A note on sums of three square-zero matrices

Clément de Seguins Pazzis∗†

May 18, 2016

Abstract

It is known that every complex trace-zero matrix is the sum of four square-zero matrices, but not necessarily of three such matrices. In this note, we prove that for every trace-zero matrix $A$ over an arbitrary field, there is a non-negative integer $p$ such that the extended matrix $A \oplus 0_p$ is the sum of three square-zero matrices (more precisely, one can simply take $p$ as the number of rows of $A$). Moreover, we demonstrate that if the underlying field has characteristic 2 then every trace-zero matrix is the sum of three square-zero matrices. We also discuss a counterpart of the latter result for sums of idempotents.

AMS Classification: 15A24; 15B33.

Keywords: Decomposition, Rational canonical form, Square-zero matrices, Fields with characteristic 2, Idempotent matrices.

1 Introduction

1.1 The problem

Let $\mathbb{F}$ be a field, and denote by $M_n(\mathbb{F})$ the algebra of all $n$ by $n$ square matrices with entries in $\mathbb{F}$. A matrix $A \in M_n(\mathbb{F})$ is called square-zero if $A^2 = 0$. Obviously, every square-zero matrix has trace zero, and hence the sum of finitely many square-zero matrices has trace zero. Conversely, by using the Jordan canonical form it is easy to split every nilpotent matrix into the sum of two-

∗Université de Versailles Saint-Quentin-en-Yvelines, Laboratoire de Mathématiques de Versailles, 45 avenue des États-Unis, 78035 Versailles cedex, France
†e-mail address: dsp.prof@gmail.com
square-zero matrices. On the other hand, it is well known that every trace-zero matrix is a sum of nilpotent matrices: More precisely, it is the sum of two such matrices if non-scalar, otherwise it is the sum of three nilpotent matrices (see e.g. [2]). Hence, every trace-zero matrix is the sum of at most six square-zero matrices, and of at most four such matrices if non-scalar. In particular, over a field of characteristic zero, every trace-zero matrix is the sum of four square-zero matrices. The result still holds over general fields, see Section 2.4 of the present article for a proof.

On the other hand, the classification of matrices that are the sum of two square-zero matrices is known [1]. Thus, it only remains to understand which matrices split into the sum of three square-zero ones. Since the set of all square-zero $n \times n$ matrices is stable under similarity, in theory one should be able to detect such matrices from their rational canonical form. Unfortunately, in practice such a classification seems out of reach. For example, it has been shown by Wang and Wu [10] that if $A \in M_n(\mathbb{F})$ is the sum of three square-zero matrices and $\mathbb{F}$ is the field of complex numbers, then $\operatorname{rk}(A - \lambda I_n) \geq \frac{n}{4}$ for every non-zero scalar $\lambda$ (the result can be generalized to an arbitrary field with characteristic not 2); Yet, it has been demonstrated in [9] that there are trace-zero matrices that satisfy this condition without being the sum of three square-zero matrices.

In this article, we shall frame the problem in a slightly different manner. If we have a trace-zero matrix $A$, maybe $A$ is not the sum of three square-zero matrices, but can we obtain such a decomposition by enlarging $A$? More precisely, can we find a positive integer $p$ such that the block-diagonal matrix $A \oplus 0_p$, that is $\begin{bmatrix} A & 0_{n \times p} \\ 0_{p \times n} & 0_p \end{bmatrix}$, is the sum of three square-zero matrices? The motivation for studying that problem stems from the infinite-dimensional setting. In a future article, its solution will help us characterize the endomorphisms of an infinite-dimensional vector space that can be decomposed as the sum of three square-zero endomorphisms.

The main strategy consists in an adaptation of the methods that were used in [5] to prove that every matrix is a linear combination of three idempotent ones. In the last section of the article, we shall prove a variation of this very result, motivated again by the case of infinite-dimensional spaces.

1.2 Main results

Here are our three main results.
Theorem 1.1. Let $F$ be an arbitrary field, and $A \in M_n(F)$ be a square matrix with trace zero. Then, $A$ is the sum of four square-zero matrices.

Theorem 1.2. Let $F$ be an arbitrary field, and $A \in M_n(F)$ be a square matrix with trace zero. Then, $A \oplus 0_n$ is the sum of three square-zero matrices of $M_{2n}(F)$.

Theorem 1.3. Let $F$ be a field with characteristic 2. Every trace-zero matrix of $M_n(F)$ is the sum of three square-zero matrices.

With a similar method, we shall obtain the following theorem:

Theorem 1.4. Let $F$ be an arbitrary field with characteristic 2, and $A \in M_n(F)$ be a square matrix whose trace belongs to $\{0, 1\}$. Then, $A \oplus 0_n$ is the sum of three idempotent matrices of $M_{2n}(F)$.

It is known that over $F_2$ every matrix is the sum of three idempotent ones (see Theorem 1 of [5]). This fails over larger fields with characteristic 2 even if we restrict to trace-zero matrices: Take indeed such a field $F$ together with a scalar $\lambda \in F \setminus \{0, 1\}$. Then, $\lambda I_n$ has trace zero yet it is not the sum of three idempotent matrices (see the remark in the middle of page 861 of [7]).

The motivation for the above results stems from the following corollaries, which will be proved in Section 7.

Corollary 1.5. Let $V$ be an infinite-dimensional vector space over $F$, and $u$ be a finite-rank endomorphism with trace zero. Then, $u$ is the sum of three square-zero endomorphisms of $V$.

Corollary 1.6. Assume that $F$ has characteristic 2. Let $V$ be an infinite-dimensional vector space over $F$, $\alpha \in F$ and $u$ be a finite-rank endomorphism of $V$ whose trace belongs to $\{0, \alpha\}$. Then, $\alpha \text{id}_V + u$ is the sum of three square-zero endomorphisms of $V$.

Corollary 1.7. Assume that $F$ has characteristic 2. Let $V$ be an infinite-dimensional vector space over $F$, $\alpha \in \{0_F, 1_F\}$, and $u$ be a finite-rank endomorphism of $V$ whose trace belongs to $\{0_F, 1_F\}$. Then, $\alpha \text{id}_V + u$ is the sum of three idempotent endomorphisms of $V$.

1.3 Two basic remarks

Throughout the article, we shall systematically use the following remarks in which, given monic polynomials $p_1, \ldots, p_d$ over $F$, a $(p_1, \ldots, p_d)$-sum is a square matrix $M$ that splits into $M = A_1 + \cdots + A_d$ where $p_i(A_i) = 0$ for all $i \in [1, d]$. In particular, the $(t^2, t^2, t^2)$-sums are the sums of three square-zero matrices.
Remark 1. If $M$ is a $(p_1, \ldots, p_d)$-sum then so is any matrix that is similar to $M$.

Remark 2. Assume that $M$ splits into a block-diagonal matrix $M = M_1 \oplus \cdots \oplus M_r$. If $M_1, \ldots, M_r$ are all $(p_1, \ldots, p_d)$-sums then so is $M$.

1.4 Structure of the article

In short, the method here is largely similar to the one that was used in [5] to prove that every matrix is a linear combination of three idempotents. In order to prove Theorem 1.2, the idea consists in showing that $A \oplus 0_n$ is similar to a block-diagonal matrix of type $N \oplus \alpha I_{2r} \oplus 0_q$ in which $\alpha \in \mathbb{F} \setminus \{0\}$, $2r \leq q$ and $N$ is well-partitioned (see Section 2.3). By subtracting a well-chosen square-zero matrix to $N$, one is able to obtain a cyclic matrix whose minimal polynomial can be chosen among the monic polynomials with the same degree and trace as the characteristic polynomial of $N$. Then, thanks to the classification of sums of two square-zero matrices (see [1] or the appendix to this article), it will follow without much effort that $N \oplus \alpha I_{2r} \oplus 0_q$ is the sum of three square-zero matrices.

Similar strategies will be used to prove Theorems 1.3 and 1.4.

The main tools for the proofs of the above theorems are laid out in Section 2: in there, we recall some useful notation and facts on cyclic matrices, we define and quickly study the notion of a well-partitioned matrix, and we give a review of the characterization of matrices that split into the sum of two square-zero matrices (in that paragraph, we include a proof of Theorem 1.1).

In Section 3 we shall prove Theorem 1.2 over fields with characteristic not 2. In Section 4 we shall prove Theorem 1.3 thereby completing the proof of Theorem 1.2 over fields with characteristic 2. Theorem 1.4 is proved in Section 5.

Section 6 is devoted to a variation of Theorem 1 of [5] on the linear combinations of three idempotent matrices. In Section 7 we derive Corollaries 1.5, 1.6 and 1.7 from Theorems 1.2, 1.3 and 1.4 respectively, and we prove a similar result for linear combinations of idempotents.

