

## PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a preprint version which may differ from the publisher's version.

For additional information about this publication click this link.

<http://hdl.handle.net/2066/159768>

Please be advised that this information was generated on 2020-11-24 and may be subject to change.

# Mathieu subspaces of codimension less than $n$ of $\text{Mat}_n(K)$

Michiel de Bondt

March 24, 2016

## Abstract

We classify all Mathieu subspaces of  $\text{Mat}_n(K)$  of codimension less than  $n$ , under the assumption that  $\text{char } K = 0$  or  $\text{char } K \geq n$ .

More precisely, we show that any proper Mathieu subspace of  $\text{Mat}_n(K)$  of codimension less than  $n$  is a subspace of  $\{M \in \text{Mat}_n(K) \mid \text{tr } M = 0\}$  if  $\text{char } K = 0$  or  $\text{char } K \geq n$ . On the other hand, we show that every subspace of  $\{M \in \text{Mat}_n(K) \mid \text{tr } M = 0\}$  of codimension less than  $n$  in  $\text{Mat}_n(K)$  is a Mathieu subspace of  $\text{Mat}_n(K)$  if  $\text{char } K = 0$  or  $\text{char } K \geq n + 1$ .

*Key words:* Mathieu subspaces, matrix algebras, radicals.

*MSC 2010:* 16S50, 16D70, 16D99.

## 1 Introduction

The notion of Mathieu subspaces has been introduced by W. Zhao in [Zha2]. The usefulness of this notion has been proven by the many notorious open problems that has been formulated in terms of it. For more information about Mathieu subspaces in general, see [Zha2], [Zha3], [Zha1] and the references therein. See also [vdE] for the connection between Mathieu subspaces and the Image Conjecture.

In this paper, we study Mathieu subspaces over a field  $K$  of  $\text{Mat}_n(K)$ : the  $n$ -dimensional matrix ring over  $K$ . But let us first give the general definition of Mathieu subspaces.

**Definition 1.1** (following [Zha3, Def. 1.1] and [Zha3, Def. 1.2]). Let  $M$  be an  $R$ -subspace ( $R$ -submodule) of an associative  $R$ -algebra  $\mathcal{A}$ . Then we call  $M$  a  $\vartheta$ -Mathieu subspace of  $\mathcal{A}$  if the following property holds for all  $a, b, c \in \mathcal{A}$  such that  $a^m \in M$  for all  $m \geq 1$ :

- (i)  $ba^m \in M$  when  $m \gg 0$ , if  $\vartheta = \textit{“left”}$ ;
- (ii)  $a^m c \in M$  when  $m \gg 0$ , if  $\vartheta = \textit{“right”}$ ;
- (iii)  $ba^m, a^m c \in M$  when  $m \gg 0$ , if  $\vartheta = \textit{“pre-two-sided”}$ ;
- (iv)  $ba^m c \in M$  when  $m \gg 0$ , if  $\vartheta = \textit{“two-sided”}$ .

As you can see, there are four different types of Mathieu subspaces. However, for Mathieu subspaces over a field  $K$  of  $\text{Mat}_n(K)$ , it is a nice exercise to show that the last two types coincide, e.g. by using theorem 6.1 in the last section. In this last section, we look at radicals of Mathieu subspaces. Those radicals play an important role in the study of Mathieu subspaces, see the references in the first paragraph of the current section.

But the radical can be taken for any subset  $S$  of the whole space  $\mathcal{A}$ , not only for Mathieu subspaces. The *radical*  $\mathfrak{r}(S)$  of  $S$  is just the set  $\{a \in \mathcal{A} \mid a^m \in S \text{ for all } m \gg 0\}$ .

In [Zha3, Th. 5.1], Zhao classifies all Mathieu subspaces of codimension one of  $\text{Mat}_n(K)$ . Zhao proved that the subspace  $H$  of  $\text{Mat}_n(K)$  consisting of all matrices with trace zero is the only candidate, and that  $H$  is indeed a  $\vartheta$ -Mathieu subspace of  $\text{Mat}_n(K)$ , if and only if  $\text{char } K = 0$  or  $\text{char } K \geq n + 1$ .

A. Konijnenberg proved in his Master's thesis [Kon] that for Mathieu subspaces of codimension two of  $\text{Mat}_n(K)$ ,  $K$ -subspaces of  $H$  are the only possible candidates, under the assumption that  $n \geq 3$ , see [Kon, Th. 3.4]. By taking a one-sided ideal, one can see that this assumption is necessary for the one-sided cases. But the following result shows that the assumption  $n \geq 3$  cannot be omitted either in the two-sided case, see also [Kon, Th. 3.10].

**Proposition 1.2.** *Let  $K$  be a field such that  $p := \text{char } K = 0$  or  $p = \text{char } K \geq n$ . Suppose that  $K \not\cong \mathbb{F}_p$  in the case where  $p \in \{n, n + 1\}$ . Take  $a = 1$  if  $p \notin \{n, n + 1\}$  and take  $a \in K \setminus \mathbb{F}_p$  if  $p \in \{n, n + 1\}$ . Then the subspace*

$$\mathcal{M} := \{M \in \text{Mat}_n(K) \mid M_{n1} = M_{n2} = \cdots = M_{n(n-1)} = 0 = \text{tr } M + aM_{nn}\}$$

*of  $\text{Mat}_n(K)$  is a two-sided Mathieu subspace of  $\text{Mat}_n(K)$ .*

*Proof.* On account of [Zha3, Lem. 4.1] and [Zha3, Cor. 4.3], it suffices to show that  $\mathcal{M}$  has no nontrivial idempotent. Hence assume that  $E \in \mathcal{M}$  is a nonzero idempotent. Since  $E_{n1} = E_{n2} = \cdots = E_{n(n-1)} = 0$ , we see that the trailing principal minor matrix  $E_{nn}$  of size 1 of  $E$  is an idempotent as well.

By looking at Jordan normal forms, we see that the rank and the trace of an idempotent matrix over  $K$  are equal within  $K$ . Hence  $\text{tr } E = \text{rk } E \in \{1, 2, \dots, n\}$  and  $E_{nn} \in \{0, 1\}$ . It follows that

$$0 = \text{tr } E + aE_{nn} \in \{1, 2, \dots, n\} + \{0, a\} \tag{1}$$

within  $K$ . If  $p = 0$  or  $p \geq n + 2$ , then  $a = 1$  and  $n + a < n + 2$ , so (1) cannot be satisfied. Hence  $p \in \{n, n + 1\}$  and  $a \notin \mathbb{F}_p$ . From (1) and  $a \notin \mathbb{F}_p$ , it follows that  $E_{nn} = 0$ , so  $\text{tr } E = \text{rk } E < n$ . This contradicts (1), because  $p \geq n$ .  $\square$

If  $K$  is closed under taking square root and  $\text{char } K \neq 2$ , then the above proposition gives all two-sided Mathieu subspaces of  $\text{Mat}_2(K)$  of (co)dimension two which are not contained in  $H$  up to linear conjugation, except that  $a$  may be any element of  $K$  such that  $\mathcal{M} \not\subseteq H$  and the left hand side of (1) is not contained in the right hand side of (1), i.e.

$$a \neq 0 \quad \text{and} \quad 0 \notin \{1, 2\} + \{0, a\} \quad \text{in } K$$

respectively, which comes down to that  $a \notin \{-2, -1, 0\}$  in  $K$ . This has been proved by Konijnenberg in [Kon, Th. 3.10].

Example 3.11 in [Kon] shows that the codimension  $n$  case seems quite difficult. Hence we shift our focus to subspaces of  $\text{Mat}_n(K)$  of codimension less than  $n$  from now on. But let us first say something about subspaces of  $H$ .

**Lemma 1.3.** *Assume  $\mathcal{M}$  is a subspace of  $\text{Mat}_n(K)$  such that  $\text{tr } M = 0$  for all  $M \in \mathcal{M}$ . Then for*

- (1)  $\text{char } K = 0$  or  $\text{char } K \geq n + 1$ ,
- (2)  $\text{char } K = 0$  or  $\text{char } K \geq n$  and  $I_n \notin \mathcal{M}$ ,
- (3) Every element of  $\mathfrak{r}(\mathcal{M})$  is nilpotent,
- (4)  $\mathcal{M}$  is a two-sided Mathieu subspace,

we have (1)  $\Rightarrow$  (2)  $\Rightarrow$  (3)  $\Rightarrow$  (4).

*Proof.* If  $\text{char } K = 0$  or  $\text{char } K \geq n + 1$ , then  $\text{tr } I_n = n \neq 0$  in  $K$ . This gives (1)  $\Rightarrow$  (2). Since (3)  $\Rightarrow$  (4) follows directly from the definition of Mathieu subspace, (2)  $\Rightarrow$  (3) remains to be proved.

So assume that  $A \in \mathfrak{r}(\mathcal{M})$ . Then there exists an  $N$  such that  $A^m \in \mathcal{M}$  for all  $m \geq N$ . Now let  $B$  be the Jordan normal form of  $A^N$  (or any other triangular matrix that is linearly conjugate to  $A^N$ ). Then  $\text{tr } B^m = \text{tr } A^{mN} = 0$  for all  $m \geq 1$ . Let  $\beta$  be the diagonal of eigenvalues of  $B$ . Then  $\sum_{i=1}^n \beta_i^m = 0$  for all  $m$ .

Using the Newton identities on the eigenvalues of  $B$ , we get that the eigenvalue polynomial of  $B$  is of the form  $t^n + (-1)^n \det B$  if  $\text{char } K = 0$  or  $\text{char } K \geq n$  (where  $\det B = 0$  in case  $\text{char } K = 0$  or  $\text{char } K \geq n + 1$ ). From the Cayley-Hamilton theorem, we deduce that  $B^n + (-1)^n (\det B) I_n = 0$ . It follows that  $B^n$  and hence also  $A^{nN}$  is a multiple of  $I_n$ . So either  $I_n \in \mathcal{M}$  or  $A^{nN} = 0$ . This gives (2)  $\Rightarrow$  (3).  $\square$

Our main theorem, theorem 1.4 below, is that we indeed have  $\mathcal{M} \subseteq H$  if the codimension is less than  $n$ , provided the base field  $K$  is large enough. Sections 3 to 5 will be devoted to the highly technical proof of theorem 1.4. But first, we will give a rough sketch of this proof in the next section.

**Theorem 1.4.** *Let  $K$  be a field. Assume  $\mathcal{M}$  is a proper Mathieu subspace of any type of  $\text{Mat}_n(K)$  of codimension less than  $\min\{n, \#K\}$ . Then  $\text{tr } M = 0$  for all  $M \in \mathcal{M}$ . In particular, every element of  $\mathfrak{r}(\mathcal{M})$  is nilpotent and  $\mathcal{M}$  is a two-sided Mathieu subspace if  $\text{char } K = 0$  or  $\text{char } K \geq n$ .*

Using theorem 1.4, we can classify all Mathieu subspaces of  $\text{Mat}_n(K)$  of codimension less than  $n$ , under the assumption that  $\text{char } K = 0$  or  $\text{char } K \geq n$ .

