

SIMILARITY OF SETS OF MATRICES
OVER A SKEW FIELD

Ph. D. Thesis

WALTER SCOTT SIZER

Bedford College, University of London

ProQuest Number: 10098309

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 10098309

Published by ProQuest LLC(2016). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code.
Microform Edition © ProQuest LLC.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

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 triangularizability of a solvable group of unipotent matrices and a determination, for any skew field D , of those finite groups all of whose representations over D are diagonalizable.

TABLE OF CONTENTS

	PAGE
Introduction	5
1. Preliminaries, and the Normal Form of a Single Matrix	7
2. Commutativity Results	17
3. Upper Triangularizing and Diagonalizing Semigroups of Matrices	39
4. Various Results on Linear Groups	55
Open Questions	67
References	69

CORRECTIONS TO "SIMILARITY OF SETS OF MATRICES OVER A
SKEW FIELD"

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

p. 10, l. 4: $w_j = v_j t - \sum_{i=1}^n v_i m_{ij}$

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: $(e_{11} + e_{21}) \begin{pmatrix} 1 & a_{12} \cdots a_{1n} \\ \bigcirc & \end{pmatrix} = \begin{pmatrix} 1 & a_{12} \cdots a_{1n} \\ 1 & a_{12} \cdots a_{1n} \\ \bigcirc & \end{pmatrix}$.

p. 63, l. 1-2: exponent e of \mathcal{O} .

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 -, left R_n -module in the natural way.

Most of our interest will be concentrated on questions of the similarity of matrices. If $\lambda \subseteq D_n$ we will frequently want to show that λ 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 ϕ_M the corresponding linear transformation of D^n ; if $\mathcal{S} \subseteq D_n$ is a set of matrices, $\phi_{\mathcal{S}} = \{\phi_S \mid S \in \mathcal{S}\}$.
2. If $\psi \in \text{End}(D^n)$ and B is a basis of D^n , $\mu(\psi, B)$ will denote the matrix representing ψ with respect to B ; if B is not explicitly given, $\mu(\psi)$ will be used to denote a representation of ψ . If $\mathcal{S} \subseteq \text{End}(D^n)$, $\mu(\mathcal{S}, B) = \{\mu(\psi, B) \mid \psi \in \mathcal{S}\}$ and $\mu(\mathcal{S}) = \{\mu(\psi) \mid \psi \in \mathcal{S}\}$.

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_λ '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 ψ 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 ψ , 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 $D * E$ --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 $v \cdot \sum \alpha_i t^i = \sum M^i v \alpha_i$, $v \in V$, $\alpha_i \in D$. 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 $\{w_j = v_j t - \sum_{i=1}^n m_{ij} v_i \mid j=1, \dots, n\}$. By results on the equivalence of matrices over a non-commutative pid, $tI - M$ is equivalent to a diagonal matrix $\text{diag}[\delta_1, \dots, \delta_n]$, where δ_i is a total divisor of δ_{i+1} , and it follows that $V \cong D[t]/\delta_1 D[t] \oplus \dots \oplus D[t]/\delta_n D[t]$. Recall that δ_i is a total divisor of δ_{i+1} means that there is an element $a_i \in D[t]$ with $a_i D[t] = D[t] a_i$ (an invariant element) such that $\delta_i | a_i$ and $a_i | \delta_{i+1}$. 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) \bar{a} for which $\delta | \bar{a}$ is called a bound for δ . Note that $\delta_1, \dots, \delta_{n-1}$ are bounded, but δ_n need not be.

Now $D[t]/\delta_1 D[t] \oplus \dots \oplus D[t]/\delta_n D[t]$ can be decomposed into a sum

$$(1.a) \quad D[t]/\alpha_1 D[t] \oplus \dots \oplus D[t]/\alpha_s D[t] \cong V,$$

where each α_i , $i=1, \dots, s-1$, is bounded and indecomposable, and α_s has no bounded non-unit factors; furthermore, the expression (1.a) will then be unique up to isomorphism. The α_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 ϕ_M as a diagonal sum of matrices

$$(1.b) \quad \mu(\phi_M) = M_1 + \dots + M_s,$$

where M_i has elementary divisor α_i for $i < s$.

The invariant elements of $D[t]$ are just those associated to polynomials in $k[t]$. Thus if α_i is bounded, say with bound $p_i(t)$, we can pick $p_i(t)$ to lie in $k[t]$. It follows that $p_i(M_i) = 0$, and M_i satisfies a non-zero polynomial with

coefficients in k . Such a matrix M_i is called algebraic. If α_s is not a unit, then since α_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 , $i < s$, which were algebraic with elementary divisors α_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 $a^* = a_1 \dots a_r a_{r+2} \dots a_q$ of a^* . We can assume by (1.A) that D contains a commutative splitting field F of a^* , so we get another atomic factorization $a^* = (t - \beta_1) \dots (t - \beta_u)$ of a^* in $F[t]$, hence in $D[t]$. Since atomic factorizations have the same number of factors and a and all the a_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 $D[t]/aD[t] \cong D[t]/bD[t]$ (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 α_i was indecomposable, hence a product of similar bounded atoms. We observe that

(1.c) if $\alpha_i = (t-a_1)\dots(t-a_m)$ is an atomic factorization of α_i (where $a \in D$) then the matrix

$$\begin{pmatrix} a_1 & 1 & & & 0 \\ & a_2 & 1 & & \\ & 0 & \ddots & \ddots & \\ & & & & 1 \\ 0 & & & & & a_m \end{pmatrix}$$

has elementary divisor α_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 α_i a bounded indecomposable element of $D[t]$. Then

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

Let $p_i(t) \in k[t]$ be the monic bound of α_i ; since α_i is indecomposable, p_i is a power of an invariant atom ([4], p. 230)

-- i.e., $p_i = (q_i)^{n_i}$, q_i a monic irreducible polynomial in

$k[t]$. We may assume (1.A) that D contains a commutative

splitting field F_1 of q_i . Let $q_i(t) = (t-c_{i1})\dots(t-c_{ir})$ be

a factorization of q_i in F_1 . We claim that $(t-c_{i1})^{n_i}$ has

bound p_i ; clearly $(t-c_{i1})^{n_i} | q_i^{n_i} = p_i$, so $(t-c_{i1})^{n_i}$ has as

bound q_i^m , $m \leq n_i$. Then we get $q_i^m = (t-c_{i1})^{n_i} s(t) =$

$(t-c_{i1})^m \prod_{j=2}^r (t-c_{ij})^m$. It is easily seen that $s(t) \in F_1[t]$, a

domain, and so we get $(t-c_{i1})^{n_i-m} s(t) = \prod_{j=2}^r (t-c_{ij})^m$. Now c_{i1}

is not a zero of the right hand side, as k is perfect, so

c_{i1} is not a zero of the left hand side, and $n_i = m$. By (1.c)

the $n_i \times n_i$ matrix $\begin{pmatrix} c_{i1} & 1 & & & 0 \\ & \vdots & \ddots & \ddots & \\ & 0 & & & 1 \\ & & & & c_{i1} \end{pmatrix} = M'_i$ has associated module

$D[t]/(t-c_{i1})^{n_i} D[t]$; by (1.d), $D[t]/(t-c_{i1})^{n_i} D[t] \approx D[t]/\alpha_i D[t]$,

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_i , and c_{i1}

to be any root of q_i in F_1 ; we can take all the F_i 's equal,

say $F_i = F$, $i=1, \dots, s-1$, 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 α_i and α_j have as bounds powers of the same $k[t]$ -irreducible polynomial, we can take $c_{i1} = c_{j1}$. 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 ([5]) in more detail. If $M \in D_n$, $\alpha \in D$, then α 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 ψ if there is a non-zero $v \in D^n$ such that $\psi v = v\alpha$; such a v is called a right eigenvector of ψ . As might be expected, we have

Theorem D: ([5], prop. 2.1) If α 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 (ψ). If α is a right eigenvalue for $M \in D_n$ and $P \in \text{GL}_n(D)$, then α 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 $\{M_\alpha \mid \alpha \in \mathcal{A}\}$ be a set of $n \times n$ matrices over a skew field D . Suppose that for every finite subset $S \subseteq \mathcal{A}$ there are an extension D_S of D and a matrix $P_S \in \text{GL}_n(D_S)$ such that $P_S^{-1}M_S P_S$ is upper triangular (diagonal) for each $s \in S$. Then there are an extension E of D and a matrix $P \in \text{GL}_n(E)$ such that $P^{-1}M_\alpha P$ is upper triangular (diagonal) for all $\alpha \in \mathcal{A}$.

Proof: Let $T = \{\text{all finite non-empty subsets of } \mathcal{A}\}$. For $S \in T$, let D_S, P_S be as in the hypotheses. Our object is to obtain a suitable ultrafilter \mathcal{F} on T so that the skew field $E = \prod_{S \in T} D_S / \mathcal{F}$ has the property of the conclusion. (For the definition of $\prod_{S \in T} D_S / \mathcal{F}$ 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 $C_1, \dots, C_m \in \mathcal{C}$, say $C_i = \{U \in T \mid X_i \subseteq U\}$, then $X_1 \cup \dots \cup X_m \in T$ and $X_1 \cup \dots \cup X_m \in \bigcap_{i=1}^m C_i$, so $\bigcap_{i=1}^m C_i$ is non-empty. It follows that \mathcal{C} is contained in an ultrafilter \mathcal{Z} on T , and we set $E = \prod_{S \in T} D_S / \mathcal{Z}$.

We have a natural isomorphism $\Pi((D_S)_n) \cong (\Pi D_S)_n$, and the induced homomorphism $(\Pi D_S)_n \rightarrow E_n$. 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_S^{-1} on the S^{th} factor, and Q its image in E_n , then $PQ = QP = I$. Let $\alpha \in \alpha$. We consider $P^{-1}M_\alpha P$ by looking at $\prod_{S \in T} P_S^{-1}M_\alpha P_S \in \Pi((D_S)_n)$. Let $Z = \{S \in T \mid P_S^{-1}M_\alpha P_S \text{ is upper triangular (diagonal)}\}$. Since \mathcal{Z} is an ultrafilter, either $Z \in \mathcal{Z}$ or the complement of Z , $Z' \in \mathcal{Z}$. If $Z \in \mathcal{Z}$, $P^{-1}M_\alpha P$ is upper triangular (diagonal) and we are done; thus we want to show $Z' \notin \mathcal{Z}$. Let $C = \{S \in T \mid \{\alpha\} \subseteq S\}$. Then $C \in \mathcal{C} \subseteq \mathcal{Z}$, and for $S_0 \in C$, $P_{S_0}^{-1}M_\alpha P_{S_0}$ is upper triangular (diagonal) by the choice of P_{S_0} . Thus $C \subseteq Z$, so $C \cap Z'$ is empty. Now $C \in \mathcal{Z}$, so if $Z' \in \mathcal{Z}$ we get $C \cap Z' = \emptyset \in \mathcal{Z}$, an impossibility. Therefore $Z' \notin \mathcal{Z}$, 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 $\{M_\alpha \mid \alpha \in \alpha\} \subseteq D_n$ be a set of commuting matrices. Then $\{M_\alpha\}$ can be simultaneously upper triangularized (by a similarity transformation leaving the center of D_n fixed).

