PACIFIC JOURNAL OF MATHEMATICS
Vol. 61, No. 1, 1975

KATO-TAUSSKY-WIELANDT COMMUTATOR RELATIONS AND CHARACTERISTIC CURVES

FERGUS GAINES

Let A and B be $n \times n$ matrices with elements in a field $\mathscr{F}$ and let $\Delta_A B = AB - BA$. Let $f_k(x) = x^{2K+1} - c_1 x^{2K-1} + c_2 x^{2K-3} + \cdots + (-1)^K c_K x$, where the c_i are in $\mathscr{F}$ and $K = k(k-1)/2$. In this paper we examine the consequences of the relation $f_k(\Delta_A)B = 0$, where $1 \leqq k < n$, and show how the replacement of A by $xA + yB$, when $k = 2$, leads to a splitting of the characteristic curve, $\det(xA+yB-zI) = 0$, into lines and conics.

1. Introduction. We open with some notation and some definitions. Let $\mathscr{F}$ be a field and let $\mathscr{F}_n$ denote the $n \times n$ matrices with elements in $\mathscr{F}$. If A and $B \in \mathscr{F}_n$, the *characteristic curve* of the pencil $xA + yB$ is the curve in the projective x, y, z-plane whose equation is $\det(xA + yB - zI) = 0$. If $A \in \mathscr{F}_n$, the operator Δ_A is given by $\Delta_A X = AX - XA$, for all X in $\mathscr{F}_n$. If $k \geqq 1$ is an integer and $K = k(k-1)/2$ and if $c_1, c_2, \cdots, c_K \in \mathscr{F}$, we let $f_k(x) = x^{2K+1} - c_1 x^{2K-3} + \cdots + (-1)^K c_K x$. Next we need some ideas usually associated with the Perron-Frobenius theory of nonnegative matrices. If $X = (x_{ij}) \in \mathscr{F}_n$, the *digraph* $\mathscr{G}(X)$ consists of vertices labeled $1, 2, 3, \cdots, n$ and there is an edge from i to j, i.e. $i \rightarrow j$, if and only if $x_{ij} \neq 0$. The matrix $X \in \mathscr{F}_n$ is *permutation-irreducible* if and only if it can not be transformed by a permutation similarity to the form $\begin{pmatrix} Y & Z \\ 0 & W \end{pmatrix}$ where Y and W are square matrices.

In [4] Taussky and Wielandt proved.

THEOREM 1. *If A and $B \in \mathscr{F}_n$ and $\alpha_1, \alpha_2, \cdots, \alpha_n$ are the eigenvalues* (in some extension field of $\mathscr{F}$) *of A, then $f_n(\Delta_A)B = 0$, where c_i is the ith elementary symmetric function of the $N = n(n-1)/2$ quantities $(\alpha_r - \alpha_s)^2$, $1 \leqq r < s \leqq n$; $i = 1, 2, \cdots, N$.*

Since $f_1(\Delta_A)B = AB - BA$, the relation $f_k(\Delta_A)B = 0$, with $1 < k < n$, for some $c_1, c_2, \cdots, c_K \in \mathscr{F}$, is a generalization of commutativity. As a generalization of matrix commutativity it is, however, quite weak, since it is still possible for A and B to satisfy such an identity (when $n = 3$) and to generate $\mathscr{F}_3$ (see the examples in §4). However, it will be shown that the relation $f_k(\Delta_A)B = 0$ imposes a restriction on the eigenvalues of A, when A and B generate $\mathscr{F}_n$. We call an expression of the form $f_k(\Delta_A)B$ a *Kato-Taussky-Wielandt commutator.*

We shall need one well-known result from graph theory which can be found, for example, in [5].