The appendix consists of a simplified proof of Botha’s characterization of sums of two square-zero matrices (Theorems 1 and 2 from [1]).

2 Some useful lemmas, and additional notation

In this section, we recall some basic results from [5].
2.1 Additional notation

Throughout the article, we choose an algebraic closure of $\mathbb{F}$ and denote it by $\overline{\mathbb{F}}$.

Similarity of two matrices $A$ and $B$ of $M_n(\mathbb{F})$ will be written $A \simeq B$.

The characteristic polynomial of a square matrix $M$ will be denoted by $\chi_M$, its trace by $\text{tr}(M)$.

Let $p(t) = t^n - \sum_{k=0}^{n-1} a_k t^k \in \mathbb{F}[t]$ be a monic polynomial with degree $n$. Its companion matrix is defined as

$$C(p(t)) := \begin{bmatrix} 0 & (0) & a_0 \\ 1 & 0 & a_1 \\ & \ddots & \ddots \\ (0) & \cdots & 0 & 1 & a_{n-1} \end{bmatrix}.$$ 

The characteristic polynomial of $C(p(t))$ is precisely $p(t)$, and so is its minimal polynomial. We define the trace of $p(t)$ as $\text{tr} \, p(t) := \text{tr} \, C(p(t)) = a_{n-1}$.

A matrix $A \in M_n(\mathbb{F})$ is called cyclic when $A \simeq C(p(t))$ for some monic polynomial $p(t)$ (and then $p(t) = \chi_A(t)$). A good cyclic matrix is a matrix of the form

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ 1 & a_{2,2} \\ & \ddots & \ddots \\ (0) & \cdots & 0 & 1 & a_{n,n} \end{bmatrix}$$

with no specific requirement on the $a_{i,j}$’s for $j \geq i$. We recall that such a matrix is always cyclic.

Finally, we denote by $H_{n,p}$ the matrix unit of $M_{n,p}(\mathbb{F})$, the set of all $n$ by $p$ matrices with entries in $\mathbb{F}$.

2.2 Two basic lemmas

The following lemma was proved in [5] (see Lemma 11):
Lemma 2.1 (Choice of polynomial lemma). Let \( A \in M_n(F) \) and \( B \in M_m(F) \) be good cyclic matrices, and \( p(t) \) be a monic polynomial of degree \( n + m \) such that \( \text{tr} \; p(t) = \text{tr}(A) + \text{tr}(B) \).
Then, there exists a matrix \( D \in M_{n,m}(F) \) such that
\[
\begin{bmatrix}
A & D \\
H_{m,n} & B
\end{bmatrix} \simeq C(p(t)).
\]

The second lemma we will need is folklore (it can be seen as an easy corollary to Roth's theorem, see [4]):

Lemma 2.2. Let \( A \in M_n(F) \), \( B \in M_p(F) \), and \( C \in M_{n,p}(F) \). Assume that \( \chi_A \) and \( \chi_B \) are coprime. Then,
\[
\begin{bmatrix}
A & C \\
0 & B
\end{bmatrix} \simeq \begin{bmatrix}
A & 0 \\
0 & B
\end{bmatrix}.
\]

2.3 Well-partitioned matrices

Definition 2.1. A square matrix \( M \) is called well-partitioned if there are positive integers \( r \) and \( s \) and monic polynomials \( p_1, \ldots, p_r, q_1, \ldots, q_s \) such that:

(i) \( M = C(p_1) \oplus \cdots \oplus C(p_r) \oplus C(q_1) \oplus \cdots \oplus C(q_s) \);

(ii) \( \deg p_i \geq 2 \) for all \( i \in [2, r] \);

(iii) \( \deg q_j \geq 2 \) for all \( j \in [1, s - 1] \);

(iv) Each polynomial \( p_i \) is coprime to each polynomial \( q_j \).

Note that the polynomials \( p_1, \ldots, p_r, q_1, \ldots, q_s \) are then uniquely determined by \( M \) (beware that in (i) we really require an equality and not just a similarity).

If in addition at most one of \( p_1 \) and \( q_s \) has degree 1, we say that \( M \) is very well-partitioned.

We convene that the void matrix (i.e. the 0 by 0 matrix) is well-partitioned.

We will need three lemmas on well-partitioned matrices. The first one was already proved in the course of the proof of Lemma 14 of [5]:

Lemma 2.3. Let \( M \in M_n(F) \) be a very well-partitioned matrix. For every monic polynomial \( R \) with degree \( n \) such that \( \text{tr}(R) \neq \text{tr}(M) \), there exists an idempotent matrix \( P \in M_n(F) \) and a scalar \( \lambda \) such that \( M - \lambda P \simeq C(R) \).
By checking the details of the proof of Lemma 14 of \[5\], one sees that the following result was also obtained:

**Lemma 2.4.** Let $M \in M_n(F)$ be a well-partitioned matrix with $m$ associated polynomials. Let $\lambda \in F \setminus \{0\}$. For every monic polynomial $R$ with degree $n$ such that $\text{tr}(R) = \text{tr}(M) - (m - 1)\lambda$, there exists an idempotent matrix $P \in M_n(F)$ such that $M - \lambda P \simeq C(R)$.

Here, we shall require the counterpart of the preceding result for square-zero matrices.

**Lemma 2.5.** Let $A \in M_n(F)$ be a well-partitioned matrix, and $R$ be a monic polynomial with degree $n$ such that $\text{tr}(R) = \text{tr}(A)$. Then, there exists a square-zero matrix $N \in M_n(F)$ such that $A - N \simeq C(R)$.

The proof is an easy adaptation of the one of Lemma 14 of \[5\]; we give it for the sake of completeness.

**Proof of Lemma 2.5.** Denote by $p_1, \ldots, p_r, q_1, \ldots, q_s$ the polynomials associated with the well-partitioned matrix $A$, and by $n_1, \ldots, n_r, m_1, \ldots, m_s$ their respective degrees. Set

$$S := \begin{pmatrix}
0_{n_1 \times n_1} & 0_{n_1 \times n_2} & \cdots & \cdots
-H_{n_2,n_1} & 0_{n_2 \times n_2} & \cdots & \cdots
(0) & \cdots & \cdots & \cdots
\end{pmatrix}$$

One checks that $S$ is square-zero (this uses the fact that $n_2 \geq 2, \ldots, n_r \geq 2, m_1 \geq 2, \ldots, m_{s-1} \geq 2$). Moreover, by setting $a = \sum_{k=1}^{s} m_k$ and $b = \sum_{k=1}^{r} n_k$, we find that there are good cyclic matrices $M'_1 \in M_b(F)$ and $M'_2 \in M_a(F)$ such that

$$A - S = \begin{bmatrix}
M'_1 & 0_{b \times a}
\end{bmatrix}.$$
On the other hand, \(\text{tr}(A - S) = \text{tr}(A) = \text{tr}(M'_1) + \text{tr}(M'_2)\). Lemma 2.4 yields a matrix \(D \in M_{b,a}(F)\) such that

\[
\begin{bmatrix}
M'_1 & D \\
H_{a,b} & M'_2
\end{bmatrix} \simeq C(R).
\]

However, \(A = A_1 \oplus A_2\) with \(A_1 \in M_b(K)\) and \(A_2 \in M_a(K)\) that have coprime characteristic polynomials. It follows that

\[A \simeq A' := \begin{bmatrix} A_1 & D \\ 0 & A_2 \end{bmatrix}.\]

However,

\[
A' - S = \begin{bmatrix} M'_1 & D \\ H_{a,b} & M'_2 \end{bmatrix} \simeq C(R).
\]

We conclude that there exists a square-zero matrix \(S'\) that is similar to \(S\) and such that \(A - S' \simeq C(R)\). □

### 2.4 A review of sums of two square-zero matrices

The classification of sums of two square-zero matrices was completed in [10] for the field of complex numbers and in [1] for general fields. We state the result:

**Theorem 2.6.** A matrix of \(M_n(F)\) can be decomposed as the sum of two square-zero matrices if and only if all its invariant factors are odd or even polynomials.

In particular, every nilpotent matrix is the sum of two square-zero matrices.

**Corollary 2.7.** Assume that \(\text{char}(F) \neq 2\). Let \(M \in M_n(F)\). Then, \(M\) is the sum of two square-zero matrices if and only if \(M\) is similar to \(-M\).

**Corollary 2.8.** Assume that \(\text{char}(F) = 2\). Let \(M \in M_n(F)\). Then, \(M\) is the sum of two square-zero matrices if and only if all the Jordan cells of \(M\) corresponding to the non-zero eigenvalues (in some algebraic closure of \(F\)) are even-sized.

As a corollary, we are now able to prove Theorem 1.1:

**Proof of Theorem 1.1.** Let \(A\) be a trace-zero matrix of \(M_n(F)\). The case when \(A = 0\) is obvious and we discard it from now on.