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

We use induction on n to show that the theorem holds in case $\phi_{M_1}, \dots, \phi_{M_m}$ leave a non-trivial subspace W of $V = D^n$ in-

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 $\phi_{M_1}, \dots, \phi_{M_m}$. Let v_1, \dots, v_n be a basis of V such that v_1, \dots, v_r are a basis of W . As W is non-trivial, $0 < r < n$. With respect to this basis, each ϕ_{M_i} has the form $M_i' = \begin{pmatrix} A_i & B_i \\ 0 & C_i \end{pmatrix}$, $A_i \in D_r$, $C_i \in D_{n-r}$. Since $M_i M_j = M_j M_i$, it follows that the A_i 's commute pairwise, as do the C_i 's. Since $0 < r < n$ we can apply the induction hypothesis twice to get $P \in GL_r(D)$, $Q \in GL_{n-r}(D)$ such that $P^{-1} A_i P$ and $Q^{-1} C_i Q$ are upper triangular for $i=1, \dots, m$. Then $\begin{pmatrix} P^{-1} & 0 \\ 0 & Q^{-1} \end{pmatrix} M_i' \begin{pmatrix} P & 0 \\ 0 & Q \end{pmatrix}$ is upper triangular for $i=1, \dots, m$, 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 α ; then for $i=1, \dots, m$, $M_1(M_i w) = M_i(M_1 w) = (M_i w)\alpha$ -- i.e., either $M_i w = 0$ or $M_i w$ is a right eigenvector of M_1 corresponding to the right eigenvalue α . Thus the space W spanned by right eigenvectors of M_1 corresponding to α is a non-zero subspace of V invariant under $\phi_{M_1}, \dots, \phi_{M_m}$. 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 α . We express the ϕ_{M_i} 's with respect to this basis, and so $\mu(\phi_{M_1}, B) = \alpha \cdot I$. Since $M_1 M_i = M_i M_1$, $i=2, \dots, m$, it follows that α centralizes the entries of $M_i'' = \mu(\phi_{M_i}, B)$, $i=2, \dots, m$. Let D_0 be the skew subfield

of D centralizing α ; then $k(\alpha) \subseteq \text{center}(D_0)$, and M_2'', \dots, 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 3: Let $\{M_\alpha \mid \alpha \in \alpha\} \subseteq D_n$ be a set of commuting matrices. If every element of $k[M_\alpha \mid \alpha \in \alpha]$ (the subring of D_n generated by

$\{M_\alpha \mid U \{b \cdot I \mid b \in k\}\}$ 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 $M \in D_n$ 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 $M_0 = \begin{pmatrix} \alpha_1 & & 0 \\ & \ddots & \\ 0 & & \alpha_n \end{pmatrix}$; since M is algebraic, so is each α_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 r_i 's, hence none are repeated.

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

Proof of proposition 3: By (1.2) it suffices to consider a finite set of matrices $\{M_1, \dots, M_r\}$. We again use induction on n , noting that the result holds trivially if

$n=1$. Now suppose $M'_1 = P^{-1}M_1P$ is diagonal with conjugate diagonal entries equal; denote $P^{-1}M_1P$ 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 $\{\phi_{M_i} \mid i=1, \dots, r\}$, 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 α , we have $M'_i \in K_n$, $i=2, \dots, r$; if we can show that every matrix in $k(\alpha)[M'_2, \dots, M'_r]$ is diagonalizable over K , we will then be done by induction on r .

Let $M \in k(\alpha)[M'_2, \dots, M'_r]$. We can clear expressions in α from the denominators of the coefficients from $k(\alpha)$ by multiplying by a polynomial $p(\alpha) \in k[\alpha]$. We then get a matrix $M' = p(\alpha)M \in k[\alpha][M'_2, \dots, M'_r]$, 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 $Q^{-1}M'_iQ$, $i=2, \dots, r$. Blocks of M'' corresponding to non-conjugate diagonal entries give a direct sum decomposition of K^n into non-trivial $k[\alpha][\phi_{M''_2}, \dots, \phi_{M''_r}]$ -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 α 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 $k[\alpha \cdot I][M''_2, \dots, M''_r]$, so $q(M'')$ is similar to an element of $k[M_1, \dots, M_r]$ and hence diagonalizable. But by inspection

$(q(M''))^n = 0$, 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 $k \subseteq F \subseteq D_n$, F a commutative field. Then F is similar to a scalar field.

Proof: Since $k \subseteq F$, $k[F] = F$. 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 α , 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 $k \subseteq F$), and any non-zero element of F is invertible. Thus $p(M) = 0$, and M is diagonalizable by our diagonalizability criterion.

Let $F' = P^{-1}FP$ 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 $i=1, \dots, n$. By (1.B), there are elements λ_i , $i=2, \dots, n$, such that $\lambda_i^{-1} m_{ii} \lambda_i = m_{11}$ for all $M \in F$. Then $\begin{pmatrix} \lambda_1^{-1} & & 0 \\ 0 & \dots & 0 \\ 0 & & \lambda_n^{-1} \end{pmatrix}_{F'} \begin{pmatrix} \lambda_1 & & 0 \\ 0 & \dots & 0 \\ 0 & & \lambda_n \end{pmatrix}$ 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 $W \subseteq D^n$ is a subspace invariant under ϕ_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, \dots, M_r . We use induction on n ; the proposition is true for $n=1$, and we assume $n > 1$ and ^{that} 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 ϕ_{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 $P^{-1}M_iP$, $i=2, \dots, r$. As before we restrict ourselves to the

skew subfield K of D centralizing α , and observe that

$\{M'_i\} \subseteq K_n$. It follows by the definition of separable matrices and (b) that since M'_2, \dots, 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 $A \in D_n$, $B \in D_m$, 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 $y\alpha - \alpha y = x$, $x\alpha = \alpha x$, then $x = 0$.

Proof: Since $x\alpha = \alpha x$, $yf(\alpha) - f(\alpha)y = f'(\alpha)x$ for any polynomial $f(t) \in k[t]$. In particular, if f is the minimal polynomial for α , 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 $i \times i$ 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

$$\begin{pmatrix} b_1 & b_2 & \dots & b_n \\ 0 & \dots & \dots & b_2 \\ \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & b_1^2 \end{pmatrix} \text{ if } i=j, \begin{pmatrix} b_1 & b_2 & \dots & b_i \\ 0 & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & b_1^2 \end{pmatrix} \text{ if } j>i, \text{ and } \begin{pmatrix} b_1 & b_2 & \dots & b_j \\ 0 & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & b_1 \end{pmatrix} \text{ if } i>j,$$

where $b_h \alpha = \alpha b_h$ for all h .

Proof: We prove the case $i \geq j$; 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 $J_i B = \begin{pmatrix} \alpha b_{11} & \dots & \alpha b_{1j} \\ \vdots & \dots & \vdots \\ \alpha b_{i1} & \dots & \alpha b_{ij} \end{pmatrix} + \begin{pmatrix} b_{21} & \dots & b_{2j} \\ \vdots & \dots & \vdots \\ 0 & \dots & 0 \end{pmatrix}$,

$$B J_j = \begin{pmatrix} b_{11} \alpha & \dots & b_{1j} \alpha \\ \vdots & \dots & \vdots \\ b_{i1} \alpha & \dots & b_{ij} \alpha \end{pmatrix} + \begin{pmatrix} 0 & b_{11} & \dots & b_{1j-1} \\ \vdots & \vdots & \dots & \vdots \\ 0 & b_{i1} & \dots & b_{ij-1} \end{pmatrix}.$$

We first show that $b_{r1} = 0$ for $r > 1$. Comparing $i-1, 1$ and $i, 1$ entries of the products, we get $b_{i1} \alpha = \alpha b_{i1}$, $b_{i-1,1} \alpha = \alpha b_{i-1,1} + b_{i1}$. 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, $j \geq s > 1$. The s^{th} columns of $J_i B$ and $B J_j$ are then $\begin{pmatrix} \alpha b_{1s} + b_{2s} \\ \vdots \\ \alpha b_{i-1s} + b_{is} \\ \alpha b_{is} \end{pmatrix}$ and $\begin{pmatrix} b_{1s} \alpha + b_{1s-1} \\ \vdots \\ b_{s-1s} \alpha + b_{s-1s-1} \\ b_{ss} \alpha \\ \vdots \\ b_{is} \alpha \end{pmatrix}$. 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 $J_i B$ and $B J_j$ we see that $b_{ss} \alpha = \alpha b_{ss}$, $s=1, \dots, j$. We next show that all entries of B commute with α . This statement is true for all sub-diagonal entries and all entries b_{rs} . Assume it is true for all b_{rs} with $0 \leq s-r \leq l-1$, where $0 < l \leq j$. Consider the $r, r+1$ entries of $J_i B$ and $B J_j$, which are $\alpha b_{rr+1} + b_{r+1r+1}$ and $b_{rr+1} \alpha + b_{rr+1-1}$ respectively. Then letting $x = \alpha b_{rr+1} - b_{rr+1} \alpha = b_{rr+1-1} - b_{r+1r+1}$, we see that the expression on the right commutes with α by assumption. Setting $y = b_{rr+1}$, we see from lemma 7 that $x=0$, which is what we wanted to show. Induction on l gives us $b_{rs} \alpha = \alpha b_{rs}$ for $r=1, \dots, i$ and $s=1, \dots, j$. But we also get $0 = x = b_{rr+1-1} - b_{r+1r+1}$, 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 $C_D(\alpha)$ (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 $A = P((\alpha_1 I + H_1) \dot{+} \dots \dot{+} (\alpha_s I + H_s)) P^{-1}$, where $H_i \in D_{n_i}$ is a diagonal sum $N_{a_1}^{(i)} \dot{+} \dots \dot{+} N_{a_{r_i}}^{(i)}$, and the α_i 's are non-conjugate elements of D ; further, at most one α_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 $n_1 \times n_1, \dots, n_s \times n_s$ 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 $a_u \times a_v$ ($1 \leq u, v \leq r_i$), 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 $B(\alpha \cdot I_j + N_j) = (\alpha \cdot I_i + N_i)B$.

