SIMILARITY OF SETS OF MATRICES
OVER A SKEW FIELD

Ph. D. Thesis

WALTER SCOTT SIZER
Bedford College, University of London
ABSTRACT

This thesis looks at various questions in matrix theory over skew fields. The common thread in all these considerations is the determination of an easily described form for a set of matrices, as simultaneously upper triangular or diagonal, for example.

The first chapter, in addition to giving some results which prove useful in later chapters, describes the work of P. M. Cohn on the normal form of a single matrix over a skew field. We use these results to show that, if the skew field \( D \) has a perfect center, then any matrix over \( D \) is similar to a matrix with entries in a commutative field.

The second chapter gives some results concerning commutativity, including the upper triangularizability of any set of commuting matrices, conditions allowing the simultaneous diagonalization of a set of commuting diagonalizable matrices, and a description, over skew fields with perfect centers, of matrices commuting with a given matrix. We end the chapter with a consideration of the problem of when a set of matrices over a skew field \( D \) is similar to a set of matrices with entries in a commutative subfield of \( D \).

The questions of simultaneously upper triangularizing and diagonalizing semigroups of matrices are considered in the third chapter. A closure operation is defined on semigroups of matrices over a skew field, and it is shown that a semigroup is upper triangularizable.
(diagonalizable) if and only if its closure is. Necessary
and sufficient conditions are then given for closed semi-
groups to be upper triangularizable (diagonalizable).

The last chapter gives a few assorted results on groups
of matrices, including the simultaneous upper triangulariza-
bility of a solvable group of unipotent matrices and a deter-
mination, for any skew field $D$, of those finite groups all
of whose representations over $D$ are diagonalizable.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>5</td>
</tr>
<tr>
<td>1. Preliminaries, and the Normal Form of a Single Matrix</td>
<td>7</td>
</tr>
<tr>
<td>2. Commutativity Results</td>
<td>17</td>
</tr>
<tr>
<td>3. Upper Triangularizing and Diagonalizing Semigroups of Matrices</td>
<td>39</td>
</tr>
<tr>
<td>4. Various Results on Linear Groups</td>
<td>55</td>
</tr>
<tr>
<td>Open Questions</td>
<td>67</td>
</tr>
<tr>
<td>References</td>
<td>69</td>
</tr>
</tbody>
</table>
CORRECTIONS TO "SIMILARITY OF SETS OF MATRICES OVER A SKEW FIELD"

(Ph. D. Thesis, Walter S. Sizer)

p. 10, l. 4: \( w_j = v_j t^{-1} \sum_{i=1}^{n} v_i^m i_j \)

p. 35, l. 13-14: inner automorphism of \( E_n \) (\( E \) an extension of \( D \)).

p. 42, l. 18: (1) and (8) follow from (2)-(7)

p. 53, l. 3: \( (e_{11} + e_{21}) \begin{pmatrix} 1 & a_2 & \cdots & a_{1n} \\ \vdots & & & \vdots \\ 0 & \cdots & 1 \\ 1 & a_{12} & \cdots & a_{1n} \end{pmatrix} = \begin{pmatrix} 1 & a_1 & \cdots & a_{1n} \\ \vdots & & & \vdots \\ 0 & \cdots & 1 \\ 1 & a_{12} & \cdots & a_{1n} \end{pmatrix} \)

p. 63, l. 1-2: exponent \( e \) of \( \varnothing \).
INTRODUCTION

The study of matrices over commutative fields has been pursued by mathematicians since the middle of the nineteenth century; the quaternion skew field was first described before 1850, with more general skew fields being described and studied since the early twentieth century; matrix rings over skew fields, as rings, have also been of interest since the early 1900’s ([2], pp. 204-205, 200-201, 251, 252); but the study of matrices over arbitrary skew fields represents a much newer development. In this thesis we look at precisely this subject.

Many of the results in this thesis—most of chapter 2 and half of chapter 4—are just analogues of well known results from the theory of linear algebra over commutative fields. Indeed, many of the proofs of these results are just adaptations of well established techniques from matrix theory. This is not to say that the results are uninteresting, however, for matrices over commutative fields possess properties not shared by matrices over skew fields, and so not all proofs, indeed not all results, can be carried over.

Some of the results reported in this thesis either go beyond what was known in the case of commutative fields or answer questions which do not make sense for matrices over commutative fields. To the author’s knowledge, the results in chapter 3 on upper triangularizing semigroups of matrices come in the former category. Under the latter heading we find the treatment in chapter 4 of finite non-
abelian groups all of whose representations are diagonalizable and the results in chapter 2 on when a set of matrices will be similar to a set of matrices with entries in a commutative field.

Detailed summaries of the contents of the various chapters are given at the beginning of each chapter.

All unoriginal work is, we trust, sufficiently credited to its rightful sources. In particular, all numbered items were a result of the author's own work. The author would, however, like to acknowledge some illuminating conversations with Dr. Warren Dicks, and would particularly like to thank Professor P. M. Cohn for his advice and enthusiasm, and especially for encouraging the author in his pursuit of this subject.

Bedford College
University of London
London, England
November, 1975
1. PRELIMINARIES, AND THE NORMAL FORM OF A SINGLE MATRIX

In this chapter we describe some basic results on skew fields and the similarity of matrices over skew fields, and give the conventions we will follow throughout the work. Many of the preliminary results are due to P. M. Cohn, and we merely state them with appropriate references, in some instances giving an indication of how they will be useful in our work.

Throughout, $D$ will be a skew field with center $k$; any further restrictions on $D$ or $k$ will be stated explicitly. The letters $E$ and $K$ will also denote skew fields, containing $k$ (or an isomorphic copy of $k$) in their centers, and $F$ will denote a commutative field containing $k$ (or an isomorphic copy of $k$). The ring of $n \times n$ matrices over a ring $R$ will be denoted $R_n$; $R^n$ will be used to denote the $n \times 1$ matrices over $R$ ("column vectors"), a right $R^n$, left $R_n$-module in the natural way.

Most of our interest will be concentrated on questions of the similarity of matrices. If $\mathcal{X} \subseteq D_n$ we will frequently want to show that $\mathcal{X}$ is similar to a set $\mathcal{J}$ of matrices with additional properties; as we are always working over arbitrary skew fields, we will not differentiate between "similar over $D$" and "similar over an extension $E$ of $D$ with $k \subseteq \text{center}(E)$", but we shall not consider similarity over extensions which do not preserve the center. For ease in notation, we will generally assume in our proofs that if two sets of matrices are similar over some extension of $D$ then we have chosen $D$ to be large.
enough so they are in fact similar over \( D \); our statements of our results, however, will generally either note parenthetically that we allow extensions preserving the center, or else will not specify any reference field.

While almost all the results discussed in this thesis are expressed in terms of matrices, we make extensive use of the fact that a matrix represents a linear transformation with respect to a particular choice of basis, and that similar matrices represent the same linear transformation with respect to different bases. We shall try to maintain the distinction between matrices and linear transformations, however, and not use terminology appropriate to linear transformations when referring to matrices. To help maintain the distinction we adopt the following conventions:

1. If \( M \in D_n \) is a matrix, we denote by \( \phi_M \) the corresponding linear transformation of \( D^n \); if \( \lambda \in D_n \) is a set of matrices, \( \phi_\lambda = \{ \phi_\mu | \mu \in \lambda \} \).

2. If \( \varphi \in \text{End}(D^n) \) and \( B \) is a basis of \( D^n \), \( \mu(\varphi, B) \) will denote the matrix representing \( \varphi \) with respect to \( B \); if \( B \) is not explicitly given, \( \mu(\varphi) \) will be used to denote a representation of \( \varphi \). If \( \lambda \in \text{End}(D^n) \), \( \mu(\lambda, B) = \{ \mu(\varphi, B) | \varphi \in \lambda \} \) and \( \mu(\lambda) = \{ \mu(\varphi) | \varphi \in \lambda \} \).

We will be concerned with matrices over skew fields, and so it is natural we should first consider skew fields themselves. Here two theorems of P. M. Cohn will prove extremely useful:

**Theorem A:** ([3], cor. 1 to th. 6.1, p.210) Let \( K \) be a skew field, \( K_\lambda (\lambda \in \Lambda) \) a family of skew fields each containing \( K \). Then the free product of the \( K_\lambda 's \) amalgamating \( K \) can be embedded in a skew field.
Theorem B: ([3], th. 6.3, p. 211) Let \( K \) be a skew field containing isomorphic skew subfields \( K_1, K_2 \). Then for any isomorphism \( \psi: K_1 \rightarrow K_2 \) there is an extension \( D \) of \( K \) such that \( \psi \) is induced by an inner automorphism of \( D \). If \( k \) is the center of \( K \), \( k \subseteq K_1 \cap K_2 \), and \( k \) is fixed by \( \psi \), then we can choose \( D \) so that \( k \subseteq \text{center}(D) \).

In general our use of theorem B will be straightforward, and we trust the reader will have no difficulty in interpreting our references to it. Perhaps, however, an additional word on a frequent context of our use of theorem A would be in order. For the most part we will have a skew field \( D \) with center \( k \), and a skew subfield \( K \) of \( D \) containing \( k \). We will know that there is an extension \( E \) of \( K \) for which a particular statement is true, and so by theorem A there will be a skew field extension \( D_0 \) of \( D \) containing \( k \) in its center ---namely the universal field of fractions of \( D \cdot E \)---for which the statement is true.

The basic starting point for considering questions of the similarity of matrices over a skew field is the paper by P. M. Cohn, "The Similarity Reduction of Matrices over a Skew Field" ([5]). We briefly summarize this theory of a single matrix over a skew field, as it is essential to what follows (all this material can be found in [5], with references to [4]; some of our definitions of terms defined in [4] will not be those of [4], but will be equivalent by propositions in [4]).

Given a matrix \( M = (m_{ij}) \in D_n \), \( V = D^n \) becomes a right \( D[t] \) module under the action \( v \cdot x = \sum v_{ij}x_i \), \( v \in V \), \( x_i \in D \). Then two matrices \( M \) and \( N \) will be similar if and only if the \( D[t] \)
modules they define are isomorphic. Now $V$ is a finitely generated module over the (non-commutative) principal ideal domain $D[t]$, the quotient of the free module $\bigoplus_{i=1}^{n} v_i D[t]$ by the module generated by $\{ w_j = \sum_{i=1}^{n} \eta_i v_i | j = 1, \ldots, n \}$. By results on the equivalence of matrices over a non-commutative PID, $M$ is equivalent to a diagonal matrix $\text{diag}[\delta_1, \ldots, \delta_n]$, where $\delta_1$ is a total divisor of $\delta_{i+1}$, and it follows that $V = D[t]/\delta_1 D[t] \oplus \cdots \oplus D[t]/\delta_n D[t]$. Recall that $\delta_1$ is a total divisor of $\delta_{i+1}$ means that there is an element $a_1 \in D[t]$ with $a_1 D[t] = D[t] a_1$ (an invariant element) such that $\delta_1 | a_1$ and $a_1 | \delta_{i+1}$. An element $\delta \in D[t]$ dividing an invariant element $a$ is said to be bounded; a minimal invariant element (under the ordering of divisibility) $\tilde{a}$ for which $\delta | \tilde{a}$ is called a bound for $\delta$. Note that $\delta_1, \ldots, \delta_{n-1}$ are bounded, but $\delta_n$ need not be.

Now $D[t]/\delta_1 D[t] \oplus \cdots \oplus D[t]/\delta_n D[t]$ can be decomposed into a sum $\bigoplus_{i=1}^{s} D[t]/\alpha_i D[t] \oplus \cdots \oplus D[t]/\alpha_s D[t] = V$, where each $\alpha_i$, $i = 1, \ldots, s-1$, is bounded and indecomposable, and $\alpha_s$ has no bounded non-unit factors; furthermore, the expression (1.a) will then be unique up to isomorphism. The $\alpha_i$'s, $i < s$, are called the elementary divisors of $M$. If we pick a basis of $D^n$ corresponding to (1.a) we get a representation of $\phi_M$ as a diagonal sum of matrices

(1.b) $\mu(\phi_M) = M_1 + \cdots + M_s$, where $M_i$ has elementary divisor $\alpha_i$ for $i \leq s$.

The invariant elements of $D[t]$ are just those associated to polynomials in $k[t]$. Thus if $\alpha_1$ is bounded, say with bound $p_1(t)$, we can pick $p_1(t)$ to lie in $k[t]$. It follows that $p_1(M_1) = 0$, and $M_1$ satisfies a non-zero polynomial with
coefficients in \( k \). Such a matrix \( M_1 \) is called **algebraic**. If \( \alpha_s \) is not a unit, then since \( \alpha_s \) has no non-unit invariant factors, it follows that for any non-zero polynomial \( p(t) \in k[t] \), \( p(M_s) \) is invertible. Such a matrix \( M_s \) is called **transcendental**; many examples of unexpected behavior in matrix theory over skew fields are a result of

**Theorem C**: ([5], cor. 1 to th. 4.2) Any transcendental matrix is similar to a scalar matrix; further, the diagonal entry of the scalar matrix can be taken to be any element of \( D \) transcendental over \( k \).

Thus \( M_s \) in (1.b) can be taken to be scalar.

We want to find an easily described form for the matrices \( M_i \), \( i \leq s \), which were algebraic with elementary divisors \( \alpha_i \). An atom in \( D[t] \) is an irreducible element; if \( a \in D[t] \) is a bounded atom, with monic bound \( a^* \in k[t] \), then we get an atomic factorization \( a^* = a_1 \ldots a_r a_{r+2} \ldots a_q \) of \( a^* \). We can assume by (1.A) that \( D \) contains a commutative splitting field \( F \) of \( a^* \), so we get another atomic factorization \( a^* = (t-\beta_1) \ldots (t-\beta_q) \) of \( a^* \) in \( F[t] \), hence in \( D[t] \). Since atomic factorizations have the same number of factors and \( a \) and all the \( \alpha_i \)'s have degree at least 1, we see that \( a \) is linear—that is to say, we can assume all bounded atoms in \( D[t] \) are linear.

Recall that two elements \( a, b \in D[t] \) are called **similar** if \( D[t]/aD[t] = D[t]/bD[t] \) (since \( D[t] \) is a domain, this notion is equivalent to the notion of GL-relatedness ([4], p. 91; cor. 1; p. 125), and the latter term is used in [5]). Now each \( \alpha_i \) was indecomposable, hence a product of similar bounded atoms. We observe that
(1.c) if \( \alpha_1 = (t-a_1) \ldots (t-a_m) \) is an atomic factorization of \( \alpha_i \) (where \( a \in D \)) then the matrix \[
\begin{pmatrix}
 a_1 & 1 & \cdots & 0 \\
 a_2 & & \ddots & \\
 0 & \ldots & \ldots & 1 \\
 \end{pmatrix}
\]
elementary divisor \( \alpha_i \) -- i.e., \( M_i \) in (1.b) can be taken to be the above matrix. (All the above is treated in [5] and [4]).

Now suppose \( k \) is perfect and \( \alpha_i \) a bounded indecomposable element of \( D[t] \). Then

(1.d) \( D[t]/\alpha_1 D[t] = D[t]/\beta D[t] \) if and only if \( \alpha_i \) and \( \beta \) have the same bound ([4], p. 231, and the definition of similar).

Let \( p_1(t) \in k[t] \) be the monic bound of \( \alpha_i \) since \( \alpha_i \) is indecomposable, \( p_1 \) is a power of an invariant atom ([4], p. 230) -- i.e., \( p_1 = (q_1)^{n_1} \), \( q_1 \) a monic irreducible polynomial in \( k[t] \). We may assume (1.A) that \( D \) contains a commutative splitting field \( F_1 \) of \( q_1 \). Let \( q_1(t) = (t-c_{1_1}) \ldots (t-c_{1_r}) \) be a factorization of \( q_1 \) in \( F_1 \). We claim that \( (t-c_{1_1})^{n_1} \) has bound \( p_1 \); clearly \( (t-c_{1_1})^{n_1} | q_1^{n_1} = p_1 \), so \( (t-c_{1_1})^{n_1} \) has as bound \( q_1^{m_1}, m_1 < n_1 \). Then we get \( q_1^{m_1} = (t-c_{1_1})^{n_1} s(t) = (t-c_{1_1})^{m_1} \prod_{j=2}^{r} (t-c_{1_j})^{m} \). It is easily seen that \( s(t) \in F_1[t], \) a domain, and so we get \( (t-c_{1_1})^{n_1} = \prod_{j=2}^{r} (t-c_{1_j})^{m} \). Now \( c_{1_1} \) is not a zero of the right hand side, as \( k \) is perfect, so \( c_{1_1} \) is not a zero of the left hand side, and \( n_1 = m_1 \). By (1.c) the \( n_1 \times n_1 \) matrix \[
\begin{pmatrix}
 c_{1_1} & 1 & \cdots & 0 \\
 0 & \ldots & \ldots & 1 \\
 \end{pmatrix}
\]
has associated module \( D[t]/(t-c_{1_1})^{n_1} D[t] \); by (1.d), \( D[t]/(t-c_{1_1})^{n_1} D[t] = D[t]/\alpha_i D[t] \), and it follows that \( M_i \) is similar to \( M_i' \). But we were only interested in the similarity class of \( M_i \), so we can take the matrix \( M_i' \).