Assume first that \(A\) is not a scalar matrix. Then, it is similar to a matrix \(A'\) with diagonal zero (see [3]), and hence it splits into the sum of two nilpotent
matrices each of which is the sum of two square-zero matrices (one splits $A'$ into the sum of a strictly upper-triangular matrix and a strictly lower-triangular matrix).

Now, assume that $A$ is a non-zero scalar matrix. Without loss of generality, we can assume that $A = I_n$. Then, as $\text{tr} A = 0$ we get that $F$ has positive characteristic $p$ and that $p$ divides $n$. Set $r(t) := t^p - t$ if $p$ is odd, otherwise set $r(t) := t^2$. In any case, both $r(t)$ and $r(t - 1)$ are even or odd, and hence both matrices $C(r(t))$ and $C(r(t - 1))$ are sums of two square-zero matrices. Yet, $I_p + C(r(t)) \simeq C(r(t - 1))$, and hence $I_p + C(r(t))$ is the sum of two square-zero matrices. We conclude that $I_p$ is the sum of four square-zero matrices. Since $A = I_p \oplus \cdots \oplus I_p$ (with $n/p$ copies of $I_p$), we conclude that $A$ is the sum of four square-zero matrices. 

The last basic result that we will need deals with the sum of an idempotent matrix and a square-zero one (see [8] for general results on the matter).

**Lemma 2.9.** Let $n \geq 1$ and $p(t) \in F[t]$ be a monic polynomial with degree $n$. Let $A$ be a rank $n$ idempotent matrix of $M_{2n}(F)$. Set $q := p(t(t - 1))$. Then, there exists a square-zero matrix $S$ such that

$$A - S \simeq C(q(t)).$$

**Proof.** Without loss of generality, we can assume that

$$A = \begin{bmatrix} I_n & 0_n \\ 0_n & I_n \end{bmatrix}.$$

Set then

$$S := \begin{bmatrix} 0_n & -C(p) \\ 0_n & 0_n \end{bmatrix}.$$

Obviously, $S^2 = 0$. By Lemma 7 of [8], we get $A - S \simeq C(q(t))$. \qed

### 3 Sums of three square-zero matrices over a field with characteristic not 2

Our aim is to prove Theorem 1.2 over fields with characteristic not 2. Throughout the section, $F$ denotes such a field.

Let $A \in M_n(F)$ be with trace zero. Let us consider the matrix $M := A \oplus 0_n$ and prove that $M$ is the sum of three square-zero matrices. The basic idea is
to find a well-chosen square-zero matrix $S$ such that $M - S$ is the sum of two square-zero matrices. Note first that replacing $M$ with a similar matrix leaves our problem invariant. Note also that $\text{tr}(M) = \text{tr}(A) = 0$.

First of all, we shall put $M$ into a simpler form that involves a well-partitioned matrix:

**Lemma 3.1.** Let $M \in M_{2n}(\mathbb{F})$. Assume that $M$ has at least $n$ Jordan cells of size 1 for the eigenvalue 0. Then, there exist non-negative integers $p, q, r$, a matrix $N \in M_p(\mathbb{F})$ and a non-zero scalar $\alpha \in \mathbb{F} \setminus \{0\}$ such that

$$M \simeq N \oplus \alpha I_{2q} \oplus 0_r, \quad r \geq 2q,$$

and $N$ is either nilpotent or well-partitioned.

**Proof.** Note that the result is obvious if $M$ is nilpotent (it suffices to take $N = M, \alpha = 1$ and $q = r = 0$ in that case). Hence, in the remainder of the proof we assume that $M$ is non-nilpotent.

Denote by $s_1(t), \ldots, s_m(t)$ the invariant factors of $M$, so that $s_i(t)$ divides $s_j(t)$ for all $i \in [1, m-1]$. For $k \in [1, m]$, split $s_k(t) = p_k(t)q_k(t)$ where $p_k(t)$ is a monic power of $t$, and $t$ does not divide $q_k(t)$. Note that $p_{i+1}(t)$ divides $p_i(t)$ for all $i \in [1, m-1]$. Since $M$ has at least $n$ Jordan cells of size 1 for the eigenvalue 0, there are at least $n$ integers $k$ such that $p_k(t) = t$, and as $M$ has $2n$ columns it follows that $\sum_{k=1}^m \deg q_k(t) \leq 2n - n = n$; in particular:

(i) There are at most $n$ integers $k$ such that $q_k(t)$ is non-constant;

(ii) One has $m \geq n$;

(iii) One has $\deg q_m(t) \leq 1$.

Now, we denote by $a$ the least integer $k$ such that $\deg p_k(t) = 1$, and we set $r := m - a$. Hence, $p_a(t) = \cdots = p_m(t) = t$ and $\deg p_k(t) > 1$ for all $k \in [1, a-1]$. Note that $r + 1$ is the number of Jordan cells of size 1 of $M$ for the eigenvalue 0, whence $r + 1 \geq n$.

From there, we split the discussion into two subcases.

**Case 1:** One has $\deg q_k(t) \neq 1$ for all $k \in [1, m]$. Since $M$ is not nilpotent, this yields an integer $b \in [1, m]$ such that $\deg q_b(t) \geq 2$ and $q_k(t) = 1$ for all $k \in [b+1, m]$. Then,

$$M \simeq C(p_a(t)) \oplus C(p_{a-1}(t)) \oplus \cdots \oplus C(p_1(t)) \oplus C(q_1(t)) \oplus \cdots \oplus C(q_b(t)) \oplus 0_r,$$




It is clear that $N$ is well-partitioned with associated polynomials $p_a(t), \ldots, p_1(t), q_1(t), \ldots, q_b(t)$, and the result follows by taking $q := 0$ and $\alpha := 1$.

**Case 2:** Assume now that $\deg q_k(t) = 1$ for some $k \in [1, m]$. We denote respectively by $c$ and $d$ the least and greatest integer $k$ such that $\deg q_k(t) = 1$. Then, for some $\alpha \in \mathbb{F} \setminus \{0\}$, we have $q_c(t) = \cdots = q_d(t) = t - \alpha$, whereas $q_k(t) = 1$ for all $k \in [d + 1, m]$, and $\deg q_k(t) \geq 2$ for all $k \in [1, c - 1]$. Hence,

$$M \simeq \bigoplus_{t \in N} C(p_a(t)) \oplus \bigoplus_{t \in N} C(p_{a-1}(t)) \oplus \cdots \oplus \bigoplus_{t \in N} C(p_1(t)) \oplus \bigoplus_{t \in N} C(q_1(t)) \oplus \cdots \oplus \bigoplus_{t \in N} C(q_c(t)) \oplus (\alpha I_{d-c}) \oplus 0_r.$$ 

Again, $N$ is well-partitioned. If $d - c$ is even, the claimed result is then obtained by noting that

$$d - c \leq -1 + \sum_{i=1}^{m} \deg q_i \leq 2n - r - 2 \leq r.$$

Assume now that $d - c$ is odd.

- If $c > 1$ we note that

$$N' := \bigoplus_{t \in N} C(p_a(t)) \oplus \bigoplus_{t \in N} C(p_{a-1}(t)) \oplus \cdots \oplus \bigoplus_{t \in N} C(p_1(t)) \oplus \bigoplus_{t \in N} C(q_1(t)) \oplus \cdots \oplus \bigoplus_{t \in N} C(q_{c-1}(t))$$

is well-partitioned and

$$M \simeq N' \oplus (\alpha I_{d-c+1}) \oplus 0_r,$$

and we conclude as in the preceding situation because now $d - c + 1 \leq n - \deg q_1(t) \leq n - 2 \leq r$.

- If $c = 1$, then $N' := \bigoplus_{t \in N} C(p_{a-1}(t)) \oplus \cdots \oplus \bigoplus_{t \in N} C(p_1(t))$ is nilpotent, $M \simeq N' \oplus (\alpha I_{d-c+1}) \oplus 0_{r+1}$, $d - c + 1 \leq 2n - (r + 1) \leq r + 1$ and $d - c + 1$ is even.