The comparable result for the case of a commutative field D --identical except for our restriction to $C_D(\alpha_i)$ --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 $d_1 d_2 \neq d_2 d_1$, $d_i x = x d_i$ ($i=1,2$); let $0, 1, a, b \in k$ be distinct. Take \mathcal{A} to be the group generated by $\begin{pmatrix} x & 0 \\ 0 & x \end{pmatrix}$, $\begin{pmatrix} x+1 & 1 \\ 0 & x+1 \end{pmatrix}$, $\begin{pmatrix} x+a & d_1 \\ 0 & x+a \end{pmatrix}$, $\begin{pmatrix} x+b & d_2 \\ 0 & x+b \end{pmatrix}$. 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 $M = \begin{pmatrix} x^{m_1}(x+1)^{m_2}(x+a)^{m_3}(x+b)^{m_4} & * \\ 0 & x^{m_1}(x+1)^{m_2}(x+a)^{m_3}(x+b)^{m_4} \end{pmatrix}$, $m_i \in \mathbb{Z}$.

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 $v = \begin{pmatrix} \alpha \\ 0 \end{pmatrix}$ 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

$p(t) = a_0 + \dots + a_r t^r$. Since M is non-singular we may assume p was chosen so that $a_0 \neq 0 \neq a_r$. 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 $m_i > 0$ we let R be the ring obtained by localizing $k[x]$ at the prime ideal $\langle x + b_i \rangle$. We can map R homomorphically onto k by mapping x to $-b_i$; then $p(y) \in R$, and as $p(y) = 0$, $p(y)$ is mapped to 0. But by inspection $p(y)$ is mapped to $a_0 \neq 0$, a contradiction. If $m_i < 0$, a similar argument using $p' = (x + b_i)^{-r m_i} p(y)$ gives us $p' \mapsto 0$ and $p' \mapsto a_r \neq 0$, a contradiction. Thus we must have $m_i = 0$ for $i = 1, \dots, 4$, 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

$P \in GL_2(D)$ 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 $\alpha \cdot IP^{-1} \mathcal{A} P \alpha^{-1} \cdot I$, so we may take $P = \begin{pmatrix} 1 & \beta \\ 0 & \gamma \end{pmatrix}$, $\gamma \neq 0$. Let $B = \begin{pmatrix} x+c & f \\ 0 & x+c \end{pmatrix}$ be a generator of \mathcal{A} (so $c \in k$). Then

$P^{-1} B P = \begin{pmatrix} 1 & -\beta \gamma^{-1} \\ 0 & \gamma^{-1} \end{pmatrix} \begin{pmatrix} x+c & f \\ 0 & x+c \end{pmatrix} \begin{pmatrix} 1 & \beta \\ 0 & \gamma \end{pmatrix} = \begin{pmatrix} x+c & x\beta + f\gamma - \beta \gamma^{-1} x\gamma \\ 0 & \gamma^{-1} x\gamma + c \end{pmatrix} \in F_2$. 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 (#) $fyey = e y f y$, or $fye = e y f$ 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 con-

tradiction 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_n 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_n which are irreducible over all extensions of D . Then we get

Theorem 11: Let $\mathcal{A} \subseteq D_n$ be an absolutely irreducible semigroup. \mathcal{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 $C_{E_n}(\mathcal{A})$ with

$$(F_0 \lambda : F_0) \leq n^2.$$

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 $k \cdot I \subseteq K \subseteq D_n$. 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 $Q^{-1}FQ = F' \cdot I$, $F \cong F' \subseteq D$. Let $K' = Q^{-1}KQ$; K' is also absolutely irreducible. Now let $M_1 = I, M_2, \dots, M_t$ 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 $d \leq t \leq n$. But the D -space spanned by m_1, \dots, m_t is the image space under K' of $e_1 D$ (e_1 being the vector with 1 in the first row and 0's elsewhere)--for if $M \in K'$, then $M = \sum_{i=1}^t M_i f_i \cdot I$, $f_i \in F'$, and $M e_1 D = \sum M_i f_i \cdot I e_1 D = \sum M_i e_1 f_i D = \sum m_i f_i D \subseteq \langle m_1, \dots, m_t \rangle$. Further, this space is K' -invariant, as $K' \langle m_1, \dots, m_t \rangle = K' K' e_1 D = K' e_1 D = \langle m_1, \dots, m_t \rangle$. As $\langle m_1, \dots, m_t \rangle \neq 0$, the irreducibility of K' gives us that $\langle m_1, \dots, m_t \rangle = D^n$, so $t \geq n$. But we saw that $t \leq n$, so $t = n$ and $\{m_1, \dots, m_t\}$ is a basis of D^n .

Let $P \in GL_n(D)$ 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 $P^{-1}P = I = P^{-1}(m_1, \dots, m_n)$, so $P = (m_1, \dots, m_n)$. We claim that $P^{-1}K'P \subseteq F'_n$. Let M_1, \dots, M_n be the above $F' \cdot I$ basis of K' . We have $M_i M_j = \sum_{l=1}^n M_l \alpha_{ijl} \cdot I$, $\alpha_{ijl} \in F'$. Now $P^{-1}M_i P P^{-1}m_j = P^{-1}M_i P e_j =$ the j^{th} column of $P^{-1}M_i P$; but $P^{-1}M_i P P^{-1}m_j = P^{-1}(\text{the } 1^{\text{st}} \text{ column of } M_i M_j) = P^{-1}(\text{the } 1^{\text{st}} \text{ column of } \sum M_l \alpha_{ijl} \cdot I) = P^{-1}(\sum M_l \alpha_{ijl}) = \sum e_l \alpha_{ijl} \in (F')^n$. Thus the j^{th}

column of $P^{-1}M_iP$ has entries in F' , so $P^{-1}M_iP \in (F')_n$. We also want to verify that, for $f \in F'$, $P^{-1}f \cdot IP \in (F')_n$. But $P^{-1}f \cdot IP = P^{-1}(f \cdot IM_1, \dots, f \cdot IM_n)$; now $f \cdot IM_i$ is the i^{st} column of $f \cdot IM_i$, and since $f \cdot IM_i \in K'$ we have $f \cdot IM_i = \sum_{j=1}^n M_j \beta_{ji} \cdot I$, $\beta_{ji} \in F'$. Thus the first column of $f \cdot IM_i$ is $\sum M_j \beta_{ji}$, and $P^{-1}f \cdot IP = P^{-1}(f \cdot IM_1, \dots, f \cdot IM_n) = P^{-1}(\sum M_j \beta_{j1}, \dots, \sum M_j \beta_{jn}) = (\beta_{rs}) \in (F')_n$. Now $P^{-1}K'P$, being generated as a ring by $P^{-1}M_iP$ ($i=1, \dots, n$) and $P^{-1}F' \cdot IP$, will be contained in $(F')_n$.

We can further assume by (1.A) that there is a skew field K_0 isomorphic to K with $F' \subseteq K_0 \subseteq D$; thus $P^{-1}K'P \subseteq (F')_n \subseteq (K_0)_n \subseteq D_n$. 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], 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 \mathcal{J} is similar to a semigroup of matrices over a commutative field F , say $P^{-1}\mathcal{J}P \subseteq C_{E_n}$. Then $F \cdot I \subseteq C_{E_n}(P^{-1}\mathcal{J}P)$ and $(F(P^{-1}\mathcal{J}P):F \cdot I) \leq n^2$. But then $F_0 = PF \cdot IP^{-1}$ is a commutative subfield of $C_{E_n}(\mathcal{J})$, and $(F_0\mathcal{J}:F_0) \leq n^2$.

if: Let $\mathcal{J} \subseteq D_n$ be an absolutely irreducible semigroup. Then $R = k[\mathcal{J}] \otimes_k D^{\text{op}}$ will be a ring and D^n a left R -module under the action $(\sum \alpha_i \cdot IS_i \otimes d_i)v = \sum \alpha_i \cdot IS_i v d_i$; the irreducibility of \mathcal{J} means D^n is an irreducible R module, so by Shur's Lemma ([9], p.5) the commuting ring of R is a skew field. The commuting ring of R , commuting with all elements $1 \otimes d$, will be a subring of D_n ; commuting with all elements $s \otimes 1$,

$s \in \mathcal{A}$, 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 $F \subseteq C_{D_n}(\mathcal{A})$ 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 $(F\mathcal{A}:F) \leq n^2$, 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 $k \cdot I \subseteq F$ and $(F\mathcal{A}:F) \leq n^2$.

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

By the Wedderburn-Artin Theorem ([9], p. 48), $F\mathcal{I} = K_r$, K a finite dimensional division algebra. We identify $F\mathcal{I}$ 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 σ of D_n and an embedding $\pi: K \rightarrow D_s$ such that

we get a commutative diagram

$$(1) \quad \begin{array}{ccc} F = K_r & \xrightarrow{\quad} & D_n \\ \pi_r \downarrow & & \sigma \downarrow \\ (D_S)_r & \xrightarrow{\sim} & D_n \end{array}, \text{ where the bottom isomorphism}$$

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 $P^{-1}f_{ii}P = \begin{pmatrix} A_i & B_i \\ C_i & H_i \end{pmatrix}$ ($i=2, \dots, r$) then since $f_{11}f_{ii} = f_{ii}f_{11} = 0$ 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 $Q \in GL_{n-s_1}(D)$ such that $Q^{-1}H_iQ$ is diagonal for $i=2, \dots, r$. 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 $i=1, \dots, r$. A further permutation of the basis vectors will now give us a block decomposition of D_n with r square blocks of dimensions

$s_1 \times s_1, s_2 \times s_2, \dots, s_r \times s_r$ down the diagonal, and the f_{ii} 's simultaneously similar to the matrices $\begin{pmatrix} 0 & \dots & 0 & 0 \\ & \ddots & & \\ & & I_0 & \\ 0 & & & 0 \end{pmatrix}$ (the

I occurring in the i^{th} diagonal block), say

$$S^{-1}f_{ii}S = \begin{pmatrix} 0 & \dots & 0 & 0 \\ & \ddots & & \\ & & I_0 & \\ 0 & & & 0 \end{pmatrix}, i=1, \dots, r. \text{ Now } f_{ij} = f_{ii}f_{ij}f_{jj}, \text{ so}$$

$$S^{-1}f_{ij}S = \begin{pmatrix} 0 & F_{ij} & 0 \\ 0 & 0 & 0 \end{pmatrix}, \text{ the only non-zero block, of dimension}$$

$s_i \times s_j$, occurring in the i^{th} row of blocks and the j^{th} column.

