Potter, Wielandt, and Drazin on the Matrix Equation $AB = \omega BA$: New Answers to Old Questions

Olga Holtz, Volker Mehrmann, and Hans Schneider

1. INTRODUCTION. The aim of this paper is to present a slice of the linear algebra of the 1950s and to give some answers to questions raised by Helmut Wielandt at that time. It was Wielandt's habit over many years to make notes on papers that interested him in what he called "diaries" (Tagebücher). These diaries were made available by his family after his death in February 2001 and are currently being transcribed. Many notes therein amount to summaries of papers, but in other cases Wielandt would add questions, ideas, or even further results. In this article we discuss one such entry, which appears on page 35 of Diary VII (1951) and will eventually be accessible in transcribed electronic form [24]. The entry concerns a paper that Wielandt reviewed for the Zentralblatt [23]. We next turn to this paper.

In 1950, H. S. A. Potter, a mathematician at Aberdeen University in Scotland, published a note in this MONTHLY [19] on the matrix equation

$$AB = \omega BA. \tag{1}$$

He called a pair of complex $n \times n$ matrices A and B satisfying (1) *quasi-commutative*. We refer to such matrices as ω-*commutative* (see section 2 for a definition of this term applicable to general rings). Otherwise we follow Potter's notation. It should be noted here that the term "quasicommutative" has also been used in a different sense (see [18]).

Potter's principal result is the following theorem [19]:

Theorem 1 (Potter). *Let A and B be complex square matrices satisfying (1), where ω is a primitive qth root of unity. Then*

$$A^q + B^q = (A + B)^q. \tag{2}$$

In his note, Potter proves Theorem 1 by deriving it from the general expansion of $(x + y)^q$ for any nonnegative integer q and ω-commutative x and y for arbitrary complex ω. This formula, which we state as (3)–(5), involves the classical q-binomial coefficients and is currently referred to as the *noncommutative q-binomial theorem* (see, for example, [1, Formula 10.0.2] or [12, Exercise 1.35], but beware: the q in the last sentence is our ω). The fact that (3)–(5) hold for ω-commutative operators is generally attributed to Schützenberger [20]. In this context we call the result described by (3)–(5) the *Potter-Schützenberger* formula. It is of considerable interest in the study of quantum groups (see, for example, [16, p. 75]). In fact, Potter's proof shows that it holds under very general conditions, which we examine in section 2.

Potter cites and applies results from the book by Turnbull and Aitken [22, p. 148], where a matrix X satisfying $AX = XC$ is called a *commutant* of A and C. There all commutants of A and C are determined under the assumption that A and C are in Jordan canonical form. If A and B are ω-commutative, then clearly B is a commutant

of A and ωA. The general question of commutants was also considered by Goddard and Schneider [14]. One might observe that all the mathematicians mentioned in this paragraph were in Scotland in the early 1950s. Figure 1 shows the participants of the 1951 Edinburgh Mathematical Society Colloquium at St. Andrews [26]. Four of the mathematicians who turn up in the present article are pictured.

Figure 1. H. S. A. Potter is third from the right in the front row; A. C. Aitken and H. W. Turnbull are tenth and eleventh from the left, respectively, in the second row (seated); H. Schneider is the fourteenth from the left in the fourth row. (This photo is used with the permission of the MacTutor History of Mathematics Archive.)

Wielandt's proof of Theorem 1 is reproduced and translated in section 3. We comment on it and present a variant form of it in section 4. Wielandt's proof uses matrix theory nontrivially and is based on an insightful observation. However, it relies heavily upon the assumption that ω is a primitive qth root of 1, and there is no obvious way of obtaining the more general Theorem 2 using his methods.

In his diary, following the proof of Potter's theorem, Wielandt also raises some questions. These include the construction of all identities satisfied by ω-commutative matrices and the determination of all irreducible pairs of ω-commutative matrices. Naturally unaware of Wielandt's question, M. P. Drazin (then in Cambridge, England) to a large extent furnished an answer to the latter question in [9].

In section 5 we present the prenormal form obtained by Drazin [9] and show that the classification problem of ω-commutative matrices is equivalent to the classification problem of pairs of commuting matrices, both under simultaneous similarity. In section 6 we demonstrate that the converse to Potter's theorem does not hold, not even for some of its weakened versions. In section 7 we determine all polynomial identities satisfied by ω-commutative matrices thus answering Wielandt's first question. Finally, in section 8 we discuss work on ω-commutative matrices preceding that of Potter and Wielandt.

