Technische Universität Berlin

Institut für Mathematik

Potter, Wielandt, and Drazin on the matrix equation $AB = \omega BA$, with some new answers to old questions

O. Holtz, V. Mehrmann and H. Schneider

Preprint 10/2003

Preprint-Reihe des Instituts für Mathematik
Technische Universität Berlin

Report 10/2003 **April 2003**

Potter, Wielandt, and Drazin on the matrix equation $AB = \omega BA$, with some new answers to old questions

Olga Holtz [*] Volker Mehrmann [†] Hans Schneider [‡]

13.04.2003

Abstract

In this partly historical and partly research oriented note, we display a page of an unpublished mathematical diary of Helmut Wielandt's for 1951. There he gives a new proof of a theorem due to H. S. A. Potter on the matrix equation $AB = \omega BA$, which is related to the q-binomial theorem, and asks some further questions, which we answer. We also describe results by M. P. Drazin and others on this equation.

1 Introduction

The aim of this paper is to present a slice of the linear algebra of the 1950's and to give some answers to questions raised then.

It was Helmut Wielandt's habit over many years to make notes in what he called diaries (Tagebücher) on papers that interested him. Many notes are essentially summaries of a paper, but in other cases Wielandt would add questions, ideas, or even further results. In this note we discuss one such entry which appears on page 35 of Diary VII (1951) which will appear in transcribed electronic form [22]. The entry concerns a paper which Wielandt reviewed for the Zentralblatt. We next turn to this paper.

In 1950, H. S. A. Potter, a mathematician at Aberdeen University in Scotland, published a note in the American Mathematical Monthly [18], on the matrix equation

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

He called a pair of complex $n \times n$ matrices A, B satisfying (1) *quasi-commutative*. We shall call matrices (1) ω-commutative, see Section 2 for a definition of this term applicable to

[*]Inst. f. Mathematik, MA 4-5, Technische Universität Berlin, D-10623 Berlin, Fed. Rep. Germany. E-mail: holtz@math.TU-Berlin.DE On leave from CS Dept., Univ. of Wisconsin, Madison, WI 53706, USA. Supported by Alexander von Humboldt Foundation.

[†]Inst. f. Mathematik, MA 4-5, Technische Universität Berlin, D-10623 Berlin, Fed. Rep. Germany. E-mail: mehrmann@math.TU-Berlin.DE Supported by DFG research grant Me790/15.

[‡]Dept. of Mathematics, Univ. of Wisconsin, Madison, WI 53706, USA. E-mail: hans@math.wisc.edu Supported by DFG research grant Me790/15.

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.

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 [17].

Potter's principal result is the following theorem:

Theorem 1 (Potter [18]) *Let A and B be complex square matrices satisfying (1) where ω is a primitive q-th root of unity. Then*

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

In his note, Potter proves his theorem 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, e.g., [1, Formula 10.0.2] or [12, Exercise 1.35]. [But care: the q in the last sentence is our ω.]. The result that (3)–(5) holds for ω-commutative operators is generally attributed to Schützenberger [19]. We shall call (3)–(5) the *Potter-Schützenberger* formula. It is of considerable interest in the study of quantum groups, see for example [15, p.75]. In fact, Potter's proof shows that it holds under very general conditions, which we examine in Section 2.

Potter refers and applies results in the book by Turnbull-Aitken [21, p.148] where a matrix X satisfying $AX = XC$ is called *commutant* of A and C. There all commutants of A and C are determined on the assumption that A and C are in Jordan canonical form. If A and B are quasi-commutative, then clearly B is a commutant of A and ωA. The general question of commutants was also considered by Goddard-Schneider [14]. One might observe that all the mathematicians mentioned in this paragraph were in Scotland in the early 1950's. Figure 1 shows the participants of the 1951 Edinburgh Mathematical Society Colloquium at St.Andrews [24]. There four mathematicians mentioned in our article are present.

Wielandt's proof of Potter's Theorem 1 is reproduced and translated in Section 3. We comment on it and give a variant in Section 4. This proof uses matrix theory non-trivially and it is based on an insightful observation. However, it heavily uses the assumption that ω is a primitive q-th 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 at Cambridge, England, essentially answers the latter question in [9].