Also $f_{ij}f_{ji} = f_{ii}$, $f_{ji}f_{ij} = f_{jj}$, 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 $\pi: K \approx f_{11} K_r f_{11}^{-1} \rightarrow S^{-1} f_{11} K_r f_{11}^{-1} S \hookrightarrow \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix} \approx D_s$. The inner automorphism σ of D_n that we want to make diagram (1) commute is one such that

$\sigma: f_{ij} \mapsto \begin{pmatrix} O & O \\ O & I_s O \end{pmatrix}$ (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 σ to be the map

$M \mapsto \begin{pmatrix} I_{F_{12}} & \dots & O \\ O & \dots & F_{1r} \end{pmatrix} S^{-1} M S \begin{pmatrix} I_{F_{12}}^{-1} & \dots & O \\ O & \dots & F_{1r}^{-1} \end{pmatrix}$; clearly we have $f_{ii} \mapsto \begin{pmatrix} O & O \\ O & I_s O \end{pmatrix}$; for $i > 1$ we have $f_{1i} \mapsto \begin{pmatrix} I_{F_{12}} & \dots & O \\ O & \dots & F_{1r} \end{pmatrix} \begin{pmatrix} 0 & \dots & 0 & F_{1i} & 0 & \dots & 0 \\ & & & O & & & \end{pmatrix} \begin{pmatrix} I_{F_{12}}^{-1} & \dots & O \\ O & \dots & F_{1r}^{-1} \end{pmatrix} = \begin{pmatrix} 0 & \dots & 0 & I & 0 & \dots & 0 \\ & & & O & & & \end{pmatrix}$. In general, $f_{ij} \mapsto \begin{pmatrix} O & O \\ O & F'_{ij} O \end{pmatrix}$, and it follows from the above remarks and the equations $f_{1j} = f_{1i} f_{ij}$ ($1 \leq i, j \leq r$) that $F'_{ij} = I$, as desired.

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

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

Example 13: Let D be the skew field of (real) quaternions, \mathfrak{L} the group generated by $\rho = \begin{pmatrix} i & 0 \\ 0 & i \end{pmatrix}$, $\mathfrak{J} = \begin{pmatrix} j & 0 \\ 0 & j \end{pmatrix}$, $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. If \mathfrak{L} is not absolutely irreducible then ρ , \mathfrak{J} , and A have a common right eigenvector w ; the only right eigenvalues of all three matrices are primitive fourth roots of 1, so we have $\rho w = w\alpha_i$, $\mathfrak{J} w = w\alpha_j$, $A w = w\alpha_a$, $\alpha_i, \alpha_j, \alpha_a$ primitive fourth roots.

of 1. Then since $\mathcal{A} = \mathcal{A}\mathcal{Q}$, $\mathcal{A} = \mathcal{A}\mathcal{J}$, $\mathcal{J} = -\mathcal{J}$, we see that $\alpha_i \alpha_a = \alpha_a \alpha_i$, $\alpha_j \alpha_a = \alpha_a \alpha_j$, $\alpha_i \alpha_j = -\alpha_j \alpha_i$. This cannot happen with primitive fourth roots of 1 in a skew field, so \mathcal{A} must be absolutely irreducible. Taking $F = \mathbb{R} \left[\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \right]$, $(F : \mathbb{R}) = 4$, so by theorem 11 \mathcal{A} is similar to a semigroup of matrices over \mathbb{C} ; in fact, taking $P = \begin{pmatrix} 1 & j \\ i & -k \end{pmatrix}$, we have $P^{-1} \mathcal{A} P = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$, $P^{-1} \mathcal{J} P = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, $P^{-1} \mathcal{A} P = \begin{pmatrix} i & 0 \\ 0 & i \end{pmatrix}$.

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 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, $A \cong F_n \cong B$, where F is a commutative field. Let $\{e_{ij} \mid 1 \leq i, j \leq n\}$ be the matrix units of D_n , and $\{f_{ij} \mid 1 \leq i, j \leq n\}$, $\{g_{ij} = f_{ij}^\psi \mid 1 \leq i, j \leq n\}$ the matrix units of A, B respectively. As in the proof of theorem 11, there are matrices $P, Q \in GL_n(D)$ with $P^{-1} f_{ij} P = e_{ij}$, $Q^{-1} g_{ij} Q = e_{ij}$, $1 \leq i, j \leq n$. We have an induced isomorphism of subfields $e_{11} P^{-1} A P e_{11} \cong e_{11} Q^{-1} B Q e_{11}$ defined by $e_{11} P^{-1} a P e_{11} = P^{-1} f_{11} a f_{11} P \mapsto Q^{-1} g_{11} a^\psi g_{11} Q = e_{11} Q^{-1} a^\psi Q e_{11}$; by (1.B) there is an element λ such that $\lambda^{-1} (e_{11} P^{-1} a P e_{11}) \lambda = e_{11} Q^{-1} a^\psi Q e_{11}$. Then we claim that $\psi: A \rightarrow B$ is induced by the inner automorphism $M \mapsto Q \lambda^{-1} \cdot I P^{-1} M P \lambda \cdot I Q^{-1}$; for if $a \in A$, $a = \sum \alpha_{ij} f_{ij}$ (where $\alpha_{ij} \in A$, and the α 's and f 's commute), so we have $a = \sum f_{i1} (\alpha_{ij} f_{11}) f_{1j}$, and $a^\psi = \sum g_{i1} (g_{11} \alpha_{ij}^\psi g_{11}) g_{1j}$. But $Q \lambda^{-1} \cdot I P^{-1} f_{ij} P \lambda \cdot I Q^{-1} = g_{ij}$, and $Q \lambda^{-1} \cdot I P^{-1} (f_{i1} \alpha_{ij} f_{11}) P \lambda \cdot I Q^{-1} = Q \lambda^{-1} \cdot I e_{11} P^{-1} \alpha_{ij} P e_{11} \lambda \cdot I Q^{-1} =$

(by the choice of λ) $Qe_{11}Q^{-1}\alpha_{ij}^\psi Qe_{11}Q^{-1} = g_{11}\alpha_{ij}^\psi g_{11}$. Combining, we see that $Q\lambda^{-1} \cdot IP^{-1}aP\lambda \cdot IQ^{-1} = Q\lambda^{-1} \cdot IP^{-1}(\sum f_{i1}(f_{11}\alpha_{ij}f_{11})f_{1j})P\lambda \cdot IQ^{-1} = \sum g_{i1}(g_{11}\alpha_{ij}^\psi g_{11})g_{1j} = a^\psi$.

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: Let $A \subseteq D_n$ 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 $B \in A$, we write $B = \begin{pmatrix} b_{11} & 0 \\ 0 & B_1 \end{pmatrix}$. We apply induction on n to get subfields F_1, F_2 of D and diagonal matrices P_1, P_2 such that $\begin{pmatrix} P_1^{-1} & 0 \\ 0 & P_2^{-1} \end{pmatrix} \begin{pmatrix} b_{11} & 0 \\ 0 & B_1 \end{pmatrix} \begin{pmatrix} P_1 & 0 \\ 0 & P_2 \end{pmatrix} \in \begin{pmatrix} F_1 & 0 \\ 0 & (F_2)_{n-1} \end{pmatrix}$ for all $B \in A$. 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 $F \subseteq D$. Now $F_2 \cong F'_2$, and $F_2, F'_2 \subseteq D$, so by (1.B) there is an element λ such that $\lambda^{-1}F_2\lambda = F'_2$. Then $\begin{pmatrix} \lambda^{-1} & 0 \\ 0 & \lambda^{-1} \end{pmatrix} \begin{pmatrix} P_1^{-1} & 0 \\ 0 & P_2^{-1} \end{pmatrix} A \begin{pmatrix} P_1 & 0 \\ 0 & P_2 \end{pmatrix} \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} \subseteq F_n$ 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 $\mathfrak{L} \mapsto \mathfrak{L}^*$ on subsemigroups of D_n and prove that a semigroup \mathfrak{L} can be upper triangularized (diagonalized) if and only if its closure \mathfrak{L}^* 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 \mathfrak{L}^* 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 \mathfrak{J} is a semigroup of matrices in upper triangular form. Then the nilpotent elements of \mathfrak{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 \mathcal{J} (on right or left) will again be nilpotent. Thus the nilpotent elements of \mathcal{J} form a (semigroup) ideal of \mathcal{J} . Trivially, as \mathcal{J} is upper triangular, every subgroup of \mathcal{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 \mathcal{J} of upper triangular matrices over a commutative field we have $\{(right) \text{ eigenvalues of } (AB)\} \subseteq \{(ab) \mid a \text{ a (right) eigenvalue of } A, b \text{ a (right) eigenvalue of } B\}$ for all A, B in \mathcal{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 $T=(t_{ij}) \in D_n$ be an upper triangular matrix. Then $\{right \text{ eigenvalues of } T\} = \{d^{-1}t_{ii}d \mid 1 \leq i \leq n, d \in D - \{0\}\}$.

Proof: Suppose α is a right eigenvalue of T . Let e_1, \dots, e_n be the standard basis of D^n , and let v be a right eigenvector of T corresponding to α . Write $v = \sum_{i=1}^r e_i \alpha_i$, where $\alpha_r \neq 0$. Then $\sum_{i=1}^r e_i \alpha_i \alpha = v \alpha = T v = \sum_{i=1}^{r-1} t_{ii} \alpha_i + e_r t_{rr} \alpha_r + \sum_{i=1}^{r-1} e_i t_{ir} \alpha_r$. Now for $i < r$ $t_{ii} \alpha_i \in \langle e_1, \dots, e_{r-1} \rangle$. 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 $\alpha = \alpha_r^{-1} t_{rr} \alpha_r$.

Suppose on the other hand that α is conjugate to some t_{ii} , $1 \leq i \leq n$. We choose i minimal such that α 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

$0 = t_{jj}x_j - x_j t_{ii} + t_{ji}$ have solutions for $x_j, j=1, \dots, i-1$. Then

$$T' = \begin{pmatrix} 1 & 0 & -x_1 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & -x_{i-1} & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 1 & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \vdots & 1 \end{pmatrix} T \begin{pmatrix} 1 & 0 & x_1 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & x_{i-1} & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 1 & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \vdots & 1 \end{pmatrix} = \begin{pmatrix} t_{11} & * & 0 & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & t_{ii} & * \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \vdots & t_{nn} \end{pmatrix}, \text{ and } T' \text{ has}$$

t_{ii} as right eigenvalue. As right eigenvalues are similarity invariants (1.D), T has t_{ii} as right eigenvalue; as conjugates of right eigenvalues are right eigenvalues (1.D), α is a right eigenvalue of T .

Summarizing the discussion prior to lemma 1 we get

Proposition 2: If a semigroup $\mathcal{J} \subseteq D_n$ is upper triangularizable, then (i) the nilpotent elements (if any) of \mathcal{J} form a (semigroup) ideal of \mathcal{J} ; (ii) any subgroup of \mathcal{J} can be upper triangularized; and (iii) for all matrices $A, B \in \mathcal{J}$, $\{\text{right eigenvalues of } (AB)\} \subseteq \{ab \mid a \text{ a right eigenvalue of } A, b \text{ a right eigenvalue of } B\}$.

