Conpseudosimilarity and Consemisimilarity over a Division Ring

Jean H. Bevis and Frank J. Hall
Department of Mathematics and Computer Science
Georgia State University
Atlanta, Georgia 30303

and

Robert E. Hartwig
Mathematics Department
North Carolina State University
Raleigh, North Carolina 27695-8205

Submitted by Richard A. Brualdi

ABSTRACT

It is shown that for $n \times n$ matrices over a division ring which is finite dimensional over its center, the notions of consimilarity, conpseudosimilarity and consemisimilarity are all equivalent, provided the conjugation is strong.

1. INTRODUCTION

In this paper we extend the results of [1] to nonelementary conjugations $(\cdot)\hat{}$ for matrices over a division ring $\mathbb{D}$ which is finite dimensional over its center.

Given a ring $\mathscr{R}$, a conjugation $(\cdot)\hat{}$ on $\mathscr{R}$ is any involutory automorphism, that is, any mapping $(\cdot)\hat{}$ which satisfies

$$a\hat{}\,\hat{} = a, \qquad (a+b)\hat{} = a\hat{} + b\hat{}, \quad \text{and} \quad (ab)\hat{} = a\hat{}\, b\hat{}$$

for any $a, b \in \mathscr{R}$. In particular the identity map is a conjugation, as is "complex" conjugation for the ring $\mathbb{C}$. There are, however, an infinite number of other conjugations on $\mathbb{C}$. We shall primarily be interested in the

655 Avenue of the Americas, New York, NY 10010 0024-3795/90/$3.50

ring $\mathscr{R} = \mathbb{D}_{n\times n}$ of $n \times n$ matrices over a division ring $\mathbb{D}$, with associated conjugation $(\cdot)^{\wedge}$.

It is easily seen that each conjugation $(\bar{\cdot})$ on $\mathbb{D}$ induces a conjugation on $\mathbb{D}_{n\times n}$ via $(A^{\wedge})_{ij} = \bar{a}_{ij}$. The converse need not be true in general, but is true for matrices over a field, as we shall demonstrate.

Throughout, we shall use $\mathbb{D}_{m\times n}$ to denote the set of all $m \times n$ matrices over $\mathbb{D}$, and shorten $\mathbb{D}_{m\times 1}$ to $\mathbb{D}^m$ and $\mathbb{D}_{1\times n}$ to $\mathbb{D}_n$. The (right row) rank of a matrix A over $\mathbb{D}$ will be denoted by $\rho(A)$. See, for example, [5, pp. 22, 51].

Given a conjugation $(\bar{\cdot})$ on $\mathbb{D}$, a semilinear map T on $\mathbb{D}^n$ (considered as a right vector space) is defined by

$$(\mathbf{x}+\mathbf{y}\alpha)T = \mathbf{x}T + \mathbf{y}T\bar{\alpha}.$$

Relative to any given basis $\mathscr{B} = (\mathbf{b}_i)$ for $\mathbb{D}^n$, the semilinear map T induces a unique matrix $A = [\mathbf{a}_1, \dots, \mathbf{a}_n]$ via

$$\mathbf{b}_i T = [\mathbf{b}_1, \dots, \mathbf{b}_n]\mathbf{a}_i, \qquad i = 1, \dots, n. \tag{1.1}$$

As usual, we denote this matrix A by $[T]_{\mathscr{B}}$.

Conversely, given a conjugation $(\bar{\cdot})$, any matrix A induces a semilinear map relative to the basis $\{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ via

$$[\mathbf{b}_1, \dots, \mathbf{b}_n]\mathbf{x}T = [\mathbf{b}_1, \dots, \mathbf{b}_n]A\bar{\mathbf{x}}. \tag{1.2}$$

Moreover, for this induced map T, its matrix representative $[T]_{\mathscr{B}}$ equals A. We shall also need the fact that if U is another semilinear map on $\mathbb{D}^n$, with matrix B (relative to the same basis and same conjugation), then

$$[TU]_{\mathscr{B}} = B\bar{A}, \tag{1.3}$$

where $(\bar{A})_{ij} = \bar{a}_{ij}$. In particular T^2 is a *linear map*, with $[T^2]_{\mathscr{B}} = A\bar{A}$.