Hence, we are reduced to proving the following result:

**Proposition 3.2.** Let $N \in \text{M}_p(\mathbb{F})$ be a well-partitioned matrix (possibly void) and $q$ be a non-negative integer, and assume that $M := N \oplus I_{2q} \oplus 0_{2q}$ has trace zero. Then, $M$ is the sum of three square-zero matrices.
Indeed, let $A \in \mathcal{M}_n(F)$ be a trace-zero matrix and set $M := A \oplus 0_n$. Then, by Lemma 3.1, we have 

$$M \simeq N \oplus \alpha I_{2q} \oplus 0_r$$

for some non-negative integers $p, q, r$ with $2q \leq r$ and $p + 2q + r = 2n$, some non-zero scalar $\alpha$ and some matrix $N \in \mathcal{M}_p(F)$ that is either nilpotent or well-partitioned. Hence, $\alpha^{-1}M \simeq \alpha^{-1} N \oplus I_{2q} \oplus 0_{2q} \oplus 0_{r-2q}$, and if we prove that $\alpha^{-1}M$ is the sum of three square-zero matrices then so is $M$ because any scalar multiple of a square-zero matrix has square zero. Moreover, $\alpha^{-1}N$ is either nilpotent or similar to a well-partitioned matrix. Finally, $0_{r-2q}$ is the sum of three square-zero matrices, and hence if Proposition 3.2 holds then we will deduce that $M$ is the sum of three square-zero matrices.

Proposition 3.2 will be established thanks to the following series of lemmas.

**Lemma 3.3.** There is a square-zero matrix $S \in \mathcal{M}_{4q}(F)$ such that 

$$I_{2q} \oplus 0_{2q} - S \simeq \bigoplus_{k=1}^{q} C((t-k)^2) \oplus C((t+k-1)^2).$$

**Proof.** Note that $I_{2q} \oplus 0_{2q}$ is the sum of $q$ copies of $I_2 \oplus 0_2$. Let $k \in [1, q]$. We claim that there exists a square-zero matrix $S_k \in \mathcal{M}_4(F)$ such that 

$$(I_2 \oplus 0_2) - S_k \simeq C((t-k)^2) \oplus C((t+k-1)^2).$$

If $2k.1_F \neq 1_F$ then this directly follows from Lemma 2.9 since 

$$C((t(t-1) - k(k-1))^2) \simeq C((t-k)^2) \oplus C((t+k-1)^2).$$

Assume now that $2k.1_F = 1_F$. Then, we know from Lemma 2.9 that there is a square-zero matrix $T_k$ such that 

$$I_1 \oplus 0_1 - T_k \simeq C((t(t-1) - k(k-1))) = C((t-k)^2).$$

Hence, 

$$(I_1 \oplus 0_1) + (I_1 \oplus 0_1) - T_k \oplus T_k \simeq C((t-k)^2) \oplus C((t+k-1)^2).$$

As $I_2 \oplus 0_2$ is similar to $(I_1 \oplus 0_1) + (I_1 \oplus 0_1)$, we obtain the claimed result.

From there, we deduce that 

$$\bigoplus_{k=1}^{q} (I_2 \oplus 0_2) - \bigoplus_{k=1}^{q} S_k \simeq \bigoplus_{k=1}^{q} C((t-k)^2) \oplus C((t+k-1)^2).$$
Since $I_{2q} \oplus 0_{2q}$ is similar to $\bigoplus_{k=1}^{q}(I_2 \oplus 0_2)$ whereas $\bigoplus_{k=1}^{q} S_k$ has square zero, the existence of the claimed matrix $S$ follows.

**Lemma 3.4.** If $q.1_F = 0_F$ then $\bigoplus_{k=1}^{q} C((t-k)^2) \oplus C((t+k-1)^2)$ is the sum of two square-zero matrices.

**Proof.** Assume that $q.1_F = 0_F$. Then,

$$\bigoplus_{k=1}^{q} C((t-k)^2) = \bigoplus_{k=0}^{q-1} C((t+k)^2) = \bigoplus_{k=1}^{q} C((t+k)^2)$$

since the polynomials $(t+q)^2$ and $t^2$ are equal over $F$. Hence,

$$\bigoplus_{k=1}^{q} C((t-k)^2) \oplus C((t+k-1)^2) \simeq \bigoplus_{k=1}^{q} C((t-k)^2) \oplus C((t+k)^2),$$

and the matrix on the right-hand side is obviously similar to its opposite. The conclusion then follows from Corollary 2.7.

From there, we can finally prove Proposition 3.2.

**Proof of Proposition 3.2.** The proof is split into two cases.

**Case 1:** $\text{tr} N = 0$.

Then, $0 = \text{tr} M = \text{tr} N + 2q.1_F = 2q.1_F$. Since $\text{char}(F) \neq 2$ we find $q.1_F = 0_F$. The above two lemmas then show that $I_{2q} \oplus 0_{2q}$ is the sum of three square-zero matrices.

- If $N$ is nilpotent then it is the sum of two square-zero matrices.

- If $N$ is well-partitioned, then Lemma 2.5 yields a square-zero matrix $S \in M_p(F)$ such that $N - S \simeq C(t^p)$ (we convene that $C(t^p) = 0$ if $p = 0$), and hence $N - S$ is the sum of two square-zero matrices.

In any case, $N$ is the sum of three square-zero matrices. Hence, so is $M = N \oplus (I_{2q} \oplus 0_{2q})$.

**Case 2:** $\text{tr}(N) \neq 0$.

In particular $p \geq 2$, $\text{tr} N = -2q.1_F$ and $q.1_F \neq 0$. Hence, by Lemma 2.5 we can find a square-zero matrix $S \in M_p(F)$ such that $N - S \simeq C(t^{p-2}(t+q)^2) \simeq$
$C(t^{p-2}) \oplus C((t+q)^2)$. On the other hand, we have a square-zero matrix $S'$ such that

$$
(I_{2q} \oplus 0_{2q}) - S' \simeq \bigoplus_{k=1}^{q} C((t-k)^2) \oplus C((t+k-1)^2).
$$

Hence, $S_0 := S \oplus S'$ has square zero and

$$
M - S_0 \simeq C(t^{p-2}) \oplus C((t+q)^2) \oplus \bigoplus_{k=1}^{q} C((t-k)^2) \oplus C((t+k-1)^2).
$$

Yet,

$$
C((t+q)^2) \oplus \bigoplus_{k=1}^{q} C((t-k)^2) \oplus C((t+k-1)^2) \simeq C(t^2) \oplus \bigoplus_{k=1}^{q} C((t-k)^2) \oplus C((t+k)^2),
$$

whence

$$
M - S_0 \simeq C(t^{p-2}) \oplus C(t^2) \oplus \bigoplus_{k=1}^{q} C((t-k)^2) \oplus C((t+k)^2),
$$

and once more Corollary 2.7 yields that $M - S_0$ is the sum of two square-zero matrices. Hence, $M$ is the sum of three square-zero matrices. \( \square \)

Thus, the proof of Theorem 1.2 is now complete.

4 Sums of three square-zero matrices over a field of characteristic 2

This section consists of a proof of Theorem 1.3.

Throughout this section, we assume that the field $F$ has characteristic 2. The basic strategy is similar to the one of the preceding section, with a few different details. First of all, we shall need the following basic lemma.

**Lemma 4.1.** Let $\alpha \in F$. Then, $\alpha I_2$ is the sum of three square-zero matrices.

**Proof.** Indeed,

$$
\alpha I_2 - \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} \alpha & 1 \\ 0 & \alpha \end{bmatrix} \simeq C((t-\alpha)^2) = C(t^2 - \alpha^2),
$$

whereas by Theorem 2.6 the matrix $C(t^2 - \alpha^2)$ is the sum of two square-zero ones. \( \square \)
The next step consists of the following lemma, which actually holds over any field.

**Lemma 4.2.** Let $A \in M_n(\mathbb{F})$. Assume that $A$ has at most one Jordan cell of size 1 for each one of its eigenvalues in $\mathbb{F}$ and that its minimal polynomial is not a power of some irreducible polynomial. Then, $A$ is similar to a well-partitioned matrix.

**Proof.** Let us choose an irreducible monic factor $p(t)$ of the minimal polynomial of $A$. Let us denote by $a_1(t), \ldots, a_r(t)$ the invariant factors of $A$, so that $a_{i+1}(t)$ divides $a_i(t)$ for all $i \in [1, r-1]$. For all $i \in [1, r-1]$, split $a_i(t) = b_i(t)c_i(t)$ where $b_i(t)$ is a power of $p(t)$ and $c_i(t)$ is coprime with $p(t)$. Set finally $u := \max\{i \in [1, r] : b_i(t) \neq 1\}$ and $v := \max\{i \in [1, r] : c_i(t) \neq 1\}$. If $p(t)$ has degree 1 then we know that $A$ has at most one Jordan cell of size 1 for the root of $p(t)$. Hence, $\deg b_i(t) \geq 2$ for all $i \in [1, u-1]$. Likewise $\deg c_i(t) \geq 2$ for all $i \in [1, v-1]$. Hence, $A' := C(b_u(t)) \oplus C(b_{u-1}(t)) \oplus \cdots \oplus C(b_1(t)) \oplus C(c_1(t)) \oplus \cdots \oplus C(c_{v-1}(t)) \oplus C(c_v(t))$ is well-partitioned, and $A \simeq A'$.