In Section 5 we take up the question of normal forms for pairs of quasi-commutative matrices. We present the pre-normal form obtained by Drazin [9] and show that the classification problem of quasi-commutative matrices is equivalent to the classification problem of pairs of commuting matrices, both under simultaneous similarity.

In Section 6 we present counterexamples showing 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 quasi-commutative matrices thus answering Wielandt's first question.

Finally, in section 8 we discuss work on quasi-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, B he proves the general formula (here stated in a slightly different but equivalent form)

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

where the c_k are determined by

$$\phi_k \phi_{q-k} c_k = \phi_q \quad , k = 0, \ldots, q, \tag{4}$$

and the ϕ_k are given by

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

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, Chapter 10] and in the theory of partitions combinatorics, see [1, Chapter 11] and [20, Section 1.3].

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

$$\begin{aligned} \omega x &= x\omega \\ \omega y &= y\omega \\ xy &= \omega y x. \end{aligned} \tag{6}$$

By Potter's argument we may obtain the following version of the Potter-Schützenberg theorem.

Theorem 2 *Let R be a ring with 1. Let $\omega \in R$ and let x and y be 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 and ϕ_k, $k = 0, \ldots, q$, are given by (4) and (5) respectively.

We observe that the coefficients c_k lie in $\overline{Z}[\omega]$ and thus there is no loss of generality by considering only the subring $\overline{Z}[\omega, x, y]$ of R.

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

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

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 q-th root of 1 R, then (8) and (9) hold. These also hold if $\overline{Z} = Z_q$ and $\omega = 1$. We may also note that in the case of an integral domain $\overline{Z}[\omega]$ a necessary condition for (10) to be satisfied is that (9) holds.

Let $K[x,y]$ be the ring in two noncommutative indeterminates x and y over a central field K. If x and y are subject to the relation $xy = \omega yx$ where $\omega \in K$ then $K[x,y]$ is today called a *quantum plane* over K, see [15, p.72].

3 Wielandt's notes

In Figure 2 we display a facsimile of page 35 of Wielandt's Diary VII, [22], dated 20 March 1951.

The transcription reads as follows:

Neuer Beweis des Satzes von H. S. A. Potter (On the latent roots of quasi-commutative matrices, Amer. Math. Monthly 57, 321–322 (1950)).

Ist $AB = \omega BA$, ω eine primitive q-te Einheitswurzel, so ist (*) $(A+B)^q = A^q + B^q$.

*Bew: Klar ist $(A+B)^q = A^q + B^q + \sum_1^{q-1} c_k A^k B^{q-k}$, wobei die Skalare c_k nicht von der speziellen Wahl von A, B (abgesehen *) abhängen. Wählt man*

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

Quasikommutative Matrizen. 20.3.51

Neuer Beweis des Satzes von H.S.A. Potter (On the latent roots of quasi-commutative matrices, Amer. Math. Monthly 57, 321-322 (1950)):
Ist $AB = \omega BA$, ω eine primitive q-te Einheitswurzel, so ist (*) $(A+B)^q = A^q + B^q$.

Bew: Klar ist $(A+B)^q = A^q + B^q + \sum_1^{q-1} c_k A^k B^{q-k}$, wobei die Skalare c_k nicht von der speziellen Wahl von A, B (abgesehen von *) abhängen. Wählt man

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

so wird $(A+B)^q = \delta E$ eine Diagonalmatrix; denn mit $T = BA^{-1}$ wird $T^{-1}AT = \omega A$, $T^{-1}BT = \omega B$, also $T^{-1}(A+B)T = \omega(A+B)$, also hat $A+B$ EWe $\varrho, \varrho\omega, \dots, \varrho\omega^{q-1}$, da nicht alle Null sind wegen Spur, und $(A+B)^q$ hat nur ϱ^q, ist ferner wie $(A+B)$ selbst auf Diagonalform transformierbar, also $(A+B)^q = \varrho^q \cdot E$. Hieraus folgt $c_1 = \dots = c_{q-1} = 0$, da andernfalls ein geeignetes Nicht-Diagonalelement von $(A+B)^q \neq 0$ wäre.

<u>Aufgabe:</u> <u>Alle</u> Identitäten für A, B aufstellen.
Ist jedes Paar A, B mit $(A+B)^q = A^q + B^q$ zerfällbar in $S^{-1}AS = \begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_r \end{pmatrix}$, $S^{-1}BS = \begin{pmatrix} B_1 & & \\ & \ddots & \\ & & B_r \end{pmatrix}$ mit

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

Bestimmung aller irreduziblen Paare quasikomm. Matrizen?
Ältere Sonderdrucke über quasikomm. Matrizen?