The concept of "consimilarity," using an entrywise conjugation, enters naturally when one changes bases. Indeed, matrices A and B represent one and the same semilinear map exactly when

$$(S^{\wedge})^{-1}BS = A \tag{1.4}$$

for some invertible S. If this happens for a conjugation $(\cdot)^{\wedge}$ on $\mathbb{D}_{n\times n}$, we say that B is *consimilar* to A. It should be clear that consimilarity depends crucially on the *particular* conjugation on $\mathbb{D}_{n\times n}$ that is being used. To emphasize this dependence, we shall replace the prefix "con" by "^," write "^-similarity" for "consimilarity," and denote the relation by $\underset{\wedge}{\approx}$.

Related to ^-similarity are ^-pseudo- and ^-semisimilarity. These may be defined as follows:

(1) B is *^-pseudosimilar* to A if $(X^g)^\wedge AX = B$ and $X^\wedge BX^G = A$ for some inner inverses X^g and X^G of X, i.e. $XX^gX = X = XX^GX$.

(2) B is *^-semisimilar* to A if $Y^\wedge AX = B$ and $X^\wedge BY = A$ for some X and Y.

If ^-is the identity map, we shall drop the "^" in each of these definitions.

Our aim in this note is to show that if $\mathbb{D}$ is finite dimensional over its center and the conjugation on $\mathbb{D}_{n\times n}$ is well behaved, then these three types of similarity are equivalent. Before we address this problem, let us first examine the question of what an arbitrary conjugation on $\mathbb{D}_{n\times n}$ looks like.

2. CONJUGATIONS ON $\mathbb{D}_{n\times n}$

If Γ is the center of $\mathbb{D}$, and $z \in \Gamma$ is nonzero, then it is easily seen that

$$A^\wedge = H^{-1}AH \qquad \text{with} \qquad H^2 = zI \tag{2.1}$$

is a conjugation on $\mathbb{D}_{n\times n}$. We shall refer to these conjugations as *weak*. If no such H and z exist for a conjugation $(\cdot)^\wedge$ on $\mathbb{D}_{n\times n}$, we call the conjugation *strong*. We shall see that our results will only go through for strong conjugations on $\mathbb{D}_{n\times n}$, with finite $(\mathbb{D}:\Gamma)$.

We shall first have to characterize *all* conjugations on $\mathbb{D}_{n\times n}$. To do this we use a result by Jacobson [5, p. 237, Theorem 8, Exercise 5].

THEOREM 1. *All conjugations on $\mathbb{D}_{n\times n}$ have the form $A^\wedge = L^{-1}\bar{A}L$ for some invertible matrix L and some automorphism $(\bar{\cdot})$ on $\mathbb{D}$ such that*

$$\bar{L}L = xI, \qquad \bar{x} = x \neq 0, \quad \textit{and} \quad \bar{\bar{d}} = xdx^{-1} \quad \textit{for all } d \in \mathbb{D}.$$

