A GENERALIZATION OF A COMMUTATOR THEOREM OF MIKUSINSKI

E. C. PAIGE

1. Introduction. In a series of papers [4]–[6] Mikusinski and Sikorski considered the following problem. Let $V$ be a vector space over a field of characteristic 0 and let $D$ be a locally algebraic linear transformation of $V$ (i.e., given any $x$ in $V$ there is a polynomial $f(\lambda) \neq 0$ over $F$ with $xf(D) = 0$). If $A = F[\lambda]$ is the polynomial ring in one variable over $F$, $V$ becomes an $A$-module under the definition $xf(\lambda) = xf(D)$ for $x$ in $V$, $f(\lambda)$ in $A$. The Mikusinski-Sikorski hypotheses on $V$ and $D$ can be phrased as follows.

I. If $f(\lambda) \in A$ has degree $n \geq 1$, the kernel of $f(D)$ has dimension $\leq n$.

II. If $f(\lambda)$, $g(\lambda)$ in $A$ have positive degrees and if the dimensions of $\text{Ker } f(D)$ and $\text{Ker } g(D)$ are $m$ and $n$ respectively, the dimension of $\text{Ker } f(D)g(D)$ is $m + n$.

Mikusinski and Sikorski [5], [6] then proved the

Theorem. If $D$ is a locally algebraic linear transformation of $V$ satisfying I and II, there is a linear transformation $T$ of $V$ with $TD - DT = I$, the identity transformation of $V$.

Mikusinski [4] also demonstrated a converse; namely he proved the

Theorem. If $D$ is a locally algebraic linear transformation of $V$ satisfying condition I and if there is a linear transformation $T$ of $V$ with $TD - DT = I$ then condition II is satisfied.

The generalizations treated in this paper may be formulated as follows. Let $D$ be a locally algebraic linear transformation of $V$; instead of the conditions listed above, the assumption will be

III. $V$ is a divisible $A$-module (i.e., given $y$ in $V$ and $f(\lambda) \neq 0$ in $A = F[\lambda]$ there is an $x$ in $V$ with $xf(\lambda) = xf(D) = y$).

The first theorem may be stated as

Theorem 1. If $D$ is a locally algebraic linear transformation of $V$ satisfying condition III, then a linear transformation, $T$, of $V$ exists with $TD - DT = I$.

The converse result established is

Theorem 2. If $D$ is a locally algebraic linear transformation on $V$
over $F$ of characteristic 0 and if a linear transformation, $T$, of $V$ exists with $TD - DT = I$, then condition III is satisfied.

The characteristic 0 hypothesis cannot be omitted in Theorem 2 as will be shown by an example due to A. A. Albert. It will be shown that these results imply a generalization of the Mikusinski-Sikorski results and that this generalization implies one obtained by Mr. James Geer in a Master's thesis [2] at the University of Virginia. The author expresses his appreciation to Professor M. Rosenblum for calling this problem to his attention.

2. Sufficiency of the condition III. As usual the $A$-module $V$ will be termed a primary $A$-module if there is an irreducible element $p = p(\lambda)$ of $A$ such that every element of $V$ is in the kernel of $(p(D))^k$ for some $k$. The following lemmas are well known [3] but are included for convenience.

Lemma 1. If $D$ is a locally algebraic linear transformation of $V$, then $V$ is the (weak) direct sum of primary $A$-modules.

Lemma 2. A direct sum of $A$-modules is divisible if and only if each summand is divisible.

Now if $V_p$ is a primary component of $V$ and if $T_p$ is a linear transformation on $V_p$ satisfying $[T_p, D] = T_pD - DT_p = I$ on $V_p$, then the direct sum $T = \sum_p T_p$ of the $T_p$ for $p$ ranging over the irreducible polynomials of $A$ clearly satisfies $[T, D] = I$ on $V$.

The previous remark and Lemmas 1 and 2 clearly reduce the problem to the case in which $V$ is a primary divisible $A$-module for the prime $p$ of $A$, and this hypothesis is maintained for the remainder of this section.

For each integer $k \geq 1$ let

$$V_k = \{ x \in V : x(p(D))^k = 0 \}$$

so that $V_k$ is the kernel of $(p(D))^k$, is an $A$-submodule of $V$ (i.e., is a $D$-invariant subspace of $V$), and satisfies $V_k \subseteq V_{k+1}$, $V_{k+1}(p(D)) \subseteq V_k$, and $\bigcup_{k=1}^\infty V_k = V$.