THEOREM 2. *$X \in \mathscr{F}_n$ is permutation-irreducible if and only if $\mathscr{G}(X)$ is strongly connected.*

2. The main theorem.

Our principal result is

THEOREM 3. *Let A and $B \in \mathscr{F}_n$ and suppose A and B generate $\mathscr{F}_n$,* i.e. *every matrix in $\mathscr{F}_n$ has the form $P(A, B)$, where $P(x, y)$ is a polynomial over $\mathscr{F}$ in the noncommuting indeterminates x and y. If the characteristic of $\mathscr{F}$ does not divide n and if, for some fixed integer k with $1 \leqq k < n$, there exist $c_1, c_2, \cdots, c_K$ in $\mathscr{F}$ so that $f_k(\Delta_A)B = 0$, then the eigenvalues of A belong to the splitting field of $f_k(x)$ over $\mathscr{F}$.*

Proof. If $k = 1$ then $AB = BA$ and, since A and B generate $\mathscr{F}_n$, we get $n = 1$. The theorem is then obvious.

If A has only one eigenvalue α, then $n\alpha =$ trace A is in $\mathscr{F}$ and, consequently, $\alpha \in \mathscr{F}$, since the characteristic of $\mathscr{F}$ does not divide n.

So assume that $1 < k < n$, that A has at least two distinct eigenvalues and that there exist $c_1, c_2, \cdots, c_K \in \mathscr{F}$ so that $f_k(\Delta_A)B = 0$. By extending $\mathscr{F}$ to its algebraic closure $\bar{\mathscr{F}}$, we may assume (via similarity) that $A = \sum_{i=1}^{r} \oplus A_i$ is in Jordan canonical form, where A_i is a direct sum of Jordan blocks, all of which have the same eigenvalue α_i and $\alpha_i \neq \alpha_j$ when $i \neq j$. We then view B as a matrix over $\bar{\mathscr{F}}$ (via the similarity above) and let $B = (B_{ij})$ be the partition of B into blocks corresponding to that of $A = \sum_{i=1}^{r} \oplus A_i$. We shall prove

LEMMA 1. *If $B_{ij} \neq 0$, then $\alpha_i - \alpha_j$ satisfies $f_k(x) = 0$.*

LEMMA 2. *$B = (B_{ij})$ is permutation-irreducible as a block matrix.*

If we assume these lemmas we can complete the proof of the theorem in a few lines. Let $\mathscr{G}(B)$ be the digraph of B viewed as a block matrix, i.e. $i \rightarrow j$ if and only if $B_{ij} \neq 0$. Then Lemma 2 and a modification of Theorem 2 (for block matrices) imply that $\mathscr{G}(B)$ is strongly connected. Thus, if α_i, α_j are distinct eigenvalues of A, there exists a sequence $i, i_1, i_2, \cdots, i_m, j$ so that $\alpha_i - \alpha_{i_1}, \alpha_{i_1} - \alpha_{i_2}, \cdots, \alpha_{i_m} - \alpha_j$ are roots of $f_k(x) = 0$. Let $\mathscr{L}$ be the splitting field of $f_k(x)$ over $\mathscr{F}$. Then

$$\alpha_i - \alpha_j = (\alpha_i - \alpha_{i_1}) + (\alpha_{i_1} - \alpha_{i_2}) + \cdots + (\alpha_{i_m} - \alpha_j) \in \mathscr{L} .$$

Let n_j be the multiplicity of α_j as an eigenvalue of A. Then

$$\sum_{j=1}^{r} n_j\alpha_i - \sum_{j=1}^{r} n_j\alpha_j \in \mathscr{L} ,$$

i.e. $n\alpha_i$ − trace $A \in \mathscr{L}$. Thus $\alpha_i \in \mathscr{L}$, $i = 1, 2, \cdots, r$, since the characteristic of $\mathscr{L}$ does not divide n.

It remains to prove the lemmas to complete the proof of the theorem.