2. POTTER'S PROOF. We begin by examining Potter's proof of Theorem 1. In the first part of the proof Potter does not assume that ω is a root of unity, and for ω-commutative matrices A and B he establishes the general formula (here stated in a slightly different but equivalent form)

$$(A + B)^q = \sum_{k=0}^{q} c_k B^k A^{q-k}, \tag{3}$$

© THE MATHEMATICAL ASSOCIATION OF AMERICA [Monthly 111

where the c_k are determined by

$$\phi_k \phi_{q-k} c_k = \phi_q \qquad (k = 0, \dots, q) \tag{4}$$

and the ϕ_k are given by

$$\phi_k = \prod_{s=1}^{k} (1 + \dots + \omega^{s-1}) \qquad (k = 0, \dots, q). \tag{5}$$

The coefficients c_k in (4), known as the *q-binomial coefficients*, were well studied in the nineteenth century in the theory of hypergeometric series (see, for example, [1, chap. 10]) and in the theory of partitions (see [1, chap. 11] and [21, sec. 1.3]).

Let R be any ring with identity 1, and let ω, x, and y be elements of R. Let $\overline{Z}$ signify the subring of R generated by 1. Thus $\overline{Z}$ is isomorphic either to the ring of integers Z or to Z_m, the ring of integers modulo m. We call x and y ω-commutative if the following three identities hold:

$$\omega x = x \omega, \qquad \omega y = y \omega, \qquad xy = \omega yx. \tag{6}$$

From Potter's argument we can obtain the following version of the Potter-Schützen-berger theorem:

Theorem 2. *Let R be a ring with 1, and let ω be an element of R. If x and y are ω-commutative elements of R, then*

$$(x + y)^q = \sum_{k=0}^{q} c_k y^k x^{q-k}, \tag{7}$$

where the c_k are given by (4) and (5).

We observe that the coefficients c_k lie in $\overline{Z}[\omega]$, so there is no loss of generality by considering only the subring $\overline{Z}[\omega, x, y]$ of R. As a consequence of Theorem 2, we obtain:

Corollary 3. *Suppose that $\overline{Z}[\omega]$ is an integral domain. Under the conditions of Theorem 2, suppose further that*

$$\phi_k \neq 0 \quad (k = 1, \dots, q-1) \tag{8}$$

but that

$$\phi_q = 0. \tag{9}$$

Then

$$(x + y)^q = x^q + y^q. \tag{10}$$

Evidently, if $\overline{Z}[\omega]$ is a field and ω is a primitive qth root of 1 in R, then (8) and (9) are satisfied. These statements also hold if $\overline{Z} = Z_q$ and $\omega = 1$. We further note that when $\overline{Z}[\omega]$ is an integral domain a necessary condition for (10) to be satisfied is that (9) be true.

A ring $K[x, y]$ in two noncommutative indeterminates x and y over a central field K, where x and y are subject to the relation $xy = \omega yx$ for some ω in K, is today called a *quantum plane* over K (see [16, p. 72]).

3. WIELANDT'S NOTES.

In Figure 2 we display a facsimile of page 35 of Wielandt's Diary VII [24], dated 20 March 1951. The following is a translation of this note:

Quasi-commutative Matrices
New proof of a theorem of H. S. A. Potter (On the latent roots of quasi-commutative matrices, Amer. Math. Monthly 57, 321–322 (1950)).
If $AB = \omega BA$, and ω is a primitive qth root of unity, then $(*) \ (A + B)^q = A^q + B^q$.

Figure 2. Page 35 of Wielandt's diary VII (reproduced with the permission of Annemarie Wielandt).

Proof: It is clear that $(A + B)^q = A^q + B^q + \sum_1^{q-1} c_k A^k B^{q-k}$, where the scalars c_k do not depend on the special choice of of A, B (except for $*$). If one chooses

$$
A = \begin{pmatrix} & 1 & & & \\ & & 1 & & \\ & & & 1 & \\ & & & & 1 \\ 1 & & & & \end{pmatrix}, \quad
B = \begin{pmatrix} 1 & & & & \\ & \omega & & & \\ & & \omega^2 & & \\ & & & \ddots & \\ & & & & \omega^{q-1} \end{pmatrix}, \quad (degree = q)
$$

then $(A + B)^q = \delta E$ is a diagonal matrix: Indeed, with $T = BA^{-1}$, it follows that $T^{-1}AT = \omega A$, $T^{-1}BT = \omega B$, and thus $T^{-1}(A + B)T = \omega(A + B)$, hence[1] $A + B$ has eigenvalues $\rho, \rho\omega, \ldots, \rho\omega^{q-1}$. Not all of them are zero due to the trace; the matrix $(A + B)^q$ has only ρ^q as eigenvalue and is, just like $(A + B)$, transformable to diagonal form, so $(A + B)^q = \rho^q E$. This implies that $c_1 = \cdots = c_{q-1} = 0$, since otherwise a particular nondiagonal element of $(A + B)^q$ would be nonzero.

