TECHNISCHE UNIVERSITÄT BERLIN

Matrices that commute with their derivative.

Research and historical note.

Olga Holtz Volker Mehrmann Hans Schneider

Preprint 2012/22

Preprint-Reihe des Instituts für Mathematik
Technische Universität Berlin
http://www.math.tu-berlin.de/preprints

Preprint 2012/22 June 2012

Matrices that commute with their derivative. Research and historical note.†

Olga Holtz‡ Volker Mehrmann‡ Hans Schneider §

04.07.2012

Abstract

We examine when a matrix whose elements are differentiable functions in one variable commutes with its derivative. This problem was discussed in a letter from Issai Schur to Helmut Wielandt written in 1934, which we found in Wielandt's Nachlass. We present this letter and its translation into English. The topic was rediscovered later and partial results were proved. However, there are many subtle observations in Schur's letter which were not obtained in later years. Using an algebraic setting, we put these into perspective and extend them in several directions. We present in detail the relationship between several conditions mentioned in Schur's letter and we focus in particular on the characterization of matrices called Type 1 by Schur. We present several examples that demonstrate Schur's observations.

1 Introduction

What are the conditions that force a matrix of differentiable functions to commute with its elementwise derivative? This problem, discussed in Schur's letter, has been rediscovered and discussed in a large number of papers [2, 3, 4, 6, 8, 10, 11, 17, 19, 20, 21, 22, 23, 26, 27, 29, 31, 32]. However, these authors were unaware of Schur's letter and did not find some of its principal results. A summary and a historical discussion of the problem and several extensions thereof are presented by Evard in [13]. There Evard also states that this problem was initially posed by Erugin [11] (translated in [12], pp. 38-39), but Schur's letter shows that this topic already appeared in Schur's lectures in the 1930s, if not earlier.

We do not know which set of functions Schur had in mind, whether Schur meant this to be the set of analytic functions or meromorphic functions in one variable. He may even have had the rational functions in one variable over the complex numbers in mind. And we do not know which arguments Schur used to reach conclusions concerning matrices of small

‡Institut für Mathematik, Technische Universität Berlin, Straße des 17. Juni 136, D-10623 Berlin, FRG. {holtz,mehrmann}@math.tu-berlin.de.

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

†Supported by *Deutsche Forschungsgemeinschaft*, through DFG project ME 790/15-2.

size at the end of his letter. Though our arguments remain close to those of Schur, we will take an algebraic approach and discuss the results in Schur's letter in differential fields. This is also the approach that was taken in [1] and in unpublished notes of Guralnick [18], where results related to ours using differential fields were discussed.

The content of the paper is as follows. In section 2 we present a facsimile of Schur's letter to Wielandt and its English translation. In section 3 we introduce some of the notation that plays a role in Schur's letter. In section 4 we discuss the results that characterize the matrices of Type 1 in Schur's letter and in section 5 we discuss the role played by diagonalizability and triangularizability of the matrix in the commutativity of the matrix and its derivative. We also present several illustrative examples.

2 A letter from Schur to Wielandt

Our paper deals with the following letter from Issai Schur to his PhD student Helmut Wielandt, see the facsimile below. Translated into English, the letter reads as follows:

Lieber Herr Doktor! *Berlin, 21.7.34*

You are perfectly right. Already for $3 \leqslant n < 6$ *not every solution of the equation* $MM' = M'M$ *has the form*

$$M_1 = \sum_\lambda f_\lambda C_\lambda, \tag{1}$$

where the C_λ *are pairwise commuting constant matrices. One must also consider the type*

$$M_2 = (f_\alpha g_\beta), \quad (\alpha, \beta = 1, \dots n), \tag{2}$$

where $f_1, \dots f_n$, $g_1, \dots, g_n$ *are arbitrary functions that satisfy the conditions*

$$\sum_\alpha f_\alpha g_\alpha = \sum_\alpha f'_\alpha g_\alpha = 0$$

and therefore also

$$\sum_\alpha f_\alpha g'_\alpha = 0.$$

In this case we obtain

$$M^2 = MM' = M'M = 0.$$

In addition we have the type

$$M_3 = \phi E + M_2, \tag{3}$$

with M_2 *of type (2). From my old notes, which I did not present correctly in my lectures, it can be deduced that for* $n < 6$ *every solution of* $MM' = M'M$ *can be completely decomposed by means of constant similarity transformations into matrices of type (1) and (3). Only from* $n = 6$ *on there are also other cases. This seems to be correct. But I have not checked my rather laborious computations (for* $n = 4$ *and* $n = 5$*).*

I concluded in the following simple manner that one can restrict oneself to the case where M *has only one characteristic root (namely* 0*): If* M *has two different characteristic*

Berlin 21. 7. 34.

Lieber Herr Doktor!

Sie haben ganz recht. Schon für $3 \leq n < 6$ hat nicht jede Lösung der Gleichung $MM' = M'M$ die Form

(1) $M_1 = \sum_\lambda f_\lambda C_\lambda$,

wobei die C_λ untereinander vertauschbare konstante Matrizen sind. Man hat noch den Typus

(2) $M_2 = (f_\alpha g_\beta)$ $(\alpha, \beta = 1, 2, \dots n)$

hervorzuheben, wo $f_1, \dots f_n, g_1, \dots g_n$ beliebige Funktionen sind, die den Bedingungen

$$\sum_\alpha f_\alpha g_\alpha = \sum_\alpha f'_\alpha g_\alpha = 0, \text{ also auch } \sum_\alpha f_\alpha g'_\alpha = 0$$

genügen. In diesem Falle wird $M^2 = MM' = M'M = 0$.