**Corollary 1.5.** *Let  $K$  be a field such that  $\text{char } K = 0$  or  $\text{char } K \geq n$ . Then for a proper  $K$ -subspace  $\mathcal{M}$  of  $\text{Mat}_n(K)$  of codimension less than  $n$ ,  $\mathcal{M}$  is a Mathieu subspace of any arbitrary type of  $\text{Mat}_n(K)$ , if and only if  $\text{tr } M = 0$  for all  $M \in \mathcal{M}$ , and either  $\text{char } K \neq n$  or  $I_n \notin \mathcal{M}$ .*

*Proof.* The ‘if’-part follows from lemma 1.3. To show the ‘only if’-part, suppose that  $\mathcal{M}$  is a proper  $K$ -subspace of codimension less than  $n$  of  $\text{Mat}_n(K)$ . Since  $\#K \geq \text{char } K \geq n$  if  $\#K < \infty$ , it follows that  $\#K \geq n$  in any case and that the codimension of  $\mathcal{M}$  is less than  $\min\{n, \#K\}$ . So  $\text{tr } M = 0$  for all  $M \in \mathcal{M}$  on account of theorem 1.4. Since  $I_n$  is not nilpotent and  $I_n$  is contained in  $\mathfrak{r}(\mathcal{M})$  as soon as it is contained in  $\mathcal{M}$ , it additionally follows from theorem 1.4 that  $I_n \notin \mathcal{M}$ , which completes the ‘only if’-part.  $\square$

## 2 Sketch of the proof of theorem 1.4

Suppose that  $\mathcal{M}$  is a subspace of codimension  $c$  of  $\text{Mat}_n(K)$ . Then the matrices  $C \in \text{Mat}_n(K)$  such that  $\text{tr} CM = 0$  for all  $M \in \mathcal{M}$ , which we call *constraints of  $\mathcal{M}$* , form a subspace of dimension  $c$  of  $\text{Mat}_n(K)$ . Write  $\mathcal{C}$  for this space of constraints of  $\mathcal{M}$ .

Write  $\square_r(C)$  for the largest submatrix above the diagonal of  $C \in \text{Mat}_n(K)$  with  $r$  rows. So  $\square_r(C)$  has  $n - r$  columns, corresponding to columns  $r + 1, r + 2, \dots, n$  of  $C$ . Notice that  $\text{tr} CM$  is the sum of the entries of the Hadamard product of  $C$  and the transpose  $M^t$  of  $M$ . So the entries of  $\square_r(C)$  act as coefficients for the entries of  $\square_r(M^t)$  and its transpose, which we call  $\square_{n-r}(M)$ . So  $\square_{n-r}(M)$  is the largest submatrix below the diagonal of  $M \in \text{Mat}_n(K)$  with  $r$  columns or  $n - r$  rows.

The reason for using the formula  $\text{tr} CM = 0$  in the definition of constraint, instead of the sum of the entries of the Hadamard product, is that when we replace  $\mathcal{M}$  by the isomorphic space  $T^{-1}\mathcal{M}T$  for some  $T \in \text{GL}_n(K)$ , the corresponding space of constraints  $\mathcal{C}$  gets replaced in a similar manner, namely by  $T^{-1}\mathcal{C}T$ .

In theorem 3.1, it is shown that  $\mathcal{M}$  has idempotents of the forms

$$\begin{pmatrix} I_r & \emptyset \\ * & \emptyset \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \emptyset & \emptyset \\ * & I_{n-r} \end{pmatrix}$$

if  $\square_r$  is injective on  $\mathcal{C}$ , after which theorem 1.4 is proved. The idea behind theorem 3.1 is more or less the following. We fix an arbitrary idempotent matrix  $E$  of one of both forms. Now for each nonzero  $C \in \mathcal{C}$ , we want to have  $\text{tr} CE = 0$ . Since  $\square_r(C)$  is not the zero matrix, we can obtain  $\text{tr} CE = 0$  for some  $C \in \mathcal{C}$  by only changing the submatrix  $\square_{n-r}(E)$  of  $E$ . The proof of theorem 3.1 shows that this can be done for all nonzero  $C \in \mathcal{C}$  simultaneously, so that  $E$  can be changed into an idempotent of  $\mathcal{M}$  by only changing the submatrix  $\square_{n-r}(E)$  of  $E$ .

The hard part of the proof of theorem 1.4 is the proof of theorem 3.3. This theorem claims that under certain conditions, among which  $I_n \notin \mathcal{C}$ , we can indeed obtain injectivity of  $\square_r$  on  $\mathcal{C}$  for some  $r$  by way of linear conjugation. This conclusion leads to a contradiction in the proof of theorem 1.4, so that  $I_n \notin \mathcal{C}$  can be ruled out. So we get  $I_n \in \mathcal{C}$ , which is equivalent to the main conclusion of theorem 1.4:  $\text{tr} M = 0$  for all  $M \in \mathcal{M}$ .

More precisely, the assertion of theorem 3.3 is the following. If we have  $I_n \notin \mathcal{C}$  besides certain conditions that are implied by those of theorem 1.4, then after replacing  $\mathcal{C}$  by  $T^{-1}\mathcal{C}T$  for an appropriate  $T \in \text{GL}_n(K)$ , there exists an  $r$  with  $1 \leq r \leq n - 1$ , such that  $\square_r(C)$  is not the zero matrix for all nonzero  $C \in \mathcal{C}$ . We will even obtain a stronger conclusion: all nonzero entries of the rightmost nonzero column of any nonzero  $C \in \mathcal{C}$  are in  $\square_r(C)$ .

More formally, let  $B' \in \text{Mat}_n(\{0,1\})$  be the binary matrix, which is 1 on some spot, if and only if some element of  $\mathcal{C}$  has a rightmost nonzero entry at that spot. Then we will obtain that all entries 1 of  $B'$  are in  $\square_r(B')$ . For that purpose, we apply a conjugation process on the space of constraints, but not on  $\mathcal{C}$  directly. In order to get the conjugation process in the way we want, we add the identity to  $\mathcal{C}$  by defining  $\mathcal{C}_n = \mathcal{C} \oplus KI_n$ , and apply the conjugation process on  $\mathcal{C}_n$ .

We can easily reason out  $I_n$  afterwards, because  $I_n$  is not affected by conjugations. Namely, if we define  $B \in \text{Mat}_n(\{0, 1\})$  as the binary matrix, which is 1 on some spot, if and only if some element of  $\mathcal{C}_n$  has a rightmost nonzero entry at that spot, then we will obtain that all entries 1 of  $B$  except  $B_{nn} = 1$  are in  $\square_r(B)$ . We have  $B_{nn} = 1$  because  $I_n \in \mathcal{C}_n$ . With that, we have the only difference with  $B'$ , so that  $B' = B - e_n e_n^t$ , where  $e_i$  is the  $i$ -th standard basis unit vector as a column vector. In theorem 4.1, we prove that we can obtain several properties for  $B$ , which we discuss below. Under the conditions of theorem 1.4, these properties imply that all entries 1 of  $B$  except  $B_{nn}$  are in  $\square_r(B)$ .

Since we want the rightmost nonzero column of nonzero  $C \in \mathcal{C}$  to have some property, we define  $\mathcal{C}_k$  as the space of  $C \in \mathcal{C}_n$  for which the rightmost nonzero column has index at most  $k$ , where the index of the rightmost nonzero column of the zero matrix is 0. So  $\mathcal{C}_0$  only contains the zero matrix, and  $\mathcal{C}_n$  is correctly defined into itself. Now we can define  $B \in \text{Mat}_n(K)$  alternatively by  $B_{ik} = 1$ , if and only if  $C_{ik} \neq 0$  for some  $C \in \mathcal{C}_k$ . This is equivalent to that  $v_i \neq 0$  for some  $v \in K^{\times n}$  which is  $k$ -th column of some matrix in  $\mathcal{C}_k$ .

So the subspace formed by the  $k$ -th columns of matrices in  $\mathcal{C}_k$ , which is isomorphic to  $\mathcal{C}_k/\mathcal{C}_{k-1}$ , plays a crucial role. We will denote this subspace of  $\mathcal{C}_k$  as  $\mathcal{C}_k e_k$ , where  $e_k$  is the  $k$ -th standard basis unit vector. More generally, we define

$$\mathcal{C}_k v := \{Cv \mid C \in \mathcal{C}_k\} \quad (v \in K^{\times n})$$

The dimension of  $\mathcal{C}_k v$  does not exceed  $n$  and neither exceeds  $\dim \mathcal{C}_k$ . If  $K$  is infinite, then we take for  $d_k$  the maximum dimension that a space of the form  $\mathcal{C}_k v$  can have. In general, we first replace  $K$  by an infinite extension field  $L$  of  $K$ , which is done by taking the tensor product over  $K$  of  $L$  and  $\mathcal{C}_k$ , and next take  $v \in L^{\times n}$ .

So  $d_k$  is the maximum dimension that a space of the form  $(L \otimes_K \mathcal{C}_k) \cdot v$  can have, where  $v \in L^{\times n}$ . Our choice of  $d_k$  is highly ambiguous, but lemma 5.2 shows that  $d_k$  is still uniquely determined. For the actual definition of  $d_k$ , which is not ambiguous, we take  $L = K(x)$  and  $v = x$ , where  $x = (x_1, x_2, \dots, x_n)$ . In any case, we have  $d_k \geq \max\{\dim \mathcal{C}_k v \mid v \in K^{\times n}\} \geq \dim \mathcal{C}_k e_k$ .

In the proof of theorem 4.1, we arrange the required properties for  $B$  by way of linear conjugation of  $\mathcal{C}_n$ . In order to do that, we additionally ensure that we get  $d_k = \dim \mathcal{C}_k e_k$  for each  $k$ . We achieve this by doing the following for  $k = n, n-1, \dots, 1$ , in that order. We first choose a  $v \in K^{\times n}$  such that  $d_k = \dim \mathcal{C}_k v$ . If  $\#K \geq d_k$ , then such a  $v$  indeed exists, and if  $\#K > d_k$ , then we can additionally choose  $v$  such that  $v_k = 1$  (see lemma 5.3). Next, we choose  $T \in \text{GL}_n(K)$  such that the  $k$ -th column  $T e_k$  of  $T$  equals  $v$ , and replace  $\mathcal{C}_n$  by  $T^{-1} \mathcal{C}_n T$ .

By choosing  $T$  properly (namely equal to the identity matrix at the right of the  $k$ -th column),  $\mathcal{C}_k$  gets replaced by  $T^{-1} \mathcal{C}_k T$ , so that  $\mathcal{C}_k e_k$  gets replaced by  $T^{-1} \mathcal{C}_k T e_k = T^{-1} \mathcal{C}_k v$ , which is isomorphic to  $\mathcal{C}_k v$  (see (ii) of lemma 5.4). This is how  $d_k = \dim \mathcal{C}_k e_k$  is obtained, but what we ignore here is the problem, that  $d_j = \dim \mathcal{C}_j e_j$  for  $j > k$  and possibly some other properties which  $B$  already satisfies, should not be affected. Such preservation problems, which we will mostly ignore in this section, makes the proof of theorem 4.1 highly technical in nature.