Figure 2: Page 35 of Wielandt's diary VII.

so wird $(A+B)^q = \delta E$ eine Diagonalmatrix; denn mit $T = BA^{-1}$ wird $T^{-1}AT = \omega A$, $T^{-1}BT = \omega B$, also $T^{-1}(A+B)T = \omega(A+B)$, also hat $A+B$ Ewe $\rho, \rho\omega, \ldots, \rho\omega^{q-1}$, da nicht alle Null sind wegen Spur, und $(A+B)^q$ hat nur ρ^q, ist ferner wie $(A+B)$ selbst auf Diagonalform transformierbar, aber $(A+B)^q = \rho^q E$. Hieraus folgt $c_1 = \ldots = c_{q-1} = 0$, da andernfalls ein geeignetes Nicht-Diagonalelement von $(A+B)^q \neq 0$ wäre.

Aufgabe: Alle Identitäten für A, B aufstellen. Ist jedes Paar A, B mit $(A+B)^q = A^q + B^q$ zerfällbar in

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

mit

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

Bestimmung aller irreduziblen Paare quasikommutativer Matrizen ?
Ältere Sonderdrucke über quasikommutative Matrizen?

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 q-th root of unity, then (*) $(A+B)^q = A^q + B^q$.

*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; since with $T = BA^{-1}$ also $T^{-1}AT = \omega A$, $T^{-1}BT = \omega B$, and thus $T^{-1}(A+B)T = \omega(A+B)$, hence $A+B$ has eigenvalues $\rho, \rho\omega, \ldots, \rho\omega^{q-1}$, since not all are zero due to the trace and $(A+B)^q$ has only ρ^q as eigenvalue and is as $(A+B)$ transformable to diagonal form, but $(A+B)^q = \rho^q E$. This implies that $c_1 = \ldots = c_{q-1} = 0$, since otherwise a particular non-diagonal element would satisfy $(A+B)^q \neq 0$.

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 means the quasi-commutativity relation rather than the relation (*) when he says 'except for (*)'. Also, he must mean that $\omega_\rho^q = 1$ since he is considering the case when 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, B, and hence the result is proved 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 a pair of matrices A, B that satisfy this condition. He then uses an argument involving eigenvalues and the diagonability of matrices to show that $c_1 = \ldots = c_{q-1} = 0$.

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 any complex numbers. Since the eigenvalues of B are the q-th roots of unity, it follows that the characteristic polynomial of sB is $\lambda^q - s^q$. Since 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

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

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 equal zero, which implies $c_k = 0$, since A and B are both nonsingular. □

Thus, the alternative proof demonstrates the following extension of Potter's Theorem.

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

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

for all $s, t \in \mathbb{C}$.

A proof in a rather similar spirit is given by R. Bhatia and L. Elsner in [4] for the following fact: Let A and B be quasicommutative, then the spectrum of $A + B$ is p-Carollian, 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 specific form given in [4, Theorem 2]. The term 'Carollian' was invented by R. Bhatia in honor of L. Carrol, initially to denote an n-tuple that contains $-x$ if it contains x, and later turned into 'p-Carollian' for n-tuples that contain all multiples of x with p-th roots of unity. It is used also in [3] and [2].

5 Normal forms for quasi-commutative matrices

Note that Wielandt asks the question on classification of irreducible quasi-commutative pairs, having in mind reductions by simultaneous similarity $A \mapsto T^{-1}AT$, $B \mapsto T^{-1}BT$, which leave the relation (1) invariant.

To study this question, we start with some preliminary observations. Suppose that

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

where α is a nonzero complex number. By the above remark on simultaneous similarity, we may assume w.l.o.g that A is in Jordan canonical form

$$A = \operatorname{diag}(J_1(\lambda_1), \dots, J_s(\lambda_s))$$

where $J_i(\lambda_i)$ is a Jordan block of size k_i corresponding to the eigenvalue λ_i, $i = 1, \dots, s$. We partition B conformably with A, where B_{ij} is a block of size $k_i \times k_j$. Using the construction of [21, p.148], see also [11], we conclude that that $B_{ij} \neq 0$ only if $\omega\lambda_i = \lambda_j$. In this case an easy computation then yields that $B_{ij} = DX$, where D is the $k_i \times k_i$ diagonal matrix $D = \operatorname{diag}(1, \omega, \dots, \omega^{k_i - 1})$ and X is a rectangular Toeplitz matrix (i.e. with equal elements on each diagonal) such that all elements in the first column below position $(1,1)$ and all elements in the last row to the left of position (k_i, l_j) are 0.

