LINEAR RELATIONS BETWEEN HIGHER ADDITIVE COMMUTATORS

O. TAUSSKY AND H. WIELANDT[1]

It has been noticed earlier [1] that the iterated commutators

$$b_1 = ab - ba, \qquad b_i = ab_{i-1} - b_{i-1}a \qquad (i = 2, 3, \cdots) \tag{1}$$

of 2×2 matrices a, b with elements in a field F (which we may assume algebraically closed) satisfy the linear relation

$$b_3 = (\alpha_1 - \alpha_2)^2 b_1 \tag{2}$$

where α_1, α_2 denote the eigenvalues of a. Although it is easily seen and well known (cf. e.g. [2]) that also for any two $n \times n$ matrices a, b, where $n > 2$, there exist linear relations between the b_i whose coefficients are independent of b no such relation seems to be explicitly known. We want to generalize (2) to $n \times n$ matrices.

The mapping $x \rightarrow Ax :\, = ax - xa$ is a linear operator on the n^2 dimensional space $F_{n \times n}$ of all $n \times n$ matrices x with elements in F. By the Hamilton-Cayley Theorem A satisfies its characteristic equation, say $h(A) = 0$. If $\alpha_1, \cdots, \alpha_n$ denote the eigenvalues of a the eigenvalues of A are known [3] to be $\alpha_i - \alpha_j$ $(i, j = 1, 2, \cdots, n)$, hence

$$\begin{aligned} h(z) &= \prod_{i,j} [z - (\alpha_i - \alpha_j)] = z^n \prod_{i<j} [z^2 - (\alpha_i - \alpha_j)^2] \\ &= z^n \cdot (z^{2N} - \delta_1 z^{2N-2} + \delta_2 z^{2N-4} - \cdots + (-1)^N \delta_n) \end{aligned} \tag{3}$$

where $N = n(n-1)/2$ and δ_k is the kth elementary symmetric function of the $(\alpha_i - \alpha_j)^2$, $1 \leq i < j \leq n$. As $h(A)b = 0$ and $A^k b = b_k$ we have

$$b_{2N+n} - \delta_1 b_{2N+n-2} + \delta_2 b_{2N+n-4} - \cdots + (-1)^N \delta_N b_n = 0. \tag{4}$$

Although this is a linear relation between the b_i it is not a generalization of (2) as (4) for $n = 2$ gives $b_4 = (\alpha_1 - \alpha_2)^2 b_2$ which is weaker than (2). This suggests that (3) is not the minimum polynomial of A. Indeed a factor z^{n-1} may be cancelled:[2]

THEOREM 1. *Let a, b be $n \times n$ matrices with elements in a field F. Define $b_1, b_2, \cdots$ by* (1) *and $\delta_1, \cdots, \delta_N$ by* (3). *Then*

Received by the editors March 8, 1961 and, in revised form, October 10, 1961.

[1] The work of this author was supported by the National Science Foundation and the Deutsche Forschungsgemeinschaft.

[2] For a specified matrix a, the minimum equation of A may even have a smaller degree than $2N+1$. This leads to the problem to determine the elementary divisors of $a \times 1 - 1 \times a$, where $\times$ denotes the Kronecker product.

(5) $$b_{2N+1} - \delta_1 b_{2N-1} + \delta_2 b_{2N-3} - \cdots + (-1)^N \delta_N b_1 = 0.$$

This theorem can be easily proved by transformation of a into diagonal form, if all α_i are distinct. The general case can be reduced to the one just mentioned by taking for a the "generic" matrix whose n^2 elements a_{ij} are independent variables. We shall not go into details as we prefer to use a more powerful method which works in arbitrary (associative) rings, not only in algebras.[3] Obviously Theorem 1 is a special case ($R = F_{n\times n}$) of

THEOREM 2. *Let R be a ring and Γ its center. Let $a, b \in R$ and define $b_1, b_2, \cdots$ by (1). Let a satisfy an equation $f(a) = 0$ of the form*

(6) $$f(x) = x^n - \gamma_1 x^{n-1} + \gamma_2 x^{n-2} - \cdots + (-1)^n \gamma_n, \qquad \gamma_\nu \in \Gamma.$$

Then

(7) $$b_{2N+1} - \delta_1 b_{2N-1} + \delta_2 b_{2N-3} - \cdots + (-1)^N \delta_N b_1 = 0$$

where $N = n(n-1)/2$ and the coefficients $\delta_k \in \Gamma$ are defined in the following way: Let $x_1, \cdots, x_n$ be independent variables, and let $\Phi_k(c_1, \cdots, c_n)$ be the unique polynomial with integer coefficients which expresses the kth elementary symmetric function d_k of the N functions $(x_i - x_j)^2$, $1 \leqq i < j \leqq n$, in terms of the elementary symmetric functions $c_1^*, \cdots, c_n^*$ of the x_i; that is, $d_k = \Phi_k(c_1^*, \cdots, c_n^*)$. Then $\delta_k = \Phi_k(\gamma_1, \cdots, \gamma_n)$.