We picked \( F_1 \) to be any splitting field of \( q_1 \) and \( c_{1_1} \) to be any root of \( q_1 \) in \( F_1 \); we can take all the \( F_i \)'s equal,
say $F_i = F$, $i = 1, \ldots, s-1$, and we can further assume that $F$ contains an element transcendental over $k$. Then by the above we get that the matrix $M$ we started with is similar to a matrix with all its entries in the commutative field $F$. Further, if $\alpha_i$ and $\alpha_j$ have as bounds powers of the same $k[t]$-irreducible polynomial, we can take $c_{ij} = c_{ji}$. Thus we get

**Proposition 1**: If $D$ is a skew field with perfect center $k$, then any matrix $M \in D_n$ is similar to a matrix $J \in F_n$, $F$ a commutative field. As matrix in $F_n$, $J$ is in Jordan canonical form, and diagonal entries of $J$ satisfying the same $k[t]$-irreducible polynomial are equal. (Now using 1.C) Further, there is at most one element appearing on the diagonal of $J$ which is transcendental over $k$, and no 1's appear on the super-diagonal above any occurrence of this element.

Throughout this thesis, whenever we speak of a matrix in normal form we mean, for matrices over skew fields with perfect centers, a matrix having the form of the matrix $J$ in proposition 1; for matrices over other skew fields the only difference will come in blocks where the diagonal entries are purely inseparable over (and not in) $k$, in which case we may no longer be able to take all diagonal entries equal. In any case, a matrix in normal form will be upper triangular with 0's and 1's on the super-diagonal and 0's above the super-diagonal, and at most one block with a transcendental entry on the diagonal and that block scalar.

Before leaving the subject of a single matrix we mention another essential concept developed in (15) in more detail. If $M \in D_n$, $\alpha \in D$, then $\alpha$ is called a right eigenvalue of $M$ if there is a non-zero vector $v \in D^n$ such that $Mv = v\alpha$. 
In this case \( v \) is a right eigenvector of \( M \). The concepts are similarly defined for linear transformations: if \( \psi \in \text{End}(D^n) \), then \( \alpha \in D \) is a right eigenvalue of \( \psi \) if there is a non-zero \( v \in D^n \) such that \( \psi v = \alpha v \); such a \( v \) is called a right eigenvector of \( \psi \). As might be expected, we have

**Theorem D:** ([5], prop. 2.1) If \( \alpha \) is a right eigenvalue for \( M \in D_n \) \((\psi \in \text{End}(D^n)) \) and \( d \in D \{0\} \), then \( d^{-1} \alpha d \) is a right eigenvalue for \( M \) \((\psi) \). If \( \alpha \) is a right eigenvalue for \( M \in D_n \) and \( P \in \text{GL}_n(D) \), then \( \alpha \) is a right eigenvalue for \( P^{-1}MP \).

We now leave the theory of a single matrix, and start looking, as we shall for the rest of this thesis, at sets of matrices.

We close this chapter with a result we will have frequent occasion to cite; this lemma allows us to reduce the problem of simultaneously upper triangularizing (diagonalizing) an arbitrary set of matrices to that of upper triangularizing (diagonalizing) a finite set, a simplification which can be made over commutative fields by just taking a maximal linearly independent subset.

**Lemma 2:** Let \( \{M_\alpha \mid \alpha \in \alpha \} \) be a set of \( n \times n \) matrices over a skew field \( D \). Suppose that for every finite subset \( S \subseteq \alpha \) there are an extension \( D_S \) of \( D \) and a matrix \( P_S \in \text{GL}_n(D_S) \) such that \( P_S^{-1}M_\alpha P_S \) is upper triangular (diagonal) for each \( \alpha \in S \). Then there are an extension \( D \) of \( \alpha \) and a matrix \( P \in \text{GL}_n(D) \) such that \( P^{-1}M_\alpha P \) is upper triangular (diagonal) for all \( \alpha \in \alpha \).

**Proof:** Let \( T = \{\text{all finite non-empty subsets of } \alpha \} \). For \( S \subseteq T \), let \( D_S, P_S \) be as in the hypotheses. Our object is to obtain a suitable ultrafilter \( \mathcal{U} \) on \( T \) so that the skew field \( E = \bigcup_{S \subseteq T} D_S / \mathcal{U} \) has the property of the conclusion. (For the definition of \( \bigcup_{S \subseteq T} D_S / \mathcal{U} \) and the verification that \( E \) is a skew field,
see ([12], p. 65)).

Write \( \mathcal{P} \) for \( \mathcal{P}(T) \), the power set of \( T \). Let
\[
\mathcal{C} = \{ A \in \mathcal{P} \mid \exists X \in T \text{ and } A = \{ U \in T \mid X \subseteq U \} \}.
\]
Then \( \mathcal{C} \) has the finite intersection property, for if \( C_1, \ldots, C_m \in \mathcal{C} \), say \( C_i = \{ U \in T \mid X_i \subseteq U \} \),
then \( X_1 \cup \cdots \cup X_m \in T \) and \( X_1 \cup \cdots \cup X_m \in \bigcap_{1 \leq i \leq m} C_i \), so \( \bigcap_{1 \leq i \leq m} C_i \) is non-empty. It follows that \( \mathcal{C} \) is contained in an ultrafilter \( \mathcal{U} \) on \( T \), and we set \( \mathcal{E} = \bigcup \mathcal{U} D_S / \mathcal{U} \).

We have a natural isomorphism \( \Pi((D_S)_n) \cong (\Pi D_S)_n \), and the induced homorphism \( (\Pi D_S)_n \to E_n \). Let \( P_0 \) be the element of \( \Pi((D_S)_n) \) which assumes the value \( P_S \) on the \( S \)-th factor, and let \( P \) be its image in \( E_n \). If \( Q_0 \) is the element of \( \Pi((D_S)_n) \) which assumes the value \( P^{-1}_S \) on the \( S \)-th factor, and \( Q \) its image in \( E_n \), then \( P_0 = Q_0 P = I \).

We consider \( P^{-1}_S M_\alpha P \) by looking at \( \Pi_\alpha P^{-1}_S M_\alpha P \in \Pi((D_S)_n) \). Let \( Z = \{ S \in T \mid \Pi_\alpha P^{-1}_S M_\alpha P \} \) is upper triangular (diagonal). Since \( \mathcal{U} \) is an ultrafilter, either \( Z \in \mathcal{U} \) or the complement of \( Z \), \( Z' \in \mathcal{U} \). If \( Z \in \mathcal{U} \), \( P^{-1}_S M_\alpha P \) is upper triangular (diagonal) and we are done; thus we want to show \( Z' \notin \mathcal{U} \). Let \( C = \{ S \in T \mid \alpha \subseteq S \} \). Then \( C \in \mathcal{U} \), and for \( S_0 \in C \), \( P^{-1}_{S_0} M_\alpha P_{S_0} \) is upper triangular (diagonal) by the choice of \( P_{S_0} \). Thus \( C \subseteq Z \), so \( C \cap Z' \) is empty. Now \( C \in \mathcal{U} \), so if \( Z \in \mathcal{U} \) we get \( C \cap Z' = \emptyset \in \mathcal{U} \), an impossibility. Therefore \( Z' \notin \mathcal{U} \), as desired.
2. COMMUTATIVITY RESULTS

In this chapter we concern ourselves for the most part with the examination of various properties of commuting matrices. We first show that, as is the case with matrices over a commutative field, any set of commuting matrices over a skew field can be simultaneously upper triangularized. In contrast to the case of matrices over commutative fields, however, a set of commuting diagonalizable matrices cannot necessarily be simultaneously diagonalized: an example is given to illustrate this fact, and proofs are given for some special cases in which simultaneous diagonalization is possible. A further parallel with conventional linear algebra appears in the description of all matrices which commute with a given matrix over a skew field with perfect center. We next give an example to disprove the appealing conjecture that commuting matrices are similar to matrices all of whose entries lie in a commutative field, and digress to give a necessary and sufficient condition for an absolutely irreducible semigroup of matrices to be similar to a semigroup with entries in a commutative field.

As indicated, our starting point is

**Theorem 1:** Let \( \{ M_\alpha \} \) be a set of commuting matrices. Then \( \{ M_\alpha \} \) can be simultaneously upper triangularized (by a similarity transformation leaving the center of fixed).

**Proof:** By (1.2) it suffices to take \( \alpha \) finite, say \( \alpha = 1, \ldots, m \); recall \( \phi_{M_1}, \ldots, \phi_{M_m} \) are the linear transformations of \( D_n^\alpha \) corresponding to \( M_1, \ldots, M_m \).

We use induction on \( n \) to show that the theorem holds in case \( \phi_{M_1}, \ldots, \phi_{M_m} \) leave a non-trivial subspace \( W \) of \( V = D_n^\alpha \)
variant: the theorem clearly holds for \(n=1\); we assume \(n>1\) and that the result is true for all \(n'<n\); further suppose \(W\) is a non-trivial subspace of \(V\) left invariant by \(\phi_{M_1}, \ldots, \phi_{M_m}\). Let \(v_1, \ldots, v_n\) be a basis of \(V\) such that \(v_1, \ldots, v_r\) are a basis of \(W\). As \(W\) is non-trivial, \(0<r<n\). With respect to this basis, each \(\phi_{M_1}\) has the form \(M_1' = \begin{pmatrix} A_1 & B_1 \\ 0 & C_1 \end{pmatrix}\), \(A_1 \in \mathbb{D}_r\), \(C_1 \in \mathbb{D}_{n-r}\). Since \(M_1 M_j = M_j M_1\), it follows that the \(A_1\)’s commute pairwise, as do the \(C_1\)’s. Since \(0<r<n\) we can apply the induction hypothesis twice to get \(P \in \text{GL}_r(D)\), \(Q \in \text{GL}_{n-r}(D)\) such that \(P^{-1} A_1 P\) and \(Q^{-1} C_1 Q\) are upper triangular for \(i=1, \ldots, m\). Then 
\[
\begin{pmatrix} P^{-1} & 0 \\ 0 & Q^{-1} \end{pmatrix} \begin{pmatrix} A_1 & B_1 \\ 0 & C_1 \end{pmatrix} \begin{pmatrix} P & 0 \\ 0 & Q \end{pmatrix}
\]

is upper triangular for \(i=1, \ldots, m\), and the theorem is true.

We now use induction on \(m\). If \(m=1\) the theorem holds as a consequence of the considerations on normal form in chapter 1. Assume \(m>1\) and the result is true for \(m-1\) commuting matrices. Since \(M_1\) can be upper triangularized we know that \(M_1\) has a right eigenvector. Let \(w\) be a right eigenvector of \(M_1\) corresponding to the right eigenvalue \(\alpha\); then for \(i=1, \ldots, m\), \(M_1(M_1 w) = M_1(M_1 w) = (M_1 w) \alpha\), i.e., either \(M_1 w = 0\) or \(M_1 w\) is a right eigenvector of \(M_1\) corresponding to the right eigenvalue \(\alpha\). Thus the space \(W\) spanned by right eigenvectors of \(M_1\) corresponding to \(\alpha\) is a non-zero subspace of \(V\) invariant under \(\phi_{M_1}, \ldots, \phi_{M_m}\). Our consideration of invariant subspaces above allows us to assume \(W = V\), and so \(V\) has a basis \(B\) consisting of right eigenvectors of \(M_1\) corresponding to the right eigenvalue \(\alpha\). We express the \(\phi_{M_1}\)’s with respect to this basis, and so \(\mu(\phi_{M_1}, B) = \alpha \cdot I\). Since \(M_1 M_i = M_i M_1\), \(i=2, \ldots, m\), it follows that \(\alpha\) centralizes the entries of \(M_1' = \mu(\phi_{M_1}, B), i=2, \ldots, m\). Let \(D_0\) be the skew subfield
of $D$, centralizing $a$; then $k(a) \subseteq \text{center}(D_0)$, and $M_2^*, \ldots, M_m^*$ are $m-1$ commuting matrices over $D_0$, so the theorem follows by induction on $m$.

In many future proofs we will use induction on the degree of the matrices involved to show that the result holds if the corresponding linear transformations leave a non-trivial subspace invariant. As the arguments would mimic the one above we will generally leave out the repetitious details.

Theorem 1 generalizes the result obtained for commuting matrices over commutative fields, and the proof is just a careful adaptation of a standard proof of the usual result ([16], th. 2, p. 14). That this approach must be pursued with caution is emphasized by the following example, which contradicts a well known result for matrices over a commutative field ([16], th. 1, p. 12):

**Example 2**, of two commuting diagonalizable matrices over a skew field, which cannot be simultaneously diagonalized.

Let $D$ be a skew field with center $k$ and containing an element $x$ transcendental over $k$. Then $M_1 = \begin{pmatrix} x & 0 \\ 0 & x \end{pmatrix}$ and $M_2 = \begin{pmatrix} x & 1 \\ 0 & x \end{pmatrix}$ are easily seen to be transcendental matrices, so are both diagonalizable by (1.C). If they could be simultaneously diagonalized then $M_2 - M_1$ would be diagonalizable; but $M_2 - M_1$ is a non-zero nilpotent matrix, so cannot be diagonalized. Therefore $M_1$ and $M_2$ cannot be simultaneously diagonalized.

The difficulty encountered in example 2 is the only impediment, however, as we see in

**Proposition 2**: Let $\{M_\alpha | \alpha \in \mathbb{C}_D \}$ be a set of commuting matrices. If every element of $k[M_\alpha | \alpha \in \mathbb{C}_D]$ (the subring of $D$ generated by
\[ M_\alpha \{U \mid b \cdot I \mid b \cdot k]\] can be diagonalized then \[ M_\alpha \] can be simultaneously diagonalized (by a similarity transformation leaving the center of \( D_n \) fixed).

Before proving proposition 3 we establish the following diagonalizability criterion for an algebraic matrix: Let \( M \in \mathbb{D}_n \) be an algebraic matrix. Then \( M \) is diagonalizable if and only if \( M \) satisfies a polynomial \( f(t) \in k[t] \) which, when written as a product of irreducible polynomials over \( k \), has no repeated factors.

only if: If \( M \) is diagonalizable, \( M \) is similar to a matrix \( M_0 = \begin{pmatrix} \alpha_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \alpha_n \end{pmatrix} \); since \( M \) is algebraic, so is each \( \alpha_i \).

Let \( r_i(t) \) be a bound in \( k[t] \) of \( t - \alpha_i \); as \( t - \alpha_i \) is an atom of \( D[t] \), each \( r_i(t) \) is an invariant atom ([4], p. 230), i.e., an irreducible polynomial in \( k[t] \). \( M_0 \), and hence \( M \), satisfy \( \text{LCM}(r_i) = f(t) \), and the irreducible factors of \( f \) are just the non-associated \( f_i \)'s, hence none are repeated.

\[ \text{If:} \quad \text{Let } f(t) \text{ be as in the hypothesis. Suppose } M \text{ has elementary divisors } \lambda_1, \ldots, \lambda_r \text{ with bounds (in } k[t]) f_1(t), \ldots, f_r(t). \text{ Since } f(M) = 0, f_1 | f \text{ for } i = 1, \ldots, r. \text{ By the definition or elementary divisors, each } \lambda_i \text{ is indecomposable, so each } r_i \text{ is a power of an irreducible polynomial } f'_1(t) \in k[t]. \text{ But } f \text{ has no repeated factors in } k[t], \text{ so } f_1 = f'_1 \text{ is irreducible. Then each } \lambda_i \text{ must be linear, as the only indecomposables bounded by an I-atom have length 1 ([4], p. 230). But then the matrix } M_0 \text{ obtained from the elementary divisors (1.c) is diagonal and similar to } M. \]

Proof of proposition 2: By (1.2) it suffices to consider a finite set of matrices \( \{M_1, \ldots, M_r\} \). We again use induction on \( n \), noting that the result holds trivially if
n=1. Now suppose $M'_1 = D^{-1} M_1 P$ is diagonal with conjugate diagonal entries equal; denote $P^{-1} M_1 P$ by $M'_1$. The commutativity of the $M_i$'s assures us that non-conjugate diagonal entries of $M'_1$ give us a decomposition of $D^n$ into a direct sum of subspaces invariant under $\{ \phi_{M_i} \}_{i=1,\ldots,r}$, so induction on $n$ allows us to assume $M'_1$ is scalar, say $M'_1 = \alpha \cdot I$. Denoting by $K$ the skew subfield of $D$ centralizing $\alpha$, we have $M'_i \in K_n$, $i=2,\ldots,r$; if we can show that every matrix in $k(\alpha)[M'_2,\ldots,M'_r]$ is diagonalizable over $K$, we will then be done by induction on $r$.

Let $M \in k(\alpha)[M'_2,\ldots,M'_r]$. We can clear expressions in $\alpha$ from the denominators of the coefficients from $k(\alpha)$ by multiplying by a polynomial $p(\alpha) \in k[\alpha]$. We then get a matrix $M' = p(\alpha) M \in k[\alpha][M'_2,\ldots,M'_r]$, and $M$ is diagonalizable over $K$ if and only if $M'$ is. Let $M'' = Q^{-1} M' Q$ be in normal form over $K$. Write $M''_i$ for $Q^{-1} M'_i Q$, $i=2,\ldots,r$. Blocks of $M''$ corresponding to non-conjugate diagonal entries give a direct sum decomposition of $K^n$ into non-trivial $k(\alpha)[\phi_{M'_2},\ldots,\phi_{M'_r}]$-invariant subspaces, and the result follows by induction on $n$, so we may assume all the diagonal entries of $M''$ conjugate. If the diagonal entries of $M''$ are transcendental over $k(\alpha)$, $M''$ is diagonal (in fact scalar) by our definition of normal form in chapter 1. If the diagonal entries of $M''$ are algebraic over $k(\alpha)$ then they all satisfy the same $k(\alpha)$-irreducible polynomial $q(t)$. Again we can clear expressions in $\alpha$ from the denominators of the coefficients of $q$, and so we assume $q(t) \in k[\alpha][t]$. Then we can think of $q(M'')$ as an element of $k[\alpha \cdot I][M''_2,\ldots,M''_r]$, so $q(M'')$ is similar to an element of $k[M'_1,\ldots,M'_r]$ and hence diagonalizable. But by inspection
(q(M"))\textsuperscript{n}=0, and it follows that q(M")=0. By our diagonalizability criterion and the irreducibility of q(t), we conclude that M" is diagonalizable over K, and the result follows by induction on r.