Proof. All automorphisms on $\mathbb{D}_{n\times n}$ have the form $\phi(A) = L^{-1}\bar{A}L$, where $(\bar{\cdot})$ is an automorphism on $\mathbb{D}$ and $(\bar{A})_{ij} = \bar{a}_{ij}$ [5]. Now $\phi(\phi(A)) = A$ is equivalent to $L^{-1}(\overline{L^{-1}\bar{A}L})L = A$, which reduces to $\bar{\bar{A}}X = XA$ with $X = \bar{L}L$. This must hold for all choices of A, and in particular for $A = E_{ij}$. Because $\bar{E}_{ij} = E_{ij}$, we see that $E_{ij}X = XE_{ij}$. This forces X to become a diagonal matrix $X = xI$. To establish the remaining conditions, observe that $\bar{L}L = xI$, which gives $\bar{L} = xL^{-1}$. Consequently $\bar{\bar{L}} = \bar{x}\cdot\overline{L^{-1}} = \bar{x}(\bar{L})^{-1} = \bar{x}Lx^{-1}$. But $\bar{\bar{A}} = XAX^{-1} = xAx^{-1}$ implies that $\bar{\bar{L}} = xLx^{-1}$. Comparing these two results shows that $x = \bar{x}$. Since $(\bar{\cdot})$ acts elementwise, it follows that $\bar{\bar{d}} = xdx^{-1}$ for all d in $\mathbb{D}$. For later use we note that $\overline{L^{-1}} = Lx^{-1}$.

We now observe that if $(\cdot)^\wedge$ is a strong conjugation, then $\bar{d}$ cannot be inner. Indeed, suppose that $\bar{d} = ydy^{-1}$ with $y = \bar{y}$. Then $\bar{A} = yAy^{-1}$ and $A^\wedge = L^{-1}yAy^{-1}L = H^{-1}AH$, where $L = yH$. Now $\bar{\bar{d}} = y^2dy^{-2} = xdx^{-1}$, and so $x^{-1}y^2d = dx^{-1}y^2$. In other words $x^{-1}y^2$ is central. Lastly, $xI = \bar{L}L = y\bar{H}yH = y(yHy^{-1})yH = y^2H^2$, which shows that $H^2 = y^{-2}xI = zI$, in which z is central. ∎

For the case of a field, a great simplification occurs.

COROLLARY 1. *All conjugations on* $\mathbb{F}_{n\times n}$ *are of the form* $A^\wedge = L^{-1}\bar{A}L$, *where* $(\bar{\cdot})$ *is a conjugation on* $\mathbb{F}$, *and*

$$\bar{L}L = xI \qquad \textit{with} \quad x = \bar{x} \neq 0. \tag{2.2}$$

Proof. Since all scalars commute, we see from Theorem 1 that $\bar{\bar{A}} = A$ for all A. In particular since $(\bar{\cdot})$ is defined entrywise, this means that $\bar{\bar{d}} = d$ for all $d \in \mathbb{F}$. Thus the automorphism $(\bar{\cdot})$ must be a conjugation on $\mathbb{F}$. Moreover, $(dI)^\wedge = L^{-1}\bar{d}IL = L^{-1}L\bar{d}I = \bar{d}I$. This shows that the conjugation $(\cdot)^\wedge$ restricted to scalar matrices does agree with $(\bar{\cdot})$. ∎

Another case of special interest is where $x = \bar{x} = \pm\bar{y}y$ for some $y \in \mathbb{D}$. In this case we may set $H = y^{-1}L$ and $A' = y^{-1}\bar{A}y$. It is easily seen that $(\cdot)'$ is a conjugation. Moreover $A^\wedge = L^{-1}\bar{A}L = H^{-1}A'H$, while $H'H = y^{-1}\bar{H}yH = (\bar{y}y)^{-1}\bar{L}L = \pm x^{-1}\bar{L}L = \pm I$, which is the "cleanest" form for a conjugation.

In particular, since every real number x has the form $x = \pm\bar{y}y$ for some y, it follows that *all* conjugations on $\mathbb{C}_{n\times n}$ are of the form

$$A^\wedge = H^{-1}\bar{A}H \qquad \text{with} \quad \bar{H} = \pm H^{-1}, \tag{2.3}$$

where $(\bar{\cdot})$ is a conjugation on $\mathbb{C}$. We now come to our main result.

3. A RELATION BETWEEN THE VARIOUS TYPES OF SIMILARITY

In the remainder of this paper we assume that $\mathbb{D}$ is finite dimensional over its center. Let us now investigate how the three types of similarity are related, thereby generalizing the results of [1].