Hierzu kommt dann noch der Typus

(3) $M_3 = \varphi E + M_2$ (M_2 vom Typus (2))

Aus meinen alten Notizen, die ich in der Vorlesung nicht richtig wiedergegeben habe, geht hervor, daß für $n < 6$ jede Lösung von $MM' = M'M$ durch eine konstante Ähnlichkeitstransformation in Matrizen vom Typus (1) oder (3) vollständig zerfällt werden kann. Erst für $n = 6$ gibt es noch andere Fälle.

Dies scheint richtig zu sein. Ich habe aber meine recht mühsamen Rechnungen (für $n=4$ und $n=5$) nicht nachgeprüft.

Daß man sich auf den Fall beschränken kann, in dem M nur die eine char. Wurzel 0 besitzt, schloß ich damals einfach so. Besitzt M zwei verschiedene char. Wurzeln, so kann man eine g. rat. Funktion N von M angeben, für die ~~M~~ $N^2 = N$ wird, ohne daß $N = \varphi E$ wird. Auch N ist mit N' vertauschbar. Aus $N^2 = N$ folgt aber $2NN' = N'$, also $2N^2N' = 2NN' = NN'$. Das gibt $2NN = N' = 0$, d. h. N ist konstant. Man kann nun auf M eine konstante Ähnlichkeitstransformation anwenden, sodaß anstelle von N eine Matrix der Form $\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$ tritt. Dies zeigt, daß M mit Hilfe einer konst. Ähnlichkeitstransf. ~~…~~ vollständig zerfällt werden kann.

Man wird auf den Typus (2) geführt, indem man den Fall $M^2 = 0$, $Rg\, M = 1$ studiert. Schon für $n = 4$ kommen noch die Fälle $M^2 = 0$, $Rg\, M = 2$, $M^3 = 0$ in Betracht.

Der Typus (1) ist vollständig dadurch charakterisiert, daß $M, M', M'', \dots$ untereinander vertauschbar sind. Dies ist nicht nur notwendig sondern auch hinreichend. Denn sind unter den n^2 Koeffizienten f_{ik} von M genau r im Gebiete der Konstanten linear unabhängig, so kann man

$$M = f_1 C_1 + \dots + f_r C_r \qquad (C_s \text{ konstante Matrix})$$

schreiben, wobei $f_1, \dots f_r$ keiner Gleichung $\sum_1 \text{const } f_\alpha = 0$ genügen. Dann wird

$$M^{(\nu)} = f_1^{(\nu)} C_1 + \dots + f_r^{(\nu)} C_r \qquad (\nu = 0, 1, \dots r-1)$$

Da die Wronskische Det. $\begin{vmatrix} f_1 & \cdots & f_r \\ f_1' & \cdots & f_r' \\ \cdots & & \cdots \end{vmatrix}$ nicht identisch verschwinden darf, erhält man Gleichungen der Form

$$C_s = \sum_{\sigma=0}^{r-1} \varphi_{s\sigma} M^{(\sigma)}.$$

Sind $M, M', \ldots M^{(r-1)}$ untereinander vertauschbar, so gilt dasselbe auch für $C_1, \ldots C_r$, M ist also vom Typus (1).

Hieraus folgt zugleich, daß M dem Typus (1) angehört wenn M^n die höchste Potenz von M ist die gleich Null wird. Im Falle $n = 3$ hat man daher nur noch den Typus (2) zu berücksichtigen.

Mit vielen Grüßen

Ihr Schur

roots, then one can determine a rational function N of M for which $N^2 = N$ but not $N = \phi E$. Also N commutes with N'. It follows from $N^2 = N$ that $2NN' = N'$, thus $2N^2N' = 2NN' = NN'$. This yields $2NN' = N' = 0$, i.e., N is constant.

Now one can apply a constant similarity transformation to M so that instead of N one achieves a matrix of the form

$$\begin{bmatrix} E & 0 \\ 0 & 0 \end{bmatrix}.$$

This shows that M can be decomposed completely by means of a constant similarity transformation.

One is led to type (2) by studying the case $M^2 = 0$, $\operatorname{rank} M = 1$. Already for $n = 4$ also the cases $M^2 = 0$, $\operatorname{rank} M = 2$, $M^3 = 0$ need to be considered.

Type (1) is completely characterized by the property that $M, M', M'', \dots$ are pairwise commuting. This is not only necessary but also sufficient. For, if among the n^2 coefficients $f_{\alpha\beta}$ of M exactly r are linearly independent over the domain of constants, then one can write

$$M = f_1 C_1 + \cdots + f_r C_r,$$

(C_s a constant matrix), where $f_1, \dots, f_r$ satisfy no equation $\sum_\alpha \mathrm{const}\ f_\alpha = 0$. Then

$$M^{(\nu)} = f_1^{(\nu)} C_1 + \cdots + f_r^{(\nu)} C_r, \quad (\nu = 1, \dots, r-1).$$

Since the Wronskian determinant

$$\left\| \begin{bmatrix} f_1 & \dots & f_r \\ f_1' & \dots & f_r' \\ & \vdots & \end{bmatrix} \right\|$$

cannot vanish identically, one obtains equations of the form

$$C_s = \sum_{\sigma=0}^{r-1} \phi_{s\sigma} M^{(\sigma)}.$$

If $M, M', M'', \dots, M^{(r-1)}$ are pairwise commuting, then the same is true also for $C_1, \dots C_r$ and thus M is of type (1). This implies furthermore that M belongs to type (1) if M^n is highest power of M that equals 0. In the case $n = 3$ one therefore only needs to consider type (2).*

With best regards
Yours, Schur

The problem addressed in Schur's letter has been studied extensively in the literature and some of his observations have been discovered independently by a large number of authors, see [13, 14] and the references therein for a historical overview. We base our presentation on the letter of Schur and put some of the later results in perspective with Schur's observations.

*We think that Schur means *lowest* in this place.

3 Notation and preliminaries

Though we do not know which set of functions Schur had in mind, his arguments work in an abstract setting, that of matrices with entries in a differential field $\mathbb{F}$. A similar approach is taken in [1] as well as in some unpublished notes of Guralnick [18].