An interesting special case where we can say more is given in

**Proposition 4:** Suppose \( \kappa \cdot I \subseteq D \), F a commutative field. Then F is similar to a scalar field.

**Proof:** Since \( \kappa \cdot I \subseteq F \), \( \kappa [F] = F \). We first want to show that every element of F is diagonalizable, and apply proposition 3. Let \( M \in F \). If \( M \) is transcendental \( M \) is diagonalizable by (1.C). If \( M \) has an algebraic part then \( M \) has an algebraic right eigenvalue \( \alpha \), and \( t-\alpha \) has as bound an irreducible polynomial \( p(t) \in k[t] \). Then \( p(M) \) annihilates a non-zero vector; but \( p(M) \in F \) (since \( \kappa \cdot I \subseteq F \)), and any non-zero element of F is invertible. Thus \( p(M)=0 \), and \( M \) is diagonalizable by our diagonalizability criterion.

Let \( F'=F^{-1}FP \) be in diagonal form. By the above paragraph, each \( M \in F' \) is either transcendental or satisfies an irreducible polynomial in \( k[t] \). Consequently the diagonal entries \( m_{ii} \) of \( M \) will all be conjugate; in fact, the \( i,i \) entries of the matrices in \( F' \) form a field isomorphic to \( F \), for \( i=1,\ldots,n \). By (1.B), there are elements \( \lambda_1, i=2,\ldots,n \), such that \( \lambda_1^{-1}m_{ii}\lambda_1=m_{ii} \) for all \( M \in F \). Then

\[
\begin{pmatrix}
\lambda_1^{-1} & 0 & \cdots & 0 \\
0 & \lambda_2^{-1} & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \lambda_n^{-1}
\end{pmatrix}
\begin{pmatrix}
1 \\
\lambda_2 \\
\vdots \\
\lambda_n
\end{pmatrix}
\begin{pmatrix}
1 \\
0 \\
\vdots \\
0
\end{pmatrix}
\]

is scalar, as desired.

Looked at from a slightly different point of view, the difficulties in example 2 arose because \( \begin{pmatrix} x & 1 \\ 0 & x \end{pmatrix} \) was diagonalizable. This phenomenon occurs only when \( x \) is transcendental
or purely inseparable over $k$ (and not in $k$). Thus we are led to make the following

**Definition:** A matrix $M \in D_n$ is called *separable* (over $k$) if $M$ is algebraic and satisfies a polynomial $f(t) \in k[t]$ whose irreducible factors have no repeated roots (in any extension).

We observe:

(a) Any matrix similar to a separable matrix is separable.

(b) A separable matrix is diagonalizable if and only if it satisfies a polynomial $f(t) \in k[t]$ with no repeated roots in any splitting field.

(c) If $M \in D_n$ is separable and $W \subset D^n$ is a subspace invariant under $\phi_M$, then any matrix representing $\phi_M|_W$ (the restriction map) will be separable.

Proofs of (a) and (c) are immediate, and (b) follows from our previous diagonalizability criterion. We are now ready to prove

**Proposition 5:** A set of commuting separable diagonalizable matrices over a skew field $D$ can be simultaneously diagonalized (by a similarity transformation leaving the center of $D_n$ fixed).

**Proof:** By (1.2) it suffices to consider finitely many matrices $M_1, \ldots, M_r$. We use induction on $n$; the proposition is true for $n=1$, and we assume $n>1$ and the proposition holds for all $n'<n$. If $M_1$ has non-conjugate right eigenvalues we get a decomposition of $D^n$ into a direct sum of proper subspaces invariant under the $\phi_{M_i}$'s, and (c) and induction on $n$ give us our result. Thus we may take $M_1$ to be similar to a scalar matrix, say $P^{-1}M_1P = \alpha \cdot I$. Denote by $M'_i$ the matrix $P^{-1}M_iP$, $i=2,\ldots,r$. As before we restrict ourselves to the
skew subfield $K$ of $D$ centralizing $\alpha$, and observe that 
\[ \{\lambda_i^j\}_{i=1}^n \in K. \]
It follows by the definition of separable matrices and (b) that since $M_1', \ldots, M_r'$ are separable and diagonalizable over $k$ they are also separable and diagonalizable over the center of $K$, and the proof is completed by applying induction to $r$.

We turn now to the question of describing all matrices which commute with a given matrix. We do this only for matrices over skew fields $D$ with perfect centers, but we start in full generality with

**Lemma 6**: Let $A \in D_m^n$, $B \in D_m^n$, and suppose $p(t) \in k[t]$ is such that $p(A) = 0$ and $p(B)$ is non-singular. Then the only $n \times m$ matrix $X$ such that $AX = XB$ is the zero matrix.

**Proof**: It is easily verified that $p(A)X = Xp(B)$. Since $p(B)$ is non-singular, $p(B)D^m = D^m$. Thus $XD^m = Xp(B)D^m = p(A)XD^m = 0$, and it follows that $X = 0$.

(This lemma is also a consequence of ([5], lem. 2.3), which is much more general than the result we need).

Our next lemma concerns the solution of certain simultaneous equations over a skew field with perfect center:

**Lemma 7**: Let $D$ be a skew field with perfect center $k$; let $\alpha \in D$ be algebraic over $k$. If $x, y \in D$ satisfy $ya = yx$, $\alpha x = a$, then $x = 0$.

**Proof**: Since $x \alpha = ax$, $yf(\alpha) - f(\alpha)y = f'(\alpha)x$ for any polynomial $f(t) \in k[t]$. In particular, if $f$ is the minimal polynomial for $\alpha$, the fact that $k$ is perfect (hence $f'(\alpha) \neq 0$) forces $x$ to be 0.

(This proof, but not the original proof of the lemma, is due to Professor Cohn).
We denote by $N_i$ the $i 	imes i$ matrix with 1's on the main super-diagonal and 0's elsewhere. Then, we have

**Lemma 8:** Let $D$ be a skew field with perfect center $k$, and suppose $\alpha \in D$ is algebraic over $k$. Let $B$ be an $i \times j$ matrix over $D$ such that $(\alpha \cdot I_i + N_i)B = B(\alpha \cdot I_j + N_j)$. Then $B$ has the form

$$
\begin{pmatrix}
    b_1 & b_2 & \cdots & b_n \\
    0 & b_2 & \cdots & b_n \\
    & 0 & \ddots & \vdots \\
    & & 0 & b_1
\end{pmatrix}
$$

if $i=j$, 

$$
\begin{pmatrix}
    b_1 & b_2 & \cdots & b_n \\
    0 & b_2 & \cdots & b_n \\
    & 0 & \ddots & \vdots \\
    & & 0 & b_1
\end{pmatrix}
$$

if $j>i$, and 

$$
\begin{pmatrix}
    b_1 & b_2 & \cdots & b_n \\
    0 & b_2 & \cdots & b_n \\
    & 0 & \ddots & \vdots \\
    & & 0 & b_1
\end{pmatrix}
$$

if $i>j$, where $b_{ih} = \alpha b_h$ for all $h$.

**Proof:** We prove the case $i=j$; the other case is proved similarly. Write $J_i$ for $\alpha \cdot I_i + N_i$ and $J_j$ for $\alpha \cdot I_j + N_j$, and set $B = (b_{rs})$. Then $J_i B = \begin{pmatrix}ab_{11} & \cdots & ab_{1j} \\
\vdots & \ddots & \vdots \\
ab_{i1} & \cdots & ab_{ij}
\end{pmatrix} + \begin{pmatrix}b_{21} & \cdots & b_{2j} \\
\vdots & \ddots & \vdots \\
b_{i1} & \cdots & b_{ij}
\end{pmatrix}$.

$B J_j = \begin{pmatrix}b_{11} & \cdots & b_{1j} \\
\vdots & \ddots & \vdots \\
b_{i1} & \cdots & b_{ij}
\end{pmatrix} + \begin{pmatrix}0 & b_{11} & \cdots & b_{1j-1} \\
\vdots & \ddots & \vdots & \vdots \\
\hat{0} & b_{i1} & \cdots & b_{ij-1}
\end{pmatrix}$. We first show that $b_{r1} = 0$ for $r>1$. Comparing $i-1,1$ and $i,1$ entries of the products, we get $b_{i1} = \alpha b_{i1}$, $b_{i-1,1} = \alpha b_{i-1,1} + b_{i1}$. By lemma 7, $b_{i1} = 0$. Given that $b_{r+1,1} = 0$, $r>1$, a similar comparison of $r,1$ and $r-1,1$ entries and application of the lemma shows that $b_{r1} = 0$.

We call an entry $b_{rs}$, $r>s$, a sub-diagonal entry. Suppose now that all sub-diagonal entries of the first $s-1$ columns of $B$ are 0, $j>s>1$. The $s^{th}$ columns of $J_i B$ and $B J_j$ are then 

\[
\begin{pmatrix}
    \alpha b_{1s} + b_{2s} \\
    \vdots \\
    \alpha b_{is}
\end{pmatrix}
\quad \text{and} \quad
\begin{pmatrix}
    b_{1s} + \alpha b_{1s-1} \\
    \vdots \\
    b_{is}
\end{pmatrix}.
\]

As above, repeated applications of lemma 7 show that the sub-diagonal entries of the $s^{th}$ column of $B$ are 0. By induction on $s$, we conclude that all sub-diagonal entries of $B$ are 0.
If we now compare the s,s entries of \( J_B \) and \( B J_j \) we see that \( b_{ss}^a = ab_{ss} \), \( s=1,\ldots,j \). We next show that all entries of \( B \) commute with \( a \). This statement is true for all sub-diagonal entries and all entries \( b_{ss} \). Assume it is true for all \( b_{rs} \) with \( 0 \leq s-r \leq l-1 \), where \( 0 \leq l \leq j \). Consider the \( r,r+1 \) entries of \( J_B \) and \( B J_j \), which are \( \alpha b_{rr+1}^a + b_{rr+1}^a \) and \( b_{rr+1}^a + \beta b_{rr+1}^a \) respectively. Then letting \( x = \alpha b_{rr+1}^a - b_{rr+1}^a \), we see that the expression on the right commutes with \( a \) by assumption. Setting \( y = b_{rr+1}^a \), we see from lemma 7 that \( x = 0 \), which is what we wanted to show. Induction on \( l \) gives us \( b_{rs}^a = ab_{rs}^a \) for \( r=1,\ldots,i \) and \( s=1,\ldots,j \). But we also get \( 0 = x = b_{rr+1}^a - b_{rr+1}^a \), and we see \( B \) has the form claimed in the lemma.

The lemma motivates the following descriptive definition: We call a matrix \( B \) in the form described in lemma 8 a **triangularly striped matrix with entries in** \( C_D(a) \) (cf. [16], pp. 26-27).

We are now ready to state

**Theorem 9:** Let \( D \) be a skew field with perfect center, and let \( A, B, \in D_n \). We know from (1.1) that \( A \) is similar to a matrix in normal form, say \( A = P((\alpha_1 I + H_1) + \cdots + (\alpha_s I + H_s))P^{-1} \), where \( H_i \in D_{\alpha_i} \) is a diagonal sum \( \sum_n^j n_{a_i}^{(i)} \), and the \( \alpha_i \)'s are non-conjugate elements of \( D \); further, at most one \( \alpha_i \) is transcendental, and if there is one such the corresponding \( H_i \) is 0.

Write \( B = P(B_{ij})P^{-1} \) in block form with square blocks of dimension \( n_i x n_i \), \( \ldots \), \( n_s x n_s \) down the diagonal. Then \( BA = AB \) if and only if \( B_{ij} = 0 \) for \( i \neq j \), and \( B_{ii} \) is divided into blocks of dimension \( a_u x a_v \) (\( 1 \leq u, v \leq r_i \)), each such block being a triangularly striped matrix with entries in \( C_D(\alpha_i) \).
Proof: only if: In this direction the proof is immediate from lemmas 6 and 8 for algebraic blocks and from lemma 6 and by inspection whenever a transcendental block is involved.

if: This follows from the observation (in the notation of lemma 8) that if $B$ is a triangularly striped matrix with entries in $C_D(\alpha)$ then $B(\alpha \cdot I_n + N_1) = (\alpha \cdot I_n + N_1)B$.

The comparable result for the case of a commutative field $D$ -- identical except for our restriction to $C_D(\alpha)$ -- can be found in ([16], th. 6, p. 28).

There is at this point the obvious question of whether two or more commuting matrices over a skew field $D$ with perfect center will be similar to matrices with entries in a commutative field. To show that this question must in general be answered in the negative, even in the case of diagonalizable matrices, we give

Example 10, of a four generator abelian subgroup of $D_2$, every element of which is diagonalizable, but which is not similar to a group of matrices with entries in a commutative subfield of (any extension of) $D$.

Let $D$ be a skew field with center $k$; suppose $x \in D$ is transcendental over $k$; suppose $d_1, d_2 \in D$ satisfy $d_1 d_2 d_1 = d_2 d_1, d_1 x = x d_1 (i=1, 2)$; let $0, 1, a, b \in k$ be distinct. Take $\mathcal{A}$ to be the group generated by $(x \ 0), (x+1 \ 1), (x+a \ d_1), (x+b \ d_2)$. Then

1) $\mathcal{A}$ is abelian, as the generators are easily seen to commute.

2) Every element of $\mathcal{A}$ is diagonalizable; this will follow from (1.C) once we show that every element of $\mathcal{A}$ except $I$
is transcendental over $k$. Any matrix $M \in \mathcal{A}$ will have the form
\[ M = \begin{pmatrix} x^{m_1}(x+1)^{m_2}(x+a)^{m_3}(x+b)^{m_4} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & x^{m_1}(x+1)^{m_2}(x+a)^{m_3}(x+b)^{m_4} \end{pmatrix}, \quad m_i \in \mathbb{Z}. \]

Let $y = x^{m_1}(x+1)^{m_2}(x+a)^{m_3}(x+b)^{m_4}$. Suppose $p(M)v = 0$ where $p(t) \in k[t]$ and $0 \neq v \in D^2$. If $v = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$ we see that $p(y)\alpha = 0$, so $p(y) = 0$; if $v = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$, $\beta \neq 0$, then $p(y)\beta = 0$ and $p(y) = 0$. Write $p(t) = a_0 + \cdots + a_n t^n$. Since $M$ is non-singular we may assume $p$ was chosen so that $a \neq 0 \neq a_n$. Suppose that $m_i \neq 0$ for some fixed $i$. For ease in notation we write $b_1 = 0, \ b_2 = 1, \ b_3 = a, \ b_4 = b$. If $m_1 > 0$ we let $R$ be the ring obtained by localizing $k[x]$ at the prime ideal $<x+b_1>$. We can map $R$ homomorphically onto $k$ by mapping $x$ to $-b_1$; then $p(y) \in R$, and as $p(y) = 0$, $p(y)$ is mapped to 0. But by inspection $p(y)$ is mapped to $a \neq 0$, a contradiction. If $m_1 < 0$, a similar argument using $p' = (x+b_1)^{-m_1}p(y)$ gives us $p' \mapsto 0$ and $p' \mapsto a \neq 0$, a contradiction. Thus we must have $m_1 = 0$ for $i = 1, \ldots, 4$, and it follows that $M = I$.

3) $\mathcal{A}$ is not similar to a subgroup of $F_2$, $F$ commutative: if $\mathcal{A}$ is similar to a subgroup of $F_2$, $F$ commutative, then by the commutative field case of theorem 1, $\mathcal{A}$ is similar to an upper triangular group of matrices over a commutative field $F$. If $P \in \text{GL}_2(D)$ is such that $P^{-1}\mathcal{A}P \subseteq F_2$ and is upper triangular, then the first column of $P$ is a common right eigenvector of the generators of $\mathcal{A}$. The only such have the form $\begin{pmatrix} \alpha \\ 0 \end{pmatrix}$, $\alpha \neq 0$, for otherwise $\begin{pmatrix} x & 0 \\ 0 & x \end{pmatrix}$ and $\begin{pmatrix} x+1 & 1 \\ 0 & x+1 \end{pmatrix} - I$ could be simultaneously diagonalized, contrary to example 2. Thus the first column of $P$ is of the form $\begin{pmatrix} \alpha \\ 0 \end{pmatrix}$. Now if $P^{-1}\mathcal{A}P$ has entries in a commutative field, so does $\alpha \cdot P^{-1}\mathcal{A}P \cdot \alpha^{-1}I$, so we may take $P = \begin{pmatrix} 1 & \beta \\ 0 & y \end{pmatrix}$, $y \neq 0$. Let $B = \begin{pmatrix} x+c & f \\ 0 & x+c \end{pmatrix}$ be a generator of $\mathcal{A}$ (so $ce \in k$). Then $P^{-1}BP = \begin{pmatrix} 1 & -\beta y^{-1} \\ 0 & y^{-1} \end{pmatrix} \begin{pmatrix} x+c & f \\ 0 & x+c \end{pmatrix} \begin{pmatrix} 1 & \beta \\ 0 & y \end{pmatrix} = \begin{pmatrix} x+cy^{-1} & x+fy^{-1}y^{-1}xy \\ 0 & y^{-1}xy+c \end{pmatrix} \in F_2$. Thus,
if \( e, f \in \{0, 1, d_1, d_2\} \),
\[
(*) \quad (x\beta + fy - \beta y^{-1}xy)(x\beta + ey - \beta y^{-1}xy) = \\
(x\beta + ey - \beta y^{-1}xy)(x\beta + fy - \beta y^{-1}xy).
\]

First we take \( e = 0 \), getting
\[
(**) \quad fy(x\beta - \beta y^{-1}xy) = (x\beta - \beta y^{-1}xy)fy, \text{ for } f \in \{0, 1, d_1, d_2\}.
\]
Using (**), (*) becomes (###) \( fy = eyfy \), or \( fye = eyf \) for \( e, f \in \{0, 1, d_1, d_2\} \). Taking \( e = 1 \) in (###) we get \( fy = yf \) for \( f \in \{d_1, d_2\} \). Now we take \( e = d_1 \), \( f = d_2 \) in (###), getting \( d_1 y d_2 = d_2 y d_1 \); but since \( y \neq 0 \) and \( d_1 y = y d_1 \) this reduces to \( d_1 d_2 = d_2 d_1 \), contradicting our choice of \( d_1, d_2 \). This contradiction arose from our assumption that \( A \) was similar to a group of matrices over a commutative field \( F \), so we get our desired result.