We note that in this way we have not obtained a canonical form for the pair (A, B) under simultaneous similarity, as in general there will be similarities that leave A invariant but change B.

Example 5 *Suppose $\alpha \neq 0$ and let*

$$A = \begin{bmatrix} \lambda & 1 & 0 & 0 & 0 \\ 0 & \lambda & 1 & 0 & 0 \\ 0 & 0 & \lambda & 0 & 0 \\ 0 & 0 & 0 & \alpha\lambda & 1 \\ 0 & 0 & 0 & 0 & \alpha\lambda \end{bmatrix}.$$

If $\lambda \neq 0$ and $\alpha \neq -1$ then B is of the form

$$B = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & x_1 & x_2 & 0 & 0 \\ 0 & 0 & \alpha x_1 & 0 & 0 \end{bmatrix}.$$

If $\lambda \neq 0$ and $\alpha = -1$, then we obtain

$$B = \begin{bmatrix} 0 & 0 & 0 & y_1 & y_2 \\ 0 & 0 & 0 & 0 & \alpha y_1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & x_1 & x_2 & 0 & 0 \\ 0 & 0 & \alpha x_1 & 0 & 0 \end{bmatrix},$$

while, if $\lambda = 0$, then

$$B = \begin{bmatrix} u_1 & u_2 & u_3 & y_1 & y_2 \\ 0 & \alpha u_1 & \alpha u_2 & 0 & \alpha y_1 \\ 0 & 0 & \alpha^2 u_1 & 0 & 0 \\ 0 & x_1 & x_2 & v_1 & v_2 \\ 0 & 0 & \alpha x_1 & 0 & \alpha v_1 \end{bmatrix}.$$

If B is nonsingular and A is not nilpotent, then every row and every column of B must contain at least one nonzero element. Thus, if λ is a nonzero eigenvalue of A so is $\alpha\lambda$. Since the number of eigenvalues is finite and A has a nonzero eigenvalue, it follows that α 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 $\alpha\lambda_i$ of equal size. Thus, we conclude that the maximal size of a Jordan block in A is the same for each nonzero eigenvalue.

Theorem 6 *Let A and B be nonsingular satisfying $AB = \alpha BA$, where $\alpha \neq 1$ is non-zero. Then α is a primitive p-th root of 1 for some p and the Jordan form of A may be written as*

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

where each K_k, $k = 1, \ldots, q$, is the direct sum of Jordan blocks of the same size belonging to $\lambda, \alpha\lambda, \ldots, \alpha^{p-1}\lambda$.

The same structure of the Jordan form holds for B.

Proof. From the assumption of the theorem, A is similar to ωA. Hence the number of Jordan blocks corresponding to any eigenvalue λ of A and their sizes coincide with the number and sizes of the Jordan blocks corresponding to the eigenvalue $\omega\lambda$. The proof for B follows by exchanging the roles of A and B. $\square$

Note further that, if $AB = \alpha BA$, then $AB \cdot B = \alpha B \cdot AB$. Thus, if A, B are nonsingular quasi-commutative matrices, then BA also satisfies the conclusions of Theorem (6).

This finally brings us to the question of Wielandt on classification of quasi-commutative pairs. It was to a large extent answered by Drazin already in 1951, although Wielandt was apparently unaware of his results. In [9], Drazin obtained the following pre-normal form for pairs of quasi-commutative matrices.