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 $\mathcal{J} \subseteq D_n$ can be simultaneously diagonalized, then (i) \mathcal{J} contains no non-zero nilpotent matrices; (ii) all subgroups of \mathcal{J} are diagonalizable; and (iii) all idempotents of \mathcal{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 non-singular matrix and a nilpotent matrix, say $P^{-1}MP = \begin{pmatrix} R & 0 \\ 0 & Q \end{pmatrix}$, R invertible, Q nilpotent. We set $A_M = P \begin{pmatrix} R & 0 \\ 0 & 0 \end{pmatrix} P^{-1}$, $N_M = P \begin{pmatrix} 0 & 0 \\ 0 & Q \end{pmatrix} P^{-1}$, $E_M = P \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} P^{-1}$, $A_M^\# = P \begin{pmatrix} R^{-1} & 0 \\ 0 & 0 \end{pmatrix} P^{-1}$. Immediately from the definitions we get the following

Proposition 4: The matrices A_M , $A_M^\#$, E_M , N_M just defined satisfy

$$(1) \quad E_M^2 = E_M$$

$$(2) \quad A_M A_M^\# = E_M = A_M^\# A_M$$

$$(3) \quad A_M E_M = E_M A_M = A_M$$

$$(4) \quad A_M^\# E_M = E_M A_M^\# = A_M^\#$$

$$(5) \quad E_M N_M = N_M E_M = 0$$

$$(6) \quad M = A_M + N_M$$

$$(7) \quad M^n = A_M^n$$

$$(8) \quad E_M M = A_M = M E_M.$$

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 $E_M = (A_M^\#)^n (A_M)^n = (A_M^\#)^n (M)^n = (A_M^\#)^n ('A_M)^n = ('A_M^\#)^n ('A_M)^n$, $E_M = (A_M^\#)^n (M)^n$, $E_M = E_M 'E_M$. Similarly, $'E_M = E_M 'E_M$, so $E_M = 'E_M$. Then $A_M = E_M M = 'E_M M = 'A_M$; $N_M = M - 'A_M = M - A_M = N_M$; and $'A_M^\# = 'A_M^\# E_M = 'A_M^\# A_M A_M^\# = E_M A_M^\# = A_M^\#$.

(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 $A_M^\#$ 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 $A_M^\#$ 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 ψ by taking any matrix M representing ψ and setting $N_\psi = \phi_{N_M}$, $A_\psi = \phi_{A_M}$; E_ψ and $A_\psi^\#$ 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 $A_M^\#$.

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 $j \leq n$ and

(1) for all $j', 1 \leq j' < j, m_{ij'} = 0$ if one but not both of $m_{ii}, m_{j',j'}$ are 0. We find a unitriangular matrix T such that $T^{-1}MT=M_1$ satisfies (1) with j replaced by $j+1$. For $i=1, \dots, j-1$, if one but not both of m_{ii}, m_{jj} are 0 we can solve the equation $0=m_{ii}x_i - x_i m_{jj} + m_{ij}$ for x_i ; if both or neither m_{ii}, m_{jj} are 0 let $x_i=0$. Then

$$M_1 = \begin{pmatrix} 1 & 0 & -x_1 & & 0 \\ & \ddots & \vdots & & \\ & & -x_{j-1} & & 0 \\ 0 & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix} M \begin{pmatrix} 1 & 0 & x_1 & & 0 \\ & \ddots & \vdots & & \\ & & x_{j-1} & & 0 \\ 0 & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix} \text{ has the property claimed. As}$$

our original matrix M satisfies (1) with $j=2$, a finite number of such steps gives us our desired matrix M' .

Now let i_1, \dots, i_r be (in ascending order) the indices such that the i_j, i_j entry of M' is 0, and let j_1, \dots, j_{n-r} be (in ascending order) the remaining indices. Let $\sigma \in S_n$ be the permutation $\begin{pmatrix} 1 \dots n-r & n-r+1 \dots n \\ j_1 \dots j_{n-r} & i_1 \dots i_r \end{pmatrix}$, and let $P \in GL_n(D)$ be the permutation matrix corresponding to the transformation $e_i \mapsto e_{\sigma(i)}$. Then $P^{-1}M'P = \begin{pmatrix} A & O \\ O & B \end{pmatrix}$, where A is upper triangular and non-singular and B is upper triangular and nilpotent. Then $A_{M'} = P \begin{pmatrix} A & O \\ O & O \end{pmatrix} P^{-1}$ 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), $E_{M'} = P^{-1} \begin{pmatrix} I & O \\ O & O \end{pmatrix} P$ is upper triangular, as is $A_M^\# = P^{-1} \begin{pmatrix} A^{-1} & O \\ O & O \end{pmatrix} P$, and so E_M and $A_M^\#$ will be too.

A subsemigroup \mathcal{J} of D_n will be called closed if whenever we have $M \in \mathcal{J}$, we also have N_M, A_M, E_M , and $A_M^\# \in \mathcal{J}$. 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 $\phi_{\mathcal{J}}$, \mathcal{J} a subsemigroup of D_n , and $\phi_{\mathcal{J}}$ is closed if and only if \mathcal{J} is closed. Clearly the intersection of closed semigroups will be closed, and we get a closure operation on subsemigroups of D_n ,

$\mathcal{J} \mapsto \mathcal{J}^* = \bigcap \{ \mathcal{I} \mid \mathcal{J} \subseteq \mathcal{I}, \mathcal{I} \text{ a closed subsemigroup of } D_n \}$. We can describe this closure of a semigroup of matrices explicitly; set $\mathcal{J}_0 = \mathcal{J}$, and define inductively (for $i \geq 1$) \mathcal{J}_i to be the subsemigroup of D_n generated by \mathcal{J}_{i-1} and $\{A_M, N_M, E_M, A_M^\# \mid M \in \mathcal{J}_{i-1}\}$.

Then we get

Lemma 7: With the above notation, $\mathcal{J}^* = \bigcup_{i=1}^{\infty} \mathcal{J}_i$.

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

$\mathcal{J} \subseteq \mathcal{J}^*$: Clearly $\mathcal{J}_0 \subseteq \mathcal{J}^*$, and if $\mathcal{J}_i \subseteq \mathcal{J}^*$ 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 $M \in \mathcal{J}$, then $M \in \mathcal{J}_i$ for some i , so $A_M, N_M, E_M, A_M^{\#} \in \mathcal{J}_{i+1}$. Since $\mathcal{J}_{i+1} \subseteq \mathcal{J}$, we see that \mathcal{J} is closed.

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 $P^{-1}\mathcal{J}_iP$ is upper triangular for each $i \in \mathbb{N}$. For $i=0$ this is our assumption; if $P^{-1}\mathcal{J}_{i-1}P$ is upper triangular ($i \geq 1$) then for any $M \in P^{-1}\mathcal{J}_{i-1}P$, A_M, N_M, E_M , and $A_M^{\#}$ are all upper triangular by lemma 6. Thus the generators of $P^{-1}\mathcal{J}_iP$ are upper triangular, and so $P^{-1}\mathcal{J}_iP$ 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 $P^{-1}\mathcal{J}_iP$ is diagonal for all i . For $i=0$ this is our assumption. Suppose $P^{-1}\mathcal{J}_{i-1}P$ is diagonal, $i \geq 1$, and let $M = (m_{ij}) \in P^{-1}\mathcal{J}_{i-1}P$. Then it follows from proposition 5 that

$N_M=0$; $A_M=M$; $E_M=(e_{ij})$, where $e_{ij} = \begin{cases} 0 & \text{if } i \neq j \\ 0 & \text{if } i=j \text{ and } m_{ii}=0 \\ 1 & \text{otherwise} \end{cases}$ and

$A_M^\#=(a_{ij})$, where $a_{ij} = \begin{cases} 0 & \text{if } i \neq j \\ 0 & \text{if } i=j \text{ and } m_{ii}=0 \\ m_{ii}^{-1} & \text{otherwise} \end{cases}$. These are all

diagonal, and together they generate $P^{-1} \mathcal{L}_i P$. Thus $P^{-1} \mathcal{L}_i P$ is diagonal, and the 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:

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

Proof: only if: Clear.

if: Suppose \mathcal{N} is an ideal of \mathcal{L} 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 \mathcal{L} -invariant subspace with quotient space \bar{V} then $u|_W$ and $\bar{u}_{\bar{V}}$ have upper triangular representations.

$\bar{u}_{\bar{V}}$: Suppose $B=\{v_1, \dots, v_n\}$ is a basis of V such that $\mu(u, B)$ is upper triangular for all $u \in \mathcal{U}$. Write V_i for $\langle v_1, \dots, v_i \rangle$, $i=1, \dots, n$. Let i_1 be minimal such that $\bar{V}_{i_1} \neq \bar{0}$, and once i_j has been determined, let i_{j+1} be minimal such

that $\bar{v}_{i,j+1} \neq \bar{v}_{i,j}$. Then the representation of $\bar{u}_{\bar{V}}$ with respect to the basis $\{\bar{v}_{i_1}, \dots, \bar{v}_{i_r}\}$ is upper triangular (where $r = \dim \bar{V}$).

$u|_W$: Let V_i be as above, $i=1, \dots, n$. Let j_1 be minimal such that $V_{j_1} \cap W \neq \{0\}$, and let $f_1 = v_{j_1+1} + \sum_{l=1}^{j_1-1} v_l d_{1l} \in V_{j_1} \cap W$. Once $f_i = v_{j_i+1} + \sum_{l=1}^{j_i-1} v_l d_{il}$ 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 $f_{i+1} = v_{j_{i+1}+1} + \sum_{l=1}^{j_{i+1}-1} v_l d_{i+1,l} \in V_{j_{i+1}} \cap W$. Then, ^{the} representation of $u|_W$ with respect to the basis $\{f_1, \dots, f_{n-r}\}$ 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{N} is empty or $\mathcal{N} = \{0\}$ then $\mathcal{I} = \mathcal{U}$ or $\mathcal{I} = \mathcal{U} \cup \{0\}$ and the theorem holds because u has an upper triangular representation by assumption. Thus we can assume that $0 \subsetneq \mathcal{N} \cap W = W$. By Levitzki's Theorem* ([11], p. 135) $W \neq V$, and since \mathcal{N} is an ideal, W is a non-trivial \mathcal{I} -invariant subspace.

The result will follow by induction on n if we can show that (1) non-nilpotents of $\mathcal{I}|_W, \bar{\mathcal{I}}_{\bar{V}}$ have ^{simultaneous} upper triangular representations, and (2) nilpotents of $\mathcal{I}|_W, \bar{\mathcal{I}}_{\bar{V}}$ form an ideal. (1) follows from our opening argument and the fact that the non-nilpotents of $\mathcal{I}|_W, \bar{\mathcal{I}}_{\bar{V}}$ are contained in $u|_W, \bar{u}_{\bar{V}}$ respectively. To prove (2), let $\phi|_W$ be nilpotent, and let $\psi \in \mathcal{I}$. We want to show that $\psi\phi|_W$ and $\phi\psi|_W$ are nilpotent. If either of ϕ, ψ come from \mathcal{N} , this follows from the fact that \mathcal{N} 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.