A *differential field* $\mathbb{F}$ is an (algebraic) field together with an additional operation, denoted by $'$ that satisfies $(a+b)' = a' + b'$ and $(ab)' = ab' + a'b$ for $a, b \in \mathbb{F}$. An element $a \in \mathbb{F}$ is called a *constant* if $a' = 0$. It is easily shown that the set of constants forms a subfield $\mathbb{K}$ of $\mathbb{F}$ with $1 \in \mathbb{K}$. Examples are provided by the rational functions over the real or complex numbers and the meromorphic functions over the complex numbers.

In what follows we consider a (differential) field $\mathbb{F}$ and matrices $M = [m_{i,j}] \in \mathbb{F}^{n,n}$. The main condition that we want to analyze is when $M \in \mathbb{F}^{n,n}$ commutes with its derivative,

$$MM' = M'M. \tag{4}$$

We introduce the following definitions.

Definition 1 *Let $\mathbb{F}$ be a differential field and let $\mathbb{H}$ be a subfield of $\mathbb{F}$. Then $M \in \mathbb{F}^{n,n}$ is called* $\mathbb{H}$-triangularizable (diagonalizable) *if there exists a nonsingular $T \in \mathbb{H}^{n,n}$ such that $T^{-1}MT$ is upper triangular (diagonal).*

As $M \in \mathbb{F}^{n,n}$, it has a minimal and a characteristic polynomial, see [16], and M is called *non-derogatory* if the characteristic polynomial is equal to the minimal polynomial, otherwise it is called *derogatory*.

In Schur's letter the following three types of matrices are considered.

Definition 2 *Let $M \in \mathbb{F}^{n,n}$. Then M is said to be of*

- Type 1 *if*

$$M = \sum_{j=1}^{k} f_j C_j,$$

where $f_j \in \mathbb{F}$, and $C_j \in \mathbb{K}^{n,n}$, for $j = 1, \ldots, k$, and the C_j are pairwise commuting;

- Type 2 *if*

$$M = fg^T,$$

with $f, g \in \mathbb{F}^n$, satisfying $f^T g = f^T g' = 0$;

- Type 3 *if*

$$M = hI + \widetilde{M},$$

with $h \in \mathbb{F}$ and $\widetilde{M}$ is of Type 2.

If the differential field $\mathbb{F}$ is algebraically closed, then $M \in \mathbb{F}^{n,n}$ is $\mathbb{F}$-triangularizable or even $\mathbb{F}$-diagonalizable for all matrices in $\mathbb{F}^{n,n}$. This is needed in some of the following statements but not in all.

Schur's letter also mentions the condition that all derivatives of M commute, i.e.,

$$M^{(i)}M^{(j)} = M^{(j)}M^{(i)} \text{ for all natural numbers } i, j. \tag{5}$$

To characterize the relationship between all these properties, we first recall several results from Schur's letter and from classical algebra.

Lemma 3 *Let $\mathbb{F}$ be a differential field with field of constants $\mathbb{K}$. Let N be an idempotent matrix in $\mathbb{F}^{n,n}$ that commutes with N'. Then $N \in \mathbb{K}^{n,n}$.*

Proof. (See Schur's letter.) It follows from $N^2 = N$ that $2NN' = N'$. Thus $2NN' = 2N^2N' = NN'$ and this implies that $0 = 2NN' = N'$. □

Another important tool in our analysis will be the following result of Frobenius [15] that Schur employs.

Theorem 4 *Consider a differential field $\mathbb{F}$ with field of constants $\mathbb{K}$. Then $y_1, \dots, y_r \in \mathbb{F}$ are linearly dependent over $\mathbb{K}$ if and only if the columns of the* Wronski matrix

$$Y = \begin{bmatrix} y_1 & y_2 & \dots & y_r \\ y_1' & y_2' & \dots & y_r' \\ \vdots & \vdots & \ddots & \vdots \\ y_1^{(n-1)} & y_2^{(n-1)} & \dots & y_r^{(n-1)} \end{bmatrix},$$

are linearly dependent over $\mathbb{F}$.

Proof. We proceed by induction over r. The case $r = 1$ is trivial.

Consider the Wronski matrix Y and the lower triangular matrix

$$Z = \begin{bmatrix} z & 0 & \dots & 0 \\ c_{2,1}z' & z & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ c_{n,1}z^{(n-1)} & c_{n,2}z^{(n-2)} & \dots & z \end{bmatrix},$$

with $c_{i,j}$ appropriate binomial coefficients such that

$$ZY = \begin{bmatrix} zy_1 & zy_2 & \dots & zy_r \\ (zy_1)' & (zy_2)' & \dots & (zy_r)' \\ \vdots & \vdots & \ddots & \vdots \\ (zy_1)^{(n-1)} & (zy_2)^{(n-1)} & \dots & (zy_r)^{(n-1)} \end{bmatrix}.$$

Since $\mathbb{F}$ is a differential field, we can choose $z = y_1^{-1}$ and obtain that

$$ZY = \begin{bmatrix} 1 & y_1^{-1}y_2 & \dots & y_1^{-1}y_r \\ 0 & (y_1^{-1}y_2)' & \dots & (y_1^{-1}y_r)' \\ \vdots & \vdots & \ddots & \vdots \\ 0 & (y_1^{-1}y_2)^{(n-1)} & \dots & (y_1^{-1}y_r)^{(n-1)} \end{bmatrix}.$$