One of these preservation problems can be solved if we can take  $v_k = 1$ , because in that case, we can choose  $T$  equal to the identity matrix outside its  $k$ -th column  $v$ . If we additionally take  $v_{k+1} = v_{k+2} = \dots = v_n = 0$ , which is also possible, then the  $j$ -th column of  $B$  will be preserved for all  $j > k$ , provided this  $j$ -th column of  $B$  is decreasing above the diagonal (see the proof of (ii) of theorem 4.1).

Once we have  $d_k = \dim \mathcal{C}_k e_k$  in theorem 4.1, we additionally have that  $B$  is increasing in every row, provided  $\#K \geq d_n$  (this is shown in (ii) of lemma 4.3, where lemma 5.2 is used to obtain the condition of lemma 4.3).

Write  $b_j$  for the number of ones in column  $j$  of  $B$ . Another property to arrange is that  $b_k = \dim \mathcal{C}_k e_k$  as well as  $d_k = \dim \mathcal{C}_k e_k$ , and lemma 4.2 tells us that for this purpose,  $\mathcal{C}_k e_k$  should be spanned by standard basis unit vectors. We do this by taking  $L \in \text{GL}_n(K)$  lower triangular, such that  $L\mathcal{C}_k e_k$  is spanned by standard basis unit vectors. Since  $L$  is lower triangular and invertible, we can prove that  $\mathcal{C}_k e_k$  is replaced by  $L\mathcal{C}_k e_k$  if  $\mathcal{C}_n$  is replaced by  $L\mathcal{C}_n L^{-1}$  (see (ii) of lemma 5.4). So  $\mathcal{C}_k e_k$  will be spanned by standard basis unit vectors after this replacement.

If  $\#K > \min\{d_{n-1}, n-1\}$ , then we must additionally obtain that every column of  $B$  is increasing above the diagonal. If  $\mathcal{C}_k e_k$  is spanned by standard basis unit vectors, then there exists a permutation  $P$  such that  $P\mathcal{C}_k e_k$  is spanned by the first  $\dim \mathcal{C}_k e_k$  standard basis unit vectors. But if we replace  $\mathcal{C}_n$  by  $P\mathcal{C}_n P^{-1}$ , then property  $d_j = \dim \mathcal{C}_j e_j$  could be affected for some  $j > k$ , as well as several properties that  $B$  satisfies.

For this reason, we take  $P$  such that only the first  $k-1$  coordinates of  $\mathcal{C}_k e_k$  are permuted. These coordinates correspond to the part above the diagonal of the  $k$ -th column of  $B$ . If we replace  $\mathcal{C}_n$  by  $P\mathcal{C}_n P^{-1}$ , then  $\mathcal{C}_k e_k$  will be replaced by  $P\mathcal{C}_k e_k$  (see (ii) of lemma 5.4). In order to preserve properties of  $B$ ,  $P$  and also  $L$  only act on coordinates  $i$  such that  $B_{ij} = 1$  for all  $j > k$ , so that  $v \mapsto Lv$  and  $v \mapsto Pv$  are isomorphisms of  $\mathcal{C}_j e_j$  for all  $j > k$  (as far as the respective  $\mathcal{C}_j e_j$  are generated by standard basis unit vectors, but that is inductively arranged).

### 3 A criterium for having idempotents

Let  $K$  be a field and  $\mathcal{M}$  be a subspace of  $\text{Mat}_n(K)$ . Let  $\mathcal{C}$  be the subspace of matrices  $C \in \text{Mat}_n(K)$  such that  $\text{tr} CM = 0$  for all  $M \in \mathcal{M}$ . Let  $x = (x_1, x_2, \dots, x_n)$  be  $n$  indeterminates. Write  $\square_r(M)$  for the submatrix consisting of the first  $r$  rows and the rightmost  $n-r$  columns of  $M$  for all  $M \in \text{Mat}_n(K)$ .

**Theorem 3.1.** *Suppose that  $1 \leq r \leq n-1$ .*

- (i) *If for all  $C \in \mathcal{C}$  such that  $\square_r(C)$  is the zero matrix, the leading principal minor matrix of size  $r$  of  $C$  has trace zero, then  $\mathcal{M}$  contains an idempotent of rank  $r$  of the form*

$$\begin{pmatrix} I_r & \emptyset \\ * & \emptyset \end{pmatrix} \quad (2)$$

- (ii) *If for all  $C \in \mathcal{C}$  such that  $\square_r(C)$  is the zero matrix, the trailing principal minor matrix of size  $n-r$  of  $C$  has trace zero, then  $\mathcal{M}$  contains an*

idempotent of rank  $n - r$  of the form

$$\begin{pmatrix} \emptyset & \emptyset \\ * & I_{n-r} \end{pmatrix} \quad (2')$$

More precisely, the dimensions of the affine spaces of idempotents in  $\mathcal{M}$  of the forms (2) and (2') respectively, are both equal to that of

$$\mathcal{N} := \left\{ M \in \mathcal{M} \mid M = \begin{pmatrix} \emptyset & \emptyset \\ \tilde{M} & \emptyset \end{pmatrix} \text{ for some } \tilde{M} \in \text{Mat}_{n-r,r}(K) \right\}$$

*Proof.* Since (ii) is similar to (i) (or take the transpose and conjugate with the reversing permutation to reduce to (i)), we only prove (i). Notice that any matrix of the form (2) is an idempotent of rank  $r$ , and that  $0 \leq \dim_K \mathcal{N} \leq (n-r)r$  because  $0 \in \mathcal{N}$ . Take  $m$  such that  $\dim_K \mathcal{N} = (n-r)r - m$ . Then there are constraints  $C_1, C_2, \dots, C_m \in \mathcal{C}$  such that

$$\mathcal{N} = \left\{ M = \begin{pmatrix} \emptyset & \emptyset \\ \tilde{M} & \emptyset \end{pmatrix} \text{ for some } \tilde{M} \in \text{Mat}_{n-r,r}(K) \mid \text{tr } C_i M = 0 \text{ for all } i \right\}$$

$\tilde{M}$  and its transpose  $\tilde{M}^t$  are submatrices of  $M$  and its transpose  $M^t$  respectively, and we have  $\tilde{M}^t = \square_r(M^t)$ . By definition of  $\mathcal{N}$ , for any constraint  $C' \in \mathcal{C}$ ,  $\square_r(C')$  is contained in the span of the corresponding submatrices of  $C_1, C_2, \dots, C_m \in \mathcal{C}$ . Hence we can write each  $C' \in \mathcal{C}$  as

$$C' = \lambda_1 C_1 + \lambda_2 C_2 + \dots + \lambda_m C_m + C^* \quad (3)$$

such that  $\square_r(C^*)$  is the zero matrix. By assumption, the leading principal minor matrix of size  $r$  of  $C^*$  has trace zero. Hence we have

$$\text{tr } C^* E = 0 \quad (4)$$

for all  $E$  of the form (2).

Since there are  $m$  independent constraints on essentially  $(n-r)r + 1$  coordinates, the dimension of the space

$$\left\{ M = \begin{pmatrix} \lambda I_r & \emptyset \\ \tilde{M} & \emptyset \end{pmatrix} \mid \lambda \in K, \tilde{M} \in \text{Mat}_{n-r,r}(K) \text{ and } \text{tr } C_i M = 0 \text{ for all } i \right\}$$

is  $(n-r)r + 1 - m$ , which is one larger than that of its subspace  $\mathcal{N}$ . Hence the dimension of its affine subspace

$$\mathcal{E} := \left\{ M = \begin{pmatrix} I_r & \emptyset \\ \tilde{M} & \emptyset \end{pmatrix} \mid \tilde{M} \in \text{Mat}_{n-r,r}(K) \text{ and } \text{tr } C_i M = 0 \text{ for all } i \right\}$$

which contains all idempotents of the form (2) in  $\mathcal{M}$ , is  $(n-r)r - m$ , just as the dimension of  $\mathcal{N}$ .

Now it remains to show that  $\mathcal{E}$  does not contain any idempotent outside  $\mathcal{M}$ . For that purpose, let  $E \in \mathcal{E}$  and suppose that there exist a  $C' \in \mathcal{C}$  such that  $\text{tr } C' E \neq 0$ . By (3) and by definition of  $\mathcal{E}$ , there exists a  $C^* \in \mathcal{C}$  such that  $\text{tr } C^* E \neq 0$  and  $\square_r(C^*)$  is the zero matrix. This contradicts (4), so a  $C'$  as above does not exist and we have  $\mathcal{E} \subseteq \mathcal{M}$ . Hence  $\mathcal{E}$  is the affine subspace of idempotents of the form (2) in  $\mathcal{M}$ .  $\square$



**Corollary 3.2.** *Assume  $I_n \notin \mathcal{C}$  and suppose that for some  $r$  with  $1 \leq r \leq n-1$  we have the following: all  $C \in \mathcal{C} \oplus KI_n$ , such that  $\square_r(C)$  is the zero matrix, are dependent of  $I_n$ . Then  $\mathcal{M}$  contains an idempotent of rank  $r$  and another one of rank  $n-r$ , such that the sum of both idempotents is unipotent.*

*Furthermore, if  $\mathcal{M}$  is a Mathieu subspace of any type, then  $\mathcal{M} = \text{Mat}_n(K)$  and  $\mathcal{C} = 0$ .*

*Proof.* Since  $I_n \notin \mathcal{C}$ , we see that all  $C \in \mathcal{C}$ , such that  $\square_r(C)$  is the zero matrix, are entirely zero by assumption. By (i) of theorem 3.1,  $\mathcal{M}$  contains an idempotent of the form

$$E := \begin{pmatrix} I_r & \emptyset \\ * & \emptyset \end{pmatrix}$$

By (ii) of theorem 3.1,  $\mathcal{M}$  contains another idempotent of the form

$$E' := \begin{pmatrix} \emptyset & \emptyset \\ * & I_{n-r} \end{pmatrix}$$

Notice that  $E + E'$  is unipotent and hence invertible. If  $\mathcal{M}$  is a left Mathieu subspace and  $A \in \text{Mat}_n(K)$ , then

$$A = AI_n = A(E + E')^{-1}E + A(E + E')^{-1}E' \in \mathcal{M}$$

because  $E^m = E \in \mathcal{M}$  and  $(E')^m = E' \in \mathcal{M}$  for all  $m \geq 1$ . Thus  $\mathcal{M} = \text{Mat}_n(K)$  and  $\mathcal{C} = 0$  in the case where  $\mathcal{M}$  is a left Mathieu subspace. The case where  $\mathcal{M}$  is a right Mathieu subspace is similar.  $\square$

Write  $x$  be the column vector  $(x_1, x_2, \dots, x_n)$ . We will show that theorem 3.3 below implies theorem 1.4.