    Problem: Determine all identities for A, B. Is every pair A, B with $(A + B)^q = A^q + B^q$ decomposable as

$$
S^{-1}AS = \begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_r \end{pmatrix}, \quad
S^{-1}BS = \begin{pmatrix} B_1 & & \\ & \ddots & \\ & & B_r \end{pmatrix}
$$

with

$$
A_\rho B_\rho = \omega_\rho B_\rho A_\rho, \quad \omega_\rho^q = E?
$$

Determination of all irreducible pairs of quasi-commutative matrices?
Earlier work on quasi-commutative matrices?

Apparently, in the first paragraph of the proof Wielandt is referring to the quasi-commutativity relation rather than to relation $(*)$ when he says "except for $(*)$." Also, he must mean that $\omega_\rho^q = 1$, for ω_ρ is a scalar.

4. WIELANDT'S PROOF AND A VARIANT.
Wielandt's proof begins with the simple but insightful remark that for ω-commutative matrices the coefficients c_k in the expansion $(A + B)^q = \sum_{k=0}^q c_k B^k A^{q-k}$ are independent of the particular matrices A and B, hence that the result will follow if he can show that the coefficients must be 0 in the case of a well-chosen pair of matrices A and B. The argument requires the linear independence of the set of matrices $B^k A^{n-k}$ $(k = 1, \ldots, q - 1)$. Though Wielandt does not say this, he chooses matrices A and B that satisfy this condition. He then uses an argument involving eigenvalues and the diagonalizability of matrices to show that $c_1 = \cdots = c_{q-1} = 0$.

We now give a variant of Wielandt's proof. Let A and B be the matrices chosen by Wielandt, and let s and t be arbitrary complex numbers. Since the eigenvalues of B are the qth roots of unity, it follows that the characteristic polynomial of sB is $\lambda^q - s^q$. Because the proper principal minors of $sA + tB$ and tB coincide and $\det(sA + tB) = (-1)^{q-1}(s^q + t^q)$, it follows that the characteristic polynomial of $sA + tB$ is $\lambda^q - (s^q + t^q)$. By the Cayley-Hamilton theorem [11] we obtain

$$
(sA + tB)^q = (s^q + t^q)I = (sA)^q + (tB)^q. \tag{11}
$$

---

[1] This implication is based on the primitivity of ω as a root of unity.

But

$$(sA + tB)^q = (sA)^q + (tB)^q + \sum_{k=1}^{q-1} c_k s^k t^{q-k} A^k B^{q-k},$$

so each matrix coefficient $c_k A^k B^{q-k}$ must be zero. Since A and B are both nonsingular, this implies that $c_k = 0$. ∎

Thus, the alternative proof demonstrates the following extension of Potter's theorem:

Proposition 4. *Let A and B be ω-commutative matrices satisfying (1), where ω is a primitive qth root of unity. Then*

$$(sA + tB)^q = (sA)^q + (tB)^q \tag{12}$$

for all complex numbers s and t.

A proof in a rather similar spirit is given by R. Bhatia and L. Elsner in [4] for the following fact: if A and B are ω-commutative, then the spectrum of $A + B$ is p-Carrollian (i.e., the eigenvalues of $A + B$ can be enumerated as

$$(\lambda_1, \ldots, \lambda_r, \omega\lambda_1, \ldots, \omega\lambda_r, \ldots, \omega^{p-1}\lambda_1, \ldots, \omega^{p-1}\lambda_r)).$$

Moreover, the same holds for all perturbations of B of the specific form given in [4, Theorem 2]. The term "Carrollian" was coined by R. Bhatia in honor of Lewis Carroll, initially to denote an n-tuple that contains $-x$ if it contains x, and later turned into "p-Carrollian" for n-tuples that contain all multiples of x with pth roots of unity. It is used in [3] and [2] as well.

5. NORMAL FORMS FOR ω-COMMUTATIVE MATRICES.

Note that Wielandt raises the question of classifying irreducible ω-commutative pairs, having in mind reductions by simultaneous similarity $A \mapsto T^{-1}AT$ and $B \mapsto T^{-1}BT$, since relation (1) is invariant under simultaneous similarity. To study this question, we start with some preliminary observations.