It follows that the columns of Y are linearly dependent over $\mathbb{F}$ if and only if the columns of

$$\begin{bmatrix} (y_1^{-1}y_2)' & \dots & (y_1^{-1}y_r)' \\ \vdots & \ddots & \vdots \\ (y_1^{-1}y_2)^{(n-1)} & \dots & (y_1^{-1}y_r)^{(n-1)} \end{bmatrix}$$

are linearly dependent over $\mathbb{F}$, which, by induction, holds if and only if $(y_1^{-1}y_2)', \dots, (y_1^{-1}y_r)'$ are linearly dependent over $\mathbb{K}$, i.e., there exist coefficients $b_2, \dots, b_r \in \mathbb{K}$, not all 0, such that

$$b_2(y_1^{-1}y_2)' + \cdots + b_r(y_1^{-1}y_r)' = 0.$$

Integrating this identity, we obtain

$$b_2(y_1^{-1}y_2) + \cdots + b_r(y_1^{-1}y_r) = -b_1$$

for some integration constant $b_1 \in \mathbb{K}$, or equivalently

$$b_1y_1 + \cdots + b_ry_r = 0. \qquad \square$$

Theorem 4 implies in particular that the columns of the Wronski matrix Y are linearly independent over $\mathbb{F}$ if and only if they are linearly independent over $\mathbb{K}$.

Remark 5 Theorem 4 is discussed from a formal algebraic point of view, which however includes the cases of complex analytic functions and rational functions over a field, since these are contained in differential fields. Necessary and sufficient conditions for Theorem 4 to hold for other functions were proved in [5] and discussed in many places, see, e.g., [7, 24] and [25, Ch. XVIII].

4 Characterization of matrices of Type 1

In this section we discuss relationships among the various properties introduced in Schur's letter and in the previous section. This will give, in particular, a characterization of matrices of Type 1.

In his letter, Schur proves the following result.

Theorem 6 *Let $\mathbb{F}$ be a differential field. Then $M \in \mathbb{F}^{n,n}$ is of Type 1 if and only if it satisfies condition (5), i.e., $M^{(i)}M^{(j)} = M^{(j)}M^{(i)}$ for all integers i, j.*

Proof. (See Schur's letter.) If M is of Type 1, then $M = \sum_{j=1}^k f_jC_j$ and the $C_j \in \mathbb{K}^{n,n}$ are pairwise commuting, which immediately implies (5). For the converse, Schur makes use of Theorem 4, since if among the n^2 coefficients $m_{i,j}$ exactly r are linearly independent over $\mathbb{K}$, then

$$M = f_1C_1 + \cdots + f_rC_r,$$

with coefficients $C_i \in \mathbb{K}^{n,n}$, where $f_1, \dots, f_r$ are linearly independent over $\mathbb{K}$. Then

$$M^{(i)} = f_1^{(i)}C_1 + \cdots + f_r^{(i)}C_r, \qquad i = 1, \dots, r-1.$$

By Theorem 4, the columns of the associated Wronski matrix are linearly independent, and hence each of the C_i can be expressed as

$$C_i = \sum_{j=0}^{r-1} g_{i,j} M^{(j)}.$$

Thus, if condition (5) holds, then the C_i, $i = 1, \ldots, r$, are pairwise commuting and thus M is of Type 1. $\square$

Using this result we immediately have the following Theorem.

Theorem 7 *Let $\mathbb{F}$ be a differential field with field of constants $\mathbb{K}$. If $M \in \mathbb{F}^{n,n}$ is non-derogatory and $MM' = M'M$, then M is of Type 1.*

Proof. If M is non-derogatory then it is well known [9] that all matrices that commute with M have the form $p(M)$ where p is a polynomial with coefficients in $\mathbb{F}$. Thus $MM' = M'M$ implies that M' is a polynomial in M. But then every derivative $M^{(j)}$ is a polynomial in M as well and thus (5) holds which by Theorem 6 implies that M is of Type 1. $\square$

The following example of a Type 2 matrix shows that one cannot easily drop the condition that the matrix is non-derogatory.

Example 8 [4, 13] Let

$$f = \begin{bmatrix} 1 \\ t \\ t^2 \end{bmatrix}, \qquad g = \begin{bmatrix} t^2 \\ -2t \\ 1 \end{bmatrix},$$

then $f^T g = 0$ and $f^T g' = 0$, hence

$$M = gf^T = \begin{bmatrix} t^2 & t^3 & t^4 \\ -2t & -2t^2 & -2t^3 \\ 1 & t & t^2 \end{bmatrix}, \tag{6}$$

is of Type 2. Since M is nilpotent with $M^2 = 0$ but $M \neq 0$ and the rank is 1, it is derogatory. One has

$$M' = \begin{bmatrix} 2t & 3t^2 & 4t^3 \\ -2 & -4t & -6t^2 \\ 0 & 1 & 2t \end{bmatrix}, \qquad M'' = \begin{bmatrix} 2 & 6t & 12t^2 \\ 0 & -4 & -12t \\ 0 & 0 & 2 \end{bmatrix},$$

and thus $MM' = M'M = 0$. By the product rule it immediately follows that $MM'' = M''M$, but

$$M'M'' = \begin{bmatrix} 4t & 0 & -4t^3 \\ -4 & 4t & 12t^2 \\ 0 & -4 & -8t \end{bmatrix} \neq M'M'' = \begin{bmatrix} -8t & -6t^2 & -4t^3 \\ 8 & 4t & 0 \\ 0 & 2 & 4t \end{bmatrix}.$$

Therefore, it follows from Theorem 6 that M is not of Type 1.

For any dimension $n \geqslant 3$, one can construct an example of Type 2 by choosing $f \in \mathbb{F}^n$, setting $F = [f, f']$ and then choosing g in the nullspace of F^T. Then fg^T is of Type 2.

Actually every nilpotent matrix M of rank one satisfying $MM' = M'M$ is of this form and hence of Type 2. This follows immediately because if $M = fg^T$ and $M^2 = 0$ then $g^T f = 0$ and hence $g^T f' + (g^T)' f = 0$. Then it follows from $MM' = M'M$ that $fg^T(f(g^T)' + f'g^T) = (g^T f')fg^T = (f(g^T)' + f'g^T)fg^T = (g^T)' f fg^T$ which implies that $g^T f' = f^T g'$ and hence $g^T f' = f^T g' = 0$.