Lemma 3. If $V$ is a primary divisible $A$-module for the prime polynomial $p(\lambda)$ of degree $m \geq 1$, there is a basis $\{ x(\alpha, k)D^j \}_{a,j,k}$ of $V$ where $\alpha$ ranges over some index set, and $0 \leq j \leq m - 1$ and $k \geq 1$ is an integer.

Since $x p = x p(D) = 0$ for all $x \in V_1$, $V_1$ is a vector space over the field $K = F[\lambda]/(p(\lambda))$ and as such has a basis $\{ x_\alpha \}_\alpha$ over $K$. Now $1, \lambda, \cdots, \lambda^{m+1}$ modulo $p(\lambda)$ form a basis for $K$ over $F$ and so well known
vector space arguments show \( \{x_\alpha \lambda^j\}_{\alpha, j} \) to be a basis of \( V_1 \) over \( F \). Now \( x_\alpha \lambda^j = x_\alpha D^j \) and so \( \{x_\alpha D^j\}_{\alpha, j} \) is a basis of \( V_1 \) over \( F \). To simplify the notation write \( x_\alpha = x_\alpha (\alpha, 1) \) and choose inductively (by the divisibility hypothesis) \( x(\alpha, k+1) \) in \( V_{k+1} \) with \( x(\alpha, k+1) p(D) = x(\alpha, k) \).

The vectors \( \{x(\alpha, k) D^j\}_{\alpha, j, k} \) are linearly independent over \( F \). For if

\[
\sum_{\alpha} \sum_{k=1}^{n+1} \sum_{j=0}^{m-1} \beta(\alpha, k, j) x(\alpha, k) D^j = 0
\]

apply \( p(D)^n \) to (2) to obtain

\[
\sum_{\alpha} \sum_{j=0}^{m-1} \beta(\alpha, n+1, j) x(\alpha, 1) D^j = 0.
\]

By the choice of \( x(\alpha, 1) \), relation (3) yields \( \beta(\alpha, n+1, j) = 0 \) and an obvious induction establishes that all \( \beta(\alpha, k, j) = 0 \). To see that the chosen vectors span \( V \), the argument proceeds from \( V_k \) to \( V_{k+1} \). In order to avoid an excessive amount of notation the step from \( V_1 \) to \( V_2 \) will be indicated. If \( x \in V_2 \), \( x p \in V_1 \) so that \( x p = \sum_{\alpha, j} \beta(\alpha, j) x(\alpha, 1) D^j \); let \( y = \sum_{\alpha, j} \beta(\alpha, j) x(\alpha, 2) D^j \) and observe that \( z = x - y \) lies in \( V_1 \) since \( y p = x p \). Thus \( z = \sum_{\alpha, j} \gamma(\alpha, j) x(\alpha, 1) D^j \) and \( x = y + z = \sum_{\alpha, j} \beta(\alpha, j) x(\alpha, 2) D^j + \sum_{\alpha, j} \gamma(\alpha, j) x(\alpha, 1) D^j \).

To conclude the proof of Theorem 1, \( T \) is explicitly constructed in terms of the basis \( \{x(\alpha, k) D^j\} \) of Lemma 3. Define

\[
\begin{align*}
x(\alpha, k) T &= x(\alpha, k + 1), \\
x(\alpha, k) DT &= x(\alpha, k + 1) D - x(\alpha, k), \\
x(\alpha, k) D^2 T &= x(\alpha, k + 1) D^2 - 2x(\alpha, k) D, \\
&\vdots \\
x(\alpha, k) D^{m-1} T &= x(\alpha, k + 1) D^{m-1} - (m - 1)x(\alpha, k) D^{m-2}.
\end{align*}
\]

It only remains to establish \( [T, D] = I \). The calculation is as follows:

\[
x(\alpha, k) D^j T D = [x(\alpha, k + 1) D^j - jx(\alpha, k) D^{j-1}] D
\]

\[
= x(\alpha, k + 1) D^{j+1} - jx(\alpha, k) D^j
\]

and

\[
x(\alpha, k) D^j (DT) = x(\alpha, k + 1) D^{j+1} - (j + 1)x(\alpha, k) D^j.
\]

Upon differencing these two results one obtains

\[
x(\alpha, k) D^j (TD - DT) = x(\alpha, k) D^j
\]

which is exactly the desired result. It should be remarked that these
calculations are valid when \( j = m - 1 \) since in this case \( D^{j+1} = D^m \) is expressible as a linear combination of lower powers of \( D \).