Theorem 7 *If A, B are $n \times n$ matrices satisfying an equation of the from $AB = \omega BA$, then either* *(i) A, B can be simultaneously reduced to triangular form by a similarity transformation, or*

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

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

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

Furthermore, Drazin also proves the following Theorem:

Theorem 8 *In Theorem 7, if (i) holds with $\omega \neq 1$, then each of AB, BA is nilpotent, and A, 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, B can be effected in such a way that A_r can be reduced, by the same similarity transformation, to the forms*

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

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

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

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

Drazin's formulas do not give a canonical form, however. Indeed, first of all some further reduction of A, B already in the form (14) is possible. We have already seen that the equation $M_1 X = X M_2$ has only the trivial solution $X = 0$, whenever the spectra of M_1 and M_2 do not intersect [11]. This implies that we can decompose an arbitrary quasi-commutative pair $(A,\, B)$ as

$$A = \mathrm{diag}(\widetilde{A}_0, \widetilde{A}_1, \widetilde{A}_2, \dots, \widetilde{A}_m), \qquad B = \mathrm{diag}(\widetilde{B}_0, \widetilde{B}_1, \widetilde{B}_2, \dots, \widetilde{B}_m),$$

where the spectrum of $\widetilde{A}_0$ consists of zero only, and the spectra of $\widetilde{A}_i$, $i = 1, \dots, m$ consist of distinct chains $\{\lambda_i (\neq 0), \omega\lambda_i, \dots, \omega^{k-1}\lambda_i\}$. (Note that ω is a primitive k-th root of unity.)

Then each of the pairs $(\widetilde{A}_i, \widetilde{B}_i)$, $i = 0, \ldots, m$ decomposes in the same way according to the spectrum of $\widetilde{B}_i$. All together, we get a block-diagonalization

$$A = \mathrm{diag}(A_0, A_1, A_2, \ldots, A_n), \qquad B = \mathrm{diag}(B_0, B_1, B_2, \ldots, B_n)$$

such that each pair (A_i, B_i) is of one of the following 4 types according to the spectra $\sigma(A_i)$ and $\sigma(B_i)$:

$$\begin{array}{lll}
\text{Type I:} & \sigma(A_i) = \{0\}, & \sigma(B_i) = \{0\}, \\
\text{Type II:} & \sigma(A_i) = \{0\}, & \sigma(B_i) = \{\mu_i(\neq 0), \omega\mu_i, \ldots, \omega^{k-1}\mu_i\}, \\
\text{Type III:} & \sigma(A_i) = \{\lambda_i(\neq 0), \omega\lambda_i, \ldots, \omega^{k-1}\lambda_i\}, & \sigma(B_i) = \{0\}, \\
\text{Type IV:} & \sigma(A_i) = \{\lambda_i(\neq 0), \omega\lambda_i, \ldots, \omega^{k-1}\lambda_i\}, & \sigma(B_i) = \{\mu_i(\neq 0), \omega\mu_i, \ldots, \omega^{k-1}\mu_i\}.
\end{array}$$

Now, by Drazin's Theorem, each pair of type II can be put in the form ((16),(15)) (notice the order of matrices) and each pair of type III or IV to the form ((15),(16)). Moreover, for a pair of type IV one can assume that all submatrices b_i in (16), except for one (b_1 say), are equal to the identity. To achieve this, simply use the transformation $T = \mathrm{diag}(I, b_2^{-1}, (b_2b_3)^{-1}, \ldots, (b_2 \cdots b_n)^{-1})$. These form being fixed, the only further similarity transformations allowed that do not destroy the identity blocks are of the form $V = \mathrm{diag}(V_1, \ldots, V_1)$ with identical diagonal blocks V_1 of the same size as the submatrix a. Therefore, the representation problem for pairs of type IV reduces to the representation problem of commuting matrix pairs (a, b_1) under simultaneous similarity. Conversely, the representation problem for commuting pairs of matrices under simultaneous similarity reduces to the representation of quasi-commuting pairs of type II, II or IV. Indeed, suppose that matrices M and N commute and are not both nilpotent. By using the transformation to Jordan canonical form and splitting the problem into subproblems, we may assume w.l.o.g. that M has only one eigenvalue. Moreover, since $M - \lambda I$ and $N - \mu I$ commute if and only if M and N commute, we may assume at least one of M or N to be nonsingular. Using M as a and and N as b_1 in (15), (16), and setting all other b_i's to be I, we obtain a quasi-commutative pair of type II, III or IV (depending on whether M, N or neither is chosen to be nilpotent). Since all transformations preserving this form of the pair (M, N) must look like $V = \mathrm{diag}(V_1, \ldots, V_1)$, the problem of representing (A, B) and that of representing (M, N) coincide. Note that this argument fails for pairs (M, N) where both matrices are nilpotent, since then it is no longer true that V has to be of the form $\mathrm{diag}(V_1, \ldots, V_1)$.