Suppose that

$$AB = \omega BA, \tag{13}$$

where ω is a nonzero complex number. By the foregoing remark on simultaneous similarity, we may assume that A is in Jordan canonical form. Notice that not all Jordan normal forms are allowed for the matrix A. For example, if B is nonsingular and A is not nilpotent, then each row and each column of B must contain at least one nonzero element. Thus, if λ is a nonzero eigenvalue of A, so is $\omega\lambda$. Since the number of eigenvalues is finite and A has a nonzero eigenvalue, it follows that ω is a root of unity. Moreover, if $J_i(\lambda_i)$ is the Jordan block of largest size in A, then using the fact that every row and column of B has at least one nonzero element, it follows that there is a block for $\omega\lambda_i$ of equal size. Thus, we conclude that the maximal size of a Jordan block in A is the same for each nonzero eigenvalue. In case both A and B are nonsingular, the Jordan forms of both are restricted in this way.

© THE MATHEMATICAL ASSOCIATION OF AMERICA [Monthly 111

Theorem 5. *If A and B are nonsingular matrices satisfying $AB = \omega BA$ with $\omega \neq 0$, then ω is a primitive pth root of 1 for some p and the Jordan form J of A may be written as*

$$J = \mathrm{diag}(K_1, \ldots, K_q),$$

where each K_i is the direct sum of Jordan blocks of the same size corresponding to the eigenvalues $\lambda_i, \omega\lambda_i, \ldots, \omega^{p-1}\lambda_i$. The Jordan form for B has the same structure.

Proof. From the assumption of the theorem we infer that A is similar to ωA. Hence the number and sizes of Jordan blocks corresponding to any eigenvalue λ of A coincide with the number and sizes of the Jordan blocks corresponding to the eigenvalue $\omega\lambda$. The proof for B follows by interchanging the roles of A and B. ∎

Assuming that A is already in its Jordan form, one can obtain a full description of all matrices B satisfying (13) using results from [22, p. 21]. However, in this way we will not have obtained a canonical form for the pair (A, B) under simultaneous similarity, for in general there will be similarities that leave A invariant but change B. Already in 1951, Drazin to a large extent answered the question of classification of ω-commutative pairs, although Wielandt was apparently unaware of his results. In [9], Drazin obtained the following prenormal form for pairs of ω-commutative matrices:

Theorem 6. *If A and B are $n \times n$ matrices satisfying an equation of the form $AB = \omega BA$, then either*

(i) *A and B can be simultaneously reduced to triangular form by a similarity transformation,*

or

(ii) *there is an integer r $(0 \leq r \leq n - 2)$ such that A and B can be reduced, by the same similarity transformation, to the forms*

$$\begin{bmatrix} S & X \\ 0 & A_r \end{bmatrix}, \quad \begin{bmatrix} T & Y \\ 0 & B_r \end{bmatrix}, \tag{14}$$

where S and T are triangular $r \times r$ matrices, and A_r and B_r are nonsingular $(n - r) \times (n - r)$ matrices.

Furthermore, Drazin also proved an additional theorem:

Theorem 7. *If (i) holds in Theorem 6 with $\omega \neq 1$, then each of AB and BA is nilpotent, and A and B have between them at least n zero eigenvalues. If, however, (i) is false, then ω is necessarily a primitive root of unity, and the order k of ω must divide $n - r$. Further, in this case ST and TS are both nilpotent, and the reduction of A and B can be effected in such a way that A_r takes the form*

$$\begin{bmatrix} a & & & \\ & \omega a & & \\ & & \ddots & \\ & & & \omega^{k-1}a \end{bmatrix}, \tag{15}$$

where a is a nonsingular square matrix of order $(n - r)/k$; then the most general form of B_r is

$$\begin{bmatrix} 0 & 0 & \ldots & 0 & b_1 \\ b_2 & 0 & \ldots & 0 & 0 \\ & & \ddots & & \\ 0 & 0 & \ldots & b_k & 0 \end{bmatrix}, \tag{16}$$

where $b_1, \ldots, b_k$ are arbitrary nonsingular matrices of order $(n - r)/k$ subject to the relations $b_i a = ab_i$ $(i = 1, 2, \ldots, k)$, S and T are triangular $r \times r$ matrices, and A_r and B_r are nonsingular $(n - r) \times (n - r)$ matrices.