3. **Necessity for characteristic zero.** In this section \( F \) will designate a field of characteristic 0 and \( V \) will be a vector space over \( F \) with two linear transformations, \( D \) and \( T \), satisfying \( [T, D] = I \). Moreover it is assumed that \( V \) is locally algebraic with respect to \( D \). Again the problem is reduced to the primary case by Lemmas 1 and 2, but it is necessary to show that the primary components of \( V \) are invariant under \( T \) before the reduction can be made.

**Lemma 4.** Let \( D, T \) be linear transformations of \( V \) satisfying \( [T, D] = I \). For any polynomial \( f(\lambda) \) in \( F[\lambda] = A \)

\[
TD^k = D^kT + kD^{k-1},
\]

\[
Tf(D) = f(D)T + f'(D)
\]

where \( f'(\lambda) \) designates the usual derivative of \( f(\lambda) \). Furthermore, if \( V_p \) is a primary component of \( V \) (relative to \( D \)) then \( V_p \) is \( T \)-invariant.

The first relation in (5) is readily established by induction and the second is an immediate consequence thereof. To see that \( V_p \) is \( T \) invariant observe first that \( V_p \) is \( D \)-invariant. Then for any \( x \in V_p \) let \( x(p(D)) = 0 \) and note \( (xT)(p(D))^r = x(p(D))^rT + x(p(D))^rD^k \)

\[
= x(p(D))^rT + rx(p(D))^{-1}p'(D) \]

which lies in \( V_p \). For the remainder of this section it is assumed that \( V \) is a primary \( A \)-module such that \( [T, D] = I \). Define the subspaces \( V_k \) by (1) again. Then the following lemma holds [3].

**Lemma 5.** If every \( y \) in \( V_1 \) has the property that for each integer \( k \geq 1 \), there is an \( x \) in \( V \) with \( y = x(p(D))^k \) then \( V \) is divisible.

To simplify the following calculations, the notation \( xf, xTf, T^k f \), etc., is used in lieu of \( x(f(D)), xTf(D), T^k f(D) \), etc. There are several steps which culminate with the verification of the hypothesis of Lemma 5. These steps are listed below where \( (f, g) = 1 \) signifies as usual that the polynomials \( f(\lambda) \) and \( g(\lambda) \) are relatively prime.

(a) If \( (f, p) = 1 \) and \( y \in V \) there is an \( x \in V \) with \( y = xf \). For if \( y \in V_k \) write \( fg +_hp^k = 1 \) so that \( yfg + yp^k = y \); the desired conclusion follows with the choice \( x = yg \).

(b) If \( y \in V_k, yf = zp^n \) where \( (f, p) = 1 \) then there is an \( x \in V \) with \( y = xp^n \). Again write \( fg + p^k = 1 \) so that \( zg = zp^n = yp^k = y[1 - p^k] = y - yp^k = y \). For the choice \( x = zg \) the conclusion \( xp^n = y \) follows.
(c) An easy induction establishes the commutativity relation

\[ T^m f = \sum_{k=0}^{m} \binom{m}{k} f^{(k)} T^{m-k} \]

where \( f^{(k)} \) designates the \( k \)th derivative of \( f \).

(d) The following known result is easily established by induction.

\[ (p^n)^{(k)} = n(n-1) \cdots (n-k+1) p^{n-k} (p')^k + p^{n-k+1} f_k(p, p', \ldots, p^{(k)}) \]

where \( f_k(p, p', \ldots, p^{(k)}) \) is an integral polynomial in \( p, p', \ldots, p^{(k)} \).

(e) For each \( y \) in \( V_1 \) there is \( x \) in \( V \) with \( y = xp^n \). For let \( z = y T^n \) and compute

\[ zp^n = y T^n p^n = y \sum_{k=0}^{n} \binom{n}{k} (p^n)^{(k)} T^{n-k} \]

by (c). By (d) above

\[ y(p^n)^{(k)} = y \left[ k! \binom{n}{k} p^{n-k} (p')^k + p^{n-k+1} f_k(p, p', \ldots, p^{(k)}) \right] = 0 \]

for \( k < n \) since \( yp = 0 \). Thus

\[ zp^n = y(p^n)^{(n)} = y \left[ n!(p')^n + p f_n(p, p', \ldots, p^{(n)}) \right] = y[n!(p')^n] \]

Since the field is of characteristic 0, the irreducibility of \( p(\lambda) \) ensures \( (p, p') = 1 \); also \( n! \not\equiv 0 \) in \( F \) and so \( (n! p', p) = 1 \) and the conclusion follows immediately from (b). The hypothesis of Lemma 5 has been established and the proof of Theorem 2 is complete.