Proof of Lemma 1. We use the relation $f_k(\Delta_A)B = 0$. Suppose $B_{ij} \neq 0$. Let b_{st} be the "first" nonzero element of B_{ij} in the following sense: if the lower left-hand corner element of B_{ij} is nonzero let it be b_{st}; otherwise let b_{st} be a nonzero element of B_{ij} so that $b_{uv} = 0$ if $u \geqq s$ and $v \leqq t$, and $(u, v) \neq (s, t)$, where, of course, we only consider those elements b_{uv} of B_{ij}. Thus

$$B_{ij} = \begin{bmatrix} & & * & & & & * \\ 0 & \cdot & \cdot & \cdot & 0 & b_{st} & \\ 0 & \cdot & \cdot & \cdot & 0 & 0 & \\ \cdot & & & & \cdot & \cdot & * \\ 0 & \cdot & \cdot & \cdot & 0 & 0 & \end{bmatrix} .$$

If b_{st} is the (s, t) element of B_{ij} we calculate the (s, t) element of the (i, j) block of $f_k(\Delta_A)B$. To simplify calculations assume that A_j has eigenvalue zero and A_i has eigenvalue $\alpha_{ij} = \alpha_i - \alpha_j$ (Subtract $\alpha_j I$ from A. Since we take commutators, this operation does not affect the end result of the calculations). The matrix $f_k(\Delta_A)B$ is a linear combination of matrices of the type $\Delta_A^m B$. The (i, j) block of $\Delta_A B$ is $A_i B_{ij} - B_{ij} A_j$. The (i, j) block of $\Delta_A^m B$ only involves B_{ij}, A_i and A_j; it consists of a linear combination of matrices of the type $A_i^c B_{ij} A_j^d$ where $c + d = m$. The (s, t) element of $A_i^c B_{ij} A_j^d$ is obtained by multiplying the tth column of $B_{ij} A_j^d$ by the tth row of A_i^c. Those elements in the tth column of $B_{ij} A_j^d$ from the sth row down are all that matter here. But these elements are zero, except when $d = 0$, since A_j has zeros on and below the main diagonal. Thus the (s, t) element of the (i, j) block of $\Delta_A^m B$ is $\alpha_{ij}^m b_{st}$. So the equation $f_k(\Delta_A)B = 0$ gives $f_k(\alpha_{ij}) = 0$, since $b_{st} \neq 0$.

Proof of Lemma 2. Suppose there exists a block permutation matrix Q, partitioned conformally with $B = (B_{ij})$, so that $Q^{-1}BQ$ has the form (†) where $m < r$. Then A and B are reduced by Q, since Q simply permutes the blocks on the diagonal of A. Thus the algebra generated by A and B over $\overline{\mathscr{F}}$ is reducible. But A and B generate $\overline{\mathscr{F}}_n$. This contradiction proves that $B = (B_{ij})$ is permutation-irreducible as a block matrix.

$$(\dagger)\qquad \begin{bmatrix} B_{11} & \cdots & B_{1m} & B_{1\ m+1} & \cdots & B_{1r} \\ \cdot & & \cdot & \cdot & & \cdot \\ \cdot & & \cdot & \cdot & & \cdot \\ B_{ml} & \cdots & B_{mm} & B_{m\ m+1} & \cdots & B_{mr} \\ 0 & \cdots & 0 & B_{m+1\ m+1} & \cdots & B_{m+1\ r} \\ \cdot & & \cdot & \cdot & & \cdot \\ \cdot & & \cdot & \cdot & & \cdot \\ 0 & \cdots & 0 & B_{r\ m+1} & \cdots & B_{rr} \end{bmatrix}$$

THEOREM 4. *Let A and B satisfy the conditions of Theorem 3 with $k = 2$ and let A have at least two distinct eigenvalues. Then there exists an ordering $\alpha_1, \alpha_2, \cdots, \alpha_r$ of the distinct eigenvalues of A so that $\alpha_1 - \alpha_2 = \alpha_2 - \alpha_3 = \cdots = \alpha_{r-1} - \alpha_r$ satisfies $x^2 - c_1 = 0$.*

Proof. As in the proof of Theorem 3, let $B = (B_{ij})$ be the block form of B (over $\mathscr{F}$) corresponding to the Jordan canonical form $\sum_{i=1}^{r} \oplus A_i$ of A. We have $f_2(\Delta_A)B = 0$, for some $c_1 \in \mathscr{F}$.