THEOREM 2. *Let $\mathbb{D}$ be finite dimensional over its center, and suppose that $(\cdot)^\wedge$ is a strong conjugation on $\mathbb{D}_{n\times n}$. Then the following are equivalent:*

(i) *B is $^\wedge$-similar to A;*
(ii) *B is $^\wedge$-pseudosimilar to A;*
(iii) *B is $^\wedge$-semisimilar to A;*
(iv) *$AA^\wedge$ is pseudosimilar to $BB^\wedge$, and $\rho[(AA^\wedge)^k A] = \rho[(BB^\wedge)^k B]$ for all $k \geqslant 0$;*
(v) *$AA^\wedge \approx BB^\wedge$, and $\rho[(AA^\wedge)^k A] = \rho[(BB^\wedge)^k B]$ for all $k \geqslant 0$.*

Proof. We shall do this in three parts. First we have four easy implications.

(i) $\Rightarrow$ (ii): Clear. Take $X = S^{-1}$, $X^g = X^G = S$.

(ii) $\Rightarrow$ (iii): Let us verify that $Y = X^g X X^G$ will do. Indeed, since $X^\wedge (X^G)^\wedge A = A = AXX^G$, $BX^g X = B$, and $AX = X^\wedge B$, we see that

$$(X^g)^\wedge X^\wedge (X^G)^\wedge AX = (X^g)^\wedge AX = B$$

while

$$X^\wedge BX^g XX^G = AXX^g XX^G = AXX^G = A.$$

(iii) $\Rightarrow$ (iv): This follows exactly as in Theorem 2.2 of [1].

(iv) $\Rightarrow$ (v): It was shown in [2, Theorem 1] that in a p.u.r. ring, pseudosimilarity and similarity are equivalent.

Before we complete the circle, several remarks are here in place.

REMARKS.

(1) The algebraic manipulations work because $(\cdot)^\wedge$, $(\cdot)^g$, and the Drazin inverse operation $(\cdot)^d$ all commute. That is,

$$(A^\wedge)^d = (A^d)^\wedge \quad \text{and} \quad (A^g)^\wedge \in \{(A^\wedge)^g\}.$$

(2) It also follows in Theorem 2 that (ii) $\Rightarrow$ (i). In fact, if $Q = (I - XX^G + XU)U^{-1}(I - X^g X + UX)$, where $XUX = X$ with U invertible, then Q is invertible and $(Q^{-1})^\wedge AQ = B$.

(3) For a weak conjugation, (v) need not imply (i). For example, let $(\bar{\cdot})$ be the identity, $z = 1$, and $H = I$. Then if

$$A = I \quad \text{and} \quad B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

we get $A\bar{A} = I = B\bar{B}$, yet A is not similar to B.

The only thing remaining is to show that (v) implies (i) for a strong conjugation.

Suppose that (v) holds and $P^{-1}AA^{\wedge}P = BB^{\wedge}$. Then $(P^{-1}AP^{\wedge})P^{\wedge\,-1}A^{\wedge}P = BB^{\wedge}$ or $A^{\sim}A^{\sim\wedge} = BB^{\wedge}$, where $A^{\sim} = P^{-1}AP^{\wedge}$. In addition, $(A^{\sim}A^{\sim\wedge})^kA^{\sim} = P^{-1}(AA^{\wedge})^kAP^{\wedge}$, so that $\rho[(A^{\sim}A^{\sim\wedge})^kA^{\sim}] = \rho[(AA^{\wedge})^kA] = \rho[(BB^{\wedge})^kB]$. Now assume that $(\cdot)^{\wedge}$ is strong and is given by Theorem 1. If we set $A' = A^{\sim}L^{-1}$ and $B' = BL^{-1}$, and recall that $\overline{L^{-1}} = Lx^{-1}$, then $A'\bar{A'} = (A^{\sim}L^{-1})(\bar{A}^{\sim}\overline{L^{-1}}) = (A^{\sim}L^{-1})(\bar{A}^{\sim}Lx^{-1}) = A^{\sim}A^{\sim\wedge}x^{-1}$, and similarly $B'\bar{B}' = BB^{\wedge}x^{-1}$. In other words $A'\bar{A'} = B'\bar{B}'$. Moreover,