In a sense we should not have expected commuting matrices to be similar to matrices over a commutative field, as matrices over a commutative field are not characterized by commutativity. Rather, \( n \times n \) matrices over a commutative subfield \( F \) of \( D \) are characterized by the fact that there is a commutative subfield \( F \cdot I \) of their centralizer in \( D \) such that the algebra they generate over this commutative field has dimension less than or equal to \( n^2 \). There might still be problems as in example 10 with reducible semigroups, so we express our result for absolutely irreducible semigroups of matrices, i.e., sub-semigroups of \( D \) which are irreducible over all extensions of \( D \). Then we get

Theorem 11: Let \( A \subset D \) be an absolutely irreducible semigroup. \( A \) is similar (over some extension \( E \) of \( D \)) to a semigroup of matrices with entries in a commutative field \( F \) if and only if there is a commutative subfield \( F_0 \) of \( C_{En}(A) \) with
Before going on to the proof of theorem 11, we need a lemma about skew subfields of $D_n$.

**Lemma 12:** Let $D$ be a skew field with center $k$, and suppose $K$ is a skew field with $k \subseteq K \subseteq D$. If $(K: \text{center } K) \leq n^2$, then either $n=1$ or $K$ is reducible.

**Proof:** Assume $K$ is absolutely irreducible, and let $F$ be a maximal commutative subfield of $K$. By proposition 4, $F$ is similar to a scalar field, say $Q^{-1}FQ=F'\cdot I$, $F=F'\subseteq D$. Let $K'=Q^{-1}KQ$; $K'$ is also absolutely irreducible. Now let $M_1=I$, $M_2, \ldots, M_t$ be a basis of $K'$ as right $F'\cdot I$ space. Since $(K: \text{center } K) \leq n^2$, $t=(K': F') \leq n$. Write $M_i=(m_i, M_i')$, where $m_i\in D^n$. The $m_i$'s span a $D$-space of dimension $d\leq t<n$. But the $D$-space spanned by $m_1, \ldots, m_t$ is the image space under $K'$ of $e_1D$ ($e_1$ being the vector with 1 in the first row and 0's elsewhere)—for if $M\in K'$, then $M=I\sum F_i^t.1, f_i\in F'$, and $M_{e_1}D=\Sigma M_i f_i.1e_1.D=\Sigma M_i e_i f_i D=\Sigma m_i f_i D=\langle m_1, \ldots, m_t \rangle$. Further, this space is $K'$-invariant, as $K'<m_1, \ldots, m_t>=K'K(1)=K'K=e_1D=\langle m_1, \ldots, m_t \rangle$. As $\langle m_1, \ldots, m_t \rangle \neq 0$, the irreducibility of $K'$ gives us that $\langle m_1, \ldots, m_t \rangle = D^n$, so $t=n$. But we saw that $t<n$, so $t=n$ and $\{m_1, \ldots, m_t\}$ is a basis of $D^n$.

Let $P\in \text{GL}_n(D)$ be such that $P^{-1}m_i=e_i$ (the vector with 1 in the $i^{th}$ row and 0's elsewhere), and take the representation $P^{-1}K'P$. Note that $P^{-1}P=I$, so $P^{-1}=(m_1, \ldots, m_n)$. We claim that $P^{-1}K'P \subseteq F'$. Let $\alpha_{ij} \in F'$. We have $M_i M_j=I+M_i \alpha_{ij}^t.1, \alpha_{ij}^t \in F'$. Now $P^{-1}M_i P^{-1} m_j = P^{-1} M_i P e_j = \alpha_{ij}^t \text{ the } j^{th} \text{ column of } P^{-1} M_i P; \text{ but } P^{-1} M_i P^{-1} m_j = P^{-1} \alpha_{ij}^t = P^{-1} (\Sigma M_i \alpha_{ij}^t) = \Sigma e_1 \alpha_{ij}^t \epsilon (F')^n$. Thus the $j^{th}$
column of \( P^{-1} \mathbf{M}_i P \) has entries in \( F' \), so \( P^{-1} \mathbf{M}_i P \in (F')_n \). We also want to verify that, for \( f \in F' \), \( P^{-1} f \cdot I P \in (F')_n \). But

\[
P^{-1} f \cdot I P = P^{-1} (f \cdot \mathbf{M}_1, \ldots, f \cdot \mathbf{M}_n);
\]

now \( f \cdot \mathbf{M}_i \) is the \( i \)-th column of \( f \cdot \mathbf{M} \), and since \( f \cdot \mathbf{M}_i \in K' \) we have

\[
f \cdot \mathbf{M}_i = \sum_j \mathbf{M}_{ij} \beta_{ij} I, \quad \beta_{ij} \in F'.
\]

Thus the first column of \( f \cdot \mathbf{M}_i \) is \( \sum_j \mathbf{M}_{ij} \beta_{ij} \), and

\[
P^{-1} f \cdot I P = P^{-1} (f \cdot \mathbf{M}_1, \ldots, f \cdot \mathbf{M}_n) = \sum_j (\sum \mathbf{M}_{ij} \beta_{ij}, \ldots, \sum \mathbf{M}_{in} \beta_{jn}) = (\beta_{rs}) \in F'_n.
\]

Now \( P^{-1} K' P \), being generated as a ring by \( P^{-1} \mathbf{M}_i P \) \((i=1, \ldots, n)\) and \( P^{-1} F' \cdot I P \), will be contained in \( F'_n \).

We can further assume by \((1.\alpha)\) that there is a skew field \( K_0 \) isomorphic to \( K \) with \( F' \subseteq K_0 \subseteq D \); thus

\[
P^{-1} K' P \subseteq (K_0)_n \subseteq D_n.
\]

Now \( K \) is finite dimensional over its center, so \((K_0)_n\) is a finite dimensional simple algebra. We have isomorphic central simple subalgebras \( K_0 \cdot I \) and \( P^{-1} K' P \) of \((K_0)_n\), so by the Skolem-Noether Theorem \([9], \text{p.} 99\) \( P^{-1} K' P \) is similar to \( K_0 \cdot I \). But \( K_0 \cdot I \) is in reduced form unless \( n=1 \), so since \( K' \) was absolutely irreducible, we conclude that \( n=1 \).

**Proof of theorem 11**: only if: Assume \( \lambda \) is similar to a semigroup of matrices over a commutative field \( F \), say

\[
P^{-1} \mathbf{J} P \subseteq F_n.
\]

Then \( F \cdot I C_{n} \cdot (P^{-1} \mathbf{J} P) \) and \( (F(P^{-1} \mathbf{J} P) : F \cdot I) \leq n^2 \). But then \( F_0 = F \cdot I P \cdot F^{-1} \) is a commutative subfield of \( C_{n} \cdot (\mathbf{J}) \), and \( (F_0 \cdot \mathbf{J} : F_0) \leq n^2 \).

**if**: Let \( \mathbf{J} \subseteq D_n \) be an absolutely irreducible semigroup. Then \( R = \mathbb{k}[\mathbf{J}] \cdot D_n^D \) will be a ring and \( D_n \) a left \( R \)-module under the action \((\mathbf{Z} \alpha_i \cdot IS_i \cdot \mathbf{d}_i) \cdot v = \mathbf{Z} \alpha_i \cdot IS_i \cdot \mathbf{d}_i \); the irreducibility of \( \mathbf{J} \) means \( D_n \) is an irreducible \( R \)-module, so by Shur's Lemma \([9], \text{p.} 5\) the commuting ring of \( R \) is a skew field. The commuting ring of \( R \), commuting with all elements \( 1 \odot d \), will be a subring of \( D_n \); commuting with all elements \( 1 \odot d \),
as \( \mathcal{A} \), it will be contained in \( C_{D_n}(\mathcal{A}) \); it is easy to see that the commuting ring of \( R \) is in fact \( C_{D_n}(\mathcal{A}) \), so \( C_{D_n}(\mathcal{A}) \) is a skew field.

If \( R \in C_{D_n}(\mathcal{A}) \) is a commutative field, then \( k[F] \) will be contained in a commutative subfield of \( C_{D_n}(\mathcal{A}) \). If in addition it had been the case that \( (F:\mathcal{A}) \leq n^2 \), the same would be true of the algebra over the field generated by \( k[F] \).

Thus we can and do assume that \( \mathcal{A} \) is an absolutely irreducible semigroup and \( F \) is a commutative subfield of \( C_{D_n}(\mathcal{A}) \) with \( k \cdot \mathcal{I} F \) and \( (F:\mathcal{A}) \leq n^2 \).

We now show that \( F \mathcal{A} \) is a simple \( F \)-algebra. Since \( F \mathcal{A} \) is a finite dimensional \( F \)-algebra, its radical \( \mathcal{R} \) is a nilpotent ideal (\cite{9}, p. 20). Because \( \mathcal{R} \) is nilpotent, \( \mathcal{R} \mathcal{D}^n = \mathcal{D}^n \), and because \( \mathcal{R} \) is an ideal of \( F \mathcal{A} \), \( F \mathcal{R} \mathcal{D}^n \subseteq \mathcal{D}^n \). Thus \( \mathcal{R} \mathcal{D}^n \) is an \( \mathcal{A} \)-invariant subspace properly contained in \( \mathcal{D}^n \). From the irreducibility of \( \mathcal{A} \) we conclude that \( \mathcal{R} \mathcal{D}^n = 0 \), whence \( \mathcal{R} = 0 \) and \( F \mathcal{A} \) is semisimple. If now \( \mathcal{J} \) is a non-zero ideal of \( F \mathcal{A} \), then by results on the structure of finite dimensional semisimple algebras (\cite{9}, p. 30) \( \mathcal{J} = F \mathcal{A} e \), where \( e \) is a non-zero central idempotent. Then \( F \mathcal{A} e \mathcal{D}^n = e F \mathcal{A} \mathcal{D}^n = e \mathcal{D}^n \), so \( e \mathcal{D}^n \) is a non-zero \( \mathcal{J} \)-invariant subspace. Since \( \mathcal{J} \) is irreducible \( e \mathcal{D}^n = \mathcal{D}^n \); then since \( e \) is idempotent, \( e = I_n \) and \( \mathcal{J} = F \mathcal{A} I_n = F \mathcal{J} \). Thus \( F \mathcal{J} \) is simple.

By the Wedderburn-Artin Theorem (\cite{9}, p. 48), \( F \mathcal{A} = K_r \), \( K \) a finite dimensional division algebra. We identify \( F \mathcal{A} \) with \( K_r \), denoting the matrix units of \( K_r \) by \( f_{ij} \) (1 \( \leq i, j \leq r \)). Thus we have \( K_r \subseteq D_n \).

We next show that \( r \mid n \), say \( n = rs \), and that there are an inner automorphism \( \sigma \) of \( D_n \) and an embedding \( \varpi : K \to D_n \) such that...
we get a commutative diagram

\[
\begin{array}{c}
\Rightarrow \\
F = K_R \quad \xleftarrow{\pi_R} \quad D_n \\
\sigma \downarrow \\
(D_n)^r \quad \xleftarrow{\sigma'} \quad D_n ^r
\end{array}
\]

where the bottom isomorphism is the obvious one. We have a decomposition of the identity \( I_n \) into a sum of orthogonal idempotents, \( I_n = \sum_{i=1}^{r} f_{ii} \).

We use induction on \( r \) to show these can be simultaneously diagonalized; for \( r=1 \) this is obvious. Otherwise we diagonalize \( f_{11} \) and get \( P^{-1} f_{11} P = \begin{pmatrix} I_{S_1} & 0 \\ 0 & 0 \end{pmatrix} \). If \( P^{-1} f_{ii} P = \begin{pmatrix} A_i & B_i \\ C_i & H_i \end{pmatrix} \) (i=2,...,r) then since \( f_{ii} f_{ii} = f_{ii} f_{ii} = 0 \) we see that \( A_i, B_i, \) and \( C_i \) are zero. The matrices \( H_i \) will be orthogonal idempotents in \( D_{n-S_1} \), and their sum will be \( I_{n-S_1} \); by induction on \( r \) we get a matrix \( Q \in \text{GL}_{n-S_1} (D) \) such that \( Q^{-1} H_i Q \) is diagonal for \( i=2,...,r \). Then it is easily seen that

\[
\begin{pmatrix} I & 0 \\ 0 & Q \end{pmatrix} P^{-1} f_{ii} P \begin{pmatrix} I & 0 \\ 0 & Q \end{pmatrix}
\]

is diagonal for \( i=1,...,r \). A further permutation of the basis vectors will now give us a block decomposition of \( D_n \) with \( r \) square blocks of dimensions \( s_1 x s_1, s_2 x s_2, ..., s_r x s_r \) down the diagonal, and the \( f_{ii} \)'s simultaneously similar to the matrices

\[
\begin{pmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix}
\]

\( i=1,...,r \). Now \( f_{ij} = f_{ij} f_{ij} f_{jj} \), so \( S^{-1} f_{ij} S = \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & 1 & \vdots \\ 0 & \cdots & 0 \end{pmatrix} \), the only non-zero block, of dimension \( s_i x s_j \), occurring in the \( i \)th row of blocks and the \( j \)th column. Also \( f_{ij} f_{ji} = f_{ii}, f_{ij} f_{ij} = f_{jj} \), so we see that \( F_{ij} F_{ji} = I_{s_i} \), \( F_{ji} F_{ij} = I_{s_j} \). Since \( F_{ij} \) and \( F_{ji} \) have entries in a skew field --but in particular a ring with invariant basis number ([4], pp. 5-6)--we see that \( s_i = s_j \) for \( 1 \leq i, j \leq r \). In particular \( r | n, \)
say n=rs.

Now K is embedded in $D_s$ by the map $\pi: K \rightarrow D_s$. The inner automorphism $\sigma$ of $D_n$ that we want to make diagram (1) commute is one such that $\sigma: f_{ij} \rightarrow \begin{pmatrix} 0 & O \\ O & I_s \end{pmatrix}$ (the non-zero block again appearing in the $i$th row of blocks and the $j$th column). We claim that, with the above notation, we can take $\sigma$ to be the map

$$M \mapsto \begin{pmatrix} I_{F_{12}} & O \\ O & F_{1r} \end{pmatrix} S^{-1} \begin{pmatrix} I_{F_{12}} & O \\ O & F_{1r} \end{pmatrix}^{-1}$$

for $i > 1$ we have $f_{ij} \mapsto \begin{pmatrix} I_{F_{12}} & O \\ O & F_{1r} \end{pmatrix} \begin{pmatrix} 0 & \cdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \cdots & 0 \end{pmatrix} \begin{pmatrix} I_{F_{12}} & O \\ O & F_{1r} \end{pmatrix} \begin{pmatrix} 0 & \cdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \cdots & 0 \end{pmatrix} = \begin{pmatrix} 0 & \cdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \cdots & 0 \end{pmatrix}$. In general, $f_{ij} \mapsto \begin{pmatrix} O & F_{ij} \end{pmatrix}$, and it follows from the above remarks and the equations $f_{ij} f_{ij} = f_{ij} f_{ij}$ ($1 \leq i, j \leq r$) that $F_{ij} = I$, as desired.

Now $r^2(K: \text{center } K) = (F_4: \text{center } F) = n^2 = r^2 s^2$, so we have $\pi K \subseteq D_s$, $(K: \text{center } K) = s^2$. Also $K \subseteq (\pi K)$. If $\pi K$ were reducible it would follow that $F_4$ was, hence that $\phi$ was, so we may assume $\pi K$ absolutely irreducible. It follows from lemma 12 that $s=1$, so $\pi K$ is commutative, and in fact $F_4 \cong K_n$ and the above similarity transformation $\sigma$ takes $F_4$ onto the nxn matrix ring over a commutative subfield $K'$ of $D$, $K' \subseteq K$.