Our discussion can be summarized as follows.

Theorem 9 *The problem of representation under simultaneous similarity for quasi-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 quasi-commutative pairs of type II, III or IV.*

We do not know whether this result also holds for pairs of type I, i.e., whether the problem of representation of quasi-commutative pairs of type I is also equivalent to the problem of representation of commuting pairs.

We now have an occasion to make a detour in the fascinating topic of simultaneous similarity of commuting matrices.

M. Gelfand and V. A. Ponomarev [13] showed that the simultaneous similarity problem of any n-tuple of matrices is equivalent to the simultaneous similarity problem for a pair of commuting matrices. The seemingly hopeless problem was later taken up by S. Friedland, who showed in [10] how to find a finite number of invariants which will characterize a 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 decomposes the variety of pairs of square matrices in finitely many subsets locally closed under simultaneous similarity (so subvarieties). For each of such subvarieties Z, he gives a rational map f from Z into a finite-dimensional vector space V, so that the pre-images under f (of points in V) consist of finitely many orbits of matrix pairs. With that, f and V depend strongly on Z, while, for a fixed f, one can give an upper bound on the number of conjugation classes in each pre-image. 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 pre-images 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., a a decision algorithm via rational computations, 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, 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 quasi-commutative matrices into blocks, we now discuss Wielandt's second question whether the converse to Potter's theorem holds for every irreducible block, i.e., whether the relation (12) where $s = t = 1$ implies

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

where ω is the q-th root of unity.

If $q = 2$, this strong version of the converse indeed holds.

Proposition 10 *A pair (A, B) is quasi-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 11 It is in general not true that if (12) holds for some q and all s, t, then (17) holds with ω the q-th root of unity.

Consider the case $n = 3$, $q = 3$ and let $\lambda \in \mathbb{C}$ be not equal to 0, -1 or one of the two primitive 3rd 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 3 distinct eigenvalues it follows that (17) does not hold. Also, if the pair (A, B) is replaced by $(\widetilde{A}, \widetilde{B}) := T(A, B)T^{-1}$, then

$$\widetilde{B}\widetilde{A} = \widetilde{E}\widetilde{A}\widetilde{B} \qquad \text{where} \quad \widetilde{E} := TET^{-1}.$$

Since E and $\widetilde{E}$ have the same spectrum, (17) does not hold for the pair $(\widetilde{A}, \widetilde{B})$ either. In other words, the pair (A, B) cannot be reduced to a direct sum of quasi-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 produce even 2×2-counterexamples, e.g., with $q = 3$.

Example 12 The pair of matrices

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}, \qquad B = \begin{bmatrix} -1 & 0 \\ 7 & -2 \end{bmatrix} \tag{21}$$

satisfies (12) with $s, t = 1, q = 3$ but, since $AB = \begin{bmatrix} 6 & -2 \\ 14 & -4 \end{bmatrix}$ and $BA = \begin{bmatrix} -1 & -1 \\ 7 & -4 \end{bmatrix}$, the pair is not quasi-commutative.

The pair of matrices

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 2 \end{bmatrix}, \qquad B = \begin{bmatrix} -2 & 2 \\ 0 & 0 \end{bmatrix} \tag{22}$$

is even triangular and (12) holds with $s, t = 1, q = 3$ but, since $AB = 0$ and $BA = \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix}$, the pair is not quasi-commutative.

One can check that the pairs in both of these examples cannot be decomposed into direct sums of quasi-commutative matrices either.

Drazin's pre-normal form for quasi-commutative matrices also suggests the question whether the converse to Potter's theorem holds at least for the pairs of matrices of the form (14)(i). The following examples demonstrate that it is not so.