$$\rho\left[(A'\bar{A}'x)^kA'\right] = \rho\left[(A^{\sim}A^{\sim\wedge})^kA^{\sim}\right] = \rho\left[(BB^{\wedge})^kB\right] = \rho\left[(B'\bar{B}'x)^kB'\right].$$

Let us now reduce our problem to the invertible and nilpotent subcases. Since $(\bar{\cdot})$ is an automorphism on $\mathbb{D}$, we may apply Fitting's decomposition [4, p. 28] to the matrices A' and B'. This states that there exist invertible Q and R such that

$$Q^{-1}A'\bar{Q} = \begin{bmatrix} V & 0 \\ 0 & \eta \end{bmatrix}, \qquad R^{-1}B'\bar{R} = \begin{bmatrix} W & 0 \\ 0 & \zeta \end{bmatrix}, \tag{3.1}$$

where V and W are invertible and where $N_k(\eta) = \eta\bar{\eta}\bar{\bar{\eta}}\cdots\eta^{(\bar{\cdot})^k}$ and $N(\zeta)$ vanish for sufficiently large k.

Using the fact that $\bar{\bar{M}} = xMx^{-1}$, it follows that $M^{(\bar{\cdot})^{2i}}M^{(\bar{\cdot})^{2i+1}} = x^iM\bar{M}x^{-i}$ and hence $N_{2l+1}(M) = (M\bar{M}x)^{l+1}x^{-l-1}$. Hence

$$N_k(\eta) = 0 \quad \text{for some } k \quad \text{iff} \quad (\eta\bar{\eta}x)^l = 0 \quad \text{for some } l, \tag{3.2a}$$

$$N_k(\zeta) = 0 \quad \text{for some } k \quad \text{iff} \quad (\zeta\bar{\zeta}x)^l = 0 \quad \text{for some } l. \tag{3.2b}$$

Moreover,

$$\begin{bmatrix} V\bar{V}x & 0 \\ 0 & \eta\bar{\eta}x \end{bmatrix} = Q^{-1}(A'\bar{A}'x)Q \quad \text{and} \quad R^{-1}(B'\bar{B}'x)R = \begin{bmatrix} W\bar{W}x & 0 \\ 0 & \zeta\bar{\zeta}x \end{bmatrix}. \tag{3.3}$$

Since $A'\bar{A}'x = B'\bar{B}'x$, it follows (over any ring) that

$$V\bar{V}x \approx W\bar{W}x \quad \text{and} \quad \eta\bar{\eta}x \approx \zeta\bar{\zeta}x. \tag{3.4}$$

In addition, since over a division ring any matrix with left and right inverses must be square, we may conclude that over our division ring, $V\bar{V}$ and $W\bar{W}$, and hence V and W, *all* have the same size. Similarly η and ζ have the same size. Furthermore, the rank conditions $\rho[(A'\bar{A}'x)^k A'] = \rho[(B'\bar{B}'x)^k B']$, ensure that

$$\rho\begin{bmatrix} (V\bar{V}x)^k V & 0 \\ 0 & (\eta\bar{\eta}x)^k \eta \end{bmatrix} = \rho\begin{bmatrix} (W\bar{W}x)^k W & 0 \\ 0 & (\zeta\bar{\zeta}x)^k \zeta \end{bmatrix}.$$

Hence $\rho[(\eta\bar{\eta}x)^k\eta] = \rho[(\zeta\bar{\zeta}x)^k\zeta]$ for all $k \geqslant 0$. Moreover, from (3.4) we see that $\rho[(\eta\bar{\eta}x)^k] = \rho[(\zeta\bar{\zeta}x)^k]$ for all $k \geqslant 0$.