**Lemma 4.3.** Let $M \in M_n(\mathbb{F})$ be a trace-zero matrix whose minimal polynomial is a power of some irreducible monic polynomial. Then, $M$ is the sum of three square-zero matrices.

**Proof.** We write the minimal polynomial of $M$ as $p(t)^r$, where $p(t)$ is an irreducible monic polynomial, and $r$ is a positive integer. If $p(t) = t$, then $M$ is nilpotent and we already know that it is the sum of two square-zero matrices. In the rest of the proof, we assume that $p(t) \neq t$. We distinguish between two cases.

**Case 1:** $\text{tr}(p(t)) = 0$.

Denote by $d$ the degree of $p(t)$. Let $k$ be a positive integer. Then, $\text{tr}(p(t)^k) = k \text{tr} p(t) = 0$. The last column of $C(p(t)^k)$ then reads $\begin{bmatrix} Y \\ 0 \end{bmatrix}$ for some $Y \in \mathbb{F}^{kd-1}$.

Setting $S := \begin{bmatrix} 0_{(kd-1)\times(kd-1)} & Y \\ 0_{1\times(kd-1)} & 0 \end{bmatrix}$, we see that $C(p(t)^k) - S$ is the transpose of a Jordan cell for the eigenvalue 0. Hence, $C(p(t)^k) - S$ is the sum of two square-zero matrices, and it follows that $C(p(t)^k)$ is the sum of three such matrices. Using the rational canonical form, we deduce that $M$ is the sum of three square-zero matrices.
Case 2: $\text{tr}(p(t)) \neq 0$.

As $\text{tr} M = 0$ and $\mathbb{F}$ has characteristic 2, evenly many invariant factors of $M$ are odd powers of $p(t)$. Hence, we have a splitting

$$M \simeq A_1 \oplus \cdots \oplus A_r \oplus B_1 \oplus \cdots \oplus B_s,$$

where each $A_i$ equals $C(p^k)$ for some even positive integer $k$, and each $B_j$ equals $C(p^k) \oplus C(p^l)$ for some pair $(k, l)$ of odd positive integers. However, for every even integer $k > 0$, we see that $p^k$ is an even polynomial (since $\mathbb{F}$ has characteristic 2) and hence $C(p^k)$ is the sum of two square-zero matrices. To conclude, we take an arbitrary pair $(k, l)$ of odd positive integers and we prove that $C(p^k) \oplus C(p^l)$ is the sum of three square-zero matrices. The matrix $S := \begin{bmatrix} 0_{kd} & 0_{kd \times ld} \\ F_{ld, kd} & 0_{ld} \end{bmatrix}$ of $M_{(k+l)d}(\mathbb{F})$ has square-zero, and one sees that $(C(p^k) \oplus C(p^l)) - S$ is a good cyclic matrix with characteristic polynomial $p^{k+l}$. Since $k + l$ is even and $\mathbb{F}$ has characteristic 2, the polynomial $p^{k+l}$ is an even one, whence $(C(p^k) \oplus C(p^l)) - S$ is the sum of two square-zero matrices. This completes the proof. 

From there, the proof of Theorem 1.3 can be completed swiftly.

Proof of Theorem 1.3. Let $M$ be a trace-zero matrix of $M_n(\mathbb{F})$. By pairing Jordan cells of size 1 of $M$ associated to the same eigenvalue in $\mathbb{F}$, we find a non-negative integer $p$, scalars $\alpha_1, \ldots, \alpha_q$ (with $p + 2q = n$) and a matrix $N \in M_p(\mathbb{F})$ such that

$$M \simeq N \oplus \alpha_1 I_2 \oplus \cdots \oplus \alpha_q I_2,$$

and $N$ has at most one Jordan cell of size 1 for each one of its eigenvalues in $\mathbb{F}$. Since $\mathbb{F}$ has characteristic 2 we find $\text{tr} N = \text{tr} M = 0$. By Lemma 4.1 the conclusion will follow should we prove that $N$ is the sum of three square-zero matrices. If the minimal polynomial of $N$ has a sole irreducible monic divisor, then this follows directly from Lemma 4.3. Assume otherwise. We get from Lemma 4.2 that $N$ is similar to a trace-zero well-partitioned matrix $N'$. By Lemma 2.5 there is a square-zero matrix $S \in M_p(\mathbb{F})$ such that $N' - S \simeq C(t^p)$, so that $N' - S$ is the sum of two square-zero matrices. Hence, $N'$ is the sum of three square-zero matrices, and we conclude that so is $N$. Hence, $M$ is the sum of three square-zero matrices. 

From there, the proof of Theorem 1.3 can be completed swiftly.
5 On sums of three idempotent matrices over a field of characteristic 2

Throughout this section, we assume that $\mathbb{F}$ has characteristic 2. Our aim is to prove Theorem 1.4. To start with, we recall the following result from [6] (Theorem 5 in that article):

**Theorem 5.1.** A matrix $A$ of $M_n(\mathbb{F})$ is the sum of two idempotent matrices if and only if, for every eigenvalue $\lambda$ of $A$ in $\mathbb{F} \setminus \{0, 1\}$, all the Jordan cells attached to $\lambda$ are even-sized.

**Lemma 5.2.** Let $\alpha \in \mathbb{F} \setminus \{0, 1\}$. Then, $\alpha I_2 \oplus 0_1$ is the sum of three idempotents.

**Proof.** Note that $M := \alpha I_2 \oplus 0_1$ is similar to $N := \begin{bmatrix} \alpha & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \alpha \end{bmatrix}$. Set $P := \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ and note that $P$ is idempotent and $N - P \simeq \begin{bmatrix} \alpha & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & \alpha \end{bmatrix} \simeq C((t - \alpha)^2(t - 1)) \oplus C((t - \alpha)^2)$. Then, by Theorem 5.1, $N - P$ is the sum of two idempotents, whence $N$ is the sum of three idempotents. We conclude that $M$ is the sum of three idempotents. \(\Box\)

**Lemma 5.3.** Let $A \in M_n(\mathbb{F})$ be a well-partitioned matrix with $\text{tr} A \in \{0, 1\}$. Then, $A$ is the sum of three idempotent matrices.

**Proof.** Denote by $m$ the number of polynomials associated with $A$. Set $\lambda := \text{tr} A - (m - 1).1_\mathbb{F} \in \{0, 1\}$. By Lemma 2.4, we can find an idempotent $P \in M_n(\mathbb{F})$ such that $A - P \simeq C(t^{n-1}(t - \lambda))$. Since $\lambda \in \{0, 1\}$, Theorem 5.1 yields that $A - P$ is the sum of two idempotent matrices. \(\Box\)

**Lemma 5.4.** Let $A \in M_n(\mathbb{F})$, and assume that the minimal polynomial of $A$ is a power of some monic irreducible polynomial $p(t)$. Assume also that $\text{tr}(A) \in \{0, 1\}$. Assume finally that if $p(t) = t - \lambda$, then $A$ has at most one Jordan cell of size 1 for the eigenvalue $\lambda$. Then, $A$ is the sum of three idempotents.

**Proof.** For every even integer $r$, we know from Theorem 5.1 that $C(p(t)^r)$ is the sum of two idempotent matrices, and its trace is obviously zero. Hence, we lose
no generality in assuming that the invariant factors of $A$ are odd powers of $p(t)$. We split the discussion into two cases.

**Case 1:** $\text{tr } p(t) \in \{0, 1\}$.

Let $k \geq 1$ be a positive integer. Denote by $d$ the degree of $p(t)$. Then, $\text{tr } p(t)^k = k \text{tr } p(t) \in \{0, 1\}$. Let $Y \in \mathbb{F}^{kd-1}$, and set $P := \begin{bmatrix} 0_{(kd-1) \times (kd-1)} & Y \\ 0_{1 \times (kd-1)} & 1 \end{bmatrix}$, which is an idempotent matrix of $M_{kd}(\mathbb{F})$. Then, $C(p^k) - P$ has trace $k \text{tr } p(t) - 1 \in \{0, 1\}$, and it is actually a companion matrix. Obviously, for any monic polynomial $q$ with degree $kd$ and trace $k \text{tr } p(t) - 1$ the vector $Y$ can be chosen so as to have $C(p^k) - P = C(q)$. Choosing $q = (t-1)t^{kd-1}$ if $k \text{tr } p(t) - 1 = 1$, and $q = t^{kd}$ if $k \text{tr } p(t) - 1 = 0$, we know from Theorem 5.1 that $C(q)$ is the sum of two idempotents, and it follows that $C(p^k)$ is the sum of three idempotents. Using the rational canonical form, we conclude that $M$ is the sum of three idempotents.