**Theorem 3.3.** *Suppose that  $I_n \notin \mathcal{C}$  and  $0 < \dim_K \mathcal{C} < n$ . Let  $\mathcal{C}_n = \mathcal{C} \oplus KI_n$  and suppose that*

$$\begin{aligned} \#K \geq r + 1 &:= \dim_{K(x)} ((K(x) \otimes_K \mathcal{C}_n) \cdot x) \\ &= \dim_{K(x)} \sum_{C \in \mathcal{C}_n} K(x) \cdot C \cdot x \end{aligned}$$

*Then we can obtain corollary 3.2 (with a corresponding  $r$ ) by way of linear conjugation (replacing  $\mathcal{M}$  by  $T^{-1}\mathcal{M}T$  and  $\mathcal{C}$  by  $T^{-1}\mathcal{C}T$  for some  $T \in \text{GL}_n(K)$ ).*

*Proof of theorem 1.4.* The primary result to show is, that  $\text{tr } M = 0$  for all  $M \in \mathcal{M}$ . This is equivalent to  $I_n \in \mathcal{C}$ , so suppose that  $I_n \notin \mathcal{C}$ . Let  $\mathcal{C}_n = \mathcal{C} \oplus KI_n$ . By assumption,  $\dim_K \mathcal{C} < \min\{n, \#K\}$ . Hence  $\dim_K \mathcal{C} < n$  and

$$\dim_{K(x)} \sum_{C \in \mathcal{C}_n} K(x) \cdot C \cdot x \leq \dim_{K(x)} \sum_{C \in \mathcal{C}_n} K(x) \cdot C = \dim_K \mathcal{C}_n \leq \#K$$

Now theorem 3.3 above gives a contradiction, so  $I_n \in \mathcal{C}$  and hence  $\text{tr } M = 0$  for all  $M \in \mathcal{M}$ .

Since  $\mathcal{M}$  is proper by assumption,  $I_n \notin \mathcal{M}$ . Hence the secondary results follow from (2)  $\Rightarrow$  (3)  $\Rightarrow$  (4) of lemma 1.3.  $\square$

## 4 A binary matrix about a filtration on the constraint space

Write  $e_i$  for the  $i$ -th standard basis unit vector as a column vector. Let  $\mathcal{C}_n$  be a  $K$ -subspace of  $\text{Mat}_n(K)$  and define

$$\mathcal{C}_k := \{C \in \mathcal{C}_n \mid Ce_{k+1} = Ce_{k+2} = \cdots = Ce_n = 0\}$$

Then  $0 = \mathcal{C}_0 \subseteq \mathcal{C}_1 \subseteq \mathcal{C}_2 \subseteq \cdots \subseteq \mathcal{C}_n$  is a filtration in the sense that we can take quotients  $\mathcal{C}_j/\mathcal{C}_{j-1}$ , which are isomorphic to  $\mathcal{C}_j e_j$ , where  $\mathcal{C}_j v := \{Cv \mid C \in \mathcal{C}_j\}$ . Define the binary matrix  $B \in \text{Mat}_n(\{0, 1\})$  by

$$B_{ij} := \dim_K e_i^t \mathcal{C}_j e_j$$

for all  $i, j$ , where

$$e_i^t \mathcal{C}_j v := \{e_i^t Cv \mid C \in \mathcal{C}_j\} = \{(Cv)_i \mid C \in \mathcal{C}_j\}$$

Write  $b_j$  for the number of ones in column  $j$  of  $B$ .

Theorem 4.1 below can be formulated in terms of the binary matrix  $B$ . We will show that it implies theorem 3.3 (and hence also theorem 1.4). The next section will be devoted to the proof of theorem 4.1.

**Theorem 4.1.** *Suppose that  $\#K \geq r + 1$ , where  $r + 1$  is as defined in theorem 3.3. By way of linear conjugation, we can obtain the following.*

- (i)  $b_j = \dim_K \mathcal{C}_j e_j$  for all  $j$ ,  $b_n = r + 1$ , and  $B$  is increasing in every row, i.e.  $B_{ij} = 0$  implies  $B_{i(j-1)} = 0$  for every  $i, j$  such that  $j > 1$ .
- (ii) If  $\#K > \min\{b_{n-1}, n - 1\}$ , then  $B$  is decreasing above the diagonal in every column, i.e.  $B_{ij} = 0$  implies  $B_{(i+1)j} = 0$  for every  $i, j$  such that  $i + 1 < j$ .
- (iii) If  $I_n \in \mathcal{C}_n$ , then  $b_n > \min\{b_{n-1}, n - 1\}$  and  $B_{(n-1)n} \geq B_{n(n-1)}$ .

*Proof of theorem 3.3.* On account of theorem 4.1, we can apply a linear conjugation on  $\mathcal{C}_n$  such that the assertions of theorem 4.1 are satisfied. By (i), we have  $\#K \geq r + 1 = b_n$  and by (iii), we have  $b_n > \min\{b_{n-1}, n - 1\}$ . Hence the condition  $\#K > \min\{b_{n-1}, n - 1\}$  in (ii) is fulfilled.

- (i) *We first show that the first  $r$  columns of  $B$  are zero.* For that purpose, take  $k$  minimal such that  $b_k \geq 1$ . On account of (i) of theorem 4.1, we even have  $b_j \geq 1$  for all  $j \geq k$ . Since  $\mathcal{C}_j/\mathcal{C}_{j-1}$  is isomorphic to  $\mathcal{C}_j e_j$  for all  $j$ , it follows from (i) of theorem 4.1 that  $b_j = \dim_K \mathcal{C}_j e_j = \dim_K \mathcal{C}_j/\mathcal{C}_{j-1}$  for all  $j$ , and

$$\begin{aligned} n &\geq \dim_K \mathcal{C} + 1 = \dim_K \mathcal{C}_n \\ &= \dim_K \mathcal{C}_1/\mathcal{C}_0 + \dim_K \mathcal{C}_2/\mathcal{C}_1 + \cdots + \dim_K \mathcal{C}_n/\mathcal{C}_{n-1} \\ &= b_1 + \cdots + b_{k-1} + b_k + \cdots + b_{n-1} + b_n \\ &\geq 0 + \cdots + 0 + 1 + \cdots + 1 + (r + 1) \\ &= n - k + r + 1 \end{aligned}$$

So  $k \geq r + 1$  and indeed  $b_1 = b_2 = \cdots = b_r = 0$ .

- (ii) We next show that  $B_{nn} = 1$  is the only nonzero entry in the last  $n - r$  rows of  $B$ . At first,  $B_{nn} = 1$  follows directly from  $I_n \in \mathcal{C}_n$ . If  $r = n - 1$ , then  $B_{nj} = 0$  for all  $j \leq n - 1$  because of (i) above, which gives the claimed result. Hence assume that  $r < n - 1$ . Since  $b_n = r + 1$ , it follows from (ii) of theorem 4.1 that  $B_{(r+1)n} = B_{(r+2)n} = \cdots = B_{(n-1)n} = 0$ . In particular  $B_{(n-1)n} = 0$ , and (iii) of theorem 4.1 subsequently gives  $B_{n(n-1)} = 0$ . By (i) of theorem 4.1, every row of  $B$  is increasing. Hence every entry in the last  $n - r$  rows of  $B$  that has not been mentioned yet is zero as well.

Take  $C \in \mathcal{C}_n$  such that  $\square_r(C)$  is the zero matrix. We must show that  $C = \lambda I_n$  for some  $\lambda \in K$ . Take  $\lambda \in K$  such that the lower right corner entry of  $C' := C - \lambda I_n$  is zero. Notice that  $\square_r(C') = \square_r(C)$ . We must show that  $C' = 0$ .

So assume that  $C' \neq 0$ . Take  $k \leq n$  maximal such that  $C'_{ik} \neq 0$  for some  $i \leq n$ . Then  $C' \in \mathcal{C}_k$  and the  $i$ -th coordinate of  $C'e_k$  is nonzero, so  $B_{ik} = 1$ . On account of (i), we have  $k \geq r + 1$ , and by the fact that  $\square_r(C')$  is the zero matrix,  $i \geq r + 1$  as well. By (ii), we even have  $i = k = n$ , so  $C'_{nn} \neq 0$ . Contradiction.  $\square$

The following lemma is not very hard, but it is used several times.

**Lemma 4.2.** *For all  $j$ , we have*

$$b_j \geq \dim_K \mathcal{C}_j e_j$$

*and equality holds, if and only if  $B_{ij}e_i \in \mathcal{C}_j e_j$  for all  $i$ , if and only if  $\mathcal{C}_j e_j$  is the linear span of standard basis unit vectors.*

*Proof.* Notice that  $\mathcal{C}_j e_j$  is the linear span of standard basis unit vectors, if and only if for all  $i$  such that  $\mathcal{C}_j e_j$  is nontrivial at the  $i$ -th coordinate, we have  $e_i \in \mathcal{C}_j e_j$ . This is equivalent to that  $B_{ij}e_i \in \mathcal{C}_j e_j$  for all  $i$ .

Let  $U$  be the linear span of the standard basis unit vectors  $e_i$  for which  $e_i \mathcal{C}_j e_j \neq \{0\}$ . Then  $U$  is a space of dimension  $b_j$  which contains  $\mathcal{C}_j e_j$ . So  $b_j \geq \dim_K \mathcal{C}_j e_j$ , and if  $b_j = \dim_K \mathcal{C}_j e_j$ , then  $\mathcal{C}_j e_j = U$  is the linear span of standard basis unit vectors.

If  $b_j > \dim_K \mathcal{C}_j e_j$ , then there must be a standard basis unit vector of  $U$  that is not contained in  $\mathcal{C}_j e_j$ , while the corresponding coordinate projection of  $\mathcal{C}_j e_j$  is nontrivial. So  $\mathcal{C}_j e_j$  is not the linear span of standard basis unit vectors if  $b_j > \dim_K \mathcal{C}_j e_j$ .  $\square$

In order to prove theorem 4.1, we will use the following lemma. The assertion that  $B_{ij} = 0$  implies  $B_{i(j-1)} = 0$  can be found in the conclusion of (ii). Taking  $k = n - 1$  in the conclusion of (iii) gives  $B_{(n-1)n} \geq B_{n(n-1)}$ , which is another assertion of theorem 4.1.

**Lemma 4.3.** *Suppose that  $\dim_K \mathcal{C}_j e_j \geq \dim_{K(x_j)} ((K(x_j) \otimes_K \mathcal{C}_j) \cdot (e_k + x_j e_j))$ . Then we have the following.*

(i)  $\dim_K \mathcal{C}_j e_j = \dim_{K(x_j)} ((K(x_j) \otimes_K \mathcal{C}_j) \cdot (e_k + x_j e_j))$ .