Drazin's formulas still do not give a canonical form. Indeed, first of all an ω-commutative pair (A, B) can be put into block diagonal form, with blocks (A_i, B_i) of one of four types, according to their spectra:

Type I: $\sigma(A_i) = \{0\}, \quad \sigma(B_i) = \{0\};$

Type II: $\sigma(A_i) = \{0\}, \quad \sigma(B_i) = \{\mu_i(\neq 0), \omega\mu_i, \ldots, \omega^{k-1}\mu_i\};$

Type III: $\sigma(A_i) = \{\lambda_i(\neq 0), \omega\lambda_i, \ldots, \omega^{k-1}\lambda_i\}, \quad \sigma(B_i) = \{0\};$

Type IV: $\sigma(A_i) = \{\lambda_i(\neq 0), \omega\lambda_i, \ldots, \omega^{k-1}\lambda_i\},$

$\sigma(B_i) = \{\mu_i(\neq 0), \omega\mu_i, \ldots, \omega^{k-1}\mu_i\}.$

Moreover, for a pair of type IV one can assume that all submatrices b_i in (16) save one (b_1, say) are equal to the identity. At that stage the problem of classification of ω-commutative matrices of type IV reduces to the problem of classification of commuting pairs (a, b_1). The same reduction can be achieved for types II and III. We can summarize this as follows (see [15] for a proof):

Theorem 8. *The problem of representation under simultaneous similarity for ω-commutative pairs is equivalent to the problem of representation under simultaneous similarity for all commuting pairs. Moreover, the latter is already equivalent to the problem of representation for ω-commutative pairs of type II, III, or IV.*

The problem of classification of commuting matrices pairs is quite fascinating and notorious. M. Gelfand and V. A. Ponomarev [13] showed that the simultaneous similarity problem of *any n-tuple* of matrices is equivalent to it. The problem was later taken up by S. Friedland, who showed in [10] how to find a finite number of invariants that will characterize an orbit of a pair (A, B) under simultaneous similarity up to a finite ambiguity, which means that these invariants may characterize a finite number of similarity orbits. For a fixed dimension d, Friedland decomposed the variety of pairs of square matrices into finitely many subvarieties locally closed under simultaneous similarity. For each such subvariety Z, he found a rational map f from Z into a finite-dimensional vector space V for which the preimages under f (of points in V) consist of finitely many orbits of matrix pairs. The map f and the space V depend strongly on Z, although, for a fixed f, there is an upper bound on the number of conjugation classes in each preimage. Friedland's method was later refined by K. Bongartz in [5]. He modified Friedland's construction (by changing Z, f, and V) so that the preimages under f are exactly the individual orbits of pairs of matrices.

In other words, given two pairs (A, B) and (C, D) of matrices, they are simultaneously similar to each other if and only if they lie in the same Z and have the same image under f. This provides, at least in principle, a complete answer to the problem of simultaneous similarity (i.e., it gives a decision algorithm, but no readily available normal forms).

We should also mention that one of the abstract versions of this problem is to find all isomorphism classes of cyclic modules of finite length over the commutative polynomial ring $R = \mathbb{C}[x, y]$. A pair of commuting $n \times n$ matrices A and B defines an R-module structure on $\mathbb{C}^n$ by letting x and y be multiplication by A and B, respectively.

6. THE CONVERSE TO POTTER'S THEOREM. Having studied the decomposition of ω-commutative matrices into blocks, we now discuss Wielandt's second question: Does the converse of Potter's theorem hold for every irreducible block (i.e., does the relation (12) with $s = t = 1$ imply that

$$AB = \omega BA \tag{17}$$

for some qth root of unity ω)?

If $q = 2$, this strong version of the converse does hold:

Proposition 9. *A pair (A, B) is ω-commutative with $q = 2$ if and only if (12) holds with $s = t = 1$.*

Proof. The condition $A^2 + B^2 = (A + B)^2$ is equivalent to $AB = -BA$. ∎

However, the converse is in general not true, even if (12) is assumed to hold for all values of s and t, as the following example shows.

Example 10. Consider the case $n = 3$ and $q = 3$. Let λ be a complex number different from $0, -1$, and the primitive third roots of unity. For the pair of matrices

$$A = \begin{bmatrix} 0 & 0 & \frac{1}{\lambda} \\ 1 & 0 & 0 \\ 0 & -\frac{\lambda+1}{\lambda} & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \tag{18}$$

we have

$$BA = EAB, \tag{19}$$

where

$$E = \begin{bmatrix} \lambda & 0 & 0 \\ 0 & -\frac{\lambda+1}{\lambda} & 0 \\ 0 & 0 & -\frac{1}{\lambda+1} \end{bmatrix}.$$

Moreover, since E is invertible, it follows that

$$A(E^{-1} + I) = -EA, \quad (E^{-1} + I)B = -BE. \tag{20}$$

But (19) and (20) imply that

$$(A + tB)^3 = A^3 + t^3 B^3$$

for all t. However, since E has three distinct eigenvalues it follows that (17) does not hold. Also, if the pair (A, B) is replaced by $(\tilde{A}, \tilde{B}) := T(A, B)T^{-1}$, then $\tilde{B}\tilde{A} = \tilde{E}\tilde{A}\tilde{B}$, where $\tilde{E} := TET^{-1}$. Since E and $\tilde{E}$ have the same spectrum, (17) does not hold for the pair $(\tilde{A}, \tilde{B})$ either. In other words, the pair (A, B) cannot be reduced to a direct sum of ω-commutative pairs. Note that in this example both matrices A and B are nonsingular.

