ with an involution $^-$;*

(b) *a nondegenerate alternative form in $n$ variables over a field $K$;*

(c) *a nondegenerate symmetric bilinear form in $n$ variables over a field $K$ of characteristic different from 2.*

*Let $\mathrm{U}_n(K,f)$, $\mathrm{Sp}_n(K,f)$, $\mathrm{O}_n(K,f)$ be, as usual, the subgroups of the general linear group $\mathrm{GL}_n(K)$ that are the isometry subgroups of the form $f$ from items* (a)–(c) *respectively, that is, the unitary, symplectic and orthogonal groups respectively, and let $\mathrm{GU}_n(K,f)$, $\mathrm{GSp}_n(K,f)$, $\mathrm{GO}_n(K,f)$ be the corresponding similitude groups. If $g$ is a unipotent element in $\mathrm{GU}_n(K,f)$ (respectively in $\mathrm{GSp}_n(K,f)$ or in $\mathrm{GO}_n(K,f)$), then $g \in \mathrm{U}_n(K,f)$ (respectively $g \in \mathrm{Sp}_n(K,f)$ or $g \in \mathrm{O}_n(K,f)$).*

Now let us turn our attention to another aspect of the problem under consideration. Namely, as we have already said, $G(J)$ is a normal subgroup of $S(C,J)$, and so $S(C,J)$ is contained in the normalizer of $G(J)$ in $R^{\times}$. Thus we are led to the following question: when does $S(C,J)$ coincide with the normalizer of $G(J)$ in $R^{\times}$? As to this question, we can prove the following simple assertion.

**Proposition 3.** *Let $C$ be a subgroup of the multiplicative group of the center of $R$. If the centralizer of the group $G = G(J)$ in $R^{\times}$ is contained in $C$, then $S = S(C,J)$ is the normalizer of $G$ in $R^{\times}$.*

Proof: Let $N$ denote the normalizer of $G$ in $R^{\times}$. It suffices to show that $N \subseteq S$. Let $h \in N$. This means that $hgh^{-1} \in G$ for any $g \in G$. By the definition of $G$, we have $J(hgh^{-1})hgh^{-1} = 1$. Since $J$ is an antiautomorphism of $R$, the last equation may be rewritten as $J(g)J(h)hg = J(h)h$. But $g \in G$, and so $J(g) = g^{-1}$. Hence $g^{-1}(J(h)h)g = J(h)h$. Thus $J(h)h$ commutes with any $g \in G$. So the assumption on the centralizer of $G$ leads us to the inclusion $J(h)h \in C$ which is amount to the relation $h \in S$. The proposition is proved. $\square$

When $K$ is a field of characteristic different from 2, it is well known that the groups $\mathrm{U}_n(K,f)$, $\mathrm{Sp}_n(K,f)$, $\mathrm{O}_n(K,f)$ with $n \geq 2$ are absolutely irreducible subgroups of $\mathrm{GL}_n(K)$ (see [6, p. 36] for a definition of this term). Consequently, their centralizers in $\mathrm{GL}_n(K)$ consist of scalar matrices only. Therefore, if $K^{\times}$ is taken as $C$, then by Proposition 3, the groups $\mathrm{GU}_n(K,f)$, $\mathrm{GSp}_n(K,f)$, $\mathrm{GO}_n(K,f)$ are the normalizers of the groups $\mathrm{U}_n(K,f)$, $\mathrm{Sp}_n(K,f)$, $\mathrm{O}_n(K,f)$ respectively. These facts were obtained first by King [8], [9], and Dye [7]. On the other hand, we can easy exhibit a ring which does not handle the requirements of Proposition 3. The next paragraph features the corresponding example.