- (ii) *If  $B_{ij} = 0$  for some  $i$ , then  $e_i^\dagger \mathcal{C}_{j-1} e_k = \{0\}$  as well.*

*In particular,  $B_{ij} = 0$  implies  $B_{i(j-1)} = 0$  if  $k = j - 1$ .*

- (iii) If  $B_{ij} = 0$  for some  $i$  and there exists a  $C' \in \mathcal{C}_j$  such that  $e_i^t C' e_k \neq 0$ , then  $C' e_j \notin \mathcal{C}_{j-1} e_k$ .

In particular, we have  $B_{kn} \geq B_{nk}$  if  $j = n$ ,  $I_n \in \mathcal{C}_n$  and  $b_k = \dim_K \mathcal{C}_k e_k$ .

*Proof.*

- (i) Let  $d := \dim_K \mathcal{C}_j e_j$ . Then we can find  $C_1, C_2, \dots, C_d \in \mathcal{C}_j$  such that  $\mathcal{C}_j e_j = KC_1 e_j \oplus KC_2 e_j \oplus \dots \oplus KC_d e_j$ . Hence the  $n \times d$  matrix

$$\left( C_1 x_j e_j \mid C_2 x_j e_j \mid \dots \mid C_d x_j e_j \right)$$

has a minor of size  $d$  which has degree  $d$ . The corresponding minor of the  $n \times d$  matrix

$$\left( C_1(e_k + x_j e_j) \mid C_2(e_k + x_j e_j) \mid \dots \mid C_d(e_k + x_j e_j) \right)$$

has degree  $d$  as well, so  $\dim_K \mathcal{C}_j e_j \leq \dim_{K(x_j)} \mathcal{C}_j(e_k + x_j e_j)$ , and (i) follows by assumption.

- (ii) By taking  $k = j - 1$ , the last claim follows from the first claim. To prove the first claim, suppose that  $i \leq n$  and that there exists a  $C_{d+1} \in \mathcal{C}_{j-1}$  such that  $e_i^t C_{d+1} e_k \neq 0$ . Then  $C_{d+1} e_j = 0$ , so we have  $C_{d+1}(e_k + x_j e_j) \in K^{\times n}$  and  $e_i^t C_{d+1}(e_k + x_j e_j) \in K^*$ . Suppose additionally that  $B_{ij} = 0$ . Then  $e_i^t \mathcal{C}_j e_j = \{0\}$ , so the  $i$ -th rows of the matrices of size  $n \times d$  in the proof of (i) are constant. It follows that the minors of these matrices in the proof of (i) do not use row  $i$ .

By expansion along the  $i$ -th row or the  $(d+1)$ -th column, which are both constant, we see that the  $n \times (d+1)$  matrix

$$\left( C_1(e_k + x_j e_j) \mid C_2(e_k + x_j e_j) \mid \dots \mid C_d(e_k + x_j e_j) \mid C_{d+1}(e_k + x_j e_j) \right)$$

has a minor of size  $d+1$  which has degree  $d$ , namely the minor of size  $d$  in the proof of (i), extended with row  $i$  and column  $d+1$ . This contradicts  $\dim_K \mathcal{C}_j e_j \geq \dim_{K(x_j)} ((K(x_j) \otimes_K \mathcal{C}_j) \cdot (e_k + x_j e_j))$ .

- (iii) We first show that the first claim implies the last claim. Take  $i = k$ ,  $j = n$  and  $C' = I_n$  in the first claim. Assuming the first claim, we see that  $B_{kn} = 0$  and  $I_n \in \mathcal{C}_n$  together imply  $e_n = I_n e_n \notin \mathcal{C}_{n-1} e_k$ . Now suppose that  $B_{kn} < B_{nk}$  and  $I_n \in \mathcal{C}_n$ . Then  $k \leq n - 1$  and  $B_{nk} = 1$ , so that  $B_{nk} e_n = e_n \notin \mathcal{C}_{n-1} e_k \supseteq \mathcal{C}_k e_k$ . From lemma 4.2, we deduce that  $b_k > \dim_K \mathcal{C}_k e_k$ . This gives the last claim.

To prove the first claim, suppose that  $B_{ij} = 0$  and there exists a  $C' \in \mathcal{C}_j$  such that  $e_i^t C' e_k \neq 0$ . By (ii), we have  $C' \notin \mathcal{C}_{j-1}$ , thus we may assume that  $C_d = C'$  in the proof of (i). Just as in the proof of (ii), we can see that the minors in the proof of (i) does not use row  $i$ , because that row is constant with respect to  $x_j$ .

Suppose additionally that  $C' e_j \in \mathcal{C}_{j-1} e_k$ . Then there exists a  $C_{d+1} \in \mathcal{C}_{j-1}$  such that  $C_{d+1} e_k = C' e_j = C_d e_j$ . Since  $x_j C_{d+1} e_k = x_j C_d e_j$  and  $x_j^2 C_{d+1} e_j \in x_j^2 \mathcal{C}_{j-1} e_j = 0$ , it follows that

$$(x_j C_{d+1} - C_d)(e_k + x_j e_j) = -C_d e_k \in K^{\times n}$$

and

$$e_i^\dagger(x_j C_{d+1} - C_d)(e_k + x_j e_j) = -e_i^\dagger C_d e_k = -e_i^\dagger C' e_k \in K^*$$

By expansion along the  $i$ -th row or the  $(d+1)$ -th column, which are both constant, we see that the  $n \times (d+1)$  matrix

$$\left( C_1(e_k + x_j e_j) \middle| C_2(e_k + x_j e_j) \middle| \cdots \middle| C_d(e_k + x_j e_j) \middle| (x_j C_{d+1} - C_d)(e_k + x_j e_j) \right)$$

has a minor of size  $d+1$  which has degree  $d$ , namely the minor of size  $d$  in the proof of (i), extended with row  $i$  and column  $d+1$ . This contradicts  $\dim_K \mathcal{C}_j e_j \geq \dim_{K(x_j)} ((K(x_j) \otimes_K \mathcal{C}_j) \cdot (e_k + x_j e_j))$ .  $\square$

## 5 Proof of theorem 4.1

The following two lemmas are not really necessary for the proof if the base field  $K$  is infinite.

**Lemma 5.1.** *Let  $K$  be a field and  $f \in K[x] = K[x_1, x_2, \dots, x_n]$  such that  $\deg f \leq d$ . Suppose that  $S \subseteq K$  such that  $f$  vanishes on  $S^{\times n}$ . Then  $f = 0$  in the following cases.*

- (i)  $\#S > d$ ,
- (ii)  $f$  is homogeneous,  $0 \in S$  and  $\#S \geq \max\{d, 2\}$ .

*Proof.* By replacing  $f(x)$  by  $f(x-s)$  for some  $s \in S$ , we may assume that  $0 \in S$  in (i) as well. Let  $\tilde{S} = S \setminus \{0\}$ .

- (i) We can write

$$f(x) = f(x_1, x_2, \dots, x_{n-1}, 0) + x_n g(x)$$

Notice that  $f(x)$  and hence also  $f(x_1, x_2, \dots, x_{n-1}, 0)$  vanishes at  $S^{\times(n-1)} \times \{0\}$ . By induction on  $n$ , we deduce that  $f(x_1, x_2, \dots, x_{n-1}, 0) = 0$ , so  $x_n g$  vanishes at  $S^{\times n}$ . Since  $x_n$  does not vanish anywhere at  $\tilde{S}^{\times n}$ , we conclude that  $g$  vanishes at  $\tilde{S}^{\times n}$ . By induction on  $d$ ,  $g = 0$ , so  $f = 0$  as well.

- (ii) If  $x_n^2 \mid f$ , then we can apply (i) on  $x_n^{-1} f$  instead of  $f$ , to obtain  $f = 0$ . The case  $\deg f \leq 1$  follows from (i) as well. So assume that  $\deg f \geq 2$  and  $x_n^2 \nmid f$ . Then  $n \geq 2$ . Take  $g(x)$  as in (i). Just as in (i),  $f(x_1, x_2, \dots, x_{n-1}, 0) = 0$  follows by induction and  $x_n g$  vanishes at  $S^{\times n}$ . Write

$$x_n g(x) = x_n g(x_1, x_2, \dots, x_{n-2}, 0, x_n) + x_{n-1} x_n h(x)$$

Notice that  $x_n g(x)$  and hence also  $x_n g(x_1, x_2, \dots, x_{n-2}, 0, x_n)$  vanishes at  $S^{\times(n-2)} \times \{0\} \times S$ . By induction on the number of variables, we deduce that  $x_n g(x_1, x_2, \dots, x_{n-2}, 0, x_n) = 0$ . Hence  $x_{n-1} x_n h(x)$  vanishes at  $S^{\times n}$ . Since  $x_{n-1} x_n$  does not vanish anywhere at  $\tilde{S}^{\times n}$ , we conclude that  $h$  vanishes at  $\tilde{S}^{\times n}$ . On account of  $\#\tilde{S} \geq d-1 > d-2 \geq \deg h$ ,  $h = 0$  follows from (i). So  $f = 0$  once again.  $\square$

Suppose that  $K$  has a  $(q-1)$ -th root of unity, e.g.  $K = \mathbb{F}_q$ . The polynomials  $x_1^{q-1} - 1$  and  $x_1^q - x_1$  show that  $\#S > d$  is necessary in (i). The polynomials  $x_1^{q-1} - x_2^{q-1}$  and  $x_1^q x_2 - x_1 x_2^q$  show that  $0 \in S$  and  $\#S \geq d$  respectively are necessary in (ii).

Notice that  $1\frac{1}{2}$  lies between the degrees of the leading and the trailing term of  $x_1^q - x_1$ . Since  $\deg(x_1^q - x_1) \geq \#\mathbb{F}_q$ , this is no coincidence, because the homogeneity condition in (ii) can be replaced by that  $1\frac{1}{2}$  is not contained in the interval that envelops the term degrees (and the proof of (ii) still applies).

**Lemma 5.2.** *Let  $L/K$  be a field extension (possibly trivial) and let  $\mathcal{V}$  be a subspace of  $\text{Mat}_{m,n}(K)$ . Define*

$$\begin{aligned} d &:= \dim_{K(x)} ((K(x) \otimes_K \mathcal{V}) \cdot x) \\ &= \dim_{K(x)} \sum_{V \in \mathcal{V}} K(x) \cdot V \cdot x \end{aligned}$$

Then we have the following.

- (i) For all  $v \in L^{\times n}$ , we have  $\dim_L ((L \otimes_K \mathcal{V}) \cdot v) \leq d$ .
- (ii) If  $\#L \geq d$ , then there exists a vector  $v \in L^{\times n}$  such that  $\dim_L ((L \otimes_K \mathcal{V}) \cdot v) = d$ .
- (iii) If  $\#L > d$ , then for each  $k \leq n$ , there exists a vector  $v \in L^{\times n}$  such that  $\dim_L ((L \otimes_K \mathcal{V}) \cdot v) = d$  and  $v_k = 1$ .