We claim that B can not have more than two nonzero off-diagonal blocks in each row or column. For if B_{ij}, B_{il} and B_{im} are nonzero off-diagonal blocks, where j, l, and m are distinct, then $\alpha_i - \alpha_j$, $\alpha_i - \alpha_l$ and $\alpha_i - \alpha_m$ satisfy the equation $x^2 - c_1 = 0$. Thus two of these α's, at least, are equal, contradicting the fact that the α's are distinct. If B_{ij} and B_{il} (resp. B_{ji} and B_{li}) are nonzero off-diagonal blocks with $j \neq l$, and B_{mi}(resp. B_{im}) is also a nonzero off-diagonal block, a similar argument proves $m = j$ or $m = l$.

Let $\mathscr{G}(B)$ be the digraph of B viewed as a block matrix. We write $i \sim j$ if $i \to j$ or $j \to i$. So if $i \sim j$ and $i \sim l$, where i, j and l are distinct, then $i \sim m$ implies either j or l is m. We claim that, by relabeling the vertices of $\mathscr{G}(B)$, we get the subgraph

$$\underset{1}{\bullet}\!-\!\!-\!\!-\!\underset{2}{\bullet}\!-\!\!-\!\!-\!\underset{3}{\bullet}\quad \cdots \quad \underset{r-1}{\bullet}\!-\!\!-\!\!-\!\underset{r}{\bullet}$$

where $i \sim i + 1$, for $i = 1, 2, \cdots, r - 1$. For let

$$\mu = \underset{1}{\bullet}\!-\!\!-\!\!-\!\underset{2}{\bullet}\!-\!\!-\!\!-\!\underset{3}{\bullet}\quad \cdots \quad \underset{s-1}{\bullet}\!-\!\!-\!\!-\!\underset{s}{\bullet}$$

be a maximal "path" in $\mathscr{G}(B)$ (on relabeling vertices), where $i \sim i + 1$ for $i = 1, 2, \cdots, s - 1$, and suppose $s \neq r$. If j is a vertex not in μ then neither $j \sim 1$ nor $j \sim s$ can hold, since μ is maximal. Since $\mathscr{G}(B)$ is strongly connected, there exists an internal vertex $i \in \mu$

and a vertex $j \notin \mu$ so that $i \sim j$. But $i \sim i+1$ and $i \sim i-1$ and, since neither or these is j, we get a contradiction. Thus $\mathscr{G}(B)$ contains the required subgraph. By Lemma 1, this means the distinct eigenvalues α_i of A can be relabeled so that $\alpha_1 - \alpha_2, \alpha_2 - \alpha_3, \cdots, \alpha_{r-1} - \alpha_r$ satisfy the equation $x^2 - c_1 = 0$. Since the α's are distinct, we get

$$\alpha_i - \alpha_{i+1} = \alpha_{i+1} - \alpha_{i+2} , \qquad 1 = 1, 2, \cdots .$$

This completes the proof of the theorem.

3. **Generalized L-property.** Let A and B be $n \times n$ matrices with elements in $\mathscr{F}$ and suppose the eigenvalues of A and B are also in $\mathscr{F}$. If there exist fixed orderings $\alpha_1, \alpha_2, \cdots, \alpha_n$ and $\beta_1, \beta_2, \cdots, \beta_n$ of the eigenvalues of A and B, respectively, so that the eigenvalues of $xA + yB$ are $x\alpha_i + y\beta_i$, for all x and y in $\mathscr{F}$, where $i = 1, 2, \cdots, n$, then A and B have *property L*. Property L has been discussed by Motzkin and Taussky [3]; it is clearly equivalent to the assertion that the characteristic curve of the pencil $xA + yB$ is the union of n lines (if $\mathscr{F}$ is big enough). In this section we discuss a condition which forces the characteristic curve to decompose into lines and conices.

Let $\mathscr{F}[x, y]$ be the integral domain of polynomials over $\mathscr{F}$ in the (commuting) indeterminates x and y. Let $\mathscr{F}(x, y)$ be its quotient field.

LEMMA 3. *Let $p(x, y, z)$ be a homogeneous polynomial in x, y and z, with coefficients in $\mathscr{F}$. Suppose*

$$p(x, y, z) = \prod_{i=1}^{r} p_i^{k_i} ,$$

where each p_i is an irreducible polynomial in z over $\mathscr{F}(x, y)$. Then each p_i is a homogeneous polynomial in x, y, and z, with coefficients in $\mathscr{F}$.