$\phi, \psi \in \mathcal{U}$. By our opening observation and our assumption on \mathcal{U} , $\phi|_W$ and $\psi|_W$ have a representation as upper triangular matrices. Since $\phi|_W$ is nilpotent, its corresponding matrix has 0's on the main diagonal, and so the matrices corresponding to $\psi\phi|_W$ and $\phi\psi|_W$ are nilpotent. The proof of (2) for $\overline{\mathcal{X}}_{\overline{V}}$ is similar.

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

Corollary 10.1: If \mathcal{L} is a subsemigroup of D_n whose nilpotent elements form an ideal and whose non-nilpotent elements commute then \mathcal{L} 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, \mathcal{L} a subsemigroup of F_n consisting only of unipotent and nilpotent matrices. Then \mathcal{L} can be upper triangularized.

Proof: First we note that the nilpotents of \mathcal{L} form an ideal of \mathcal{L} : if $N \in \mathcal{L}$ is nilpotent and $S \in \mathcal{L}$ then $NS \in \mathcal{L}$ is singular. Thus NS cannot be unipotent (unipotent matrices are non-singular), and since every element of \mathcal{L} is either unipotent or nilpotent, NS must be nilpotent. Similarly SN is nilpotent, so the nilpotents of \mathcal{L} form an ideal.

Also, the unipotents form a subsemigroup, as unipotents are units in F_n , so no product of unipotents in \mathcal{L} 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 \mathcal{J} is a closed semigroup of linear transformations of $V=D^n$ and W is an \mathcal{J} -invariant subspace, then $\mathcal{J}|_W$ and $\overline{\mathcal{J}}|_W$ are closed.

Proof: Let $\phi \in \mathcal{J}$; A_ϕ , E_ϕ , N_ϕ , and A_ϕ^\pm are all in \mathcal{J} (since \mathcal{J} is closed), and satisfy (1)-(8) of proposition 4. Thus $A_\phi|_W$, $E_\phi|_W$, $N_\phi|_W$, and $A_\phi^\pm|_W$ are all in $\mathcal{J}|_W$ 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 $\mathcal{J}|_W$ is closed. Similarly for $\overline{\mathcal{J}}|_W$.

We can now state and prove

Theorem 12: Let $\mathcal{J} \subseteq D_n$ be a closed semigroup (cf. p. 44).

Then \mathcal{J} can be upper triangularized if and only if the following conditions are satisfied: (i) the nilpotents of \mathcal{J} form a (semigroup) ideal of \mathcal{J} ; (ii) every subgroup of \mathcal{J} can be upper triangularized; (iii) for any idempotents $A, B \in \mathcal{J}$, $\{\text{right eigenvalues of } (AB)\} \subseteq \{0, 1\}$.

Proof: only if: By proposition 2.

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

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

(iii) $\phi_{\mathcal{J}}|_W$: If $\phi_M|_W$ is idempotent, $\phi_M|_W = E_{\phi_M}|_W = \phi_{E_M}|_W$, and $\phi_{E_M}|_W \in \phi_{\mathcal{J}}|_W$. Thus if $\phi_M|_W$, $\phi_{M_0}|_W$ are idempotent, $\{\text{right eigenvalues of } \phi_M|_W \phi_{M_0}|_W\} =$

{right eigenvalues of $\phi_{E_M}|_W \phi_{E_{M_0}}|_W$ } = {right eigenvalues of $\phi_{E_M E_{M_0}}|_W$ } \subseteq {right eigenvalues of $\phi_{E_M E_{M_0}}$ } \subseteq {0, 1}. Thus any representations of $\phi_{\mathcal{L}}|_W$ satisfy (iii). A similar proof works for $\overline{\phi_{\mathcal{L}}}$.

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

We first need to make two observations about idempotent matrices; they follow immediately if we take a representation of ϕ_E in the form $\begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}$:

(a) if $B \in D_S$ and $B = EBE$ for an idempotent matrix E then

$$E_B = EE_B E;$$

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

We return to the task of finding a group \mathcal{H} as described above. By the proof that $\phi_{\mathcal{L}}|_W$ satisfies (iii), there is an idempotent $E \in \mathcal{L}$ such that $\phi_E|_W = 1_{\mathcal{G}}$; pick E to be such an idempotent of minimal rank, and let \mathcal{H} be the group of units of $E\mathcal{L}E$. Let $G_0 \in \mathcal{G}$, and let $G \in \mathcal{L}$ be such that $\phi_G|_W = G_0$. Then $\phi_{EGE}|_W = G_0$; set $H = EGE$. Now $\phi_H|_W \in \mathcal{G}$, so $\exists K \in \mathcal{L}$ with $\phi_H|_W \phi_K|_W =$

$1_{\mathcal{G}} = \phi_E|_W$. Then $\phi_{H^n}|_W \phi_{K^n}|_W = \phi_E|_W$, and by proposition 4 (7) and (3) $\phi_{E_H}|_W \phi_E|_W = \phi_{E_H}|_W \phi_{H^n}|_W \phi_{K^n}|_W = \phi_{H^n}|_W \phi_{K^n}|_W = \phi_E|_W$. But by

(a), $\phi_{E_H} = \phi_{E_H} \phi_E$, so we see that $\phi_{E_H}|_W = \phi_E|_W$. Now $\text{rank}(E_H) \leq \text{rank } H = \text{rank}(HE) \leq \text{rank } E$; by the minimality of rank E ,

$\text{rank}(E_H) = \text{rank } E$. Also, by (a), $E_H = EE_H E$, so by (b) $E = E_H$.

Then $HA_H^{\#} = E_H = E = A_H^{\#}H$, so $H \in \mathcal{H}$, and the proof is complete.

A similar argument works for $\overline{\phi_{\mathcal{L}}}$.

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 μ of $\phi_{\mathcal{J}}$ such that $\mu(\phi_E) = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$. Let $S \in E\mathcal{J}E$ be non-nilpotent, so $E_S \neq 0$. By (a), $\text{rank } E_S \leq \text{rank } E$, so by the minimality of rank E and (a) and (b) above, $E_S = E$. As above, it follows that S is a unit of $E\mathcal{J}E$. Then $E\mathcal{J}E \subseteq \mathcal{J}$ 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, $E\mathcal{J}E$ can be upper triangularized; it follows that we can find a representation μ' of $\phi_{\mathcal{J}}$ such that $\mu'(\phi_E) = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$ and, for $T \in \mathcal{J}$, $\mu'(\phi_T) = \begin{pmatrix} T_1 & T_2 \\ T_3 & T_4 \end{pmatrix}$, $T_1 \in D_r$, T_1 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 $0 \neq W \neq V$, and W is a $\phi_{\mathcal{J}}$ -invariant subspace, so we are done.

We thus assume $\mu'(\phi_E) = e_{11}$. Let $\mu'(\phi_X) = (x_{ij})$ 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

$$\mu'(\phi_{XE}) = \begin{pmatrix} x_{11} & & \\ \vdots & & \\ x_{n1} & & \end{pmatrix} \begin{pmatrix} 1 & & \\ & 0 & \\ & & 0 \end{pmatrix} \in \mu'(\phi_{\mathcal{J}}). \text{ Since } \phi_{XE} \text{ has rank 1, either } \phi_{XE}$$

is nilpotent or $\phi_{XE} = A_{\phi_{XE}}$.

Case 1: ϕ_{XE} is nilpotent; then $x_{11} = 0$. Since $0 \neq \begin{pmatrix} x_{21} \\ \vdots \\ x_{n1} \end{pmatrix} \in D^{n-1}$, $\exists P \in GL_{n-1}(D)$ such that $P^{-1} \begin{pmatrix} x_{21} \\ \vdots \\ x_{n1} \end{pmatrix} = \begin{pmatrix} 1 \\ \vdots \\ 0 \end{pmatrix}$; we take the representation

$\mu^2(\phi_g) = (1 + P^{-1})\mu'(\phi_g)(1 + P)$. Then $\mu^2(\phi_E) = e_{11}$, $\mu^2(\phi_{XE}) = e_{21}$.

Let $Y = (y_{ij}) \in \mu^2(\phi_g)$. Then $\mu^2(\phi_E)Y = \begin{pmatrix} y_{11} \cdots y_{1n} \\ \circ \end{pmatrix} \in \mu^2(\phi_g)$, and

$\mu^2(\phi_E)Y\mu^2(\phi_{XE}) = \begin{pmatrix} y_{12} & 0 & \cdots & 0 \\ \circ & & & \end{pmatrix} \in \mu^2(\phi_g)$. Since XE is nilpotent

and the nilpotents of \mathcal{g} 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 $\mu^2(\phi_g)$ has 1,2 entry 0, and we get a non-trivial ϕ_g -invariant subspace as before.

Case 2: $\phi_{XE} = A_{\phi_{XE}}$, so $x_{11} \neq 0$. Now by (b), $E_{XE} = E_{XE}E$, so $\mu'(\phi_{E_{XE}}) = \mu'(\phi_{E_{XE}})\mu'(\phi_E) =$ (say) $\begin{pmatrix} f_{11} & & \\ \vdots & & \\ f_{n1} & & \circ \end{pmatrix}$. Since $\mu'(\phi_{E_{XE}})$ is

idempotent and non-zero, $f_{11} = 1$; since $E_{XE}XE = XE$, $f_{i1}x_{11} = x_{i1}$; then because $x_{i1} \neq 0$ for some $i > 1$, $f_{i1} \neq 0$ for some $i > 1$. Again, taking an appropriate choice of basis, we get a representation μ^2 of ϕ_g such that $\mu^2(\phi_E) = e_{11}$, $\mu^2(\phi_{E_{XE}}) = e_{11} + e_{21}$. Now let $\mu^2(\phi_Y) = (y_{ij}) \in \mu^2(\phi_g)$. We will again show that $y_{12} = 0$, and so get a non-trivial ϕ_g -invariant subspace as before.

Now $\mu^2(\phi_{EY}) = \begin{pmatrix} y_{11} \cdots y_{1n} \\ \circ \end{pmatrix}$; if $EY = 0$, $y_{12} = 0$, so we need only

consider the case where EY has rank 1. Again, either EY is nilpotent or $EY = A_{EY}$.

Case a: EY is nilpotent. Then $y_{11} = 0$; also

$\mu^2(\phi_{E_{XE}})\mu^2(\phi_{EY}) = \begin{pmatrix} 0 & y_{12} \cdots y_{1n} \\ 0 & y_{12} \cdots y_{1n} \\ \circ & \end{pmatrix}$ must be nilpotent by assumption (i), so $y_{12} = 0$.