If we assume that $s = t = 1$ in (12), then we can even produce 2×2 counter-examples, for example, with $q = 3$ (see [15]). Drazin's prenormal form for ω-commutative matrices also raises the question of whether the converse to Potter's theorem at least holds for pairs of matrices of the form (14)(i). However, there are again examples demonstrating that this is not the case [15].

On the other hand, if a pair (A, B) of nonsingular block $k \times k$ matrices of the form (15)–(16) with some ω satisfies (12), then necessarily ω is a primitive qth root of unity, equality (17) holds, and q divides k. Indeed, suppose that a pair (A, B) is in the form (15)–(16), where a and b_j are nonsingular commuting matrices and $\omega \neq 0$. If the pair (A, B) satisfies (12), then the matrix coefficient of $s^{q-1}t$ must be zero, that is

$$A^{q-1}B + A^{q-2}BA + \cdots + BA^{q-1} = 0. \tag{21}$$

On the other hand, by direct calculation,

$$BA^j = E(\omega^j)A^j B \quad (j = 0, \ldots, q-1),$$

where

$$E(\alpha) := \operatorname{diag}\left(\alpha^{k-1}I, \frac{1}{\alpha}I, \frac{1}{\alpha}I, \ldots, \frac{1}{\alpha}I\right).$$

Since $E(\alpha)$ is block-diagonal for any $\alpha \neq 0$, the matrices A and $E(\omega^j)$ commute. These rules now allow us to interchange A^j and B in (21) to obtain

$$A^{q-1}B + A^{q-2}BA + \cdots + BA^{q-1} = \left(I + E(\omega) + E(\omega^2) + \cdots + E(\omega^{q-1})\right)A^{q-1}B = 0.$$

Since A and B are nonsingular, this implies that the matrix factor in front of $A^{q-1}B$ must be zero, hence that $1 + 1/\omega + 1/\omega^2 + \cdots + 1/\omega^{q-1} = 0$ (i.e., $\omega^q = 1$). Next, the coefficient of st^{q-1} must also be zero, which can be similarly shown to imply that $\omega^k = 1$ (although the interchange rules for B^j and A are a bit more involved). This implies, in turn, that (17) holds. Now, ω is a primitive root of order q', with q' dividing q. Then relation (17), which we just established, implies that the matrices $A^{q'}$ and $B^{q'}$ commute and that

$$(sA + tB)^q = \left((sA)^{q'} + (tB)^{q'}\right)^{q/q'} = (sA)^q + (tB)^q,$$

which is possible only when $q' = q$. Since the relations $\omega^q = \omega^k = 1$ imply that $\omega^{\gcd(q,k)} = 1$, we also conclude that q must divide k.

But even the commutativity of the blocks a and b_i in (15)–(16) does not follow automatically from the relation (12), as we also show in [15]. In view of these counterexamples, it seems natural to pose the more general problem of characterizing all classes of matrices for which the equivalence of (12) and (17) holds.

Finally, recall that Potter derived from (17) infinitely many identities (3)–(5). Wielandt, in effect, asked whether just one of these identities, with q such that $\omega^q = 1$,

© THE MATHEMATICAL ASSOCIATION OF AMERICA [Monthly 111

already implies ω-commutativity. From this point of view, it is not very surprising that the answer to his question turns out to be no. Notice, however, that (3) with $q = 2$ is exactly equivalent to (17). We believe the value $q = 2$ is the only one for which the expansion given by (3)–(5) implies ω-commutativity (17).

7. IDENTITIES SATISFIED BY ω-COMMUTATIVE MATRICES.

The first question Wielandt asked was which identities are satisfied by ω-commutative matrices. We now show that the polynomial identities $f(x, y) = 0$ that hold for all ω-commutative matrices have $f(x, y)$ in the ideal in $\mathbb{C}[x, y]$ generated by the polynomial $g(x, y) := xy - \omega yx$.

Theorem 11. *Let $\mathbb{C}[x, y]$ denote the ring of polynomials in noncommuting indeterminates x and y over the field $\mathbb{C}$, and let $\mathcal{I}$ denote the ideal of $\mathbb{C}[x, y]$ generated by the polynomial $g(x, y) := xy - \omega yx$, where $\omega^q = 1$. Then $f(x, y)$ belongs to $\mathcal{I}$ if and only if the condition (17) implies that $f(A, B) = 0$ in $\mathbb{C}^{q \times q}$.*

Proof. One direction is obvious: any polynomial $f(x, y)$ from $\mathcal{I}$ satisfies $f(A, B) = 0$ for all ω-commutative matrices A and B.