Before noting a corollary we give an example illustrating theorem 11:

**Example 13:** Let $D$ be the skew field of (real) quaternions, $\mathcal{Q}$ the group generated by $\mathcal{Q} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $\mathcal{J} = \begin{pmatrix} j & 0 \\ 0 & j \end{pmatrix}$, $\mathcal{A} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. If $\mathcal{Q}$ is not absolutely irreducible then $\mathcal{Q}$, $\mathcal{J}$, and $\mathcal{A}$ have a common right eigenvector $w$; the only right eigenvalues of all three matrices are primitive fourth roots of 1, so we have $\mathcal{Q} w = w \alpha_1$, $\mathcal{J} w = w \alpha_j$, $\mathcal{A} w = w \alpha_1$, $\alpha_1$, $\alpha_j$, $\alpha_a$ primitive fourth roots.
of 1. Then since \( \mathcal{A} = \mathcal{B} \), \( \mathcal{A} \mathcal{B} = \mathcal{A} \mathcal{B} \mathcal{A} \mathcal{B} \), we see that \( \alpha_1 \alpha_2 = \alpha_2 \alpha_1 \), 
\( \alpha_3 \alpha_4 = \alpha_4 \alpha_3 \), \( \alpha_5 \alpha_6 = -\alpha_6 \alpha_5 \). This cannot happen with primitive fourth roots of 1 in a skew field, so \( \mathcal{A} \) must be absolutely irreducible. Taking \( F = \mathbb{R} \left[ \begin{array}{c} 0 \\ 1 \end{array} \right] \), \( (F, F) = 4 \), so by theorem 11, \( \mathcal{A} \) is similar to a semigroup of matrices over \( C \); in fact, taking \( P = \left[ \begin{array}{c} 1 \\ 0 \\ -1 \end{array} \right] \), we have \( P^{-1} P = \left( \begin{array}{c} 1 \\ 0 \\ -1 \end{array} \right) \), \( P^{-1} P = \left( \begin{array}{c} 1 \\ 0 \\ -1 \end{array} \right) \), 
\( P^{-1} P = \left( \begin{array}{c} 1 \\ 0 \\ 1 \end{array} \right) \).

We note as a consequence of theorem 11 a result on extending isomorphisms of simple subalgebras of \( D_n \):

**Corollary 11.1:** Let \( A, B \) be isomorphic absolutely irreducible (hence simple) \( k \)-subalgebras of \( D_n \) of dimension less than or equal to \( n^2 \) over their respective centers. Then any \( k \)-algebra isomorphism \( \psi: A \rightarrow B \) is induced by an inner automorphism of \( \mathcal{E}_n \) (\( E \) an extension of \( D \)).

**Proof:** Let \( \psi: A \rightarrow B \) be a fixed isomorphism of \( k \)-algebras. From the proof of theorem 11, \( A = P \cdot B \cdot P^{-1} \), where \( P \) is a commutative field. Let \( \{ e_{i,j} \} \) be the matrix units of \( D_n \), and \( \{ f_{i,j} \} \) the matrix units of \( A, B \) respectively. As in the proof of theorem 11, there are matrices \( P, Q \in GL_n(D) \) with \( P^{-1} f_{i,j} P = e_{i,j}, \) \( Q^{-1} g_{i,j} Q = e_{i,j}, \) \( 1 \leq i, j \leq n \).

We have an induced isomorphism of subfields \( e_{11} P^{-1} A P e_{11} \) and \( e_{11} Q^{-1} B Q e_{11} \) defined by, \( e_{11} P^{-1} a P e_{11} = P^{-1} f_{11} a f_{11} P \) and \( Q^{-1} g_{11} a Q e_{11} = e_{11} Q^{-1} a Q e_{11} \); by (1.2) there is an element \( \lambda \) such that \( \lambda^{-1} (e_{11} P^{-1} a P e_{11}) \lambda = e_{11} Q^{-1} a Q e_{11} \). Then we claim that \( \psi: A \rightarrow B \) is induced by the inner automorphism

\[ M \mapsto \lambda^{-1} \cdot IP^{-1} MPL \cdot IQ^{-1} \; \text{for if} \; a \in A, \; a = \sum_{i,j} a_{ij} f_{ij} \; \text{(where} \; a_{ij} \in A, \; \text{and the} \; a's \; \text{and} \; f's \; \text{commute)}, \; \text{so we have} \; a = \sum_{i,j} f_{ij} (f_{11} a_{ij} f_{11}) f_{ij}, \] 

and \( \psi = \sum_{i,j} \left( g_{11} a_{ij} \psi e_{11} g_{11} \right) e_{11} \). But \( Q^{-1} \cdot IP^{-1} f_{11} P A \cdot IQ^{-1} = g_{11} \), and \( Q^{-1} \cdot IP^{-1} (f_{11} a_{1j} f_{11}) P A \cdot IQ^{-1} = Q^{-1} \cdot IP^{-1} a_{1j} P e_{11} P A \cdot IQ^{-1} = \)
(by the choice of \( \lambda \)) \( Qe_{11}Q^{-1}a_{ij}Qe_{11}Q^{-1} = e_{11}a_{ij}e_{11} \). Combining, we see that \( Q\lambda^{-1}IP^{-1}aP\lambda IP^{-1} = Q\lambda^{-1}(\Sigma_{i,j}a_{ij}r_{ij})P\lambda IP^{-1} = e_{11}a_{ij}e_{11} = a \).

We conclude this chapter with a result describing one additional case in which a set of commuting matrices will be similar to a set of matrices over a commutative field.

**Proposition 14.1:** Let \( \mathcal{A} \subseteq D_n \) be a set of commuting diagonal matrices. Then there are a commutative subfield \( F \) of (an extension of) \( D \) and an invertible diagonal matrix \( P \) such that \( P^{-1}AP \subseteq F_n \).

**Proof:** We prove this proposition by induction on \( n \); for \( n=1 \) we have a set of commuting elements of \( D \) which generate (as field) a commutative subfield of \( D \); taking \( P=1 \) we are done.

Assume \( n>1 \) and the result is true for \( n' < n \). We write \( A \) in block form: if \( B \in \mathcal{A} \), we write \( B = \begin{pmatrix} B_{11} & 0 \\ 0 & B_{ii} \end{pmatrix} \). We apply induction on \( n \) to get subfields \( F_1, F_2 \) of \( D \) and diagonal matrices \( P_1, P_2 \) such that \( \begin{pmatrix} P_1^{-1} & 0 \\ 0 & P_2^{-1} \end{pmatrix} \begin{pmatrix} P_{11} & 0 \\ 0 & P_{ii} \end{pmatrix} \begin{pmatrix} P_1 & 0 \\ 0 & P_2 \end{pmatrix} = \begin{pmatrix} F_1 & 0 \\ 0 & F_2 \end{pmatrix} \) for all \( B \in \mathcal{A} \). By commutative field theory (cf. [10], ch. 4, sec. 11) we can embed isomorphic copies \( F_1' \) and \( F_2' \) of \( F_1, F_2 \), respectively, in a commutative field \( F \). We identify \( F_1 \) with \( F_1' \), and by (1.A) we may assume \( F_2 \subseteq D \). Now \( F_2 \subseteq F_2' \subseteq F_2 \subseteq D \), so by (1.B) there is an element \( \lambda \) such that \( \lambda^{-1}F_2 \lambda = F_2' \). Then \( \begin{pmatrix} 1 & 0 \\ 0 & \lambda^{-1} \end{pmatrix} \begin{pmatrix} P_1^{-1} & 0 \\ 0 & P_2^{-1} \end{pmatrix} \begin{pmatrix} P_1 & 0 \\ 0 & P_2 \end{pmatrix} = \begin{pmatrix} F_1 & 0 \\ 0 & F_2' \end{pmatrix} \) as desired.

Combining this result with proposition 5 we get

**Corollary 14.1:** A set of commuting separable diagonal-
zable matrices over a skew field $D$ is similar to a set of diagonal matrices with entries in a commutative subfield $F$ of (some extension of) $D$. 
3. UPPER TRIANGULARIZING AND DIAGONALIZING

SEMIGROUPS OF MATRICES

In the previous chapter we considered, among other questions, the questions of upper triangularizing and diagonalizing sets of commuting matrices. In this chapter we look more closely at the questions of simultaneously upper triangularizing and diagonalizing sets of matrices, directing our attention to arbitrary semigroups of matrices. Our approach will be divided into three stages: first we note some necessary conditions for a semigroup of matrices to be simultaneously upper triangularizable (diagonalizable); then we define a closure operation $\mathcal{J} \rightarrow \mathcal{J}^*$ on subsemigroups of $D_n$ and prove that a semigroup $\mathcal{J}$ can be upper triangularized (diagonalized) if and only if its closure $\mathcal{J}^*$ can be; and finally we show that our previously noted necessary conditions are actually sufficient for a closed semigroup to be upper triangularizable (diagonalizable). Unfortunately, one of these necessary and sufficient conditions will be that all subgroups of $\mathcal{J}^*$ can be upper triangularized (diagonalized), so the utility of these results is limited by our ability to answer the corresponding questions for groups. It perhaps bears pointing out, since the author is not aware that the corresponding results were known for matrices over commutative fields, that the theorems and proofs of this chapter remain valid if "skew field" is everywhere replaced by "commutative field".

Suppose $\mathcal{J}$ is a semigroup of matrices in upper triangular form. Then the nilpotent elements of $\mathcal{J}$ are those
with all entries on the main diagonal equal to 0, and the product of such an element with any other element of \( J \) (on right or left) will again be nilpotent. Thus the nilpotent elements of \( J \) form a (semigroup) ideal of \( J \). Trivially, as \( J \) is upper triangular, every subgroup of \( J \) is upper triangular. Third, for an upper triangular matrix \( M \) over a commutative field, the (right) eigenvalues of \( M \) are precisely the diagonal entries of \( M \); thus for a semigroup \( J \) of upper triangular matrices over a commutative field we have [(right) eigenvalues of \((AB)\)] \( \subseteq \) [(right) eigenvalue of \( A \), b a (right) eigenvalue of \( B \)] for all \( A, B \) in \( J \). Not surprisingly, the same is true for a semigroup of upper triangular matrices over a skew field; this fact follows from

**Lemma 1:** Let \( T=\left(t_{ij}\right)_{i,j} \in D_n \) be an upper triangular matrix. Then \( \{ \text{right eigenvalues of } T \} = \{ d^{-1}t_{ii}d \mid 1 \leq i \leq n, d \in D \setminus \{0\} \} \).

**Proof:** Suppose \( \alpha \) is a right eigenvalue of \( T \). Let \( e_1, \ldots, e_n \) be the standard basis of \( D^n \), and let \( v \) be a right eigenvector of \( T \) corresponding to \( \alpha \). Write \( v = \sum_{i=1}^{r} e_i \alpha_i \), where \( \alpha_i \neq 0 \). Then \( \sum_{i=1}^{r} e_i \alpha_i \alpha = v \alpha = T v = \sum_{i=1}^{r} e_i \alpha_i t_{ii} + e_i t_{rr} \alpha_i + e_i t_{ir} \alpha_r \).

Now for \( i < r \), \( e_i \in \langle e_1, \ldots, e_{r-1} \rangle \). We see from the above equations and the fact that the \( e_i \)'s are independent that \( \alpha r \alpha = t_{rr} \alpha_r \), and so \( \alpha = \alpha^{-1} t_{rr} \alpha_r \).

Suppose on the other hand that \( \alpha \) is conjugate to some \( t_{ii} \), \( 1 \leq i \leq n \). We choose \( i \) minimal such that \( \alpha \) is conjugate (over some extension of \( D \)) to \( t_{ii} \), say \( \alpha = d^{-1}t_{ii}d \). Then if \( t_{ii} \) is algebraic over the center \( k \) of \( D \) and \( j < i \), \( t_{jj} \) is not a root of the \( k \)-irreducible polynomial satisfied by \( t_{ii} \).

It follows from ([6], th. 3.2(ii)) that the equations
\[ O = t_{ij}x_j - x_j t_{ii} + t_{ji} \]

have solutions for \( x_j, j = 1, \ldots, i-1 \). Then

\[
T' = \begin{bmatrix}
0 & -x_1 & \cdots & 0 \\
\vdots & \ddots & \ddots & \vdots \\
0 & \cdots & -x_i & 0 \\
0 & \cdots & 0 & 0 \\
\end{bmatrix}
\]

and \( T' \) has \( t_{ii} \) as right eigenvalue. As right eigenvalues are similarity invariants \((1.1)\), \( T \) has \( t_{ii} \) as right eigenvalue; as conjugates of right eigenvalues are right eigenvalues \((1.1)\), \( \alpha \) is a right eigenvalue of \( T \).

Summarizing the discussion prior to lemma 1, we get

**Proposition 2:** If a semigroup \( J \subseteq D_n \) is upper triangularizable, then (i) the nilpotent elements (if any) of \( J \) form a (semigroup) ideal of \( J \); (ii) any subgroup of \( J \) can be upper triangularized; and (iii) for all matrices \( A, B \in J \),

\[
\text{right eigenvalues of } (AB) \subseteq \text{right eigenvalues of } A, \text{ b a right eigenvalue of } B.
\]

We observe that condition (iii) in many cases is a weak condition: if \( A \) and \( B \) have non-central right eigenvalues, then the right hand side contains products of whole conjugacy classes, and may in fact be all of \( D \); however, we shall be interested in condition (iii) for idempotent matrices \( A \) and \( B \), in which case the right hand side is a subset of \( \{0, 1\} \).

Finding necessary conditions for a semigroup of matrices to be diagonalizable is easier, and we have

**Proposition 3:** If a semigroup \( J \subseteq D_n \) can be simultaneously diagonalized, then (i) \( J \) contains no non-zero nilpotent matrices; (ii) all subgroups of \( J \) are diagonalizable; and (iii) all idempotents of \( J \) are central.

The conditions of proposition 3 and a weakening of
those in proposition 2 are sufficient for a large class of semigroups, but to define this class we must digress.

If \( M \in D_n \) is any matrix, then we saw in chapter 1 that \( M \) is similar to a diagonal sum of a nonsingular matrix and a nilpotent matrix, say \( P^{-1}M^q = \begin{pmatrix} R & 0 \\ 0 & Q \end{pmatrix} \), \( R \) invertible, \( Q \) nilpotent. We set \( A_M = P\left(\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}\right)P^{-1} \), \( N_M = P\left(\begin{pmatrix} 0 & 0 \\ 0 & Q \end{pmatrix}\right)P^{-1} \), \( E_M = P(0,0)P^{-1} \), \( A_M^* = P\left(\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}\right)P^{-1} \). Immediately from the definitions we get the following

**Proposition 4:** The matrices \( A_M, A_M^*, E_M, N_M \) just defined satisfy

\[
\begin{align*}
(1) \quad & E_M^2 = E_M \\
(2) \quad & A_M A_M^* = E_M = A_M^* A_M \\
(3) \quad & A_M E_M = E_M A_M = A_M \\
(4) \quad & A_M^* E_M = E_M A_M^* = A_M^* \\
(5) \quad & E_M N_M = N_M E_M = 0 \\
(6) \quad & M = A_M + N_M \\
(7) \quad & M^n = A_M^n \\
(8) \quad & E_M^* = A_M = M E_M^* 
\end{align*}
\]

In fact, (1) and (8) follow from (2)-(7), but we do not need that result. We do need to show, however, that \( A_M, E_M, A_M^*, N_M \) are independent of the matrices \( P, R, \) and \( Q \) which we used to define them. This fact follows from

**Proposition 5:** If \( A_M, A_M^*, E_M, N_M \in D_n \) satisfy (1)-(8) of proposition 4 then \( A_M = A_M^*, A_M^* = A_M^*, E_M = E_M, \) and \( N_M = N_M \).

**Proof:** We have \( E_M = (A_M^*)^n (A_M)^n = (A_M^{*})^n (M)^n = (A_M^*)^n (A_M)^n \), \( E_M = (A_M^*)^n (A_M)^n \). Similarly, \( E_M = E_M^* E_M \), so \( E_M = E_M^* \). Then \( A_M E_M = E_M A_M = A_M \), \( N_M = N_M - A_M = M - A_M = N_M \); and \( A_M^* E_M = A_M^* \) \( E_M = A_M^* A_M = A_M = A_M^* \).

(The first part of this proof is lemma 4 of ([13])).

Since the decomposition defined above depends only
on $M$, we call $A_M$ the non-singular part of $M$, $N_M$ the nilpotent part of $M$, $E_M$ the idempotent associated to $M$, and $A_M^\mu$ the relative inverse of $M$. (Note that $M=A_M+N_M$ is not a generalization of the Jordan decomposition for matrices over a commutative field, and $A_M^\mu$ does not correspond to the usual generalized inverse of a matrix over a commutative field).

In view of proposition 5 we can define the nilpotent and non-singular parts of a linear transformation $\psi$ by taking any matrix $M$ representing $\psi$ and setting $N_\psi=\phi_{N_M}$, $A_\psi=\phi_{A_M}$; $E_\psi$ and $A_\psi^\mu$ are similarly defined.

Because of our interest in upper triangularizing matrices, we need

**Lemma 6:** With the above notation, (a) $M$ is upper triangular if and only if both $N_M$ and $A_M$ are; (b) if $M$ is upper triangular, so are $E_M$ and $A_M^\mu$.