As the mapping $x \to ax - xa$ is a derivation in the ring R one can suspect that Theorem 2 belongs to the theory of derivations and can be proved by means of that theory (cf. e.g. [2]). However, Theorem 2 can also be considered as a special case of a theorem which refers neither to derivations nor even to rings but to double modules:

THEOREM 3. *Let M be a left R_1-module and a right R_2-module where R_1 and R_2 are two rings over a common subring Γ which is in the center of both R_1 and R_2. Let $\gamma x = x\gamma$ for all $\gamma \in \Gamma$, $x \in M$. Let $a_1 \in R_1$, $a_2 \in R_2$, $b \in M$ and define $b_1 = a_1 b - b a_2$, $b_i = a_1 b_{i-1} - b_{i-1} a_2$ $(i = 2, 3, \cdots)$. Let $f(a_1) = f(a_2) = 0$ where f is a polynomial of the form (6). Then (7) holds.*

Theorem 2 is what becomes of Theorem 3 if we specialize $R_1 = R_2 = M = R$ and $a_1 = a_2 = a$. On the other hand, Theorem 3 is a consequence of the fact that Lagrange's resolvents work in arbitrary commutative rings, not only in fields (where they usually are employed). More specifically we need

THEOREM 4. *Let C be a commutative ring, and let $a', a'' \in C$ satisfy*

[3] We are indebted to the referee for pointing out this possibility.

$f(a')=f(a'')=0$ *where* f *is a polynomial of the form*

$$f(x) = x^n - \gamma_1 x^{n-1} + \gamma_2 x^{n-2} - \cdots + (-1)^n \gamma_n, \qquad \gamma_\nu \in C.$$

Put

$$g(z) = z^{2N+1} - \delta_1 z^{2N-1} + \delta_2 z^{2N-3} - \cdots + (-1)^N \delta_N z$$

where $\delta_k = \Phi_k(\gamma_1, \cdots, \gamma_n)$, $k=1, \cdots, N$, *is defined as in Theorem* 2. *Then* $g(a'-a'')=0$.

Theorem 3 follows from Theorem 4 if we choose for a' the endomorphism $x \rightarrow a_1 x$ of M, for a'' the endomorphism $x \rightarrow x a_2$ of M, and for C that subring of the endomorphism ring of M which is generated by a', a'' and the n endomorphisms $x \rightarrow \gamma_\nu x = x\gamma_\nu$.

We prove Theorem 4 by means of an elementary algebraic identity:

LEMMA. *Let* $x, y, z, c_1, c_2, \cdots, c_n$ *be independent variables over the field* Q *of rational numbers, and let* Z *denote the ring of rational integers. Put*

$$f(x) = f(x; c_1, \cdots, c_n) = x^n - c_1 x^{n-1} + c_2 x^{n-2} - \cdots + (-1)^n c_n$$

and

$$g(z) = g(z; c_1, \cdots, c_n) = z^{2N+1} - \Phi_1 z^{2N-1} + \cdots + (-1)^N \Phi_N z$$

where $\Phi_k \in Z[c_1, \cdots, c_n]$, $k=1, \cdots, N$, *is defined as in Theorem* 2. *Then there exist polynomials* $p, q \in Z[x, y, c_1, \cdots, c_n]$ *such that*

$$g(x - y) = pf(x) + qf(y) \tag{8}$$

identically in $x, y, c_1, \cdots, c_n$.

In order to prove the lemma we divide $g(x-y)$, *qua* polynomial in x, by $f(x)$. As the highest coefficient in $f(x)$ is 1 no denominators arise, and we obtain

$$g(x - y) = pf(x) + \sum_{\mu=1}^{n-1} r_\mu x^\mu$$

where $p \in Z[x, y, c_1, \cdots, c_n]$ and $r_\mu \in Z[y, c_1, \cdots, c_n]$. Now we divide each r_μ, *qua* polynomial in y, by $f(y)$ and obtain

$$g(x - y) = pf(x) + qf(y) + \sum_{\mu,\nu}^{n-1} r_{\mu\nu} x^\mu y^\nu \tag{9}$$

where $q \in Z[x, y, c_1, \cdots, c_n]$ and $r_{\mu\nu} \in Z[c_1, \cdots, c_n]$. In order to prove $r_{\mu\nu}=0$ we take independent variables $x_1, \cdots, x_n$ and denote their elementary symmetric functions by $c_1^*, \cdots, c_n^*$. By definition of $g(z)$ we have (cf. (3))

$$g(z; \overset{*}{c}_1, \cdots, \overset{*}{c}_n) = z \prod_{i \neq j} [z - (x_i - x_j)],$$

hence

$$g(x_i - x_j; \overset{*}{c}_1, \cdots, \overset{*}{c}_n) = 0 \qquad (i, j = 1, \cdots, n).$$

This together with (9) shows that the polynomial

$$\overset{*}{r}(x, y) = \sum_{\mu,\nu}^{n-1} \overset{*}{r}_{\mu\nu} x^\mu y^\nu, \qquad \overset{*}{r}_{\mu\nu} = r_{\mu\nu}(\overset{*}{c}_1, \cdots, \overset{*}{c}_n)$$

has the n^2 zeros $x = x_i$, $y = x_j$ in the field $Q(x_1, \cdots, x_n)$. As the degrees of r^* with respect to x and y are both less than n we find $r_{\mu\nu}^* = 0$, hence $r_{\mu\nu} = 0$ as $c_1^*, \cdots, c_n^*$ are algebraically independent. Hence (9) reduces to (8), and the lemma is proved. Theorem 4 follows from the lemma by specializing $x = a'$, $y = a''$ and $c_i = \gamma_i$ in (8).

References

1. T. Kato and O. Taussky, *Commutators of A and A**, J. Washington Acad. Sci. **46** (1956), 38–40.

2. S. A. Amitsur, *Derivations in simple rings*, Proc. London Math. Soc. **7** (1957), 87–112.

3. C. C. MacDuffee, *The theory of matrices*, Springer, Berlin, 1933; pp. 83–84, 89.

California Institute of Technology
University of Tübingen, Germany