Consequently our problem has been split into the invertible and nilpotent subproblems:

(1) Given $V\bar{V}x \approx W\bar{W}x$ with V and W invertible, show V is consimilar to W with respect to the conjugation $(\bar{\cdot})$.

(2) Given $\eta\bar{\eta}x \approx \zeta\bar{\zeta}x$ with (3.2) valid and $\rho[(\eta\bar{\eta}x)^k\eta] = \rho[(\zeta\bar{\zeta}x)^k\zeta]$ for all $k \geqslant 0$, show η is consimilar to ζ with respect to the conjugation $(\bar{\cdot})$.

The nilpotent case is easily disposed of. Indeed, if T and U are the semilinear maps induced on $\mathbb{D}^m$ by $(\eta,(\bar{\cdot}))$ and $(\zeta,(\bar{\cdot}))$, respectively, then by (3.2) we know that the matrix representations $N(\eta)$ and $N(\zeta)$ of T^k and U^k vanish. Hence both T and U are nilpotent semilinear maps. Next, using the Jordan chain decomposition of $\mathbb{D}^m$, for a nilpotent semilinear map we obtain $S^{-1}\eta\bar{S} = \text{diag}[J_{n_1}(0),\dots,J_{n_k}(0)] = M$, where $J(0) = [\delta_{i,j+1}]$ and S is some invertible matrix. Likewise, $K^{-1}\zeta\bar{K} = \text{diag}[J_{m_1}(0),\dots,J_{m_r}(0)] = M'$. Since $S^{-1}(\eta\bar{\eta}x)S = M^2x$ and $S^{-1}[(\eta\bar{\eta}x)^k\eta]\bar{S} = M^{2k+1}x$, and a similar result holds for $\zeta\bar{\zeta}x$, it follows from the rank conditions that $\rho(M^k) = \rho(M'^k)$ for all $k \geqslant 0$. Thus $M = M'$, ensuring that η and ζ are consimilar.

We now come to the invertible case. Suppose $V\bar{V}x \approx W\bar{W}x$ and both are invertible. Since $(\cdot)^\wedge$ is strong, and $\bar{\bar{d}} = xdx^{-1}$ is the smallest power of $(\bar{\cdot})$ that is inner, it follows by Theorem 34 of [4] (with $\mu = 1$, $r = 2$, and $\xi^s = \bar{\xi}$) that V and W are consimilar.

Combining the two subcases, we may conclude that $Y^{-1}A'\bar{Y} = B'$ for some invertible Y, and hence $Y^{-1}A^\sim\hat{Y} = B$. This in turn ensures that A and B are $\hat{}$-similar, since $\hat{}$-similarity is an equivalence relation. ■

We close with an open problem. What extra conditions are needed for a *weak* conjugation, to allow us to go from (v) to (i)?

The authors wish to thank the referee for several valuable comments and references.

REFERENCES

1 J. Bevis, F. Hall, and R. E. Hartwig, Pseudoconsimilarity and semiconsimilarity of complex matrices, *Linear Algebra Appl.* 90:73–80 (1987).
2 F. Hall, R. E. Hartwig, J. Katz, and M. Newman, Pseudo-similarity and partial unit regularity, *Czechoslovak Math. J.* 33(108):361–372 (1983).
3 Y. P. Hong and R. A. Horn, A canonical form for matrices under consimilarity, *Linear Algebra Appl.* 102:143–168 (1988).
4 N. Jacobson, *The Theory of Rings*, Amer. Math. Soc., Providence, 1943, p. 28.
5 N. Jacobson, *Lectures in Abstract Algebra*, Vol. 2, Von Nostrand, Princeton, N.J., 1953.

Received 5 December 1988; final manuscript accepted 24 March 1989