**Proof:** (a) if: This is clear, as $M=N_M+A_M$.

only if: Suppose $M=(m_{ij})$ is upper triangular. Since $N_M=M-A_M$, it suffices to show that $A_M$ is upper triangular. We first show that there is a uni-triangular matrix $Q$ such that $Q^{-1}MQ=M'$ has $i,j$ entry 0 if one but not both of $m_{ii}$, $m_{jj}$ are 0. Suppose $j<n$ and

(1) for all $j', 1\leq j' < j$, $m_{jj'}=0$ if one but not both of $m_{ii}$, $m_{jj'}$ are 0. We find a uni-triangular matrix $T$ such that $T^{-1}MT=M'$ satisfies (1) with $j$ replaced by $j+1$. For $i=1,\ldots,j-1$, if one but not both of $m_{ii}$, $m_{jj}$ are 0 we can solve the equation $C=m_{ii}x_i-x_i^m_{jj}+m_{ij}$ for $x_i$; if both or neither $m_{ii}$, $m_{jj}$ are 0 let $x_i=0$. Then

$$M_i=\begin{pmatrix} 1 & 0 & -x_i & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \cdots \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 1 \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 1 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \cdots \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 1 \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 1 \end{pmatrix}$$

has the property claimed. As
our original matrix $M$ satisfies (1) with $j=2$, a finite num-
ber of such steps gives us our desired matrix $M'$.

Now let $i_1, \ldots, i_r$ be (in ascending order) the indices
such that the $i_j, i_j$ entry of $M'$ is 0, and let $j_1, \ldots, j_{n-r}$
be (in ascending order) the remaining indices. Let $\sigma \in S_n$ be
the permutation $\left(1, \ldots, n-r, n-r+1, \ldots, n\right)$, and let $P \in \text{GL}_n(\mathbb{D})$ be
the permutation matrix corresponding to the transformation
$e_1 \rightarrow e_\sigma(1)$. Then $P^{-1}M'P = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$, where $A$ is upper triangular
and non-singular and $B$ is upper triangular and nilpotent.

Then $A_{M'} = P \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix} P^{-1}$ is upper triangular by the way $P$ was
defined, and so (since $Q$ is upper triangular) $A_{M'} = Q^{-1}A_M$, $Q$
is upper triangular, as desired.

(b) In the notation of the proof of (a),

$E_{M'} P^{-1} \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} P$
is upper triangular, as is $A_{M'} = P^{-1} \begin{pmatrix} A^{-1} & 0 \\ 0 & 0 \end{pmatrix} P$,

and so $E_M$ and $A_{M'}$ will be too.

A subsemigroup $\mathcal{S}$ of $D_n$ will be called closed if whenever
we have $M \in \mathcal{S}$, we also have $N_M$, $A_M$, $E_M$, and $A_{M'} \in \mathcal{S}$. Any
semigroup similar to a closed semigroup will be closed (by
proposition 5), so we can make the analogous definition for
a semigroup of linear transformations—in fact, any such
semigroup will be of the form $\phi$, $\mathcal{S}$ a subsemigroup of $D_n$,
and $\phi$ is closed if and only if $\mathcal{S}$ is closed. Clearly the
intersection of closed semigroups will be closed, and we
get a closure operation on subsemigroups of $D_n$,

$\mathcal{S} \rightarrow \mathcal{S}^* = \cap \{ J \mid J \subseteq \mathcal{S}, J$ a closed subsemigroup of $D_n \}$. We can des-
cribe this closure of a semigroup of matrices explicitly;
set $\mathcal{S}_0 = \mathcal{S}$, and define inductively (for $i \geq 1$) $\mathcal{S}_i$ to be the sub-
semigroup of $D_n$ generated by $\mathcal{S}_{i-1}$ and $\{ A_M, N_M, E_M, A_{M'} \mid M \in \mathcal{S}_{i-1} \}$.

Then we get
Lemma 7: With the above notation, $A^* = \bigcup_{i=1}^{\infty} A_i$. 

Proof: Write $\mathcal{J}$ for $\bigcup_{i=1}^{\infty} A_i$.

$\mathcal{J} \subseteq \mathcal{J}^*$: Clearly $\mathcal{J} \subseteq \mathcal{J}^*$, and if $\mathcal{J} \subseteq \mathcal{J}^*$ then $\mathcal{J}_{i+1} \subseteq \mathcal{J}^*$. Thus $\mathcal{J} \subseteq \mathcal{J}^*$ by induction on $i$.

$\mathcal{J}^* \subseteq \mathcal{J}$: It suffices to show that $\mathcal{J}$ is a closed semigroup. But as a union of a tower of subsemigroups of $D_n$, $\mathcal{J}$ is clearly a semigroup. Further, if $\mathcal{M} \in \mathcal{J}$, then $\mathcal{M} \in A_i$ for some $i$, so $A_{\mathcal{M}}$, $N_{\mathcal{M}}$, $F_{\mathcal{M}}$, $A_{\mathcal{M}}^r A_i$, and $A_{\mathcal{M}}^r$ are all upper triangular by lemma 6. Thus the generators of $A_{\mathcal{M}}^r A_i$ are upper triangular, and so $A_{\mathcal{M}}^r A_i$ will be. Induction on $i$ completes the proof.

One desirable property of the closure operation, from our point of view, is given in

Theorem 8: Let $\mathcal{J}$ be a subsemigroup of $D_n$. Then $\mathcal{J}$ can be upper triangularized if and only if $\mathcal{J}^*$ can be.

Proof: if: This is clear, as $\mathcal{J} \subseteq \mathcal{J}^*$.

only if: Suppose $P^{-1} \mathcal{J} P$ is upper triangular. By lemma 7 it suffices to show that $P^{-1} A_i P$ is upper triangular for each $i \in N$. For $i=0$ this is our assumption; if $P^{-1} A_i^{-1} P$ is upper triangular ($i>1$) then for any $M \in P^{-1} A_i^{-1} P$, $A_i$, $N_i$, $F_i$, and $A_i^r$ are all upper triangular by lemma 6. Thus the generators of $P^{-1} A_i P$ are upper triangular, and so $P^{-1} A_i P$ will be. Induction on $i$ completes the proof.

As for diagonalizability we have

Proposition 9: Let $\mathcal{J}$ be a subsemigroup of $D_n$. Then $\mathcal{J}$ can be diagonalized if and only if $\mathcal{J}^*$ can be.

Proof: if: Clear.

only if: Again, if $P^{-1} \mathcal{J} P$ is diagonal, it suffices to show that $P^{-1} A_i P$ is diagonal for all $i$. For $i=0$ this is our assumption. Suppose $P^{-1} A_i^{-1} P$ is diagonal, $i>1$, and let $M=(m_{ij}) \in P^{-1} A_i^{-1} P$. Then it follows from proposition 5 that
\( N_M = 0; A_M = M; E_M = (e_{ij}), \) where \( e_{ij} = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{otherwise} \end{cases} \) and

\( A^+(M) = (a_{ij}), \) where \( a_{ij} = \begin{cases} 0 & \text{if } i \neq j \\ 0 & \text{if } i = j \text{ and } m_{ii} = 0; \end{cases} \) These are all diagonal, and together they generate \( P^{-1}J_1P. \) Thus \( P^{-1}J_1P \)

is diagonal, and the \textit{proposition} follows by induction on \( i. \)

Before going on, it might be helpful to note that, as a special case, a semigroup with zero (=0) consisting only of idempotent matrices is closed. It was this example which led to theorem 12. But in order to prove theorem 12, where we get a converse to proposition 2 for closed semigroups, we shall need two preliminary results. The first of these is of some interest in itself:

\textbf{Theorem 10:} Let \( J \) be a semigroup of linear transformations of \( D^n. \) Denote by \( \mathcal{N} \) the nilpotent elements of \( J \) and by \( \mathcal{U} \) the non-nilpotent elements of \( J. \) Then \( J \) has an upper triangular representation if and only if \( \mathcal{N} \) is a (semigroup) ideal of \( J \) or \( \mathcal{N} \) is empty, and \( \mathcal{U} \) has an upper triangular representation.

\textbf{Proof:} only if: Clear.

if: Suppose \( \mathcal{N} \) is an ideal of \( J \) or \( \mathcal{N} \) is empty, and that \( \mathcal{U} \) has an upper triangular representation. Write \( V \) for \( D^n; \) we first show that if \( W \) is an \( J \)-invariant subspace with quotient space \( \overline{V} \) then \( u\big|_W \) and \( \overline{u} \) have upper triangular representations.

\( \overline{u}\big|_\overline{V}: \) Suppose \( B = \{v_1, \ldots, v_n\} \) is a basis of \( V \) such that \( \mu(u, B) \) is upper triangular for all \( u \in J. \) Write \( V_i \) for \( \langle v_1, \ldots, v_i \rangle, i = 1, \ldots, n. \) Let \( i_i \) be minimal such that \( \overline{V}_{i_i} \neq 0, \) and once \( i_j \) has been determined, let \( i_{j+1} \) be minimal such
that $\bar{V}_{i,j+1} = \bar{V}_{i,j}$. Then the representation of $\bar{W}_V$ with respect to the basis $\{\bar{V}_1, \ldots, \bar{V}_r\}$ is upper triangular (where $r = \dim V$).

$U|_W$: Let $V_i$ be as above, $i=1,\ldots,n$. Let $j_i$ be minimal such that $V_{j_i} \cap W \neq \{0\}$, and let $f_i = v_{j_i} + v_{j_i} \varepsilon_1 \varepsilon_1 \varepsilon v_{j_i} \cap W$. Once $f_i = v_{j_i} + v_{j_i} \varepsilon_1 \varepsilon_1 \varepsilon v_{j_i} \cap W$ has been selected, let $j_{i+1}$ be minimal such that $V_{j_{i+1}} \cap W \neq V_{j_i} \cap W$, and let $f_{i+1} = v_{j_{i+1}} + v_{j_{i+1}} \varepsilon_1 \varepsilon_1 \varepsilon v_{j_{i+1}} \cap W$. Then the representation of $U|_W$ with respect to the basis $\{f_1, \ldots, f_{n-1}\}$ is upper triangular.

We return to the main proof and use induction on $n$. If $n=1$ the result is clear. Assume $n>1$ and the result is true for all $n'<n$.

If $\mathcal{H}$ is empty or $\mathcal{H} = \{0\}$ then $J=U$ or $J=W \cup \{0\}$ and the theorem holds because $U$ has an upper triangular representation by assumption. Thus we can assume that $Q \subseteq V=W$. By Levitzki's Theorem* ([11], p. 135) $W \not\subseteq V$, and since $\mathcal{H}$ is an ideal, $W$ is a non-trivial $J$-invariant subspace.

The result will follow by induction on $n$ if we can show that (1) non-nilpotents of $J|_W$, $T_V$ have upper triangular representations, and (2) nilpotents of $J|_W$, $T_V$ form an ideal. (1) follows from our opening argument and the fact that the non-nilpotents of $J|_W$, $T_V$ are contained in $U|_W$, $T_V$ respectively. To prove (2), let $\phi|_W$ be nilpotent, and let $\psi \in J$. We want to show that $\psi \phi|_W$ and $\psi \psi|_W$ are nilpotent. If either of $\phi, \psi$ come from $\mathcal{H}$, this follows from the fact that $\mathcal{H}$ is an ideal and the restriction of a nilpotent transformation is nilpotent, so we can assume

*Levitzki's Theorem: A semigroup of nilpotent matrices over a skew field can be simultaneously upper triangularized.
ϕ, ψ ∈ U. By our opening observation and our assumption on
U, ϕ|₆ and ψ|₆ have a representation as upper triangular
matrices. Since ϕ|₆ is nilpotent, its corresponding matrix
has 0's on the main diagonal, and so the matrices corre-
sponding to ψϕ|₆ and ϕψ|₆ are nilpotent. The proof of (2)
for 1₆ is similar.

We digress briefly to note just two applications of
theorem 10:

**Corollary 10.1**: If 1 is a subsemigroup of Dₙ whose nil-
potent elements form an ideal and whose non-nilpotent ele-
ments commute then 1 can be upper triangularized.

Recall that a unipotent matrix is a matrix of the form
I+N, N nilpotent.

**Corollary 10.2**: Let F be a commutative field, 1 a subsemi-
group of Fₙ consisting only of unipotent and nilpotent ma-
trices. Then 1 can be upper triangularized.

**Proof**: First we note that the nilpotents of 1 form an
ideal of 1: if N ∈ 1 is nilpotent and S ∈ 1 then NS ∈ 1 is sin-
gular. Thus NS cannot be unipotent (unipotent matrices
are non-singular), and since every element of 1 is either
unipotent or nilpotent, NS must be nilpotent. Similarly SN
is nilpotent, so the nilpotents of 1 form an ideal.

Also, the unipotents form a subsemigroup, as unipotents
are units in Fₙ, so no product of unipotents in 1 could be
nilpotent. The corollary then follows by theorem 10 and
Kolchin's Theorem* ([11], p. 100).

We return now to the goal of proving the converse of
proposition 2 for closed semigroups. We continue with

*Kolchin's Theorem*: A semigroup of unipotent matrices over
a commutative field can be simultaneously upper triangularized.
Lemma 11: If \( J \) is a closed semigroup of linear transformations of \( V = D^n \) and \( W \) is an \( J \)-invariant subspace, then \( J|_W \) and \( \overline{J}_W \) are closed.

Proof: Let \( \phi \in J \); \( A_\phi, E_\phi, N_\phi \), and \( A_\phi^\# \) are all in \( J \) (since \( J \) is closed), and satisfy (1)-(8) of proposition 4. Thus \( A_\phi|_W, E_\phi|_W, N_\phi|_W \), and \( A_\phi^\#|_W \) are all in \( J|_W \) and satisfy (1)-(8) of proposition 4, so by proposition 5 they are (respectively) the non-singular part of \( \phi|_W \), the idempotent associated to \( \phi|_W \), the nilpotent part of \( \phi|_W \), and the relative inverse of \( \phi|_W \). Thus we see that \( J|_W \) is closed. Similarly for \( \overline{J}_W \).

We can now state and prove

Theorem 12: Let \( J \subseteq D^N \) be a closed semigroup (cf. p. 44). Then \( J \) can be upper triangularized if and only if the following conditions are satisfied: (i) the nilpotents of \( J \) form a (semigroup) ideal of \( J \); (ii) every subgroup of \( J \) can be upper triangularized; (iii) for any idempotents \( A, B \in J \), \{right eigenvalues of \((AB)\}\subseteq\{0,1\}.

Proof: only if: By proposition 2.

if: Let \( J \subseteq D^N \) be a closed semigroup satisfying (i), (ii), (iii). We first show that if \( W \) is a \( J \)-invariant subspace of \( V = D^N \), then any representations of \( \phi_J|_W \) and \( \overline{\phi}_J|_W \) satisfy (i), (ii), (iii).

(i) \( \phi_J|_W \): If \( \phi_M|_W \) is nilpotent then \( \phi_M|_W = \phi_M|_W = \phi_M|_W \), and \( \phi_{NM}|_W \phi_J|_W \) as \( J \) is closed. Then for any \( S \in J \), \( \phi_S|_W \phi_M|_W = \phi_S|_W \phi_M|_W = \phi_{SNM}|_W \) is nilpotent by our assumptions on \( J \). A similar argument works for \( \phi_M|_W \phi_S|_W \), so any representation of \( \phi_J|_W \) satisfies (i). A similar proof works for \( \phi_J|_W \).

(iii) \( \phi_J|_W \): If \( \phi_M|_W \) is idempotent, \( \phi_M|_W = \phi_M|_W = \phi_M|_W \), and \( \phi_{EM}|_W = \phi_{EM}|_W \). Thus if \( \phi_M|_W, \phi_{MO}|_W \) are idempotent, \{right eigenvalues of \( \phi_M|_W \phi_{MO}|_W \} = \{0,1\}.\}
The right eigenvalues of $\phi_{E_M}^L | \phi_{E_M}^L | W = \phi_{E_M}^L | W \leq \phi_{E_M}^L | W \leq \{0, 1\}$. Thus any representations of $\phi_j^L | W$ satisfy (iii). A similar proof works for $\phi_j^N$.

(ii) $\phi_j^L | W$: Let $G$ be a subgroup of $\phi_j^L | W$. We will find a subgroup $H$ of $\mathcal{L}$ such that, for any $G \in G$, $G = \phi_H^L | W$ for some $H \in H$. Since $H$ is upper triangularizable ($\mathcal{L}$ satisfies (ii)), $\phi_H^L | W$ will have an upper triangular representation by the opening argument of theorem 10, and so $G$ will also.

We first need to make two observations about idempotent matrices; they follow immediately if we take a representation of $\phi_E$ in the form $(I \ 0)$:

(a) if $B \in D_8$ and $B = EB^E$ for an idempotent matrix $E$ then $E_B = E_B E$;

(b) if $E, F \in D_n$ are idempotents of the same rank and $F = E F E$ then $E = F$.

We return to the task of finding a group $H$ as described above. By the proof that $\phi_j^L | W$ satisfies (iii), there is an idempotent $E \in \mathcal{L}$ such that $\phi_E^L | W = 1_G$; pick $E$ to be such an idempotent of minimal rank, and let $H$ be the group of units of $E \mathcal{L}$. Let $G \in G$, and let $G \in H$ be such that $\phi_G^L | W = G$. Then $\phi_{E_G}^E | W = G$. Now $\phi_H^L | W \in G$, so $\exists \in \mathcal{L}$ with $\phi_H^L | W \phi_K^L | W = 1_G$.