To establish the converse, first recall that the condition $\omega^q = 1$ implies that there exists a pair of *nonsingular* matrices A and B in $\mathbb{C}^{q \times q}$ satisfying (17). Since the pair (sA, tB) also satisfies (17) for any scalars s and t, we see that $f(sA, tB) = 0$. Now interchange A and B in $f(sA, tB)$ using relation (17) as many times as necessary to obtain a polynomial in the form

$$f_1(sA, tB) := \sum_{i,j} c_{ij} s^i t^j A^i B^j.$$

The polynomials $f(x, y)$ and $f_1(x, y)$ differ by some element of $\mathcal{I}$. Now, since $f_1(sA, tB) = 0$ and since s and t are independent scalars, each term $c_{ij} s^i t^j A^i B^j$ in the sum must be zero. But as both A and B are nonsingular, this shows that $c_{ij} = 0$. Thus $f_1(x, y)$ is the zero polynomial, placing $f(x, y)$ in $\mathcal{I}$. ∎

8. FURTHER HISTORICAL COMMENTS.

We now address the last question posed by Wielandt, the question about work predating that of H. S. A. Potter. In fact, M. P. Drazin briefly addressed this issue in [9]. Specifically, Drazin cites Cayley's paper [6], where the case $\omega = -1$ was considered, and the works of F. Cecioni [7], S. Cherubino [8], and T. Kurosaki [17] devoted to the general case. (Biographies of the two Italian mathematicians can be found at [25].) Cecioni's paper is a memoir summarizing and extending results on ω-commutative matrices known at that time. He gives a condition on a matrix A that is necessary and sufficient for the equation $AX = \omega XA$ to have a nonzero solution X, describes the structure of an arbitrary solution along the lines of Turnbull and Aitken [22, p. 148], and stops one step short of arriving at the formulas (15)–(16) for an ω-commutative pair (A, B) with AB nonsingular. A slightly different prenormal form is derived by Cherubino [8], who also describes the structure of the algebra of matrices commuting with a given matrix. The pair (15)-(16) also appears in Kurosaki [17], even in reduced form (with all b_j except for one equal to the identity), although not in a formal statement. Kurosaki's main result [17, Theorem 4] is a description of the group of all nonsingular matrices P satisfying the equation $AP = cPA$ for some c (depending on P) and a fixed nonsingular matrix A. Drazin, on the other hand, is apparently mostly interested in simultaneous triangularization of an ω-commutative pair, which leads him to derive, in his remarkably short paper [9], the prenormal form described in section 5.