**Case 2:** $\text{tr } p(t) \not\in \{0, 1\}$.

As $\text{tr } A \in \{0, 1\}$, there are evenly many invariant factors of $A$ (remember that all of them are odd powers of $p(t)$). To conclude, it suffices to take an arbitrary pair $(k, l)$ of odd positive integers, with $k \leq l$, and to prove that $B := C(p(t)^k) + C(p(t)^l)$ is the sum of three idempotents, unless $k = l = 1$ and $p = t - \lambda$ for some $\lambda \in \mathbb{F} \setminus \{0, 1\}$. Assume indeed that we do not simultaneously have $k = l = 1$ and $p = t - \lambda$ for some $\lambda \in \mathbb{F} \setminus \{0, 1\}$.

Denote by $d$ the degree of $p(t)$. For a positive integer $i$, denote by $D_i$ the diagonal matrix of $M_i(\mathbb{F})$ with all diagonal entries zero except the last one, which equals 1. Assume first that $d \geq 2$. Set $S_1 := D_d \oplus \cdots \oplus D_d$ (with $k + l$ copies of $D_d$) and

$$S_2 := \begin{bmatrix} 0_{kd} & 0_{kd \times ld} \\ F_{ld, kd} & 0_{ld} \end{bmatrix}.$$

Since $d \geq 2$, one sees that $S_1 + S_2$ is idempotent. On the other hand, one sees that $B - S_1 - S_2$ is a good cyclic matrix with characteristic polynomial $(p(t) + t^{d-1})^{k+l}$. Since $k + l$ is even, we deduce from Theorem 5.1 that $B - (S_1 + S_2)$ is the sum of two idempotent matrices.

Assume now that $d = 1$, so that $p(t) = t - \lambda$ for some $\lambda \in \mathbb{F}$. Then, $S := \begin{bmatrix} D_k & 0_{k \times l} \\ F_{l, k} & D_l \end{bmatrix}$ is idempotent because $l > 1$. Moreover, one sees that $B - S$ is a good cyclic matrix with characteristic polynomial $(t - \lambda)^{k+l-2}(t - \lambda + 1)^2$, and again by Theorem 5.1 it is the sum of two idempotent matrices.
In any case, we conclude that \( C(p(t)^k) \oplus C(p(t)^l) \) is the sum of three idempotent matrices, which completes the proof.

**Corollary 5.5.** Let \( A \in M_n(\mathbb{F}) \), and assume that \( A \) has at most one Jordan cell of size 1 for each one of its eigenvalues in \( \mathbb{F} \). Assume also that \( \text{tr} A \in \{0, 1\} \). Then, \( A \) is the sum of three idempotent matrices.

**Proof.** If the minimal polynomial of \( A \) is not a power of some irreducible monic polynomial, then Lemma 4.2 shows that \( A \) is similar to a well-partitioned matrix, and by Lemma 5.3 \( A \) is the sum of three idempotent matrices.

Otherwise, Lemma 5.4 shows that \( A \) is the sum of three idempotent matrices.

Now, we are ready to complete the proof of Theorem 1.4.

Let \( A \in M_n(\mathbb{F}) \) be with \( \text{tr} A \in \{0, 1\} \). By pairing Jordan cells of size 1 associated to the same eigenvalue in \( \mathbb{F} \), we find a decomposition

\[
A \simeq N \oplus \alpha_1 I_2 \oplus \cdots \oplus \alpha_q I_2
\]

in which \( \alpha_1, \ldots, \alpha_q \) are scalars, and \( N \in M_{n-2q}(\mathbb{F}) \) has at most one Jordan cell of size 1 for each one of its eigenvalues in \( \mathbb{F} \). Note that \( \text{tr} N = \text{tr} A \in \{0, 1\} \). By Corollary 5.5, \( N \) is the sum of three idempotent matrices. On the other hand, for all \( k \in [1, q] \), we know from Lemma 5.2 that \( \alpha_k I_2 \oplus 0_1 \) is the sum of three idempotent matrices. Hence, \( A \oplus 0_q \) is the sum of three idempotent matrices, QED.

### 6 A note on linear combinations of idempotent matrices

We recall the following terminology from [5]:

**Definition 6.1.** Let \( \alpha_1, \ldots, \alpha_k \) be scalars. A matrix \( M \in M_n(\mathbb{F}) \) is called an \((\alpha_1, \ldots, \alpha_k)\)-composite when there are idempotents \( P_1, \ldots, P_k \) of \( M_n(\mathbb{F}) \) such that \( M = \sum_{i=1}^{k} \alpha_i P_i \).

Our aim here is to prove the following result:

**Theorem 6.1.** Let \( \mathbb{F} \) be an arbitrary field and let \( \alpha \in \mathbb{F} \setminus \{0\} \). Then, for all \( A \in M_n(\mathbb{F}) \), there exist scalars \( \beta, \gamma \) such that \( A \oplus \alpha I_n \) is an \((\alpha, \beta, \gamma)\)-composite.
The motivation for proving Theorem 6.1 is the following corollary, which will be derived in the next section:

**Corollary 6.2.** Let $V$ be an infinite-dimensional vector space over $\mathbb{F}$. Let $u$ be a finite-rank endomorphism of $V$, and let $\alpha \in \mathbb{F}$. Then, $f := \alpha \text{id}_V + u$ is a linear combination of three idempotent endomorphisms of $V$.

Now, we prove Theorem 6.1. As we will see, it is a mostly straightforward adaptation of the line of reasoning from [5].

Our first lemma is obtained by following the proof of Lemma 15 of [5]:

**Lemma 6.3.** Let $(\alpha, \beta, \gamma) \in \mathbb{F}^3$ be such that $\alpha \neq \beta$ and $\alpha \neq 0$, and let $(q, r, s) \in \mathbb{N}^3$ be with $s > 0$. Then, there exist a monic polynomial $p(t) \in \mathbb{F}[t]$ of degree $s$ such that $\text{tr} p(t) \neq \gamma$, together with a scalar $\alpha'$ such that $C(p(t)) \oplus \alpha I_q \oplus \beta I_r$ is an $(\alpha, \alpha')$-composite.

Similarly, the next lemma is obtained by following the details of the proof of Lemma 13 of [5] (indeed, in the notation of that proof we have $\alpha = \alpha_1$ for $A := M \oplus \alpha I_n$):

**Lemma 6.4.** Let $M \in M_n(\mathbb{F})$ and $\alpha \in \mathbb{F} \setminus \{0\}$. If $M$ is not triangularizable with sole eigenvalue $\alpha$, then there exist a triple $(p, q, r)$ of non-negative integers such that $p + q + r = 2n$, a scalar $\beta \neq \alpha$ and a matrix $N \in M_p(\mathbb{K})$ such that $M \oplus \alpha I_n \simeq N \oplus \alpha I_q \oplus \beta I_r$ and either $N = 0$ or $N$ is very well-partitioned.

Now, we can finish the proof of Theorem 6.1:

**Proof of Theorem 6.1.** Let $A \in M_n(\mathbb{F})$. Assume that $A$ is triangularizable with sole eigenvalue $\alpha$. Hence, $M := A \oplus \alpha I_n$ is also triangularizable with sole eigenvalue $\alpha$. Then, $M - \alpha I_{2n}$ is nilpotent, and hence by Proposition 15 of [6] it is a $(1, -1)$-composite. Thus, $M$ is an $(\alpha, 1, -1)$-composite.

Assume now that $A$ is not triangularizable with sole eigenvalue $\alpha$. Then, by Lemma 6.4 there exist a scalar $\beta \neq \alpha$, a triple $(p, q, r)$ of non-negative integers, and a matrix $N \in M_p(\mathbb{K})$ such that $A \oplus \alpha I_n \simeq A' := N \oplus \alpha I_q \oplus \beta I_r$ and either $N = 0$ or $N$ is very well-partitioned. If $N = 0$ then $A \oplus \alpha I_n$ is an $(\alpha, \beta)$-composite, and hence an $(\alpha, \beta, 1)$-composite. Assume now that $N$ is very well-partitioned. By Lemma 6.3 there exist a monic polynomial $u(t) \in \mathbb{F}[t]$ of

\footnote{Note that the statement of Lemma 13 of [5] is partly incorrect because it does not take into account the possibility that the matrix $A$ be diagonalizable with exactly two eigenvalues. Nevertheless, this has no impact on the validity of the rest of the results from [5].}
degree $p$ and a scalar $\alpha'$ such that $\text{tr } u(t) \neq \text{tr } N$ and $C(u(t)) \oplus \alpha I_q \oplus \beta I_r$ is an $(\alpha, \alpha')$-composite. By Lemma 2.3, there exist a scalar $\delta$ and an idempotent $P \in M_p(\mathbb{F})$ such that $N - \delta P \simeq C(u(t))$. Set $\tilde{P} := P \oplus 0_{q+r}$, which is idempotent. Then,

$$A' - \delta \tilde{P} \simeq C(u(t)) \oplus \alpha I_q \oplus \beta I_r$$

is an $(\alpha, \beta)$-composite, and hence $A'$ is an $(\alpha, \beta, \delta)$-composite. Therefore, $A \oplus \alpha I_n$ is an $(\alpha, \beta, \delta)$-composite. This completes the proof.