Then $\phi_H^L | W \phi_K^L | W = \phi_E^L | W$, and by proposition 4 (7) and (3) $\phi_E^L | W \phi_E^L | W = \phi_E^L | W$, and by proposition 4 (7) and (3) $\phi_E^L | W \phi_E^L | W = \phi_E^L | W$, and by proposition 4 (7) and (3) $\phi_E^L | W \phi_E^L | W = \phi_E^L | W$. But by (a), $\phi_E^L | W = \phi_E^L | W$, and by (a), $\phi_E^L | W = \phi_E^L | W$. Now rank $\phi_H^L | W = 1_G$, rank $H = \operatorname{rank}(H) \leq \operatorname{rank} E$; by the minimality of rank $E$, rank $\phi_H^L | W = \operatorname{rank} E$. Also, by (a), $E_H = E_H E$, so by (b) $E = E_H$. Then $H_{E_H} = E_H = E_{E_H}^H$, so $H \in H$, and the proof is complete.

A similar argument works for $\phi_j^N$. 

-50-
We return to the proof of the main theorem. If \( n=1 \) we have nothing to prove. We show that for \( n>1 \), \( V \) is reducible, and the result will then follow by lemma 11, our opening remarks, and induction on \( n \).

Assume \( n>1 \). By Levitzki's Theorem (p. 47) we can assume \( \mathcal{J} \) contains a non-nilpotent matrix, so, as \( \mathcal{J} \) is closed, \( \mathcal{J} \) contains a non-zero idempotent. Let \( E \in \mathcal{J} \) be an idempotent of minimal positive rank \( r \), and take a representation \( \mu \) of \( \phi_\mathcal{J} \) such that \( \mu(\phi_\mathcal{J}) = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \). Let \( S \in \mathcal{J} \) be non-nilpotent, so \( E \subseteq S \). By \( (a) \), rank \( E < \text{rank } E \), so by the minimality of rank \( E \) and \( (a) \) and \( (b) \) above, \( E = S \). As above, it follows that \( S \) is a unit of \( E \mathcal{J} \). Then \( E \mathcal{J} \mathcal{J} \) is a semigroup whose nilpotent elements form a semigroup ideal and whose non-nilpotent elements form a group, which, as a subgroup of \( \mathcal{J} \) can be put in upper triangular form (by \( (ii) \)). By theorem 10, \( E \mathcal{J} \mathcal{J} \) can be upper triangularized; it follows that we can find a representation \( \mu' \) of \( \phi_\mathcal{J} \) such that 
\[
\mu'(\phi_\mathcal{J}) = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}
\]
and, for \( T \in \mathcal{J} \), 
\[
\mu'(\phi_T) = \begin{pmatrix} T_1 & T_2 \\ T_3 & T_4 \end{pmatrix}, \quad T_1 \in D_r, \quad T_4
\]
upper triangular. If \( r>1 \), every element of \( \mu'(\phi_\mathcal{J}) \) has 2,1 entry 0, and the subspace \( W \) of \( V \) generated by the first basis vector and its images under \( \phi_\mathcal{J} \) will have 0 projection on the space spanned by the second basis vector. Thus \( 0 \not\subseteq W, \) and \( W \) is a \( \phi_\mathcal{J} \)-invariant subspace, so we are done.

We thus assume \( \mu'(\phi_\mathcal{J}) = e_{11} \). Let \( \mu'(\phi_X) = (x_{i j}) \) be such that \( x_{i1} \neq 0 \) for some \( i>1 \) (such exist, or the subspace generated by the first basis vector and its images would be a non-trivial \( \phi_\mathcal{J} \)-invariant subspace). Then
\[
\mu'(\phi_{XB}) = \begin{pmatrix} x_{11} \\ \vdots \\ x_{n1} \end{pmatrix} e_{11} \mu'(\phi_\mathcal{J}).
\]
Since \( \phi_{XB} \) has rank 1, either \( \phi_{XB} \)
is nilpotent or $\phi_{XE}=A_{XE}$.

**Case 1:** $\phi_{XE}$ is nilpotent; then $x_{11}=0$. Since $P^{-1}\begin{pmatrix} x_{21} \\ x_{n1} \end{pmatrix} \in \mathbb{D}^{n-1}$, $\exists P \in \text{GL}_{n-1}(D)$ such that $P^{-1}\begin{pmatrix} x_{21} \\ x_{n1} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$; we take the representation $\mu^2(\phi_{E})=(1 \cdot P^{-1}) \cdot (1 \cdot P)$. Then $\mu^2(\phi_{E})=e_{11}$, $\mu^2(\phi_{XE})=e_{21}$.

Let $Y=(y_{ij}) \in \mathbb{E}^2$. Then $\mu^2(\phi_{E})Y=\begin{pmatrix} y_{11} \cdots y_{1n} \end{pmatrix} \in \mathbb{E}^2(\phi_{E})$, and $\mu^2(\phi_{E})Y \mu^2(\phi_{XE})=\begin{pmatrix} y_{12} 0 \cdots 0 \end{pmatrix} \in \mu^2(\phi_{E})$. Since $XE$ is nilpotent and the nilpotents of $\mathbb{E}$ form an ideal, $\mu^2(\phi_{E})Y \mu^2(\phi_{XE})$ is nilpotent and $y_{12}=0$. As $Y$ was arbitrary, we see that every element of $\mu^2(\phi_{E})$ has $1,2$ entry $0$, and we get a non-trivial $\phi_{E}$-invariant subspace as before.

**Case 2:** $\phi_{XE}=A_{XE}$, so $x_{11} \neq 0$. Now by (b), $E_{XE}=E_{XE}E$, so $\mu'(\phi_{XE})=\mu'(\phi_{XE})\mu'(\phi_{E})=(\text{say})\begin{pmatrix} 11 \\ n1 \end{pmatrix}$. Since $\mu'(\phi_{XE})$ is idempotent and non-zero, $f_{11}=1$; since $E_{XE}E=E_{XE}$, $f_{11}x_{11}=x_{11}$; then because $x_{11} \neq 0$ for some $i>1$, $f_{11} \neq 0$ for some $i>1$. Again, taking an appropriate choice of basis, we get a representation $\mu^2$ of $\phi_{E}$ such that $\mu^2(\phi_{E})=e_{11}$, $\mu^2(\phi_{XE})=e_{11}+e_{21}$. Now let $\mu^2(\phi_{E})=(y_{ij}) \in \mu^2(\phi_{E})$. We will again show that $y_{12}=0$, and so get a non-trivial $\phi_{E}$-invariant subspace as before.

Now $\mu^2(\phi_{E})=\begin{pmatrix} y_{11} \cdots y_{1n} \end{pmatrix}$; if $E_{E}=0$, $y_{12}=0$, so we need only consider the case where $E_{E}$ has rank $1$. Again, either $E_{E}$ is nilpotent or $E_{E}=A_{E_{E}}$.

**Case a:** $E_{E}$ is nilpotent. Then $y_{11}=0$; also $\mu^2(\phi_{XE})=\begin{pmatrix} 0 & y_{12} & \cdots & y_{1n} \\ 0 & y_{12} & \cdots & y_{1n} \end{pmatrix}$ must be nilpotent by assumption (i), so $y_{12}=0$.

**Case b:** $E_{E}=A_{E_{E}}$, and $y_{11} \neq 0$. Again, $\mu^2(\phi_{E})$ is of the form $\begin{pmatrix} a_{11} & \cdots & a_{1n} \end{pmatrix}$, and $a_{11}=1$. Further, if $y_{12} \neq 0$, we conclude
from the equation $\mathbf{EY}_E = \mathbf{EY}$ that $a_{12} \neq 0$. Thus, to show $y_{12} = 0$
it is enough to show $a_{12} = 0$. Let $H = \mu^2(\phi_E)\mu^2(\phi_E) = (e_{11} + e_{21}) \begin{pmatrix} 1 & a_{12} & \cdots & a_1 & a_n \\ 1 & a_{12} & \cdots & a_1 & a_n \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{12} & a_{12} & \cdots & 1 & a_n \\ a_{12} & a_{12} & \cdots & 1 & a_n \end{pmatrix}$. $H$ is a product of idempotents, so by (iii) the only right eigenvalues of $H$
are 0 and/or 1. Now $H \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}(1+a_{12})$, so $1+a_{12} = 0$ or $1+a_{12} = 1$.

If $1+a_{12} = 1$, $a_{12} = 0$ as desired. If $1+a_{12} = 0$, $a_{12} = -1$ and $H^2 = 0$.

As the nilpotents of $\mathcal{I}$ form an ideal (by (i)), $\mu^2(\phi_E)H = e_{11} \begin{pmatrix} 1 & a_{12} & \cdots & a_1 & a_n \\ 1 & a_{12} & \cdots & a_1 & a_n \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{12} & a_{12} & \cdots & 1 & a_n \\ a_{12} & a_{12} & \cdots & 1 & a_n \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \end{pmatrix}$ must be nilpotent, which it clearly is not. Thus $1+a_{12} \neq 0$, and the proof is complete.

We state explicitly as a corollary the example which motivated the theorem:

**Corollary 12.1:** If $\mathcal{E} \subset D_n$ is a semigroup consisting of idempotent matrices then $\mathcal{E}$ can be put in upper triangular form.

**Proof:** Clearly $\mathcal{E}^* = \{0\} \cup \mathcal{E}$ will satisfy conditions (i), (ii), (iii) of theorem 12.

The converse of proposition 3 for closed semigroups is much easier:

**Proposition 13:** Let $\mathcal{I} \subset D_n$ be a closed semigroup of matrices. Then the matrices of $\mathcal{I}$ can be simultaneously diagonalized if and only if the following three conditions hold: (i) $\mathcal{I}$ contains no non-zero nilpotent matrices; (ii) all subgroups of $\mathcal{I}$ can be diagonalized; (iii) all idempotents of $\mathcal{I}$ are central.

**Proof:** only if: Proposition 3.

if: Let $\mathcal{I}$ be a closed subsemigroup of $D_n$ satisfying (i), (ii), (iii). If $W$ is a $\phi_\mathcal{I}$-invariant subspace then $\phi_\mathcal{I}|_W$
will be closed and satisfy (i), (ii), (iii) (the proofs are similar to those in theorem 12). If $E \in \mathcal{A}$ is an idempotent of minimal positive rank, we take a representation $\mu$ of $\phi_2$ such that $\mu(\phi_E) = \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}$. Let $S E$ and write $\mu(\phi_S) = \begin{pmatrix} S_1 & S_2 \\ S_3 & S_4 \end{pmatrix}$.

Then $\mu(\phi_{ESE}) = \begin{pmatrix} S_1 & 0 \\ 0 & 0 \end{pmatrix}$; since $E$ is central, $\mu(\phi_{ESE}) = \mu(\phi_{SE}) = \begin{pmatrix} S_1 & 0 \\ S_3 & 0 \end{pmatrix} = \mu(\phi_{ES}) = \begin{pmatrix} S_1 & S_2 \\ 0 & 0 \end{pmatrix}$, so $S_2$ and $S_3$ are both $0$. Thus $\phi_2$ leaves $EV$ and $(I-E)V$ invariant. The minimality of rank $E$ allows us to show that $E \in \mathcal{E}$ is a group (with zero), and so can be diagonalized by assumption. On the subspace $(I-E)V \subseteq V$ left invariant by $\phi_2$, we can get a diagonal representation of the restrictions by induction on $n$; combining, we get a diagonal representation for $\phi_2$.

*Note we do not claim $I-E \in \mathcal{A}$, only that $\phi_2$ leaves $(I-E)V$ invariant.*
4. VARIOUS RESULTS ON LINEAR GROUPS

In light of the results of the previous chapter it is appropriate that we consider the questions of upper triangularizing and diagonalizing groups of matrices over a skew field. Results here are much more fragmentary and less closely related. First we prove that a solvable group of unipotent matrices over a skew field can be simultaneously upper triangularized (cf. Kolchin's Theorem, p. 48). We also get a partial analogue for nilpotent groups of a well known theorem of Mal'cev in the theory of linear groups (cf. [14], p. 75)—that a solvable group of matrices over a commutative field has a subgroup of finite index which can be upper triangularized. And lastly we determine, for a given skew field D, those finite groups for which every representation over D can be diagonalized.

We recall that if \( G=G^{(0)} \) is a group then \( G^{(1)}=G' \), the subgroup generated by \( \{g_1g_2g_1^{-1}g_2^{-1} | g_1, g_2 \in G \} \), is a normal subgroup of \( G \) called the derived group. Once \( G^{(1)} \) has been defined, \( G^{(i+1)} \) is defined to be \( (G^{(i)})' \); in this way we get a series of normal subgroups of \( G \), called the derived series. If there is an integer \( n \) such that \( G^{(n)}=\{1\} \), \( G \) is called solvable; if \( G^{(n-1)}\neq\{1\} \), \( G \) is said to be solvable of length \( n \) (cf. [14], p. 45). We emphasize that if \( G \) is a solvable group of length \( n>0 \) then \( G^{(n-1)} \) is a non-trivial abelian normal subgroup of \( G \).

We can now state and prove

**Theorem 1:** If \( Q \subset D_n \) is a solvable group of unipotent matrices (see p. 48), then \( Q \) can be put in upper triangular form.
Proof: We use induction on \( n \); if \( n = 1 \) the result is clear. Assume \( n > 1 \) and the theorem is true for all \( n' < n \). Since the restriction of \( \phi_q \) to an invariant subspace and the induced group on the quotient space will be solvable groups of unipotent transformations, our induction assumption allows us to assume that \( \phi_q \) leaves no non-trivial invariant subspaces.

What we now prove is that if \( \phi_q \) is irreducible then \( G = \{I\} \) and \( n = 1 \); this will complete the proof of the theorem. Assume \( \phi_q \) irreducible, and let \( A \) be any abelian normal subgroup of \( G \). By (2.1) we can put \( A \) in upper triangular form, so the matrices in \( A \) have a common right eigenvector. Let \( w \) be any common right eigenvector of the matrices in \( A \), let \( G \in G \), and let \( M \in A \). We now show that \( Gw \) is a right eigenvector of \( M \)--i.e., \( G \) maps common right eigenvectors of \( A \) into common right eigenvectors of \( A \). Since \( A \) is normal, \( MG = GM' \) for some \( M' \in A \). Then \( M(Gw) = G(M'w) = (Gw)\alpha_{M'} \), where \( \alpha_{M'} \) is the right eigenvalue of \( M' \) corresponding to the right eigenvector \( w \). Since \( G \) is a group and \( w \neq 0 \), \( Gw \neq 0 \), so \( Gw \) is a right eigenvector of \( M \). Thus the set of right eigenvectors of \( A \) is invariant under \( \phi_q \), and the subspace \( W \) they generate will be a \( \phi_q \)-invariant subspace. We saw that \( W \neq 0 \); by our assumption on the irreducibility of \( G \) we conclude that \( W = D^n \). It follows that the matrices in \( A \) can be simultaneously diagonalized. But the only diagonal unipotent matrix is \( I \), so \( A = \{I\} \). Thus the only abelian normal subgroup of \( G \) is \( \{I\} \); since \( G \) is solvable it follows that \( G = \{I\} \); then since \( G \) is irreducible, \( n = 1 \) and the result follows.
We note a corollary of possible interest to students of linear groups:

**Corollary 1.1:** Let $\mathcal{G} \leq \mathfrak{D}_n$ be a group of unipotent matrices. If $\mathcal{G}$ is locally solvable then $\mathcal{G}$ is solvable.

**Proof:** If $\mathcal{G}$ is locally solvable then by theorem 1 and (1.2) $\mathcal{G}$ can be upper triangularized. But an upper triangular group of unipotent matrices is easily seen to be solvable.

We recall that, for any group $G$, the upper central series of $G$ is defined to be the series

$$\{1\} = \mathfrak{z}_0(G) = \mathfrak{z}_1(G)^n = \ldots = \mathfrak{z}_k(G)^n = \ldots,$$

where for $i > 0$

$$\mathfrak{z}_i(G) = \{ xy \in G | \text{ for all } y \in G, \ xyx^{-1}y^{-1} \in \mathfrak{z}_{i-1}(G) \}.$$ $G$ is called nilpotent if there is an integer $m$ such that $\mathfrak{z}_m(G) = G$; if $\mathfrak{z}_{m-1}(G) \neq G$, $G$ is said to be nilpotent of class $m$. Any subgroup or homomorphic image of a nilpotent group is nilpotent (cf. [15], pp. 140-142).

Our interest will be centered on finitely generated nilpotent groups. By ([7], p. 153), any subgroup of a finitely generated nilpotent group is finitely generated, so it follows that subgroups and homomorphic images of finitely generated nilpotent groups are finitely generated and nilpotent.

We can now prove

**Proposition 2:** Let $\mathfrak{G} \leq \mathfrak{D}_n$ be a finitely generated nilpotent group of algebraic matrices. Then $\mathfrak{G}$ has a subgroup $\mathcal{N}$ of finite index which can be upper triangularized.

**Proof:** We use induction on $n$ to show that it suffices to take $\mathfrak{G}$ irreducible. If $n=1$ the result is trivial, so we assume $n>1$ and the proposition is true for $n'<n$. Suppose
also that $W$ is a non-trivial $\phi_1$-invariant subspace.

Since $\phi_G|_W$ and $\phi_1|_V$ are homomorphic images of $\phi_1$, they will be finitely generated nilpotent groups. Also, for $G \in \mathcal{Q}$, $\phi_G|_W$ and $\phi_1|_V$ satisfy the same polynomial $p_G(t) \in \mathbb{K}[t]$ satisfied by $G$, and so (any representations of) these transformations will be algebraic. Thus we can apply our induction assumption to (representations of) $\phi_1|_W$ and $\phi_1|_V$ to find a subgroup $\phi_{H_0}$ of $\phi_1|_W$ of finite index such that $\phi_{H_0}$ has an upper triangular representation, and a subgroup $\phi_{K_0}$ of $\phi_1|_V$ of finite index such that $\phi_{K_0}$ has an upper triangular representation. Let $H, K$ be maximal subgroups of $\mathcal{Q}$ such that $\phi_H|_W = \phi_{H_0}$, $\phi_K|_V = \phi_{K_0}$; then $(\mathcal{Q}:H)<\infty$, $(\mathcal{Q}:K)<\infty$, and so $H \cap K$ is a subgroup of $\mathcal{Q}$ of finite index which can be upper triangularized. Thus we can assume that $\phi_1$ leaves no non-trivial invariant subspaces.