*Proof.* Let  $D := \dim_K \mathcal{V}$  and take a basis  $V_1, V_2, \dots, V_D$  of  $\mathcal{V}$ . Since  $D = \dim_{K(x)}(K(x) \otimes_K \mathcal{V})$ ,  $V_1, V_2, \dots, V_D$  is also a basis of  $K(x) \otimes_K \mathcal{V}$ . Hence  $V_1 x, V_2 x, \dots, V_D x$  is a spanning set of  $(K(x) \otimes_K \mathcal{V}) \cdot x$ . After an appropriate renumbering of the  $V_i$ 's, we have that  $V_1 x, V_2 x, \dots, V_d x$  is a basis of  $(K(x) \otimes_K \mathcal{V}) \cdot x$ .

- (i) Take any  $v \in L^{\times n}$ . Notice that  $V_1, V_2, \dots, V_D$  is also a basis of  $K(v) \otimes_K \mathcal{V}$ . Hence  $V_1 v, V_2 v, \dots, V_D v$  is a spanning set of  $(K(v) \otimes_K \mathcal{V}) \cdot v$ . Suppose that we have a subset  $\{W_1, W_2, \dots, W_{d+1}\}$  of  $\{V_1, V_2, \dots, V_D\}$  such that  $W_1 v, W_2 v, \dots, W_{d+1} v$  are independent over  $K(v)$ . Then the matrix with columns  $W_1 v, W_2 v, \dots, W_{d+1} v$  has a minor of size  $d+1$  that does not vanish. The corresponding minor of the matrix with columns  $W_1 x, W_2 x, \dots, W_{d+1} x$  does not vanish either, so  $W_1 x, W_2 x, \dots, W_{d+1} x$  are independent over  $K(x)$ . This contradicts the definition of  $d$ , so if we reduce  $V_1 v, V_2 v, \dots, V_D v$  to a basis, we get  $d' \leq d$  vectors  $W_1 v, W_2 v, \dots, W_{d'} v$ .
- (ii) If  $d = 0$ , then we can take  $v$  arbitrary on account of (i), so assume that  $d \geq 1$ . The matrix with columns  $V_1 x, V_2 x, \dots, V_d x$  contains a minor  $h(x) \neq 0$  of size  $d$  which has degree  $d$  and is homogeneous. Suppose that  $\#L \geq d$ . By (ii) of lemma 5.1, there exists a vector  $v \in L^{\times n}$  such that  $h(x)$  does not vanish at  $v$ . Hence the matrix with columns  $V_1 v, V_2 v, \dots, V_d v$  has a minor of size  $d$  that does not vanish either. This gives (ii).
- (iii) If  $d = 0$ , then we can take  $v = (1, 1, \dots, 1)$  on account of (i), so assume that  $d \geq 1$ . Suppose that  $\#L > d$  and take any  $k \leq n$ . Take  $h(x)$  as in the proof of (ii). By (ii) of lemma 5.1, there exists a vector  $v \in L^{\times n}$  such that  $x_k h(x)$  does not vanish at  $v$ . Hence we can deduce the conclusion of (ii) once again. Since we have  $v_k \neq 0$  in addition, we can obtain  $v_k = 1$  by dividing  $v$  by  $v_k$ , because  $h$  is homogeneous.  $\square$

From now on in this section, we assume that  $\mathcal{C}_n$  is a subspace of  $\text{Mat}_n(K)$ , and define

$$\mathcal{C}_k := \{C \in \mathcal{C}_n \mid Ce_{k+1} = Ce_{k+2} = \cdots = Ce_n = 0\}$$

for all  $k < n$ , where  $e_i$  is the  $i$ -th standard basis unit vector.

Define

$$\begin{aligned} d_k &:= \dim_{K(x)} \left( (K(x) \otimes_K \mathcal{C}_k) \cdot x \right) \\ &= \dim_{K(x)} \sum_{C \in \mathcal{C}_k} K(x) \cdot C \cdot x \end{aligned}$$

Notice that  $0 = d_0 \leq d_1 \leq d_2 \leq \cdots \leq d_n = r + 1$ , where  $r + 1$  is as in theorems 3.3 and 4.1.

Lemma 5.2 leads to the following corollary.

**Corollary 5.3.** *If  $\#K \geq d_k$ , then there exists a  $v \in K^{\times n}$  with  $v_{k+1} = v_{k+2} = \cdots = v_n = 0$ , such that  $d_k = \dim_K \mathcal{C}_k v$ . If  $\#K > d_k$ , then we can additionally take  $v_k = 1$ .*

*Proof.* The existence of a vector  $v$  as claimed, except that  $v_{k+1} = v_{k+2} = \cdots = v_n = 0$ , follows from (ii) and (iii) of lemma 5.2 respectively. Since columns  $k + 1, k + 2, \dots, n$  of  $\mathcal{C}_k$  are zero, we can indeed take  $v_{k+1} = v_{k+2} = \cdots = v_n = 0$ .  $\square$

Unlike  $d_n = r + 1$ ,  $d_k$  is not invariant under linear conjugation in general. But  $d_k$  is indeed invariant under conjugation with lower triangular linear maps for every  $k$ , because of (i) of the following lemma.

**Lemma 5.4.** *Suppose that the last  $n - k$  columns of  $T \in \text{GL}_n(K)$  match those of a lower triangular matrix. Then we have the following changes when we replace  $\mathcal{C}_n$  by  $T^{-1}\mathcal{C}_nT$ .*

- (i)  $\mathcal{C}_k e_k$  gets replaced by  $T^{-1}\mathcal{C}_k T e_k$  and  $d_k$  stays the same.
- (ii) If the  $k$ -th column  $T e_k$  of  $T$  is zero above the diagonal, then  $\mathcal{C}_k e_k$  gets replaced by  $T^{-1}\mathcal{C}_k e_k$  and  $\dim_K \mathcal{C}_k e_k$  stays the same.

Furthermore, we have the following for all  $j > k$  when we replace  $\mathcal{C}_n$  by  $T^{-1}\mathcal{C}_nT$ .

- (iii)  $\mathcal{C}_j e_j$  gets replaced by  $T^{-1}\mathcal{C}_j e_j$  and  $d_j$  and  $\dim_K \mathcal{C}_j e_j$  stay the same.
- (iv) If  $B_{ij} = 1$  implies  $T e_i = e_i$  for every  $i$ , then  $B_{ij}$  will not change for any  $i$ .
- (v) If  $B_{ij} = 0$  implies  $e_i^t T = e_i^t$  for each  $i$  and  $b_j = \dim_K \mathcal{C}_j e_j$ , then  $B_{ij}$  will not change for any  $i$ .

*Proof.* Since  $T$  is lower triangular at the last  $n - k$  columns, the last  $n - k$  columns of  $C \in \text{Mat}_n(K)$  are zero, if and only if the last  $n - k$  columns of  $CT$  are zero, if and only if the last  $n - k$  columns of  $T^{-1}CT$  are zero. Hence  $\mathcal{C}_k$  gets replaced by  $T^{-1}\mathcal{C}_kT$  when we replace  $\mathcal{C}_n$  by  $T^{-1}\mathcal{C}_nT$ .

- (i) Since  $\mathcal{C}_k$  gets replaced by  $T^{-1}\mathcal{C}_kT$ , the first claim is obvious. For a vector  $v \in K(x)^{\times n}$ , let  $\phi(v) = T^{-1}v|_{x=Tx}$ , where  $|_{x=f(x)}$  means substituting  $x$  by  $f(x)$ . Then  $\phi^{-1}(v) = Tv|_{x=T^{-1}x}$ , so  $\phi$  is an isomorphism between the spaces  $(K(x) \otimes_K \mathcal{C}_k) \cdot x$  and  $T^{-1} \cdot (K(x) \otimes_K \mathcal{C}_k) \cdot Tx$ . In particular, the dimensions of these spaces are equal, which gives the second claim.

(ii) Since  $\mathcal{C}_k e_k$  and  $T^{-1}\mathcal{C}_k e_k$  are isomorphic, the second claim follow from the first. Hence by (i), it suffices to show that  $\mathcal{C}_k T e_k = \mathcal{C}_k e_k$ . For that purpose, assume that  $T e_k$  is zero above the  $k$ -th coordinate. Since  $\mathcal{C}_k$  in turn is zero at the right of the  $k$ -th column, only the  $k$ -th column of  $\mathcal{C}_k$  and the  $k$ -th coordinate of  $T e_k$  contribute to the product  $\mathcal{C}_k \cdot T e_k$ , i.e.  $\mathcal{C}_k \cdot T e_k = \mathcal{C}_k e_k \cdot e_k^\dagger T e_k$ .

The  $k$ -th coordinate  $e_k^\dagger T e_k$  of  $T e_k$  is nonzero, because  $T \in \text{GL}_n(K)$  is lower triangular at the last  $n - k + 1$  columns. So we can cancel  $e_k^\dagger T e_k$  to obtain  $\mathcal{C}_k T e_k = \mathcal{C}_k e_k$ .

(iii) Since  $T$  is lower triangular at the last  $n - j + 1$  columns, the desired results follow from (ii), (i) and (ii) respectively.

(iv) Assume that  $B_{ij} = 1$  implies  $T e_i = e_i$  for all  $i$ . We prove that  $B_{ij}$  will not change for any  $i$  by showing that  $\mathcal{C}_j e_j$  stays the same. By (iii),  $\mathcal{C}_j e_j$  gets replaced by  $T^{-1}\mathcal{C}_j e_j$ , so it suffices to show  $(T^{-1} - I_n)\mathcal{C}_j e_j = 0$ . If  $B_{ij} = 0$ , then the  $i$ -th coordinate of  $\mathcal{C}_j e_j$  is zero. If  $B_{ij} = 1$ , then the  $i$ -th column of  $T^{-1} - I_n = T^{-1}(I_n - T)$  is zero by assumption. So

$$(T^{-1} - I_n)\mathcal{C}_j e_j = (T^{-1} - I_n)I_n \mathcal{C}_j e_j \subseteq \sum_{i=1}^n (T^{-1} - I_n)e_i \cdot e_i^\dagger \mathcal{C}_j e_j = 0$$

indeed.