5 Triangularizability and Diagonalizability

In his letter Schur claims that it is sufficient to consider the case that $M \in \mathbb{F}^{n,n}$ is triangularizable with only one eigenvalue. This follows from his argument in case the matrix has its eigenvalues in $\mathbb{F}$, which could be guaranteed by assuming that $\mathbb{F}$ is algebraically closed, because then for every matrix in $\mathbb{F}^{n,n}$ the characteristic polynomial splits into linear factors. If one considers a specific matrix $M \in \mathbb{F}^{n,n}$, then it suffices that this matrix is $\mathbb{F}$-diagonalizable or even $\mathbb{F}$-triangularizable .

Using Lemma 3, we can strengthen a classical result in matrix theory, for matrices $M \in \mathbb{F}^{n,n}$ that commute with their derivative M'.

Lemma 9 *Let $\mathbb{F}$ be a differential field with field of constants $\mathbb{K}$, and suppose that $M \in \mathbb{F}^{n,n}$ satisfies $MM' = M'M$. Then there exists an invertible matrix $T \in \mathbb{K}^{n,n}$ such that*

$$T^{-1}MT = \operatorname{diag}(M_1, \dots, M_k), \tag{7}$$

where the minimal polynomial of each M_i is a power of a polynomial that is irreducible over $\mathbb{F}$.

Proof. Let the minimal polynomial of M be $\mu(\lambda) = \mu_1(\lambda) \cdots \mu_k(\lambda)$, where the $\mu_i(\lambda)$ are powers of pairwise distinct polynomials that are irreducible over $\mathbb{F}$. Set

$$p_i(\lambda) = \mu(\lambda)/\mu_i(\lambda), \qquad i = 1, \dots, k.$$

Since the polynomials $p_i(\lambda)$ have no common factor, there exist polynomials $q_i(\lambda)$, $i = 1, \dots, k$, such that the polynomials $\epsilon_i(\lambda) = p_i(\lambda)q_i(\lambda)$, $i = 1, \dots, k$, satisfy

$$\epsilon_1(\lambda) + \cdots + \epsilon_k(\lambda) = 1.$$

Setting $E_i = \epsilon_i(M)$, $i = 1, \dots, k$ and using the fact that $\mu(M) = 0$ yields that

$$\begin{aligned} E_1 + \cdots + E_k &= I, && (8)\\ E_iE_j &= 0, \qquad i, j = 1, \dots, k, \quad i \neq j, && (9)\\ E_i^2 &= E_i, \qquad i = 1, \dots, k. && (10) \end{aligned}$$

Since the E_i are polynomials in M and $MM' = M'M$, it follows that the E_i commute with E_i', $i = 1, \dots k$. Hence, by Lemma 3, $E_i \in \mathbb{K}^{n,n}$, $i = 1, \dots, k$. By (8), (9), and (10), $\mathbb{K}^n$ is a direct sum of the ranges of the E_i and we obtain that, for some nonsingular $T \in \mathbb{K}^{n,n}$,

$$T^{-1} E_i T = \operatorname{diag}(0, I_i, 0), \qquad i = 1, \dots, k,$$

where the I_i are identity matrices of the size equal to the dimension to the range of E_i. Since each E_i commutes with M, setting $\widetilde{M}_i = E_i M = E_i M E_i$, we obtain that

$$T^{-1} \widetilde{M}_i T = \operatorname{diag}(0, M_i, 0), \qquad i = 1, \dots, k.$$

Moreover, since $\sum_1^k \operatorname{diag}(0, M_i, 0) = T^{-1} M T$, we obtain (7).

Now observe that $\mu_i(\widetilde{M}_i) = E_i \mu_i(M) = \epsilon_i(M)\mu_i(M) = 0$, since $\epsilon_i(\lambda)\mu_i(\lambda) = \mu(\lambda) q_i(\lambda)$. Hence $\mu_i(M_i) = 0$ as well. We assert that $\mu_i(\lambda)$ is the minimal polynomial of M_i, for if $r(M_i) = 0$ for a proper factor $r(\lambda)$ of $m_i(\lambda)$ then $r(M)\Pi_{j \neq i} \mu_j(M) = 0$, contrary to the assumption that $\mu(\lambda)$ is the minimal polynomial of M. □

Lemma 9 has the following corollary, which has been proved in a different way in [1] and [18].

Corollary 10 *Let $\mathbb{F}$ be a differential field with field of constants $\mathbb{K}$. If $M \in \mathbb{F}^{n,n}$ satisfies $MM' = M'M$ and is $\mathbb{F}$-diagonalizable, then M is $\mathbb{K}$-diagonalizable.*

Proof. In this case, the minimal polynomial of M is a product of distinct linear factors and hence, the minimal polynomial of each M_i occurring in the proof of Theorem 6 is linear. Therefore, each M_i is a scalar matrix. □

We also have the following Corollary.

Corollary 11 *Let $\mathbb{F}$ be a differential field with field of constants $\mathbb{K}$. If $M \in \mathbb{F}^{n,n}$ satisfies $MM' = M'M$ and is $\mathbb{F}$-diagonalizable, then M is of Type 1.*

Proof. By Corollary 10, $M = T^{-1} \operatorname{diag}(m_1, \dots, m_n) T$ with $m_i \in \mathbb{F}$ and nonsingular $T \in \mathbb{K}^{n,n}$. Hence

$$M = \sum_{i=1}^{n} m_i T^{-1} E_{i,i} T$$

where $E_{i,i}$ is a matrix that has a 1 in position (i, i) and zeros everywhere else. Since all the matrices $E_{i,i}$ commute, M is of Type 1. □

For matrices that are just triangularizable the situation is more subtle. We have the following theorem.