Suppose a field $K$ contains more than two elements. Let $R$ be the ring $M_2(K) \oplus M_2(K)$, the direct sum of two copies of the ring $M_2(K)$ of all matrices of degree 2 over $K$. We write elements in $R$ as pairs $(a,b)$ with $a,b \in M_2(K)$. Let $I$ be an involution on $M_2(K)$ such that

$$I\left(\begin{pmatrix} x & y \\ z & t \end{pmatrix}\right) = \begin{pmatrix} t & -y \\ -z & x \end{pmatrix} \quad \text{for all} \quad \begin{pmatrix} x & y \\ z & t \end{pmatrix} \in M_2(K).$$

Define an involution $J$ on $R$ as follows: $J(a,b) = (I(a), I(b))$ for $a, b \in M_2(K)$. Let $C$ be the subgroup of $R^\times = \mathrm{GL}_2(K) \times \mathrm{GL}_2(K)$, the direct product of two copies of $\mathrm{GL}_2(K)$, consisting of all elements $(r1_2, r1_2)$, where $r$ runs through $K^\times$ and $1_2$ denotes the identity matrix in $M_2(K)$. Then $G(J) = \mathrm{SL}_2(K) \times \mathrm{SL}_2(K)$ and $S(C,J)$ consists of all pairs $(g_1, g_2) \in \mathrm{GL}_2(K) \times \mathrm{GL}_2(K)$ such that the determinants of $g_1$ and $g_2$ are equal. Nevertheless, the normalizer of $G(J)$ in $R^\times$ coincides with the group $\mathrm{GL}_2(K) \times \mathrm{GL}_2(K)$ which is different from $S(C,J)$ for $K$ contains more than two elements.

We close this paper by making the remark that the converse of Proposition 3 is false. For instance, let us consider a commutative and associative two-dimensional algebra $R$ over a field $P$ with a basis $1, v$ such that 1 is the identity element of $R$ and $v^2 = 0$. It should be clear that $R$ is an algebra with the involution $J$ defined by $J(a + bv) = a - bv$ for all $a, b \in P$. Let $C$ be the subgroup of $R$ consisting of all squares in $P$, that is, $C = \{a^2 \mid a \in P^\times\}$. Then $S(C,J) = R^\times$ and $G(J) = \{\pm 1 + bv \mid b \in P\}$. Observe that the normalizer and the centralizer of $G(J)$ in $R^\times$ coincide with $R^\times$ because $R$ is commutative. However $R^\times \nsubseteq C$ and this is just that we intend to show.

## References


[1] Bashkirov E.L., *Linear groups that contain a root subgroup*, Siberian Math. J. **37** (1996), no. 5, 754–759.

[2] Bashkirov E.L., *Irreducible linear groups of degree four over a quaternion division algebra that contain a subgroup diag*$(T_3(K, \Phi_0), 1)$, J. Algebra **287** (2005), no. 2, 319–350.

[3] Bashkirov E.L., *Irreducible linear groups of degree four over a quaternion division algebra that contain a root subgroup*, Comm. Algebra **34** (2006), no. 6, 1931–1948.

[4] Bashkirov E.L., *Completely reducible linear groups over a quaternion division algebra that contain a root subgroup*, Comm. Algebra **35** (2007), no. 3, 1019–1054.

[5] Dieudonné J., *La Géométrie des Groups Classiques*, Ergebnisser der Mathematik, Springer, Berlin-New York, 1997.

[6] Dixon J.D., *The Structure of Linear Groups*, Van Nostrand Reinhold Company, London, 1971.

[7] Dye R.H., *Maximal subgroups of* $\mathrm{GL}_{2n}(K)$*,* $\mathrm{SL}_{2n}(K)$*,* $\mathrm{PGL}_{2n}(K)$ *and* $\mathrm{PSL}_{2n}(K)$ *associated with symplectic polarities*, J. Algebra **66** (1980), no. 1, 1–11.

[8] King O.H., *On subgroups of the special linear group containing the special orthogonal group*, J. Algebra **96** (1985), no. 1, 178–193.

[9] King O.H., *On subgroups of the special linear group containing the special unitary group*, Geom. Dedicata **19** (1985), no. 3, 297–310.

[10] O'Meara O.T., *Symplectic Groups*, American Mathematical Society, Providence, R.I., 1978.

[11] Zalesskiĭ A.E., Serežkin V.N., *Linear groups generated by transvections*, Izv. Akad. Nauk SSSR. Ser. Mat. **40** (1976), no. 1, 26–49.



Smolensk State University Przhevalsky st. 4 Smolensk 214000 Russia

*E-mail*: bashkirov57@mail.ru


(*Received* October 12, 2007, *revised* May 2, 2008)