Example 13 Let $n = q = 3$, let ω be either of the two primitive 3-rd roots of unity and let x_i, $i = 1, 3$ be arbitrary nonzero numbers. Consider the triangular matrices

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \omega & 0 \\ 0 & 0 & \omega^2 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & x_1 & x_2 \\ 0 & 0 & x_3 \\ 0 & 0 & 0 \end{bmatrix}. \tag{23}$$

Then $A^3 = I$, $B^3 = 0$, and $(A + tB)^3 = I$ for any t. But

$$AB = \begin{bmatrix} 0 & \omega x_1 & \omega^2 x_2 \\ 0 & 0 & \omega^2 x_3 \\ 0 & 0 & 0 \end{bmatrix}, \quad BA = \begin{bmatrix} 0 & x_1 & x_2 \\ 0 & 0 & \omega x_3 \\ 0 & 0 & 0 \end{bmatrix}, \tag{24}$$

so the products AB and BA are not scalar multiples of each other. Since the matrix B is similar to a single Jordan block of size 3 corresponding to the eigenvalue 0, the pair (A, B) is also seen to be irreducible.

On the other hand, if a pair of block $r \times r$-matrices of the form (15)-(16) with some ω satisfies (12), then necessarily $r = q$, $\omega^q = 1$, and (17) holds. Indeed, suppose a pair (A, B) is block $r \times r$ in the form (15)-(16). Then, by direct calculation, it satisfies (17), and, comparing the determinants on both sides, we obtain $\omega^r = 1$. So, we only need to establish that $q = r$. The relations (12) and (17) together imply that $\omega^q = 1$, hence $q = rm$ for some natural number m. So, the pair $(\widetilde{A} := A^r, \widetilde{B} := B^r)$ satisfies the relation

$$\widetilde{A}^m + \widetilde{B}^m = (\widetilde{A}^m + \widetilde{B}^m). \tag{25}$$

But the matrices $\widetilde{A}$ and $\widetilde{B}$ commute, hence (25) implies that $m = 1$. Thus, $q = r$.

However, the commutativity of the blocks a and b_i in (15)-(16) does not follow automatically from the relation (12), so here again the converse to the Potter's result fails, as we show next.

Example 14 Let $n = 6$, $q = 3$ and let

$$A = \begin{bmatrix} A_1 & 0 & 0 \\ 0 & \omega A_1 & 0 \\ 0 & 0 & \omega^2 A_1 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & B_1 & 0 \\ 0 & 0 & B_1 \\ B_1 & 0 & 0 \end{bmatrix}, \tag{26}$$

where

$$A_1 = \begin{bmatrix} 1 & 1+\frac{\omega}{1-\omega^2} \\ 0 & \omega^2 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 1 \\ 0 & \omega^2 \end{bmatrix}.$$

Then $(sA+tB)^3 = (sA)^3 + (tB)^3$ for all s, t, but $B_1A_1 = JA_1B_1$, with

$$J = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix},$$

so, in particular, A_1 and B_1 do not commute and hence the matrices A, B are not quasi-commutative either.

In view of these counterexamples, it seems natural to pose the more general problem to characterize all classes of matrices for which the equivalence of (12) and (17) holds.

7 Identities satisfied by quasi-commutative matrices

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

Theorem 15 *Let $\mathbb{C}[x,y]$ denote the ring of polynomials in non-commuting indeterminates x, y over the field $\mathbb{C}$ and let $\mathcal{I}$ denote the ideal of $\mathbb{C}[x,y]$ generated by the polynomial $g(x,y) = xy - \omega yx$ with $\omega^q = 1$. Then*

$$f(x,y) \in \mathcal{I} \tag{27}$$

if and only if the condition (17) implies $f(A,B) = 0$ in $\mathbb{C}^{q\times q}$.

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

To show the converse, first recall that the condition $\omega^q = 1$ implies that there exist a pair of *nonsingular* matrices A, $B \in \mathbb{C}^{q\times q}$ satisfying (17). Since the pair (sA, tB) also satisfies (17) for any scalars s, t, we get

$$f(sA, tB) = 0.$$

Now interchange A and B using the relation (17) as many times as to obtain a polynomial in the form

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

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 s t are independent scalars, each term $c_{i,j}s^it^jA^iB^j$ in the sum must equal zero. But as both A and B are nonsingular, this shows that $c_{i,j} = 0$. Thus, $f_1(x,y)$ is the zero polynomial and hence $f(x,y) \in \mathcal{I}$. □