Theorem 12 *Let $\mathbb{F}$ be a differential field with an algebraically closed field of constants $\mathbb{K}$. If $M \in \mathbb{F}^{n,n}$ is Type 1, then M is $\mathbb{K}$-triangularizable.*

Proof. It is well known that any finite set of pairwise commutative matrices with elements in an algebraically closed field may be simultaneously triangularized, see e.g., [28, Theorem 1.1.5]. Under this assumption on $\mathbb{K}$, if M is Type 1, then it follows that the matrices $C_i \in \mathbb{K}^{n,n}$ in the representation of M are simultaneously triangularizable by a matrix $T \in \mathbb{K}^{n,n}$. Hence T also triangularizes M. □

Theorem 12 implies that Type 1 matrices have n eigenvalues in $\mathbb{F}$ if $\mathbb{K}$ is algebraically closed and it further immediately leads to a Corollary of Theorem 7.

Corollary 13 *Let $\mathbb{F}$ be a differential field with field of constants $\mathbb{K}$. If $M \in \mathbb{F}^{n,n}$ is non-derogatory, satisfies $MM' = M'M$ and if $\mathbb{K}$ is algebraically closed, then M is $\mathbb{K}$-triangularizable.*

Proof. By Theorem 7 it follows that M is Type 1 and thus the assertion follows from Theorem 12. □

Example 8 again shows that it is difficult to drop some of the assumptions, since this matrix is derogatory, not of Type 1, and not $\mathbb{K}$-triangularizable.

One might be tempted to conjecture that any $M \in \mathbb{F}^{n,n}$ that is $\mathbb{K}$-triangularizable and satisfies (4) is of Type 1 but this is so only for small dimensions and is no longer true for large enough n, as we will demonstrate below. Consider small dimensions first.

Proposition 14 *Consider a differential field $\mathbb{F}$ of functions with field of constants $\mathbb{K}$. Let $M = [m_{i,j}] \in \mathbb{F}^{2,2}$ be upper triangular and satisfy $M\,M' = M'\,M$. Then M is of Type 1.*

Proof. Since $M\,M' = M'\,M$ we obtain

$$m_{1,2}(m'_{1,1} - m'_{2,2}) - m'_{1,2}(m_{1,1} - m_{2,2}) = 0,$$

which implies that $m_{1,2} = 0$ or $m_{1,1} - m_{2,2} = 0$ or both are nonzero and $\frac{d}{dt}(\frac{m_{1,1}-m_{2,2}}{m_{1,2}}) = 0$, i.e., $c_1 m_{1,2} + c_2(m_{1,1} - m_{2,2}) = 0$ for some nonzero constants c_1, c_2.

If $m_{11} = m_{22}$ or $m_{12} = 0$, then M, being triangular, is obviously of Type 1. Otherwise

$$M = m_{1,1} I + m_{1,2} \begin{bmatrix} 0 & 1 \\ 0 & c_2/c_1 \end{bmatrix}.$$

and hence again of Type 1 as claimed. □

Proposition 14 implies that 2×2 $\mathbb{K}$-triangularizable matrices satisfying (4) are of Type 1.

Proposition 15 *Consider a differential field $\mathbb{F}$ with an algebraically closed field of constants $\mathbb{K}$. Let $M = [m_{i,j}] \in \mathbb{F}^{2,2}$ satisfy $M\,M' = M'\,M$. Then M is of Type 1.*

Proof. If M is $\mathbb{F}$-diagonalizable, then the result follows by Theorem 10. If M is not $\mathbb{F}$-diagonalizable, then it is non-derogatory and the result follows by Theorem 12. □

Example 16 In the 2×2 case, any Type 2 or Type 3 matrix is also of Type 1 but not every Type 1 matrix is Type 3.

Let $M = \phi I_2 + f g^T$ with

$$\phi \in \mathbb{F}, \quad f = \begin{bmatrix} f_1 \\ f_2 \end{bmatrix}, \quad g = \begin{bmatrix} g_1 \\ g_2 \end{bmatrix} \in \mathbb{F}^2$$

be of Type 3, i.e., $f^T g = f'^T g = f^T g = 0$.

If $f_2 = 0$, then M is upper triangular and hence by Proposition 14, M is of Type 1. If $f_2 \neq 0$, then with

$$T = \begin{bmatrix} 1 & -f_1/f_2 \\ 0 & 1 \end{bmatrix}.$$

we have

$$TMT^{-1} = \phi I_2 + \begin{bmatrix} 0 & 0 \\ f_2 g_1 & 0 \end{bmatrix} = \phi I_2 + f_2 g_1 \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix},$$

since $f_1 g_1 + f_2 g_2 = 0$, and hence M is of Type 1.

However, if we consider

$$M = \phi I_2 + f \begin{bmatrix} 0 & c \\ 0 & d \end{bmatrix}$$

with ϕ, f nonzero functions and c, d nonzero constants, then M is Type 1 but not Type 3.

Proposition 17 *Consider a differential field $\mathbb{F}$ of functions with field of constants $\mathbb{K}$. Let $M = [m_{i,j}] \in \mathbb{F}^{3,3}$ be $\mathbb{K}$-triangularizable and satisfy $M M' = M' M$. Then M is of Type 1.*

Proof. Since M is $\mathbb{K}$-triangularizable, we may assume that it is upper triangular already and consider different cases for the diagonal elements. If M has three distinct diagonal elements, then it is $\mathbb{K}$-diagonalizable and the result follows by Corollary 10. If M has exactly two distinct diagonal elements, then it can be transformed to a direct sum of a 2×2 and 1×1 matrix and hence the result follows by Proposition 14. If all diagonal elements are equal, then $M = m_{11} I + m_{13} E_{13} + \widetilde{M}$, where $\widetilde{M} = \widetilde{m}_{12} E_{12} + \widetilde{m}_{2,3} E_{2,3}$ also satisfies (4). Then it follows that $\widetilde{m}_{12} \widetilde{m}'_{23} = \widetilde{m}'_{12} \widetilde{m}_{23}$. If either $\widetilde{m}_{1,2} = 0$ or $\widetilde{m}_{2,3} = 0$, then we immediately have again Type 1, since $\widetilde{M}$ is a direct sum of a 2×2 and a 1×1 problem.