Proof. $\mathscr{F}[x, y]$ is a unique factorization domain (UFD). Since a UFD is integrally closed ([2], p. 84), the coefficients of the powers of z in p_i must be polynomials in x and y.

Suppose p_i is not homogeneous in x, y and z. Let $M(q)$ (resp. $m(q)$) be the maximum (resp. minimum) degree of the monomials in a polynomial q. Then $M(p_i) > m(p_i)$ and M (resp. m) has the property that $M(q_1q_2) = M(q_1) + M(q_2)(m(q_1q_2) = m(q_1) + m(q_2))$ for polynomials q_1 and q_2. Hence $M(p) > m(p)$, which is false. The Lemma is proved.

We now apply the results of §2 to $xA + yB$ and B.

THEOREM 5. *Let A and $B \in \mathscr{F}_n$, where $\mathscr{F}$ is an infinite field* *whose characteristic does not divide n. If A and B generate $\mathscr{F}_n$ and if, for each x and y in $\mathscr{F}$, there exists c_1 in $\mathscr{F}$ so that*

$$f_2(\Delta_{xA+yB})B = 0 \ ,$$

then the characteristic polynomial $p(x, y, z)$ of $xA + yB$ splits into linear and quadratic homogeneous factors with coefficients in $\mathscr{F}$.

Proof. Without loss, we may assume that $n \geqq 3$. Let $X = xA + yB$. If $\Delta_X B = 0$, for $x \neq 0$, then $AB = BA$, which implies $n = 1$. So $\Delta_X B \neq 0$ (for $x \neq 0$) and the relation $f_2(\Delta_X)B = 0$ imply that c_1 is a rational function of x and y. Since $\mathscr{F}$ is infinite, we may replace x and y by two algebraically independent indeterminates and the relation $f_2(\Delta_X)B = 0$ still holds. X and B clearly generate $\mathscr{F}(x, y)_n$, and thus we may apply Theorem 3. Since $f_2(w) = w^3 - c_1 w$, each eigenvalue of $X = xA + yB$ satisfies an equation of degree at most 2 over $\mathscr{F}(x, y)$. Lemma 3 is now used to complete the proof.

COROLLARY. *Let $\mathscr{F}$ be an algebraically closed field of characteristic zero or greater than n. If A and $B \in \mathscr{F}_n$ and if, for each x and $y \in \mathscr{F}$, there exists $c_1 \in \mathscr{F}$ so that*

$$f_2(\Delta_{xA+yB}) \in \mathscr{J} \ ,$$

where $\mathscr{J}$ is the radical of the algebra generated by A and B over $\mathscr{F}$, then the characteristic polynomial $p(x, y, z)$ of $xA + yB$ splits into linear and quadratic homogeneous factors in x, y and z with coefficients in $\mathscr{F}$.

Proof. If A_1 and B_1 are the representatives of A and B respectively, in an irreducible representation of the algebra generated by A and B over $\mathscr{F}$, then $f_2(\Delta_{X_1})B_1 = 0$, where $X_1 = xA_1 + yB_1$. Also A_1 and B_1 generate a complete matrix algebra, since $\mathscr{F}$ is algebraically closed. A and B may be transformed by a similarity into block upper triangular form, where the corresponding diagonal blocks generate irreducible matrix algebras. The conclusion follows.

REMARK. Nothing we have said so far forces the characteristic curve of $xA + yB$ to contain a line. If $\mathscr{F}$ has characteristic zero or greater than n and $f_2(\Delta_{xA+yB})B = 0$ for some $c_1 \in \mathscr{F}(x, y)$, where x and y are algebraically independent over $\mathscr{F}$, and if $xA + yB$ has an *odd* number of distinct eigenvalues, then at least one of the eigenvalues has the form $x\alpha + y\beta$, where $\alpha, \beta \in \mathscr{F}$. For let $z_1, z_2, \cdots, z_r$ be the distinct eigenvalues of $xA + yB$, with r odd. If $r = 1$ the result is trivial. Let $r \geqq 3$; then, by Theorem 4, we may assume the eigenvalues are ordered so that $z_1 - z_2 = z_2 - z_3 = \cdots z_{r-1} - z_r$ satisfies $w^2 - c_1 = 0$. By the condition on the characteristic of $\mathscr{F}$, the irreducible factors of $p(x, y, z)$ — the characteristic polynomial of $xA + yB$ — are separable. Hence $\sum_{i=1}^{r} z_i \in \mathscr{F}(x, y)$. Now