Case b: $EY = A_{EY}$, and $y_{11} \neq 0$. Again, $\mu^2(\phi_{E_{EY}})$ is of the form $\begin{pmatrix} a_{11} \cdots a_{1n} \\ \circ \end{pmatrix}$, and $a_{11} = 1$. Further, if $y_{12} \neq 0$, we conclude

from the equation $EY\mu_{EY} = EY$ that $a_{12} \neq 0$. Thus, to show $y_{12} = 0$

it is enough to show $a_{12} = 0$. Let $H = \mu^2(\phi_{E_{XE}})\mu^2(\phi_{E_{EY}}) =$

$$(e_{11} + e_{21}) \begin{pmatrix} 1 & a_{12} \cdots a_{1n} \\ & \bigcirc \end{pmatrix} = \begin{pmatrix} 1 & a_{12} \cdots a_{1n} \\ 1 & a_{12} \cdots a_{1n} \\ & \bigcirc \end{pmatrix}. \quad H \text{ is a product of}$$

idempotents, so by (iii) the only right eigenvalues of H are 0 and/or 1. Now $H \begin{pmatrix} 1 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} (1 + a_{12})$, 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 \mathfrak{J} form an ideal (by (i)), $\mu^2(\phi_E)H =$

$$e_{11} \begin{pmatrix} 1 & a_{12} \cdots a_{1n} \\ 1 & a_{12} \cdots a_{1n} \\ & \bigcirc \end{pmatrix} = \begin{pmatrix} 1 & a_{12} \cdots a_{1n} \\ & \bigcirc \end{pmatrix} \text{ must be nilpotent, which it}$$

clearly is not. Thus $\underline{1 + a_{12} \neq 0}$, and the proof is complete.

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

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

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

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

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

Proof: only if: Proposition 3.

if: Let \mathfrak{J} be a closed subsemigroup of D_n satisfying (i), (ii), (iii). If W is a $\phi_{\mathfrak{J}}$ -invariant subspace then $\phi_{\mathfrak{J}}|_W$

will be closed and satisfy (i), (ii), (iii) (the proofs are similar to those in theorem 12). If $E \in \mathfrak{J}$ is an idempotent of minimal positive rank, we take a representation μ of $\phi_{\mathfrak{J}}$ such that $\mu(\phi_E) = \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}$. Let $S \in \mathfrak{J}$ 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 $\phi_{\mathfrak{J}}$ leaves EV and $(I-E)V$ invariant.* The minimality of rank E allows us to show that $E\mathfrak{J}E$ is a group (with zero), and so can be diagonalized by assumption. On the subspace $(I-E)V \subseteq V$ left invariant by $\phi_{\mathfrak{J}}$ we can get a diagonal representation of the restrictions by induction on n ; combining, we get a diagonal representation for $\phi_{\mathfrak{J}}$.

* Note we do not claim $I-E \in \mathfrak{J}$, only that $\phi_{\mathfrak{J}}$ 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 $\{g_1 g_2 g_1^{-1} g_2^{-1} \mid g_1, g_2 \in G\}$, is a normal subgroup of G called the derived group. Once $G^{(i)}$ 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 $Q \subseteq D_n$ 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 $\phi_{\mathcal{G}}$ 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 $\phi_{\mathcal{G}}$ leaves no non-trivial invariant subspaces.

What we now prove is that if $\phi_{\mathcal{G}}$ is irreducible then $\mathcal{G}=\{I\}$ and $n=1$; this will complete the proof of the theorem. Assume $\phi_{\mathcal{G}}$ irreducible, and let \mathcal{A} be any abelian normal subgroup of \mathcal{G} . By (2.1) we can put \mathcal{A} in upper triangular form, so the matrices in \mathcal{A} have a common right eigenvector. Let w be any common right eigenvector of the matrices in \mathcal{A} , let $G \in \mathcal{G}$, and let $M \in \mathcal{A}$. We now show that Gw is a right eigenvector of M --i.e., \mathcal{G} maps common right eigenvectors of \mathcal{A} into common right eigenvectors of \mathcal{A} . Since \mathcal{A} is normal, $MG=GM'$ for some $M' \in \mathcal{A}$. 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 \mathcal{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 \mathcal{A} is invariant under $\phi_{\mathcal{G}}$, and the subspace W they generate will be a $\phi_{\mathcal{G}}$ -invariant subspace. We saw that $W \neq 0$; by our assumption on the irreducibility of \mathcal{G} we conclude that $W=D^n$. It follows that the matrices in \mathcal{A} can be simultaneously diagonalized. But the only diagonal unipotent matrix is I , so $\mathcal{A}=\{I\}$. Thus the only abelian normal subgroup of \mathcal{G} is $\{I\}$; since \mathcal{G} is solvable it follows that $\mathcal{G}=\{I\}$; then since \mathcal{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 $\mathcal{G} \subseteq D_n$ 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 $\{1\} = \mathcal{Z}_0(G) \triangleleft \mathcal{Z}_1(G) \triangleleft \dots \triangleleft \mathcal{Z}_i(G) \triangleleft \dots$, where for $i > 0$ $\mathcal{Z}_i(G) = \{x \in G \mid \text{for all } y \in G, xyx^{-1}y^{-1} \in \mathcal{Z}_{i-1}(G)\}$. G is called nilpotent if there is an integer m such that $\mathcal{Z}_m(G) = G$; if $\mathcal{Z}_{m-1}(G) \neq G$, 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 $\mathcal{G} \subseteq D_n$ be a finitely generated nilpotent group of algebraic matrices. Then \mathcal{G} has a subgroup \mathcal{H} of finite index which can be upper triangularized.

Proof: We use induction on n to show that it suffices to take $\phi_{\mathcal{G}}$ irreducible. If $n=1$ the result is trivial, so we assume $n > 1$ and the proposition is true for $n' < n$. Suppose

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

We now use induction on the class of nilpotency m of q . If $m \leq 1$ 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 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 $(q:H)$ will be finite, and the proposition proved.

We must first show that we can assume that $\mathcal{Z}_1(q)$ consists of scalar matrices with diagonal entries in the center of D ; suppose $\mathcal{Z}_1(q) = \langle M_1, \dots, M_s \rangle$, where M_1, \dots, M_{r-1} are central scalar matrices ($1 \leq r \leq s$). Suppose $P^{-1}M_rP$ is in

normal form; since M_r is central in \mathfrak{g} , the right eigenvectors of M_r generate a $\phi_{\mathfrak{g}}$ -invariant subspace, so since $\phi_{\mathfrak{g}}$ is irreducible, we conclude that $P^{-1}M_rP$ is diagonal. Now non-conjugate diagonal entries of $P^{-1}M_rP$ give rise to non-trivial $\phi_{\mathfrak{g}}$ -invariant subspaces, and we see (from our definition of normal form) that $P^{-1}M_rP$ is scalar, say $P^{-1}M_rP = \alpha \cdot I$. The centrality of M_r in \mathfrak{g} means that the matrices $P^{-1}MP$, $M \in \mathfrak{g}$, have their entries in the skew subfield of D centralizing α , 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 $\mathfrak{J}_1(\mathfrak{g})$ consists of central scalar matrices.

Let $A \in \mathfrak{J}_2(\mathfrak{g})$; we now show that $(\mathfrak{g}:C_{\mathfrak{g}}(A)) < \infty$. We do this by showing A has only finitely many conjugates in \mathfrak{g} . Let $B \in \mathfrak{g}$; since $A \in \mathfrak{J}_2(\mathfrak{g})$, $B^{-1}AB = A\lambda_B \cdot I$, where $\lambda_B \cdot I \in \mathfrak{J}_1(\mathfrak{g})$ is a central scalar matrix. Let μ be a right eigenvalue of A , with right eigenvector v . Then $A(Bv) = BA(\lambda_B \cdot Iv) =$ (since λ_B is central) $BAv\lambda_B = Bv(\mu\lambda_B)$, 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 α of A , $p(\alpha) = 0$. In particular, $p(\mu\lambda_B) = 0$ for every $B \in \mathfrak{g}$. Thus $\{\mu\lambda_B | B \in \mathfrak{g}\} \subseteq \{\text{roots of } p(t) \text{ in } k(\mu)\}$; since $k(\mu)$ is commutative, $p(t)$ has at most degree p roots in $k(\mu)$, so $\{\mu\lambda_B | B \in \mathfrak{g}\}$ is finite. As $\mu \neq 0$ (A , being in \mathfrak{g} , is invertible), $\{\lambda_B | B \in \mathfrak{g}\}$ is finite. But all conjugates of A are of the form $A\lambda_B \cdot I$, so A has only finitely many conjugates.

Now $\mathfrak{J}_2(\mathfrak{g})$ is finitely generated, say $\mathfrak{J}_2(\mathfrak{g}) = \langle M_1, \dots, M_t \rangle$. Then $C_{\mathfrak{g}}(\mathfrak{J}_2(\mathfrak{g})) = \bigcap_{i=1}^t C_{\mathfrak{g}}(M_i)$, and since each $C_{\mathfrak{g}}(M_i)$ has finite

index in \mathfrak{g} , it follows that $(\mathfrak{g} : C_{\mathfrak{g}}(\mathfrak{J}_2(\mathfrak{g}))) < \infty$. Recall we were looking for a nilpotent subgroup \mathcal{H}_1 of \mathfrak{g} of class at most $m-1$; we take $\mathcal{H}_1 = C_{\mathfrak{g}}(\mathfrak{J}_2(\mathfrak{g}))$. Then

$\mathfrak{J}_2(\mathfrak{g}) \cap \mathcal{H}_1 = \mathfrak{J}_2(\mathfrak{g}) \cap C_{\mathfrak{g}}(\mathfrak{J}_2(\mathfrak{g}))$ is central in $C_{\mathfrak{g}}(\mathfrak{J}_2(\mathfrak{g})) = \mathcal{H}_1$, so $\mathfrak{J}_2(\mathfrak{g}) \cap \mathcal{H}_1 \subseteq \mathfrak{J}_1(\mathcal{H}_1)$. We claim that for $i \geq 1$, $\mathfrak{J}_{i+1}(\mathfrak{g}) \cap \mathcal{H}_1 \subseteq \mathfrak{J}_i(\mathcal{H}_1)$; this has been established for $i=1$, so assume it is true for $i-1 \geq 1$. Then $\mathfrak{J}_i(\mathcal{H}_1) = \{x \in \mathcal{H}_1 \mid \forall y \in \mathcal{H}_1, xyx^{-1}y^{-1} \in \mathfrak{J}_{i-1}(\mathcal{H}_1)\} \supseteq \{x \in \mathcal{H}_1 \mid \forall y \in \mathcal{H}_1, xyx^{-1}y^{-1} \in \mathcal{H}_1 \cap \mathfrak{J}_i(\mathfrak{g})\} \supseteq \mathcal{H}_1 \cap \{x \in \mathfrak{g} \mid \forall y \in \mathfrak{g}, xyx^{-1}y^{-1} \in \mathfrak{J}_i(\mathfrak{g})\} = \mathcal{H}_1 \cap \mathfrak{J}_{i+1}(\mathfrak{g})$. It follows that $\mathfrak{J}_{m-1}(\mathcal{H}_1) \supseteq \mathcal{H}_1 \cap \mathfrak{J}_m(\mathfrak{g}) = \mathcal{H}_1$, so \mathcal{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 $\{M_1, \dots, M_t\} \subseteq \mathfrak{J}_i(\mathfrak{g})$ such that $C_{\mathfrak{g}}(\mathfrak{J}_i(\mathfrak{g})) = C_{\mathfrak{g}}(\{M_1, \dots, M_t\})$, could be dropped, as any maximal (necessarily finite) D -linearly independent subset of $\mathfrak{J}_i(\mathfrak{g})$ 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 \mathfrak{G} such that any subgroup of D_n isomorphic to \mathfrak{G} can be simultaneously diag-

onalized. We seek in the rest of this chapter to determine such finite groups Q ; we discover that for some skew fields there are such non-abelian groups.