8 Further historical comments

We now address the last question asked by Wielandt, namely on the work preceding that of H. S. A. Potter. This question turns out to be also briefly answered by M. P. Drazin 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 [16] devoted to the general case. [Biographies of the two Italian mathematicians can be found at [23].] Cecioni's paper is a memoir summarizing and extending results on quasi-commutative matrices known at that time. He proves a condition on a matrix A necessary and sufficient for the equation $AX = \omega XA$ to have a nonzero solution X, describes the structure of an arbitrary solution similarly to Turnbull-Aitken [21, p.148], and stops one step before arriving at the formulas (15)-(16) for a quasi-commutative pair (A, B) with AB nonsingular. A slightly different pre-normal form is derived by Cherubino [8]; he also describes the structure of the algebra of matrices commuting with a given matrix. The pair (15)-(16) appears also in Kurosaki [16], even in the reduced form (with all b_j's except for one equal to the identity), although not in a formal statement. Kurosaki's main result ([16, 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 in [9] is apparently more interested in simultaneous triangularization of a quasi-commutative pair, hence obtains, in his remarkably short paper, yet another reduced form.

Acknowledgements

We are indebted to H. S. A. Potter and M. P. Drazin for further references, to S. Friedland, L. Levy, and K. Bongartz for discussions of the simultaneous similarity problem, to R. Askey for information on the q-binomial formulae, to M. Lorenz for drawing our attention to the connection to quantum planes, and to M. Benzi and N. Laghi for their help with the works of Cecioni and Cherubino.

References

[1] G. E. Andrews, R. Askey and R. Roy. Special functions. *Encyclopedia of Mathematics and its Applications, 71.* Cambridge University Press, Cambridge, 1999.

[2] R. Bhatia. Matrix analysis. Graduate Texts in Mathematics, 169. Springer-Verlag, New York, 1997.

[3] R. Bhatia, R. and L. Elsner, *Symmetries and variation of spectra.* Canad. J. Math. 44 (1992), no. 6, 1155–1166.

[4] R. Bhatia and L. Elsner. *The q-binomial theorem and spectral symmetry.* Indag. Mathem., N.S., 4 (1), 11–16, 1993.

[5] K. Bongartz. *A remark on Friedland's stratification of varieties of modules.* Comm. Algebra 23 (1995), no. 6, 2163–2165.

[6] A. Cayley, *A memoir on the theory of matrices.* Philos. Trans. 148 (1858), 17–37; Coll. Works (Cambridge, 1889), 2, 475–496.

[7] F. Cecioni, *Sull' equazione fra matrici $AX = \varepsilon XA$.* Ann. Univ. Toscane, 14 (1931), fasc. 2, 1–49.

[8] S. Cherubino, *Sulle omografie 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, 7–10 (1951).

[10] S. Friedland. *Simultaneous similarity of matrices.* Advances Math. 50 (1983), 189-265.

[11] F. R. Gantmacher. The Theory of Matrices, vol.II; New York, Chelsea Publ., 1959.

[12] G. Gasper and M. Rahman. Basic Hypergeometric Series. *Encyclopedia of Mathematics and its Applications, 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, 551–553 (1955).

[15] C. Kassel. Quantum Groups. Springer Verlag, New York, 1995.

[16] T. Kurosaki. *Über die mit einer Kollineation vertauschbaren Kollineationen.* Proc. Imp. Acad. Tokyo, 17 (1941), 24–28.

[17] N. McCoy. *On quasicommutative matrices.* Trans. Amer. Math. Soc. 36, 327–340 (1934).

[18] H. S. A. Potter. *On the latent roots of quasi-commutative matrices.* Am. Math. Monthly, 57, 321-322 (1950). Review by H. Wielandt, Zbl. M. 40, p. 35. MR review 11,637e by W. Ledermann.

[19] M. P. Schützenberger, Une interprétation de certaines solutions de l'quation fonctionnelle: $F(x+y) = F(x)F(y)$. C. R. Acad. Sci. Paris 236, (1953). 352–353.

[20] R. P. Stanley, Enumerative Combinatorics. Vol. I. With a foreword by Gian-Carlo Rota. The Wadsworth & Brooks/Cole Mathematics Series. Wadsworth & Brooks/Cole Advanced Books & Software, Monterey, CA, 1986.

[21] H. W. Turnbull and A. C. Aitken, An Introduction to the Theory of Canonical Matrices. 1st Edn., Blackie, 1932.

[22] H. Wielandt, *Diary VII, 1951.* To appear in transcribed form.

[23] History of Italian Mathematics *Matematica - Storia - Indice*, `http://matematica.uni-bocconi.it/indice.html`

[24] 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`