7 Application to the decomposition of endomorphisms of an infinite-dimensional space

In this section, we prove Corollaries 1.5, 1.6, 1.7 and 6.2.

7.1 Proof of Corollary 1.5

We can choose a finite-dimensional linear subspace $W$ of $V$ such that $\text{Im } u \subset W$ and $\text{Ker } u + W = V$. Denote by $u'$ the endomorphism of $W$ induced by $u$, and set $n := \dim W$. Since $V$ is infinite-dimensional we can split $V = W \oplus W' \oplus W_0$, where $\dim W' = n$ and $W' \oplus W_0 \subset \text{Ker } u$. Denote by $u''$ the endomorphism of $W \oplus W'$ induced by $u$. Choose some square matrix $A$ that represents $u'$. Then, $\text{tr}(A) = \text{tr}(u) = 0$ since $\text{Im } u \subset W$. By Theorem 1.2, $A \oplus 0_n$ is the sum of three square-zero matrices, yielding square-zero endomorphisms $a, b, c$ of $W \oplus W'$ such that $u'' = a + b + c$. Extending $a, b, c$ into endomorphisms $\tilde{a}, \tilde{b}, \tilde{c}$ of $V$ which vanish everywhere on $W_0$, we obtain that $\tilde{a}, \tilde{b}, \tilde{c}$ have square zero and $u = \tilde{a} + \tilde{b} + \tilde{c}$.

7.2 Proof of Corollary 1.6

Set $f := \alpha \text{id}_V + u$. We can choose a finite-dimensional linear subspace $W$ of $V$ such that $\text{Im } u \subset W$ and $\text{Ker } u + W = V$. Set $n := \dim W$. Set $p := 0$ if $\text{tr } u = n\alpha$, and $p := 1$ if $\text{tr } u = (n+1)\alpha$. Denote by $u'$ the endomorphism of $W$ induced by $u$. Since $V$ is infinite-dimensional we can split $V = W \oplus W' \oplus W_0$, where $\dim W' = n + p$ and $W' \oplus W_0 \subset \text{Ker } u$. Choose some square matrix $A$ that represents the endomorphism $u''$ of $W \oplus W'$ induced by $u$. Then, $\text{tr } A = \text{tr}(u'') = \text{tr}(u)$ since $\text{Im } u \subset W$. Hence, $\text{tr}(A + \alpha I_{n+p}) = \text{tr}(u) + (n+p)\alpha = 0$, and $A + \alpha I_{n+p}$ represents $\alpha \text{id}_{W \oplus W'} + u''$. By Theorem 1.3 we obtain square-zero endomorphisms $a, b, c$ of $W \oplus W'$ such that $f|_{W \oplus W'} = a + b + c$. Next, as $W_0$ is
infinite-dimensional, we can split it into $W_0 = \bigoplus_{i \in I} P_i$ in which $P_i$ has dimension 2 for all $i \in I$. Let $i \in I$. Then, $f_{P_i} = \alpha \operatorname{id}_{P_i}$ for all $i \in I$, which has trace 0. By Theorem 1.3, there are square-zero endomorphisms $a_i, b_i, c_i$ of $P_i$ such that $f_{P_i} = a_i + b_i + c_i$. Define $\tilde{a}$ as the endomorphism of $V$ that coincides with $a$ on $W \oplus W'$, and with $a_i$ on $P_i$ for all $i \in I$. Likewise, define $\tilde{b}$ and $\tilde{c}$ from the data of $b_i, (b_i)_{i \in I}$ and $c_i, (c_i)_{i \in I}$, respectively. Then, one checks that $\tilde{a}, \tilde{b}, \tilde{c}$ have square zero and $f = \tilde{a} + \tilde{b} + \tilde{c}$.

7.3 Proof of Corollary 1.7

Set $f := \alpha \operatorname{id}_V + u$. We can choose a finite-dimensional linear subspace $W$ of $V$ such that $\operatorname{Im} u \subset W$ and $\ker u + W = V$. Denote by $f'$ the endomorphism of $W$ induced by $f$, and set $n := \dim W$. Since $V$ is infinite-dimensional we can split $V = W \oplus W' \oplus W_0$, where $\dim W' = n$ and $W' \oplus W_0 \subset \ker u$. Choose some square matrix $A$ that represents $f'$. Then, $\operatorname{tr}(A) = \operatorname{tr}(u) + n\alpha \in \{0, 1\}$ since $\operatorname{Im} u \subset W$. The matrix $A \oplus \alpha I_n$ represents $f|_{W \oplus W'}$, and its trace belongs to $\{0, 1\}$. Hence, by Theorem 1.2, the endomorphism $f|_{W \oplus W'}$ is the sum of three idempotent endomorphisms $p, q, r$ of $W \oplus W'$.

Let us extend $p, q$ into endomorphisms $\tilde{p}, \tilde{q}$ of $V$ which vanish everywhere on $W_0$. Let us extend $r$ into an endomorphism $\tilde{r}$ of $V$ whose restriction to $W_0$ is $\alpha \operatorname{id}_{W_0}$. Then, $\tilde{p}, \tilde{q}, \tilde{r}$ are idempotent endomorphisms of $V$ and $f = \tilde{p} + \tilde{q} + \tilde{r}$.

7.4 Proof of Corollary 6.2

If $\alpha = 0$ then the result is a straightforward consequence of Theorem 1 of [5]. Assume now that $\alpha \neq 0$.

We can find a finite-dimensional linear subspace $W$ of $V$ such that $\operatorname{Im} u \subset W$ and $W + \ker u = V$. Set $n := \dim W$. Hence, $W$ is stable under $f$, and we can denote by $f'$ the endomorphism of $W$ induced by $f$. We split $V = W \oplus W' \oplus W_0$, where $W' \oplus W_0 \subset \ker u$ and $\dim W = \dim W'$. Denote respectively by $f_1$ and $f_0$ the endomorphisms of $W + W'$ and $W_0$ induced by $f$, and note that $f_0 = \alpha \operatorname{id}_{W_0}$. Choose a square matrix $A \in M_n(\mathbb{F})$ that represents $f'$. Then, $A \oplus \alpha I_n$ represents $f_1$. By Theorem 6.1, there are scalars $\beta, \gamma$ and idempotent matrices $P, Q, R$ of $M_{2n}(\mathbb{F})$ such that $A \oplus \alpha I_n = \alpha P + \beta Q + \gamma R$. This yields idempotent endomorphisms $p, q, r$ of $W + W'$ such that $f_1 = \alpha p + \beta q + \gamma r$. Hence, $f = \alpha (\tilde{p} \oplus \operatorname{id}_{W_0}) + \beta (q \oplus 0_{W_0}) + \gamma (r \oplus 0_{W_0})$. Therefore, $f$ is a linear combination of three idempotent endomorphisms of $V$. 

22
A Appendix. On sums of two square-zero matrices

This appendix consists in a short proof of Botha’s theorem (that is, Theorem 2.6). First of all, the sufficient conditions:

Lemma A.1. Let \( n \) be a positive integer. Then, \( C(t^n) \) is the sum of two square-zero matrices.

**Proof.** Define \( A = (a_{i,j}) \) and \( B = (b_{i,j}) \) in \( M_n(\mathbb{F}) \) by \( a_{i,j} = 1 \) if \( i = j + 1 \) and \( i \) is even, and \( a_{i,j} = 0 \) otherwise, and \( b_{i,j} = 1 \) if \( i = j + 1 \) and \( i \) is odd, and \( b_{i,j} = 0 \) otherwise. One checks that \( A^2 = B^2 = 0 \), while \( A + B = C(t^n) \). \[\Box\]

**Lemma A.2.** Let \( p \) be a non-constant monic polynomial. Then,

\[
C(p(t^2)) \simeq \begin{bmatrix} 0_n & C(p(t)) \\ I_n & 0_n \end{bmatrix}.
\]

**Proof.** Indeed, denote by \( (e_1, \ldots, e_{2n}) \) the standard basis of \( \mathbb{F}^{2n} \) and by \( u \) the endomorphism of \( \mathbb{F}^{2n} \) associated with \( \begin{bmatrix} 0_n & C(p(t)) \\ I_n & 0_n \end{bmatrix} \) in it. One checks that the matrix of \( u \) in the basis \( (e_1, e_{n+1}, e_2, e_{n+2}, \ldots, e_n, e_{2n}) \) is \( C(p(t^2)) \). \(\Box\)

**Corollary A.3.** For all \( M \in M_n(\mathbb{F}) \), the invariant factors of

\[
D(M) := \begin{bmatrix} 0_n & M \\ I_n & 0_n \end{bmatrix}
\]

are even polynomials.