In general, if Q is a finite group of diagonal $n \times n$ matrices over a skew field D , then Q is embeddable in a direct product of n copies of the multiplicative group $D^* = D - \{0\}$, and the projection of Q 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 Q are abelian.

Thus we assume for the rest of our considerations that D has characteristic 0, and so \mathbb{Q} (the rational number field) is contained in the center k of D . If Q is a finite group, of order m , say, then we can embed the group ring $\mathbb{Q}Q$ in \mathbb{Q}_m by the regular representation. We denote by $\{M_G | G \in Q\}$ the matrices of this representation corresponding to the elements of Q . We thus have $\mathbb{Q}Q = \mathbb{Q}[M_G | G \in Q] \subseteq D_m$ for every skew field D of characteristic 0, and clearly the group of matrices $\{M_G | G \in Q\} \cong Q$ can be diagonalized if and only if $\mathbb{Q}Q$ can be. If Q has a non-normal subgroup H , then $\frac{1}{|H|} \sum_{h \in H} M_h$ is a non-central idempotent of $\mathbb{Q}Q$, and it follows that $\mathbb{Q}Q$ cannot be diagonalized.

Thus the only possible non-abelian groups Q of finite order with all representations over D diagonalizable have all their subgroups normal. Such groups are called Hamiltonian, and have the form $H \times O \times T$, where H is the quaternion group of order 8, O an abelian group of odd order, and T

a group of exponent 2 ([7], p. 190).

We consider now a case where a group of the form $H \times \mathcal{O}$, 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 $\rho = \begin{pmatrix} A & \cdot & O \\ O & \cdot & A \end{pmatrix}$, $\eta = \begin{pmatrix} B & \cdot & O \\ O & \cdot & B \end{pmatrix}$, $\eta = \begin{pmatrix} O & I_4 & 0 & \dots & 0 \\ \vdots & 0 & I_4 & \dots & \vdots \\ 0 & \vdots & \vdots & I_4 & \vdots \\ I_4 & 0 & \dots & \vdots & 0 \end{pmatrix}$, where

$$A = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{pmatrix}. \quad \text{If } \mathcal{G} = \langle \rho, \eta \rangle, \text{ then } \mathcal{G} \cong H \times C_n,$$

where C_n is cyclic of order n . Further, ρ and η have as right eigenvalues only primitive fourth roots of 1, and η has a primitive n^{th} root of 1 as right eigenvalue. Thus if \mathcal{G} can be simultaneously diagonalized, some diagonal entry, call it n_{ii} , of the matrix corresponding to η will be a primitive n^{th} root of 1. If we denote by i_{ii} , j_{ii} the corresponding diagonal entries of the matrices corresponding to ρ , η respectively, then from the above observations and the multiplication table of \mathcal{G} it follows that $\langle i_{ii}, j_{ii}, n_{ii} \rangle \cong H \times C_n$. Thus if \mathcal{G} can be diagonalized $H \times C_n$ 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 3: Let D be a skew field of characteristic 0,

and denote by k its center. Let $H \times \mathcal{O} \times \mathcal{T}$ (as above) be a

*In Amitsur's notation, $H \times C_n = G_{4n, 2n+1}$; theorem 4(2b) is applicable.

Hamiltonian group, where 2 has odd order modulo the exponent e of \mathcal{O} . If there is a skew field E containing k in its center and containing a subgroup isomorphic to $C_e \times \mathcal{H}$, then any group \mathcal{G} of $n \times n$ matrices isomorphic to $\mathcal{H} \times \mathcal{O} \times \mathcal{T}$ 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 $W \subseteq D^n$ is a $\phi_{\mathcal{G}}$ -invariant subspace, $\phi_{\mathcal{G}}|_W$ 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 $\phi_{\mathcal{G}}$ is irreducible.

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

By (1.A) we may assume $E \subseteq D$. Let $\{\omega, i, j\} \subseteq D$ generate a group isomorphic to $\mathcal{H} \times C_e$, with $\langle \omega \rangle \cong C_e$ and $\langle i, j \rangle \cong \mathcal{H}$. Let $M \in \mathcal{G}$ have order e . We can diagonalize M to get a matrix M' whose diagonal entries come from the field $Q(\omega)$, and we can take conjugate diagonal entries equal. As M is in the center of \mathcal{G} our assumption on proper invariant subspaces allows us to assume M' is the scalar matrix $\omega \cdot I$. Denote by \mathcal{G}' the

corresponding representation of ϕ_q . The centrality of M' in \mathfrak{q}' means that \mathfrak{q}' is contained in the matrix ring over the skew subfield D_0 of D centralizing ω , so now we restrict ourselves to considering matrices over D_0 . If $M_0 \in \mathfrak{q}$ 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 ω . Now distinct powers of ω will be non-conjugate over D_0 , and by the centrality of M_0 and the irreducibility of ϕ_q we conclude that M''_0 is scalar (hence $M''_0 = M'_0$). Thus all matrices in \mathfrak{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 $Q \cong H$ and the unique element of Q of order 2 is $-I$, or equivalently, of showing that if $n > 1$ ϕ_q is reducible. Let $\mathcal{I}, \mathcal{J} \in \mathfrak{q}$ be generators satisfying $\mathcal{I}^2 = -I = \mathcal{J}^2$, $\mathcal{I}\mathcal{J} = -\mathcal{J}\mathcal{I}$. We first show that $k(\omega)[\mathcal{I}, \mathcal{J}] \cong k(\omega, i, j)$. Any element in $k(\omega)[\mathcal{I}, \mathcal{J}]$ can be written in the form $a \cdot I + b \cdot \mathcal{I} + c \cdot \mathcal{J} + d \cdot \mathcal{I}\mathcal{J}$, $\{a, b, c, d\} \subseteq k(\omega)$, and any element of $k(\omega, i, j)$ can be expressed $a \cdot 1 + b \cdot i + c \cdot j + d \cdot ij$, $\{a, b, c, d\} \subseteq k(\omega)$. The map $a \cdot 1 + b \cdot i + c \cdot j + d \cdot ij \mapsto a \cdot I + b \cdot \mathcal{I} + c \cdot \mathcal{J} + d \cdot \mathcal{I}\mathcal{J}$ is clearly a k -algebra homomorphism onto, and as the domain is simple, it is an isomorphism. Thus $k(\omega)[\mathcal{I}, \mathcal{J}]$ 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 \mathfrak{q} , is reducible. _____

In particular, it follows from Amitsur's work ([1]) that if $k=Q$ then any group of the sort described in proposition 3 can be diagonalized. More generally, this will

be true if $k(\omega)$, ω a primitive e^{th} root of 1, is any commutative 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 quaternion algebra will be a skew field, the inverse of a non-zero element $a \cdot 1 + b \cdot i + c \cdot j + d \cdot ij$ being $(a^2 + b^2 + c^2 + d^2)^{-1}(a \cdot 1 - b \cdot i - c \cdot j - d \cdot ij)$.

The first part of the document discusses the importance of maintaining accurate records of all transactions. It emphasizes that proper record-keeping is essential for the integrity of the financial system and for the ability to detect and prevent fraud. The text also highlights the need for transparency and accountability in all financial activities.

The second part of the document outlines the specific procedures and controls that should be implemented to ensure the accuracy and reliability of financial data. This includes the use of standardized accounting practices, the implementation of internal controls, and the regular review and audit of financial records.

OPEN QUESTIONS

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

1. Amitsur, S. A. "Finite Subgroups of Division Rings",
Transactions of the American Math. Soc., vol. 80,
pp. 361-386.
2. Bell, E. T. The Development of Mathematics, McGraw-Hill
(London/N.Y.C.) 1945.
3. Cohn, P. M. "The Embedding of FIRs in Skew Fields",
Proceedings of the London Math. Soc., vol. 23,
pp. 193-213.
4. Cohn, P. M. Free Rings and Their Relations, Academic
Press (London/N.Y.C.) 1971.
5. Cohn, P. M. "The Similarity Reduction of Matrices over
a Skew Field", Math. Zeit., vol. 132, pp. 151-163.
6. Cohn, P. M. "The Range of Derivations on a Skew Field
and the Equation $ax-xb=c$ ", Journal of the Indian
Math. Soc., vol. 37, pp. 61-69.
7. Hall, Marshall, Jr. The Theory of Groups, Macmillan &
Co. (N.Y.C.) 1959.
8. Herstein, I. N. "Finite Multiplicative Subgroups in
Division Rings", Pacific Jour. Math., vol. 1,
pp. 121-126.
9. Herstein, I. N. Non-commutative Rings (A.M.A. Carus
Monographs) John Wiley & Sons, Inc. (N.Y.C.)
1968.
10. Jacobson, N. Lectures in Abstract Algebra III, Van
Nostrand (London/Princeton, N.J.) 1964.
11. Kaplansky, I. Fields and Rings, University of Chicago
Press (Chicago) 1972.

12. Kegel, O., and B. A. F. Wehrfritz Locally Finite Groups, North Holland (Amsterdam) 1973.
13. Kimura, Naoki "Maximal Subgroups of a Semigroup", Kōdai Math. Sem. Rep., 1954, pp. 85-88.
14. Robinson, D. J. S. Finiteness Conditions and Generalized Soluble Groups I, Springer-Verlag (Berlin/N.Y.C.) 1972.
15. Scott, W. R. Group Theory, Prentice Hall, Inc. (London/Englewood Cliffs, N.J.) 1964.
16. Suprunenko, D. A., and R. I. Tyshkevich Commutative Matrices, Academic Press (London/N.Y.C.) 1968.
17. Zalesskii, A. E. "The Structure of Certain Classes of Matrix Groups over a Division Ring", Siberian Math. Jour., vol. 8, pp. 978-988.