If both are nonzero, then $\widetilde{M}$ is non-derogatory and the result follows by Theorem 7. In fact, in this case $\widetilde{m}_{1,2} = c \widetilde{m}_{2,3}$ for some $c \in \mathbb{K}$ and therefore

$$M = m_{11} I + m_{13} E_{13} + \widetilde{m}_{2,3} \begin{bmatrix} 0 & c & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix},$$

which is clearly of Type 1. $\square$

In the 4×4 case, if the matrix is $\mathbb{K}$-triangularizable, then we either have at least two different eigenvalues, in which case we have reduced the problem again to the case of dimensions smaller than 4, or there is only one eigenvalue, and thus w.l.o.g. M is nilpotent. If M is non-derogatory then we again have Type 1. If M is derogatory then it is the direct sum of blocks of smaller dimension. If these dimensions are smaller than 3, then we are again in the Type 1 case. So it remains to study the case of a block of size 3 and a block of size 1. Since M is nilpotent, the block of size 3 is either Type 1 or Type 2. In both cases the complete matrix is also Type 1 or Type 2, respectively.

The following example shows that $\mathbb{K}$-triangularizability is not enough to imply that the matrix is Type 1.

Example 18 Consider the 9×9 block matrix

$$\hat{M} = \begin{bmatrix} 0 & M & 0 \\ 0 & 0 & M \\ 0 & 0 & 0 \end{bmatrix},$$

where M is the Type 2 matrix from Example 8. Then M is nilpotent upper triangular and not of Type 1, 2, or 3, the latter two facts due to its $\mathbb{F}$-rank being 2.

Already in the 5×5 case, we can find examples that are none of the (proper) types.

Example 19 Consider $M = T^{-1} \operatorname{diag}(M_1, M_2)T$ with $T \in \mathbb{K}^{n,n}$, $M_1 \in \mathbb{F}^{3,3}$ of Type 2 (e.g., as in Example 8) and $M_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. Then clearly M is not of Type 1 and it is not of Type 2, since it has an $\mathbb{F}$-rank larger than 1. By definition it is not of Type 3 either. Clearly examples of any size can be constructed by building direct sums of smaller blocks.

Schur's letter states that for $n \geqslant 6$ there are other types. The following example demonstrates this.

Example 20 Let M be the Type 2 matrix in Example 8 and form the block matrix

$$A = \begin{bmatrix} M & I \\ 0 & M \end{bmatrix}.$$

Direct computation shows $AA' = A'A$ but $A'A'' \neq A''A$. Furthermore $A^3 = 0$ and A has $\mathbb{F}$-rank 3. Thus A is neither Type 1, Type 2 nor Type 3 (the last case need not be considered, since A is nilpotent). We also note that $\operatorname{rank}(A'') = 6$. We now assume that $\mathbb{K}$ is algebraically closed and we show that A is not $\mathbb{K}$-similar to the direct sum of Type 1 or Type 2 matrices.

To obtain a contradiction we assume that (after a $\mathbb{K}$-similarity) $A = \operatorname{diag}(A_1, A_2)$ where A_1 is the direct sum of Type 1 matrices (and hence Type 1) and A_2 is the direct sum of Type 2 matrices that are not Type 1. Since A is not Type 1, A_2 cannot be the empty matrix. Since the minimum size of a Type 2 matrix that is not Type 1 is 3 and its rank is 1 it follows that A cannot be the sum of Type 2 matrices that are not Type 1. Hence the size of A_1 must be larger or equal to 1 and, since A_1 is nilpotent, it follows that $\operatorname{rank}(A_1) < \operatorname{size}(A_1)$. Since A_1 is $\mathbb{K}$-similar to a strictly triangular matrix, it follows that $\operatorname{rank}(A_1'') < \operatorname{size}(A_1)$. Hence $\operatorname{rank}(A'') = \operatorname{rank}(A_1'') + \operatorname{rank}(A_2'') < 6$, a contradiction.

Example 21 If the matrix $M = \sum_{i=0}^{r} C_i t^i \in \mathbb{F}^{n,n}$ is a polynomial with coefficients $C_i \in \mathbb{K}^{n,n}$, then from (4) we obtain a specific set of conditions on sums of commutators that have to be satisfied. For this we just compare coefficients of powers of t and obtain a set of quadratic equations in the C_i, which has a clear pattern. e.g., in the case $r = 2$, we obtain the three conditions $C_0C_1 - C_1C_0 = 0$, $C_0C_2 - C_2C_0 = 0$ and $C_1C_2 - C_2C_1 = 0$, which shows that this is Type 1. For $r = 3$ we obtain the first nontrivial condition $3(C_0C_3 - C_3C_0) + (C_1C_2 - C_2C_1) = 0$.

We have implemented a Matlab routine for Newton's method to solve the set of quadratic matrix equations in this case and ran it for many different random starting coefficients C_i of different dimensions n. Whenever Newton's method converged (which it did in most of the cases) it converged to a matrix of Type 1. Even in the neighborhood of a Type 2 matrix it converged to a Type 1 matrix. This suggests that the matrices of Type 1 are generic in the set of matrices satisfying 4. (A copy of the Matlab routine is available from the authors upon request.)