**Proof.** Note first that if \( M = A_1 \oplus \cdots \oplus A_p \) for some square matrices \( A_1, \ldots, A_p \), then \( D(M) \simeq D(A_1) \oplus \cdots \oplus D(A_p) \). Moreover, for every invertible matrix \( P \in GL_n(\mathbb{F}) \), we see that

\[
D(PMP^{-1}) = QD(M)Q^{-1}
\]

where \( Q = P \oplus P \in GL_{2n}(\mathbb{F}) \). Using this last remark, we see that no generality is lost in assuming that \( M = C(p_1) \oplus \cdots \oplus C(p_r) \), where \( p_1, \ldots, p_r \) are monic polynomials, and \( p_{r+1} \) divides \( p_i \) for all \( i \) from 1 to \( r - 1 \). Then,

\[
D(M) \simeq D(C(p_1)) \oplus \cdots \oplus D(C(p_r)).
\]
and by Lemma A.2, this entails
\[ D(M) \simeq D(C(p_1(t^2))) \oplus \cdots \oplus D(C(p_r(t^2))). \]

Obviously, \( p_{i+1}(t^2) \) divides \( p_i(t^2) \) for all \( i \) from 1 to \( r - 1 \), whence the monic polynomials \( p_1(t^2), \ldots, p_r(t^2) \) are the invariant factors of \( D(M) \).

\[ \square \]

**Corollary A.4.** For every even monic polynomial \( p(t) \), the matrix \( C(p(t)) \) is the sum of two square-zero matrices.

**Proof.** Indeed,
\[
\begin{bmatrix}
0_n & C(p(t)) \\
I_n & 0_n
\end{bmatrix} = \begin{bmatrix}
0_n & 0_n \\
I_n & 0_n
\end{bmatrix} + \begin{bmatrix}
0_n & C(p(t)) \\
0_n & 0_n
\end{bmatrix}
\]
is the sum of two square-zero matrices, and hence the result follows directly from Lemma A.2.

\[ \square \]

**Corollary A.5.** For every even or odd monic polynomial \( p(t) \), the matrix \( C(p(t)) \) is the sum of two square-zero matrices.

**Proof.** Let \( p(t) \) be an odd monic polynomial. Then, \( p(t) = t^kq(t) \) for some positive integer \( k \) and some even polynomial \( q(t) \) such that \( q(0) \neq 0 \). Then, \( t^k \) and \( q(t) \) are coprime, whence \( C(p(t)) \simeq C(t^k) \oplus C(q(t)) \). By Lemma A.1 and Corollary A.4, we conclude that \( C(p(t)) \) is the sum of two square-zero matrices.

\[ \square \]

**Remark 3.** This last corollary can also be obtained with a similar proof as for Lemma A.1 (see the proof of Lemma 2 of [1]). Yet, since we shall use Lemma A.2 once more later on, it was more efficient to prove Corollary A.5 as we have done.

By Remarks 1 and 2, we conclude that a matrix is the sum of two square-zero matrices if each one of its invariant factors is odd or even. Now, we turn to the necessary conditions.

**Lemma A.6.** Let \( a \) and \( b \) be square-zero endomorphisms of a vector space over \( F \). Then, \( a \) and \( b \) commute with \( (a + b)^2 \).

**Proof.** Indeed, \( (a + b)^2 = ab + ba \), whence \( a(a + b)^2 = aba = (a + b)^2a \), and likewise \( b \) commutes with \( (a + b)^2 \).

\[ \square \]
Corollary A.7. Let \( u \) be an endomorphism of a finite-dimensional vector space over \( \mathbb{F} \). Denote by \( u_i \) the invertible part of \( u \) in its Fitting decomposition. If \( u \) is the sum of two square-zero operators, then so is \( u_i \).

Proof. Denote by \( n \) the dimension of the domain of \( u \). Since \( 2n \geq n \), we know that \( \text{Im} \ u^{2n} \) is the domain of \( u_i \). Assume now that \( u = a + b \) for some square-zero endomorphisms \( a \) and \( b \) of the domain of \( u \). By Lemma A.6, both \( a \) and \( b \) commute with \( u^2 \), whence both of them stabilize \( \text{Im}(u^2)^n \). Denoting by \( a' \) and \( b' \) the endomorphisms of \( \text{Im} \ u^{2n} \) that are induced by \( a \) and \( b \), respectively, we see that \((a')^2 = (b')^2 = 0 \) and \( u_i = a' + b' \).

Lemma A.8. Let \( u \) be an automorphism of a finite-dimensional vector space \( V \) over \( \mathbb{F} \). Assume that \( u \) is the sum of two square-zero endomorphisms of \( V \). Then, the invariant factors of \( u \) are even polynomials.

Proof. Let \( a \) and \( b \) be square-zero endomorphisms of \( V \) such that \( u = a + b \). Since \( \text{Im} \ a \subset \text{Ker} \ a \) and \( \text{Im} \ b \subset \text{Ker} \ b \), we have \( \dim \text{Ker} \ a \geq \frac{\dim V}{2} \) and \( \dim \text{Ker} \ b \geq \frac{\dim V}{2} \). On the other hand, since \( a + b \) is injective we have \( \text{Ker} \ a \cap \text{Ker} \ b = \{0\} \). It follows that \( n \geq \dim \text{Ker} \ a + \dim \text{Ker} \ b \). We deduce that \( \dim \text{Ker} \ a = \dim \text{Ker} \ b = \frac{\dim V}{2} \), that \( \text{Ker} \ a = \text{Im} \ a \), \( \text{Ker} \ b = \text{Im} \ b \), and \( V = \text{Ker} \ a \oplus \text{Ker} \ b \).

Choose a basis \((e_1, \ldots, e_n)\) of \( \text{Ker} \ b \). Then, \((a(e_1), \ldots, a(e_n))\) is a basis of \( \text{Im} \ a = \text{Ker} \ a \), whence \((e_1, \ldots, e_n, a(e_1), \ldots, a(e_n))\) is a basis of \( V \). In that basis, the matrices that represent \( a \) and \( b \) are, respectively,

\[
A = \begin{bmatrix} 0_n & 0_n \\ I_n & 0_n \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 0_n & M \\ 0_n & 0_n \end{bmatrix}
\]

for some \( M \in M_n(\mathbb{F}) \).

Hence,

\[
A + B = \begin{bmatrix} 0_n & M \\ I_n & 0_n \end{bmatrix}.
\]

By Corollary A.3, all the invariant factors of \( A + B \) are even polynomials, which proves our result.

We can finish the proof of Theorem 2.6. Let \( u \) be an endomorphism of a finite-dimensional vector space which splits into the sum of two square-zero endomorphisms. Then, each invariant factor of \( u \) is the product of a power of \( t \) (possible \( t^0 \)) with either 1 or an invariant factor of the invertible part \( u_i \) in the Fitting decomposition. By Corollary A.7 and Lemma A.8, each invariant factor of \( u_i \) is an even polynomial, whence every invariant factor of \( u \) is an even or an odd polynomial.

25
References

[1] J.D. Botha, Sums of two square-zero matrices over an arbitrary field, Linear Algebra Appl. 436 (2012) 516–524.

[2] S. Breaz, G. Călugăreanu, Sums of nilpotent matrices, Linear Multilinear Algebra, in press, doi: 10.1080/03081087.2016.1167816

[3] P.A. Fillmore, On similarity and the diagonal of a matrix, Amer. Math. Monthly 76 (2) (1969) 167–169.

[4] W. Roth, The equations $AX - YB = C$ and $AX - XB = C$ in matrices, Proc. Amer. Math. Soc. 3 (1952) 392–396.

[5] C. de Seguins Pazzis, On decomposing any matrix as a linear combination of three idempotents, Linear Algebra Appl. 433 (2010) 843–855.

[6] C. de Seguins Pazzis, On linear combinations of two idempotent matrices over an arbitrary field, Linear Algebra Appl. 433 (2010) 625–636.

[7] C. de Seguins Pazzis, On sums of idempotent matrices over a field of positive characteristic, Linear Algebra Appl. 433 (2010) 856–866.

[8] C. de Seguins Pazzis, Sums of two triangularizable quadratic matrices over an arbitrary field, Linear Algebra Appl. 436 (2012) 3293–3302.

[9] K. Takahashi, On sums of three square-zero matrices, Linear Algebra Appl. 306 (2000) 45–57.

[10] J.-H. Wang, P.Y. Wu, Sums of square-zero operators, Studia Math. 99 (1991) 115–127.