A counterexample for finite characteristic is readily given. For example if \( F = GF(3) \) and \( V \) has basis \( x_1, x_2, x_3 \) over \( F \) define \( T \) by \( x_1 T = x_2, x_2 T = x_3, x_3 T = 0 \) and \( D \) by \( x_1 D = 0, x_2 D = x_1 \) and \( x_3 D = 2x_2 \).

An easy check shows \( TD - DT = I \) and \( D \) is surely singular. It is clear that if \( V \) were divisible as an \( A \)-module, \( D \) would have to map \( V \) onto itself and so \( V \) cannot be divisible. Closely related to these results is a result of Albert and Muckenhoupt [1] which states that if \( S \) is a linear transformation of the finite dimensional vector space \( V \) over \( F \) it is a commutator, i.e., \( S = TU - UT \) for linear transformations \( U, T \) of \( V \) if and only if Trace \( S = 0 \).

4. Results of Mikusinski and Geer. In [2] Mr. Geer gave a generalization of Mikusinski's result. Using Theorem 1 it is easy to prove a result which includes both of their results and is stated as
Theorem 3. Let $D$ be a locally algebraic linear operator on $V$ satisfying
\((i)\) for each irreducible $p(\lambda)$ in $A$ the kernel of $p(D)$ is finite dimensional,
\((ii)\) for each irreducible $p(\lambda)$, $\dim \ker p^n = n(\dim \ker p)$. Then a linear operator $T$ on $V$ exists with $[T, D] = I$.

The only hypothesis of Theorem 1 that must be verified is the divisibility condition and by Lemmas 1 and 2 it suffices to verify this condition for each primary component of $V$. Obviously the restriction of $D$ to a primary component also satisfies (i) and (ii) and so it may be assumed that $V$ is primary for some prime $p = p(\lambda)$. The subspaces $V_k$ are again defined by (1) so that $V_{k+1} \subset V_k$ and the kernel of $p(D)|_{V_{k+1}}$ is clearly $V_1$ so $V_{k+1}/V_1$ is $A$-isomorphic to $V_k/p$. Thus $\dim V_{k+1} - \dim V_1 = \dim V_k$ but by (ii) $\dim V_{k+1} = (k+1)(\dim V_1)$ and so $\dim V_{k+1} = k(\dim V_1)$ which is $\dim V_k$ by (ii). Therefore, $V_{k+1} \subset V_k$ together with the dimension count given shows $V_{k+1} = V_k$ and the condition of Lemma 5 is verified and $V$ is divisible as desired.

For completeness the converse of Mikusinski is deduced from Theorem 2 in the following form.

Theorem 4. Let $D$ be a linear operator on the vector space $V$ over the field $F$ of characteristic 0 such that $D$ is locally algebraic on $V$. Suppose that for each irreducible $p(\lambda)$ in $A = F[\lambda]$, $\dim \ker p$ is finite and suppose that a linear operator $T$ of $V$ exists with $[T, D] = I$, then condition II is satisfied and $\dim \ker p(\lambda)$ is finite for every $\lambda$ of positive degree.

By Theorem 2, $V$ must be a divisible $A$-module and so is each primary component of $V$ by Lemmas 1 and 2. If $S$ is a primary component, let $S_k = \ker p^k$ so that the divisibility property of $S$ ensures $S_{k+1}p = S_k$. Since $\ker p(D)|_{S_{k+1}} = S_1$ the isomorphism theorem yields $S_{k+1}/S_1 A$-isomorphic to $S_{k+1}p = S_k$. Thus $\dim S_{k+1} = \dim S_k + \dim S_1$ and $\dim S_1$ is finite by hypothesis; an obvious induction argument establishes $\dim S_{k+1} = (k+1) \dim S_1 = (k+1) \dim \ker p$ as desired.

Next, observe that if $f(\lambda)$, $g(\lambda)$ are relatively prime then $\ker fg = \ker f + \ker g$. For surely $\ker f + \ker g \subset \ker fg$ holds; consequently, write $1 = fh + gk$ for $h$, $k$ in $A$ so that $x$ in $\ker fg$ can be written as $x = xfh + xgk$ where $xfh \in \ker g$ and $xgk \in \ker f$ is obvious. This shows $\ker f + \ker g = \ker fg$; finally if $x \in (\ker f) \cap (\ker g)$ then $x = xfh + xgk = 0$ and so the sum is direct. The desired conclusion is now an obvious consequence of the preceding results and the unique factorization in $A$. 
References


5. ———, *Sur l'espace lineaire avec derivation*, Studia Math. 16 (1957), 113–123.


University of Virginia