We now use induction on the class of nilpotency $m$ of $\mathcal{Q}$. If $m=1$ $\mathcal{Q}$ is abelian and the proposition holds by (2.1). We thus assume $m>1$ and the result is true for all nilpotent subgroups of $D_n$ of class less than $m$. We shall find a subgroup $H_1$ of finite index in $\mathcal{Q}$ of class less than $m$. $H_1$ will be finitely generated, and clearly every element of $H_1$ will be algebraic. We can thus apply induction on $m$ to get a subgroup $H$ of $H_1$, $(H_1:H)<\infty$, such that $H$ can be upper triangularized; but then $(\mathcal{Q}:H)$ will be finite, and the proposition proved.

We must first show that we can assume that $\phi_1(\mathcal{Q})$ consists of scalar matrices with diagonal entries in the center of $D$; suppose $\phi_1(\mathcal{Q}) = \langle M_1, \ldots, M_s \rangle$, where $M_1, \ldots, M_{r-1}$ are central scalar matrices ($1<r<s$). Suppose $P^{-1}M_rP$ is in
normal form; since $M_r$ is central in $\mathcal{Q}$, the right eigenvectors of $M_r$ generate a $\phi_q$-invariant subspace, so since $\phi_q$ is irreducible, we conclude that $P^{-1}M_rP$ is diagonal. Now non-conjugate diagonal entries of $P^{-1}M_rP$ give rise to nontrivial $\phi_q$-invariant subspaces, and we see (from our definition of normal form) that $P^{-1}M_rP$ is scalar, say $P^{-1}M_rP=\alpha I$. The centrality of $M_r$ in $\mathcal{Q}$ means that the matrices $P^{-1}MP$, $M \in \mathcal{Q}$, have their entries in the skew subfield of $D$ centralizing $\alpha$, and once we restrict ourselves to this skew field $P^{-1}M_rP$ is a central scalar matrix. We use induction on $r$ (starting with $r=1$) to justify our assumption that $J_1(\mathcal{Q})$ consists of central scalar matrices.

Let $A \in J_2(\mathcal{Q})$; we now show that $(\mathcal{Q}:C_2(A))<\infty$. We do this by showing $A$ has only finitely many conjugates in $\mathcal{Q}$. Let $B \in \mathcal{Q}$; since $A \in J_2(\mathcal{Q})$, $B^{-1}AB=\lambda_B I$, where $\lambda_B I \in J_1(\mathcal{Q})$ is a central scalar matrix. Let $\mu$ be a right eigenvalue of $A$, with right eigenvector $v$. Then $A(Bv)=B(Av)\mu=(\text{since } \lambda_B \text{ is central}) BV\lambda_B=BV(\mu \lambda_B)$, so $\mu \lambda_B$ is also a right eigenvalue of $A$. Now since $A$ is algebraic, there is a polynomial $p(t) \in k[t]$ such that $p(A)=0$. We see that for any right eigenvalue $\alpha$ of $A$, $p(\alpha)=0$. In particular, $p(\mu \lambda_B)=0$ for every $B \in \mathcal{Q}$. Thus $\{\mu \lambda_B | B \in \mathcal{Q}\} \subseteq \{\text{roots of } p(t) \text{ in } k(\mu)\}$; since $k(\mu)$ is commutative, $p(t)$ has at most degree $p$ roots in $k(\mu)$, so $\{\mu \lambda_B | B \in \mathcal{Q}\}$ is finite. As $\mu \neq 0$ ($A$, being in $\mathcal{Q}$, is invertible), $\{\lambda_B | B \in \mathcal{Q}\}$ is finite. But all conjugates of $A$ are of the form $A \lambda_B I$, so $A$ has only finitely many conjugates.

Now $J_2(\mathcal{Q})$ is finitely generated, say $J_2(\mathcal{Q})=<M_1,\ldots,M_t>$. Then $C_\mathcal{Q}(J_2(\mathcal{Q}))=\bigcap_{i=1}^t C_\mathcal{Q}(M_i)$, and since each $C_\mathcal{Q}(M_i)$ has finite
index in \( q \), it follows that \( (q: C_q(\mathcal{Z}_2(q))) = \infty \). Recall we were looking for a nilpotent subgroup \( H_1 \) of \( q \) of class at most \( m-1 \); we take \( H_1 = C_q(\mathcal{Z}_2(q)) \). Then 
\[
\mathcal{Z}_2(q) \cap H_1 = \mathcal{Z}_2(q) \cap C_q(\mathcal{Z}_2(q)) = H_1, 
\]
so 
\[
\mathcal{Z}_2(q) \cap H_1 = \mathcal{Z}_1(H_1). 
\]
We claim that for \( i \geq 1 \), \( \mathcal{Z}_{i+1}(q) \cap H_1 \leq \mathcal{Z}_i(H_1) \); this has been established for \( i=1 \), so assume it is true for \( i-1 \geq 1 \). Then 
\[
\mathcal{Z}_1(H_1) = \{ x \in H_1 | \forall y \in H_1, xyx^{-1}y^{-1} \in \mathcal{Z}_{i-1}(H_1) \} = \{ x \in H_1 | \forall y \in H_1, xyx^{-1}y^{-1} \in \mathcal{Z}_1(q) \} \subseteq H_1 \cap \{ x \in q | \forall y \in q, xyx^{-1}y^{-1} \in \mathcal{Z}_1(q) \} = H_1 \cap \mathcal{Z}_{i+1}(q). 
\]
It follows that \( \mathcal{Z}_{m-1}(H_1) \supseteq H_1 \cap \mathcal{Z}_m(q) = H_1 \), so \( H_1 \) is nilpotent of class at most \( m-1 \), as desired.

We remark that, if \( D \) had been commutative, we would have been able to show that any nilpotent group of matrices has a subgroup of finite index which can be upper triangularized; in that case every matrix would be algebraic anyway, and the assumption about finite generation, which was used to get a finite subset \( \{ M_1, \ldots, M_t \} \subseteq \mathcal{Z}_1(q) \) such that 
\[
C_q(\mathcal{Z}_1(q)) = C_q(\{ M_1, \ldots, M_t \}),
\]
could be dropped, as any maximal (necessarily finite) \( D \)-linearly independent subset of \( \mathcal{Z}_1(q) \) would have this property.

We now leave our study of simultaneously upper triangularizing matrices and return to the question of the simultaneous diagonalization of matrices.

If \( F \) is a commutative field, a group of matrices over \( F \) can be simultaneously diagonalized if and only if it is abelian and every matrix in the group is diagonalizable. Over a skew field, however, a group of diagonal matrices need not be abelian, and one might ask, for a given skew field \( D \), if there are non-abelian groups \( q \) such that any subgroup of \( D_n \) isomorphic to \( q \) can be simultaneously diag-
onalized. We seek in the rest of this chapter to determine such finite groups \( G \); we discover that for some skew fields there are such non-abelian groups.

In general, if \( G \) is a finite group of diagonal \( n \times n \) matrices over a skew field \( D \), then \( G \) is embeddable in a direct product of \( n \) copies of the multiplicative group \( D^* = D - \{0\} \), and the projection of \( G \) onto each coordinate will be a finite subgroup of \( D^* \). Since by ([8]) the only finite multiplicative subgroups of a skew field of characteristic \( p > 0 \) are cyclic, we see that if the characteristic of \( D \) is different from 0 the only such finite groups \( G \) are abelian.

Thus we assume for the rest of our considerations that \( D \) has characteristic 0, and so \( Q \) (the rational number field) is contained in the center \( k \) of \( D \). If \( G \) is a finite group, of order \( m \), say, then we can embed the group ring \( QG \) in \( Q_m \) by the regular representation. We denote by \( \{M_G | G \in G\} \) the matrices of this representation corresponding to the elements of \( G \). We thus have \( QG = Q[\{M_G | G \in G\}] \subseteq D_m \) for every skew field \( D \) of characteristic 0, and clearly the group of matrices \( \{M_G | G \in G\} \approx Q \) can be diagonalized if and only if \( QG \) can be. If \( G \) has a non-normal subgroup \( \mathcal{N} \), then

\[
\sum_{h \in \mathcal{N}} M_h
\]

is a non-central idempotent of \( QG \), and it follows that \( QG \) cannot be diagonalized.

Thus the only possible non-abelian groups \( G \) of finite order with all representations over \( D \) diagonalizable have all their subgroups normal. Such groups are called Hamiltonian, and have the form \( \mathcal{N} \times \mathcal{O} \times \mathcal{T} \), where \( \mathcal{N} \) is the quaternion group of order 8, \( \mathcal{O} \) an abelian group of odd order, and \( \mathcal{T} \)
a group of exponent 2 ([7], p. 190).

We consider now a case where a group of the form \( \mathcal{H} \times \mathcal{O} \), \( \mathcal{H} \) the quaternion group, \( \mathcal{O} \) a cyclic group of odd order, have representations which are not diagonalizable. Let \( n \) be an odd natural number (greater than 1), and consider the \( 4n \times 4n \) matrices

\[
\varphi = \begin{pmatrix} A & 0 \\ 0 & A \end{pmatrix}, \quad \eta = \begin{pmatrix} 0 & I_4 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & I_4 \\ I_4 & 0 & \cdots & 0 \end{pmatrix},
\]

where

\[
A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}.
\]

If \( \varphi = \langle \varphi, \eta \rangle \), then \( \varphi = \mathcal{H} \times \mathcal{C}_n \), where \( \mathcal{C}_n \) is cyclic of order \( n \). Further, \( \varphi \) and \( \eta \) have as right eigenvalues only primitive fourth roots of \( 1 \), and \( \eta \) has a primitive \( n \)-th root of \( 1 \) as right eigenvalue. Thus if \( \varphi \) can be simultaneously diagonalized, some diagonal entry, call it \( n_{ij} \), of the matrix corresponding to \( \eta \) will be a primitive \( n \)-th root of \( 1 \). If we denote by \( i_{ij}, j_{ij} \) the corresponding diagonal entries of the matrices corresponding to \( \varphi, \eta \) respectively, then from the above observations and the multiplication table of \( \varphi \) it follows that \( \langle i_{ij}, j_{ij}, n_{ij} \rangle = \mathcal{H} \times \mathcal{C}_n \). Thus if \( \varphi \) can be diagonalized \( \mathcal{H} \times \mathcal{C}_n \) can be embedded in a skew field of characteristic 0. By Amitsur's results on finite subgroups of skew fields of characteristic 0 ([1], th. 7(2))*, this can happen if and only if 2 has odd order modulo \( n \).

In light of the above examples we are left with the following result:

**Proposition 2:** Let \( D \) be a skew field of characteristic 0, and denote by \( k \) its center. Let \( \mathcal{H} \times \mathcal{O} \times \mathcal{C}_n \) (as above) be a

*In Amitsur's notation, \( \mathcal{H} \times \mathcal{C}_n = \mathcal{C}_{4n, 2n+1} \); theorem 4(2b) is applicable.
Hamiltonian group, where 2 has odd order modulo the exponent e of G. If there is a skew field E containing k in its center and containing a subgroup isomorphic to G × H, then any group G of nxn matrices isomorphic to H × G × T can be simultaneously diagonalized.

Proof: By a result of Zalesskii ([17], lem. 1, p. 980) any finite group of matrices over a skew field of characteristic 0 is completely reducible. In particular, in view of (2.1), any finite abelian group of matrices over such a skew field will be diagonalizable. Now if W ⊆ D^n is a g^-invariant subspace, φ_g|_W will either be abelian, hence any representation of it will be diagonalizable, or any representation of it will be the sort of group described in the proposition. Thus by Zalesskii's result and induction on n we can assume φ_g is irreducible.

Any matrix M ∈ g of order 2 can be put in the form 
\begin{pmatrix}
I_\sigma & 0 \\
O^\sigma & -I_{n-\sigma}
\end{pmatrix};
if 0 < s < n, the centrality of M in g will give a decomposition of D^n into a direct sum of g^-invariant subspaces, so we may assume that the only matrix of order 2 in g is -I, which will be fixed by all similarity transformations. In other words, we may take g = H × G, ignoring the other matrices of order 2.

By (1.A) we may assume E ⊆ D. Let {ω, i, j} ⊆ D generate a group isomorphic to H × G, with <ω> = C_\omega and <i, j> = H. Let M ∈ g have order e. We can diagonalize M to get a matrix M' whose diagonal entries come from the field Q(ω), and we can take conjugate diagonal entries equal. As M is in the center of g our assumption on proper invariant subspaces allows us to assume M' is the scalar matrix ω · I. Denote by g' the
corresponding representation of \( \phi_f \). The centrality of \( M' \) in \( q' \) means that \( q' \) is contained in the matrix ring over the skew subfield \( D_0 \) of \( D \) centralizing \( \omega \), so now we restrict ourselves to considering matrices over \( D_0 \). If \( M_0 \in q' \) is any other matrix of odd order, we can diagonalize \( M_0' \) over \( D_0 \), and since the order of \( M_0' \) divides \( e \), the matrix \( M_0'' \) we get will have as diagonal entries powers of \( \omega \). Now distinct powers of \( \omega \) will be non-conjugate over \( D_0 \), and by the centrality of \( M_0 \) and the irreducibility of \( \phi_f \), we conclude that \( M_0'' \) is scalar (hence \( M_0''=M_0' \)). Thus all matrices in \( q' \) of odd order are scalar, and if we work over \( D_0 \) they will remain scalar.

We are thus reduced to the problem of diagonalizing a group \( q \) of matrices in \((D_0)^n\), where \( q \approx \# \) and the unique element of \( q \) of order 2 is \(-1\), or equivalently, of showing that if \( n>1 \) \( \phi_f \) is reducible. Let \( \mathcal{X}, y \in q \) be generators satisfying \( \mathcal{X}^2=-1=\mathcal{Y}^2 \), \( \mathcal{X} \neq \mathcal{Y} \). We first show that \( k(\omega)[\mathcal{X}, \mathcal{Y}]=k(\omega, i, j) \). Any element in \( k(\omega)[\mathcal{X}, \mathcal{Y}] \) can be written in the form \( a \cdot I + b \cdot \mathcal{X} + c \cdot \mathcal{Y} + d \cdot \mathcal{X} \mathcal{Y} \), \( \{a, b, c, d\} \subseteq k(\omega) \), and any element of \( k(\omega, i, j) \) can be expressed \( a \cdot I + b \cdot i + c \cdot j + d \cdot ij \), \( \{a, b, c, d\} \subseteq k(\omega) \). The map \( a \cdot I + b \cdot i + c \cdot j + d \cdot ij \mapsto a \cdot I + b \cdot \mathcal{X} + c \cdot \mathcal{Y} + d \cdot \mathcal{XY} \) is clearly a \( k \)-algebra homomorphism onto, and as the domain is simple, it is an isomorphism. Thus \( k(\omega)[\mathcal{X}, \mathcal{Y}] \) is a skew field; it is in fact four dimensional over its center \( k(\omega) \cdot I \), so by (2.12) either \( n=1 \) or it, and hence \( q \), is reducible.

In particular, it follows from Amitsur's work ([1]) that if \( k=\mathbb{Q} \) then any group of the sort described in proposition 3 can be diagonalized. More generally, this will
be true if $k(\omega)$, $\omega$ a primitive $e^{th}$ root of 1, is any commu-
tative field in which 0 is not a sum of four or fewer non-
zero squares; for then the algebra generated over $k(\omega)$ by
elements 1, i, j, ij with the multiplication of the quater-
nion algebra will be a skew field, the inverse of a non-zero
element $a \cdot i + b \cdot i + c \cdot j + d \cdot ij$ being

$$(a^2 + b^2 + c^2 + d^2)^{-1}(a \cdot i - b \cdot i - c \cdot j - d \cdot ij).$$
OPEN QUESTIONS

1. a. Let $k \subseteq K \subseteq D_n$, $K$ a skew field. Is $K$ similar to a scalar skew field? (cf. 2.4).
   
b. Let $k \subseteq K \subseteq D_n$, $n > 1$. Is $K$ reducible? (cf. 2.12).
   
c. Characterize absolutely irreducible semigroups of matrices.
   
d. If $A_1, A_2$ are isomorphic simple algebras, $k \subseteq A_1 \subseteq D_n$, are $A_1$ and $A_2$ similar? (cf. 2.11.1)

2. a. (Kaplansky) Can every semigroup of unipotent matrices be upper triangularized? (cf. Kolchin's Theorem, p. 48; also 4.1).
   
b. Is every group of unipotent matrices solvable? (cf. 4.1).

3. a. Does 4.2 generalize to solvable groups of matrices?
   
b. Can the finiteness conditions in 4.2 (the requirements that the matrices be algebraic and the group finitely generated) be weakened or dropped?

4. a. (Zalcstein) Is a periodic subgroup of $D_n$ locally finite?
   
b. Is a subgroup of $D_n$ of bounded exponent locally finite? (These are germane to question 2 for the case of characteristic $p > 0$).
REFERENCES