(v) Assume that  $b_j = \dim_K \mathcal{C}_j e_j$ . By (iii),  $\dim_K \mathcal{C}_j e_j$  will stay the same, so by lemma 4.2,  $b_j$  cannot decrease. So if some  $B_{ij}$  changes, there will be an  $i$  such that  $B_{ij}$  changes from 0 to 1, which we assume from now on. We additionally assume that  $B_{ij} = 0$  implies  $e_i^\dagger T = e_i^\dagger$ , so that  $e_i^\dagger T^{-1} = e_i^\dagger$  as well. By (iii),  $B_{ij} = \dim_K e_i^\dagger \mathcal{C}_j e_j$  gets replaced by  $\dim_K e_i^\dagger T^{-1} \mathcal{C}_j e_j = \dim_K e_i^\dagger \mathcal{C}_j e_j = B_{ij}$ . So  $B_{ij}$  will stay the same, which is a contradiction.  $\square$

*Proof of theorem 4.1.* If  $\dim_K \mathcal{C}_j e_j = d_j$  for some  $j$ , then by (i) of lemma 5.2 with  $L = K(x_j)$  and  $v = e_k + x_j e_j$ , the condition of lemma 4.3 is satisfied for every  $k$ . Hence we will additionally arrange that  $\dim_K \mathcal{C}_j e_j = d_j$  for all  $j$  by way of conjugation. As soon as we have  $\dim_K \mathcal{C}_j e_j = d_j$  for some  $j \geq 2$ , it follows from (ii) of lemma 4.3 that  $B_{ij} = 0$  implies  $B_{i(j-1)} = 0$  for all  $i$ , so we do not need to show that any more.

(i) (*Pass 1*) We start with obtaining  $\dim_K \mathcal{C}_j e_j = d_j$  for all  $j$ . Suppose inductively that  $\dim_K \mathcal{C}_j e_j = d_j$  for all  $j > k$  already. Since  $d_k \leq d_n = r + 1 \leq \#K$ , it follows from corollary 5.3 that there exists a  $v \in K^{\times n}$  with  $v_{k+1} = v_{k+2} = \dots = v_n = 0$ , such that  $d_k = \dim_K \mathcal{C}_k v$ . Take  $T \in \text{GL}_n(K)$  such that  $T e_k = v$  and  $T e_j = e_j$  for all  $j > k$ . Then  $T$  is as in lemma 5.4.

Now replace  $\mathcal{C}_n$  by  $T^{-1}\mathcal{C}_n T$ . By (i) of lemma 5.4,  $d_k$  will not change, and  $\mathcal{C}_k e_k$  will become  $T^{-1}\mathcal{C}_k T e_k = T^{-1}\mathcal{C}_k v$ . Since  $T^{-1}\mathcal{C}_k v$  is isomorphic to  $\mathcal{C}_k v$ ,  $\dim_K \mathcal{C}_k e_k$  will become  $\dim_K \mathcal{C}_k v = d_k$ . By (iii) of lemma 5.4,  $\dim_K \mathcal{C}_j e_j = d_j$  will not be affected for any  $j > k$ .

So we can obtain  $\dim_K \mathcal{C}_j e_j = d_j$  for all  $j$  inductively. The other claims of (i) follow as soon as we have  $b_j = \dim_K \mathcal{C}_j e_j = d_j$  for all  $j$ . We will arrange that by way of another induction pass.



(Pass 2) Suppose inductively that  $b_j = \dim_K \mathcal{C}_j e_j = d_j$  for all  $j > k$  already. We will obtain  $b_k = \dim_K \mathcal{C}_k e_k$  by way of a conjugation with a lower triangular matrix  $T$ . Just as above, the validity of  $\dim_K \mathcal{C}_j e_j = d_j$  for every  $j > k$  will not be affected. But the validity of  $\dim_K \mathcal{C}_j e_j = d_j$  will not be affected for any other  $j$  either, because  $T$  is lower triangular at the last  $n$  columns, see the proof of (iii) of lemma 5.4.

Take a basis of  $\mathcal{C}_k e_k$  such that the positions of the first nonzero coordinates of the basis vectors are all different. Next, take  $T \in \text{GL}_n$  lower triangular, such that every column of  $T$  is either one of those basis vectors of  $\mathcal{C}_k e_k$  (with its first nonzero coordinate on the diagonal of  $T$ ) or a standard basis unit vector (with its only nonzero coordinate on the diagonal of  $T$ ), in such a way that all those basis vectors of  $\mathcal{C}_k e_k$  are included.

Then  $T^{-1}$  maps those basis vectors of  $\mathcal{C}_k e_k$  to standard basis unit vectors (with corresponding positions of the first nonzero coordinate), so that  $T^{-1} \mathcal{C}_k e_k$  is spanned by standard basis unit vectors. By lemma 4.2,  $b_k$  will become equal to  $\dim_K \mathcal{C}_k e_k$  when  $\mathcal{C}_k e_k$  gets replaced by  $T^{-1} \mathcal{C}_k e_k$ . Now replace  $\mathcal{C}_n$  by  $T^{-1} \mathcal{C}_n T$ . By (ii) of lemma 5.4,  $\mathcal{C}_k e_k$  will indeed be replaced by  $T^{-1} \mathcal{C}_k e_k$ , so that  $b_k$  will become  $\dim_K \mathcal{C}_k e_k$ . Furthermore,  $\dim_K \mathcal{C}_k e_k$  and  $d_k$  will not change, so we indeed get  $b_k = \dim_K \mathcal{C}_k e_k = d_k$ .

We prove that  $b_j = \dim_K \mathcal{C}_j e_j$  will not be affected by this conjugation for any  $j > k$ , by showing that  $B_{ij}$  will not change for any  $i$  and any  $j > k$ . By (v) of lemma 5.4, it suffices to show that  $B_{ij} = 0$  implies  $e_i^t T = e_i^t$ . So assume that  $B_{ij} = 0$ . Since the  $i$ -th row of  $B$  is increasing, we have  $B_{ik} = 0$  as well. Hence the  $i$ -th coordinate of any vector of  $\mathcal{C}_k e_k$  is zero. By construction of  $T$ , we have  $e_i^t T = e_i^t$  indeed. So we can decrease  $k$  and proceed.

- (ii) (1 pass) We start with the first step of the first induction pass in (i), to obtain  $\dim_K \mathcal{C}_n e_n = d_n$ . As opposed to the double pass construction in (i), we will use a single pass construction here to fulfill the claims of (i) and (ii) and the additional claim that  $\dim_K \mathcal{C}_j e_j = d_j$  for all  $j$ , provided  $\#K > \min\{d_{n-1}, n-1\}$  after the first step of the first induction pass of (i) to obtain  $\dim_K \mathcal{C}_n e_n = d_n$ .

If  $\#K \leq \min\{d_{n-1}, n-1\}$  after the first step of the first induction pass in (i), then we proceed with the double pass construction of (i), to obtain  $b_{n-1} = \dim_K \mathcal{C}_{n-1} e_{n-1} = d_{n-1}$ . Since  $d_{n-1}$  does not change any more after the first step of the first induction pass of (i), we get  $\#K \leq \min\{b_{n-1}, n-1\}$ , which implies (ii).

So assume that  $\dim_K \mathcal{C}_n e_n = d_n$  and  $\#K > \min\{d_{n-1}, n-1\}$ . As long as  $d_k = n$ , we can proceed as in the first induction pass of (i) to obtain  $b_j = \dim_K \mathcal{C}_j e_j = d_j$  for all  $j \geq k$ , because by lemma 4.2, we have  $b_k = n$  automatically if  $\dim_K \mathcal{C}_k e_k = n$ . So suppose that  $d_k \leq n-1$  and that  $b_j = \dim_K \mathcal{C}_j e_j = d_j$  for all  $j > k$ .

(Step 1) We will first obtain  $\dim_K \mathcal{C}_k e_k = d_k$ . If  $k = n$ , then we have already obtained  $\dim_K \mathcal{C}_k e_k = d_k$ . So assume that  $k \leq n-1$ . Then  $d_k \leq \min\{d_{n-1}, n-1\} < \#K$ . It follows from corollary 5.3 that there exists a  $v \in K^{\times n}$  with  $v_{k+1} = v_{k+2} = \dots = v_n = 0$ , such that  $d_k = \dim_K \mathcal{C}_k v$  and additionally  $v_k = 1$ . Make  $T \in \text{GL}_n$  by replacing the  $k$ -th column of  $I_n$  by  $v$ .

Just as in the first pass of (i), we will obtain  $\dim_K \mathcal{C}_k e_k = d_k$  when we replace  $\mathcal{C}_n$  by  $T^{-1}\mathcal{C}_n T$ . Furthermore,  $d_k$  will not change, and neither will  $d_j$  and  $\dim_K \mathcal{C}_j e_j$  for any  $j > k$ . But as opposed to (i) and the case  $k = n$ , we have to show that  $b_j = \dim_K \mathcal{C}_j e_j$  will be preserved for all  $j > k$ , and that the rightmost  $n - k$  columns of  $B$  will stay decreasing above the diagonal. We do that by showing that the rightmost  $n - k$  columns of  $B$  will be preserved. For that purpose, take any column index  $j > k$ .

Since  $T$  is just the identity matrix outside column  $k$ , it follows from (iv) of lemma 5.4 that  $Be_j$  will stay the same in case  $B_{kj} = 0$ . Hence assume that  $B_{kj} = 1$ . Then the induction assumption tells us that even  $B_{1j} = B_{2j} = \dots = B_{kj} = 1$ . Since the last  $n - k$  rows of  $T$  are the same as those of  $I_n$ , it follows from (v) of lemma 5.4 that  $Be_j$  will stay the same again. So let us proceed with replacing  $\mathcal{C}_n$  by  $T^{-1}\mathcal{C}_n T$ .

(Step 2) The next thing to arrange is that  $b_k = \dim_K \mathcal{C}_k e_k$ , which can be done in the same manner as in the second induction pass of (i).

(Step 3) At last, we must make the  $k$ -th column of  $B$  decreasing above the diagonal. For that purpose, take  $s < k$  maximal, such that  $B_{sk} = 1$ . Then there exists a permutation matrix  $P$ , which matches the identity matrix outside the leading principal minor matrix of size  $s$ , such that  $PBe_k$  is decreasing above the  $k$ -th coordinate. Take  $T = P^{-1}$ . Then  $Pe_j = e_j = P^{-1}e_j = Te_j$  for all  $j \geq k$ , so  $T$  satisfies both the condition of lemma 5.4 and the additional condition of (ii) of lemma 5.4.

Now replace  $\mathcal{C}_n$  by  $P\mathcal{C}_n P^{-1} = T^{-1}\mathcal{C}_n T$ . By (i), (ii) and (iii) of lemma 5.4,  $\dim_K \mathcal{C}_j e_j$  and  $d_j$  will not change for any  $j \geq k$ . By (ii) of lemma 5.4,  $\mathcal{C}_k e_k$  will be replaced by  $T^{-1}\mathcal{C}_k e_k = P\mathcal{C}_k e_k$ , and  $Be_k$  will be replaced by  $PBe_k$  along with it. So  $Be_k$  will become decreasing above the  $k$ -th coordinate and  $b_k$  stays the same.