6 Conclusion

We have presented a letter of Schur that contains a major contribution to the question when a matrix with elements that are functions in one variable commutes with its derivative. Schur's letter precedes many partial results on this question, which is still partially open. We have put Schur's result in perspective with later results and extended it in an algebraic context to matrices over a differential field. In particular, we have presented several results that characterize Schur's matrices of Type 1.

Acknowledgements

We thank Carl de Boor for helpful comments on a previous draft of the paper.

References

[1] W.A. Adkins, J.-C. Evard, and R.M. Guralnik, *Matrices over differential fields which commute with their derivative,* Linear Algebra Appl., 190:253–261, 1993.

[2] V. Amato, *Commutabilità del prodotto di una matrice per la sua derivata,* (Italian), Matematiche, Catania, 9:176–179, 1954.

[3] G. Ascoli, *Sulle matrici permutabili con la propria derivata*, (Italian), Rend. Sem. Mat. Univ. Politec. Torino, 9:245–250, 1950.

[4] G. Ascoli, *Remarque sur une communication de M. H. Schwerdtfeger*, (French), Rend. Sem. Mat. Univ. Politec. Torino, 11:335–336, 1952.

[5] M. Bocher, *Certain cases in which the vanishing of the Wronskian is a sufficient condition for linear dependence*, Trans. American Math. Soc. 2:139–149, 1901.

[6] Ju. S. Bogdanov and G. N. Čebotarev, *On matrices which commute with their derivatives.* (Russian) Izv. Vysš. Učebn. Zaved. Matematika 11:27–37, 1959.

[7] A. Bostan and P. Dumas, *Wronskians and linear independence*, American Math. Monthly, 117:722–727, 2010.

[8] J. Dieudonné, *Sur un théoreme de Schwerdtfeger*, (French), Ann. Polon. Math., XXIX:87–88, 1974.

[9] M.P. Drazin, J.W. Dungey, and K.W. Gruenberg, *Some theorems on commutative matrices*, J. London Math. Soc., 26:221–228, 1951.

[10] I.J. Epstein, *Condition for a matrix to commute with its integral*, Proc. Amererican Math. Soc., 14:266-270, 1963.

[11] N.P. Erugin, *Privodimyye sistemy (Reducible systems)*, Trudy Fiz.-Mat. Inst. im. V.A. Steklova XIII, 1946.

[12] N.P. Erugin, *Linear Systems of Ordinary Differential Equations with Periodic and Quasi-periodic Coefficients*, Academic, New York, 1966.

[13] J.-C. Evard, *On matrix functions which commute with their derivative,* Linear Algebra Appl., 68:145–178, 1985.

[14] J.-C. Evard, *Invariance and commutativity properties of some classes of solutions of the matrix differential equation* $X(t)X'(t) = X'(t)X(t)$, Linear Algebra Appl., 218:89–102, 1995.

[15] F.G. Frobenius, *Über die Determinante mehrerer Functionen einer Variabeln*, (German), J. Reine Angew. Mathematik, 77:245–257, 1874. In Gesammelte Abhandlungen I, 141–157, Springer Verlag, 1968.

[16] F.R. Gantmacher, *Theory of Matrices.* Vol. 1, Chelsea, New York, 1959.

[17] S. Goff, *Hermitian function matrices which commute with their derivative*, Linear Algebra Appl., 36:33–46, 1981.

[18] R. Guralnick, *Private communication*, 2005.

[19] M.J. Hellman, *Lie algebras arising from systems of linear differential equations*, Dissertation, Dept. of Mathematics, New York Univ., 1955.

[20] L. Kotin and I.J. Epstein, *On matrices which commute with their derivatives*, Lin. Multilinear Algebra, 12:57–72, 1982.

[21] L.M. Kuznecova, *Periodic solutions of a system in the case when the matrix of the system does not commute with its integral*, (Russian), Differentsial'nye Uravneniya (Rjazan'), 7:151–161, 1976.

[22] J.F.M. Martin, *On the exponential representation of solutions of systems of linear differential equations and an extension of Erugin's theorem*, Doctoral Thesis, Stevens Inst. of Technology, Hoboken, N.J., 1965.

[23] J.F.P. Martin, *Some results on matrices which commute with their derivatives*, SIAM J. Ind. Appl. Math., 15:1171–1183, 1967.

[24] G.H. Meisters, *Local linear dependence and the vanishing of the Wronskian*, American Math. Monthly, 68:847–856, 1961.

[25] T. Muir, *A treatise on the theory of determinants*, Dover, 1933.

[26] I.V. Parnev, *Some classes of matrices that commute with their derivatives* (Russian), Trudy Ryazan. Radiotekhn. Inst., 42:142–154, 1972.

[27] G. N. Petrovskii, *Matrices that commute with their partial derivatives*, Vestsi Akad. Navuk BSSR Minsk Ser. Fiz. i Mat. Navuk, 4:45-47, 1979.

[28] H. Radjavi and P.Rosenthal, *Simultaneous triangulation*, Springer 2000.

[29] N.J. Rose, *On the eigenvalues of a matrix which commutes with its derivative*, Proc. American Math. Soc., 4:752–754, 1965.

[30] I. Schur, *Letter from Schur to Wielandt*, 1934.

[31] H. Schwerdtfeger, *Sur les matrices permutables avec leur dérivée*, (French), Rend. Sem. Mat. Univ. Politec. Torino, 11:329–333, 1952.

[32] A. Terracini. *Matrici permutabili con la propia derivata*, (Italian), Ann. Mat. Pura Appl. 40:99–112, 1955.

[33] E.I. Troickii, *A study of the solutions of a system of three linear homogeneous differential equations whose matrix commutes with its integral*, (Russian), Trudy Ryazan. Radiotekhn. Inst., 53:109–115, 188-189, 1974.

[34] I.M. Vulpe, A theorem due to Erugin, Differential Equations, 8:1666–1671, 1972, transl. from Differentsial'nye Uravneniya, 8:2156–2162, 1972.