REFERENCES

1. G. E. Andrews, R. Askey, and R. Roy, *Special functions*, Encyclopedia of Mathematics and its Applications, no. 71, Cambridge University Press, Cambridge, 1999.
2. R. Bhatia, *Matrix Analysis*, Springer-Verlag, New York, 1997.
3. R. Bhatia and L. Elsner, Symmetries and variation of spectra, *Canad. J. Math.* **44** (1992) 1155–1166.
4. ———, The q-binomial theorem and spectral symmetry, *Indag. Mathem. N.S.* **4** (1993) 11–16.
5. K. Bongartz, A remark on Friedland's stratification of varieties of modules, *Comm. Algebra* **23** (1995) 2163–2165.
6. A. Cayley, A memoir on the theory of matrices, *Philos. Trans.* **148** (1858) 17–37; also in *Collected Works*, vol. 2, Cambridge University Press, Cambridge, 1889, pp. 475–496.
7. F. Cecioni, Sull' equazione fra matrici $AX = \varepsilon XA$, *Ann. Univ. Toscane* **14** (1931) 1–49.
8. S. Cherubino, Sulle omagrafic permutabili, *Rend. Semin. mat. Roma (4)* **2** (1938) 14–46.
9. M. P. Drazin, A reduction for the matrix equation $AB = \varepsilon BA$, *Proc. Camb. Philos. Soc.* **47** (1951) 7–10.
10. S. Friedland, Simultaneous similarity of matrices, *Advances Math.* **50** (1983) 189–265.
11. F. R. Gantmacher, *The Theory of Matrices*, vol. 2, Chelsea, New York, 1959.
12. G. Gasper and M. Rahman, *Basic Hypergeometric Series*, Encyclopedia of Mathematics and its Applications, no. 35, Cambridge University Press, Cambridge, 1990.
13. M. Gelfand and V. A. Ponomarev, Remarks on the classification of a pair of commuting linear transformations in a finite dimensional vector space, *Functional Anal. Appl* **3** (1969) 325-326.
14. L. S. Goddard and H. Schneider, Pairs of matrices with a non-zero commutator, *Proc. Camb. Philos. Soc.* **51** (1955) 551–553.
15. O. Holtz, V. Mehrmann, and H. Schneider, Potter, Wielandt, and Drazin on the matrix equation $AB = \omega BA$, with some new answers to old questions (preprint 10-2003, Institute of Mathematics, Technische Universität Berlin, 2003; available online at http://www.math.tu-berlin.de/preprints/).
16. C. Kassel, *Quantum Groups*, Springer-Verlag, New York, 1995.
17. T. Kurosaki, Über die mit einer Kollineation vertauschbaren Kollineationen, *Proc. Imp. Acad. Tokyo* **17** (1941) 24–28.
18. N. McCoy, On quasicommutative matrices, *Trans. Amer. Math. Soc.* **36** (1934) 327–340.
19. H. S. A. Potter, On the latent roots of quasi-commutative matrices, this MONTHLY **57** (1950) 321-322.
20. M. P. Schützenberger, Une interprétation de certaines solutions de l'Équation fonctionnelle: $F'(x + y) = F(x)F(y)$, *C. R. Acad. Sci. Paris* **236** (1953) 352–353.
21. R. P. Stanley, *Enumerative Combinatorics*, vol. 1, Wadsworth & Brooks/Cole, Monterey, CA, 1986.
22. H. W. Turnbull and A. C. Aitken, *An Introduction to the Theory of Canonical Matrices*, Blackie, London, 1932.
23. H. Wielandt, Review of [19], *Zbl. M.* **40** (November, 1951) 5.
24. ———, Diary VII, 1951 (to appear in transcribed form).
25. History of Italian Mathematics *Matematica - Storia - Indice*, http://matematica.uni-bocconi.it/indice.html
26. The MacTutor History of Mathematics Archive, Photograph of the participants of the Edinburgh Mathematical Society 1951 Colloquium at St.Andrews, http://www-history.mcs.st-and.ac.uk/history/Societies/EMS_colloq/EMS_1951.html

OLGA HOLTZ received her Diploma in applied mathematics from Chelyabinsk State Technical University, Russia, in 1995 and her Ph.D. in mathematics from the University of Wisconsin-Madison in 2000. She is primarily interested in matrix analysis and approximation theory. Her nonmathematical interests include classical piano and voice as well as ballroom dance and ballet. She spent the 2002–2003 academic year at TU Berlin as a Humboldt research fellow and is now an assistant professor at the University of California-Berkeley.
Department of Mathematics, University of California-Berkeley, 821 Evans Hall, Berkeley, CA 94720, USA
holtz@Math.Berkeley.EDU

VOLKER MEHRMANN received his Diplom in mathematics in 1979, his Ph.D. in 1982, and his Habilitation in 1987 from the University of Bielefeld, Germany. He spent research years at Kent State University

in 1979–1980, at the University of Wisconsin in 1984–1985, and at IBM Research Center in Heidelberg in 1988–1989. After spending the years 1990–1992 as a visiting full professor at the RWTH Aachen, he was a full professor at TU Chemnitz from 1993 to 2000. Since then he has been a full professor at TU Berlin. His research interests are in the areas of numerical mathematics/scientific computing, applied and numerical linear algebra, control theory, and differential algebraic equations.

Institut für Mathematik, MA 4-5, Technische Universität Berlin, D-10623 Berlin, Germany.
mehrmann@math.TU-Berlin.DE

HANS SCHNEIDER, born in 1927 in Vienna, Austria, left that country rather precipitously to take up the life of a Scottish schoolboy in 1939. In due course, he entered Edinburgh University, obtaining an M.A. (1948) and a Ph.D. (1952) with a thesis on nonnegative matrices under that genius, A. C. Aitken. This led to a lifelong addiction to linear algebra. After seven years at Queen's University, Belfast, he migrated to the University of Wisconsin, and now boasts of being J. J. Sylvester Professor Emeritus. He has published about 140 research papers, joint with about half that many coauthors, and has produced three physical children (joint with his wife Miriam) and seventeen mathematical children, of all of whom he is proud. He was the first president of the International Linear Algebra Society, and his stint as editor-in-chief of *Linear Algebra and its Applications* is as yet unfinished after thirty-two years. He likes to walk barefoot on Lanikai beach in the light of the full moon and wishes it would happen more often.

Department of Mathematics, University of Wisconsin, Madison, WI 53706, USA.
hans@math.wisc.edu