In order to prove that  $Be_j$  will stay decreasing above the  $j$ -th coordinate and that  $b_j$  will be maintained, for all  $j > k$ , we show that  $B_{ij}$  stays the same for all  $i$  and all  $j > k$ . By (v) of lemma 5.4, it suffices to show that  $B_{ij} = 0$  implies  $e_i^t T = e_i^t$ . If  $i > s$ , then  $e_i^t P = e_i^t = e_i^t P^{-1} = e_i^t T$ , so we may assume that  $i \leq s$ . By (ii) of lemma 4.3, which is valid as long as  $j > k$ , we have  $1 = B_{sk} = B_{s(k-1)} = \dots = B_{sj}$ . Since  $i \leq s < k < j$  and  $Be_j$  is decreasing above the  $j$ -th coordinate,  $B_{ij} = 1$  is satisfied as well as  $B_{sj} = 1$ . Hence  $B_{ij} = 0$  implies  $e_i^t T = e_i^t$  once again. So we can decrease  $k$  and proceed.

- (iii) Assume that  $I_n \in \mathcal{C}_n$ . Since  $b_{n-1} = \dim_K \mathcal{C}_{n-1} e_{n-1}$  on account of (i), we deduce from (iii) of lemma 4.3 that  $B_{(n-1)n} \geq B_{n(n-1)}$ . So  $b_n > \min\{b_{n-1}, n - 1\}$  remains to be proved. Hence assume that  $b_n \leq n - 1$ . Then there exists an  $i$  such that  $B_{in} = 0$ . By (ii) of lemma 4.3, we have  $e_i^t \mathcal{C}_{n-1} = \{0\}$ . So

$$e_i^t \cdot \mathcal{C}_{n-1} \cdot (e_i + x_{n-1} e_{n-1}) = 0 \neq 1 = e_i^t \cdot I_n \cdot (e_i + x_{n-1} e_{n-1})$$

Consequently, we deduce from (i) of lemma 4.3 that

$$\begin{aligned}
b_{n-1} &= \dim_K \mathcal{C}_{n-1} e_{n-1} \\
&\leq \dim_{K(x_{n-1})} (K(x_{n-1}) \otimes_K \mathcal{C}_{n-1}) \cdot (e_i + x_{n-1} e_{n-1}) \\
&< \dim_{K(x_{n-1})} (K(x_{n-1}) \otimes_K (\mathcal{C}_{n-1} + KI_n)) \cdot (e_i + x_{n-1} e_{n-1}) \\
&\leq \dim_{K(x_{n-1})} (K(x_{n-1}) \otimes_K \mathcal{C}_n) \cdot (e_i + x_{n-1} e_{n-1})
\end{aligned}$$

From (i) of lemma 5.2, it follows that the right hand side does not exceed  $\dim_{K(x)} ((K(x) \otimes_K \mathcal{C}_n) \cdot x)$ . So  $b_{n-1} < d_n$ . Since we arranged  $b_n = \dim_K \mathcal{C}_n e_n = d_n$ , we have  $b_n > b_{n-1} \geq \min\{b_{n-1}, n-1\}$ .  $\square$

The double pass construction in (i) of the proof of theorem 4.1 is needed because the first induction pass may affect  $b_j = \dim_K \mathcal{C}_j e_j$ . In the first step in (ii) of the proof of theorem 4.1, we additionally have  $v_k = 1$ , so that we can choose the transformation matrix  $T$  more conveniently than in the first induction pass in (i) of the proof of theorem 4.1. Consequently,  $b_j = \dim_K \mathcal{C}_j e_j$  will not be affected in the first step in (ii) of the proof of theorem 4.1, so that a double pass construction is not necessary there.

If  $n = 3$  and  $\mathcal{C}$  is the space over  $\mathbb{F}_2$  which is spanned by

$$\begin{pmatrix} \bar{0} & \bar{1} & \bar{0} \\ \bar{0} & \bar{1} & \bar{0} \\ \bar{0} & \bar{0} & \bar{0} \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \bar{0} & \bar{0} & \bar{0} \\ \bar{0} & \bar{1} & \bar{1} \\ \bar{0} & \bar{0} & \bar{0} \end{pmatrix}$$

then a computer calculation reveals that  $\mathcal{C}$  does not satisfy the claim of theorem 3.3. This is because the condition of lemma 4.3 cannot be met. We use lemma 5.1 to obtain this condition, but that requires a subset of cardinality three of  $\mathbb{F}_2$ .

## 6 The radical of a Mathieu subspace of $\text{Mat}_n(K)$

The results about the ideal  $I$  in the preamble of the theorem below are well-known. We have added the proof of these results for completeness only.

**Theorem 6.1.** *Assume  $\mathcal{M}$  is a  $K$ -subspace of  $\text{Mat}_n(K)$ . As  $\{0\} \subseteq \mathcal{M}$ , we can take a left ideal  $I$  of  $\text{Mat}_n(K)$  which is contained in  $\mathcal{M}$  and maximal as such. Then  $I$  is unique, has dimension  $nk$  for some  $k \leq n$  and there exist a  $T \in \text{GL}_n(K)$  such that*

$$IT = T^{-1}IT = \{M \in \text{Mat}_n(K) \mid Me_{k+1} = Me_{k+2} = \cdots = Me_n = 0\}$$

*Furthermore,  $I$  is a principal left ideal which is generated by an idempotent, and the following statements are equivalent.*

- (1)  $\mathcal{M}$  is a left Mathieu subspace of  $\text{Mat}_n(K)$ ,
- (2)  $I$  contains all idempotents of  $\mathcal{M}$ ,
- (3)  $\tau(\mathcal{M}) = \tau(I)$ .

*Proof.* Since  $\mathcal{M}$  is a  $K$ -subspace of  $\text{Mat}_n(K)$ , the sum of two left ideals contained in  $\mathcal{M}$  is again contained in  $\mathcal{M}$ . Since  $\mathcal{M}$  is a finite  $K$ -subspace of  $\text{Mat}_n(K)$ , we can deduce that  $I$  is unique.

Take  $M \in I$  of maximum rank  $k$ , and  $T \in \text{GL}_n(K)$  such that the last  $n - k$  columns of  $MT$  are zero. Since the first  $k$  columns of  $MT$  are independent of the last  $n - k$  columns, the subspace of  $A \in \text{Mat}_n(K)$  such that  $Ae_{k+1} = Ae_{k+2} = \cdots = Ae_n = 0$  is generated by  $MT$  and therefore contained in  $IT$ . If  $IT$  contains another matrix, then we get a contradiction with the maximality of  $k$ , because  $IT$  is a left ideal of  $M_n(K)$ .

Furthermore,  $IT = T^{-1}IT$  is a principal left ideal which is generated by

$$\begin{pmatrix} I_k & \emptyset \\ \emptyset & \emptyset \end{pmatrix}$$

Hence  $I$  is a principal left ideal which is generated by an idempotent as well. So it remains to show the following.

(2)  $\Rightarrow$  (1) This follows from [Zha3, Th. 4.2].

(3)  $\Rightarrow$  (2) Suppose that  $\tau(\mathcal{M}) = \tau(I)$ . Since each idempotent of  $\mathcal{M}$  is contained in  $\tau(\mathcal{M})$  and every idempotent in  $\tau(I)$  is already in  $I$ ,  $I$  contains all idempotents of  $\mathcal{M}$ .

(1)  $\Rightarrow$  (3) This follows from [Zha3, Lm. 4.9] or [Zha3, Th. 4.10].  $\square$

**Corollary 6.2.** *Suppose that  $\mathcal{M}$  is a left Mathieu subspace of  $\text{Mat}_n(K)$ , such that  $0 < n^2 - \dim_K \mathcal{M} < n$ . Then  $\mathcal{M}$  is even a two-sided Mathieu subspace of  $\text{Mat}_n(K)$  and  $\#K > 2$ .*

*Proof.* Take  $I$  as in theorem 6.1. We first prove that  $\mathcal{M}$  is even two-sided. On account of [Zha3, Th. 4.2], it suffices to show that  $\mathcal{M}$  has no nontrivial idempotent, which by (1)  $\Rightarrow$  (2) of theorem 6.1 comes down to that  $I$  has no nontrivial idempotents. Since theorem 6.1 additionally tells us that  $I$  is generated by a single idempotent, we just have to show that  $I = (0)$ .

So assume that  $I \neq (0)$ . On account of theorem 6.1,  $I$  has dimension  $nk$ , where  $1 \leq k \leq n - 1$  because  $0 < n^2 - \dim_K \mathcal{M}$ . Furthermore, we may assume that  $I = \{M \in \text{Mat}_n(K) \mid Me_{k+1} = Me_{k+2} = \cdots = Me_n = 0\}$ .

The space  $\mathcal{V}$  defined by

$$\left\{ M \in \mathcal{M} \mid M = \begin{pmatrix} \emptyset & \tilde{M} \\ \emptyset & \lambda I_{n-k} \end{pmatrix} \text{ for some } \tilde{M} \in \text{Mat}_{k, n-k}(K) \text{ and a } \lambda \in K \right\}$$

is the intersection of  $\mathcal{M}$  with a space of dimension  $k(n - k) + 1$ . Since the codimension of  $\mathcal{M}$  is less than  $n \leq k(n - k) + 1$ , we have  $\dim_K \mathcal{V} \geq 1$ , so  $\mathcal{V}$  has a nonzero element  $M$ . If  $\lambda \neq 0$  for  $M$ , then we take  $E = \lambda^{-1}M$ . If  $\lambda = 0$  for  $M$ , then we make  $E$  from  $M$  by replacing the leading principal minor matrix of size  $k$  by  $I_k$ , so that  $E - M \in I$ . In both cases,  $E$  is an idempotent of  $\mathcal{M}$  which is not contained in  $I$ . This contradicts (1)  $\Rightarrow$  (2) of theorem 6.1, so  $I = (0)$  indeed.

Next, we show that  $\#K > 2$ . Since the subspace of diagonal matrices of  $\text{Mat}_n(K)$  has dimension  $n$  and  $\mathcal{M}$  has codimension less than  $n$ ,  $\mathcal{M}$  contains at least two diagonal matrices, of which one, say  $E$ , is nonzero. If  $\#K = 2$ , then  $E$  is an idempotent and we have  $E \in I$  because  $\mathcal{M}$  is a left Mathieu subspace of  $\text{Mat}_n(K)$ . This contradicts  $I = (0)$ , so  $\#K > 2$ .  $\square$

## References

- [vdE] Arno van den Essen. The amazing image conjecture. arXiv:1006.5801, 2010.
- [Kon] Aart Konijnenberg. Mathieu subspaces of finite products of matrix rings. Master's thesis, Radboud University of Nijmegen, Toernooiveld, 6525 ED Nijmegen, The Netherlands, 2012. Directed by A.R.P. van den Essen.
- [Zha1] Wenhua Zhao. A generalization of Mathieu subspaces to modules of associative algebras. *Cent. Eur. J. Math.*, 8(6):1132–1155, 2010.
- [Zha2] Wenhua Zhao. Generalizations of the image conjecture and the Mathieu conjecture. *J. Pure Appl. Algebra*, 214(7):1200–1216, 2010.
- [Zha3] Wenhua Zhao. Mathieu subspaces of associative algebras. *J. Algebra*, 350:245–272, 2012.