$$(1/r) \sum_{i=1}^{r} z_i = z_r + ((r-1)/2)c_1^{1/2} ,$$

since $z_i - z_{i+1} = c_1^{1/2}$, $i = 1, 2, \cdots, r-1$. Since r is odd, $(r-1)/2 = s$ is an integer and $z_{r-s} = z_r + ((r-1)/2)c_1^{1/2}$ is in $\mathscr{F}(x, y)$. By Lemma 3, $z_{r-s} = x\alpha + y\beta$, where α and $\beta \in \mathscr{F}$.

4. Examples.

EXAMPLE 1. This example illustrates the main results of the paper. Let $\mathscr{F}$ be a field of characteristic not 2 or 3. Let

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} .$$

Then A and B generate $\mathscr{F}_3$. If $X = xA + yB$, where x and y are algebraically independent over $\mathscr{F}$, then X and B generate $\mathscr{F}(x, y)_3$ and, if $c_1 = \sqrt{2x^2 + y^2}$, then $f_2(\Delta_X)B = 0$. The characteristic polynomial of $xA + yB$ is

$$(x + 2y - z)(z^2 - (2x + 4y)z - x^2 + 4xy + 3y^2)$$

(cf. Theorem 5). The eigenvalues of $xA + yB$ are

$$z_2 = x + 2y$$

$$z_1 = x + 2y + \sqrt{2x^2 + y^2} , \quad z_3 = x + 2y - \sqrt{2x^2 + y^2} .$$

Clearly

$$z_1 - z_2 = z_2 - z_3 = \sqrt{2x^2 + y^2}$$

(cf. Theorem 4). We see that

$$z_1, z_2, z_3 \in \mathscr{F}(x, y, \sqrt{2x^2 + y^2})$$

(cf. Theorem 3).

EXAMPLE 2. The example we give here is a counterexample to Theorem 3 and Theorem 5, when the condition on the characteristic of $\mathscr{F}$ is not satisfied. Let $\mathscr{F}$ have characteristic 3 and let

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

Then A and B generate $\mathscr{F}_3$ and, if $X = xA + yB$ and $c_1 = 0$, then $f_2(\Delta_X)B = 0$. Now $f_2(w) = w^3$ and the characteristic polynomial of $xA + yB$ is $xy^2 - z^3 = ((xy^2)^{1/3} - z)^3$. Theorems 3 and 5 clearly fail here.

Acknowledgments. The results in this paper are contained in the last chapter of the author's Ph. D. thesis (Caltech, 1966), which was directed by Olga Taussky Todd. Sincere thanks are offered to her for much helpful advice; thanks are also due to E. C. Dade. A special word of thanks to Ed Bender for a helpful conversation on the material of this paper.

REFERENCES

1. F. Gaines, *Kato-Taussky-Wielandt commutator relations,* Linear Algebra and its Applications, **I** (1968), 127-138.
2. N. Jacobson, *Lectures in Abstract Algebra, I,* Van Nostrand, Princeton, 1951.
3. T. S. Motzkin and O. Taussky, *Pairs of matrices with property L, I,* Trans. Amer. Math. Soc., **73** (1952), 108-114; II, ibid., **80** (1955), 387-401.
4. O. Taussky and H. Wielandt, *Linear relations between higher additive commutators,* Proc. Amer. Math. Soc., **13** (1962), 732-735.
5. R. S. *Varga, Matrix Iterative Analysis,* Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1962.

Received March 15, 1975. Research supported in part by a grant from the National Science Foundation.

UNIVERSITY COLLEGE DUBLIN 4, IRELAND