ELSEVIER

Linear Algebra and its Applications 306 (2000) 45–57

LINEAR ALGEBRA AND ITS APPLICATIONS

www.elsevier.com/locate/laa

On sums of three square-zero matrices ☆

K. Takahashi [1]

Department of Mathematics, Hokkaido University, Sapporo 060, Japan

Received 23 September 1999; accepted 15 October 1999

Submitted by R.A. Brualdi

Abstract

Wang and Wu characterized matrices which are sums of two square-zero matrices, and proved that every matrix with trace-zero is a sum of four square-zero matrices. Moreover, they gave necessary or sufficient conditions for a matrix to be a sum of three square-zero matrices. In particular, they proved that if an $n \times n$ matrix A is a sum of three square-zero matrices, the dim $\ker(A - \alpha I) \leqslant 3n/4$ for any scalar $\alpha \neq 0$. Proposition 1 shows that this condition is not necessarily sufficient for the matrix A to be a sum of three square-zero matrices, and characterizes sums of three square-zero matrices among matrices with minimal polynomials of degree 2.

Let $\mathbb{C}^{n\times m}$ denote the space of all $n \times m$ complex matrices. The block diagonal matrix with diagonal blocks $A_1, \ldots, A_m$ is denoted by $A_1 \oplus \cdots \oplus A_m$, and if $A_1 = \cdots = A_m = A$, then we write $A_1 \oplus \cdots \oplus A_m = A^{(m)}$. Let I_n denote the $n \times n$ identity matrix.

Proposition 1. *Let A be an* $n \times n$ *matrix with* $\operatorname{tr} A = 0$ *and assume that A is similar to*

$$\begin{bmatrix} \beta & 0 \\ 0 & \alpha \end{bmatrix}^{(m)} \oplus \alpha I_r,$$

☆ All correspondence to: Pei Yuan Wu, Department of Applied Mathematics, National Chiao Tung University, 1001 Ta Hsueh Road, Hsinchu, Taiwan, ROC.

E-mail address: pywu@cc.nctu.edu.tw (P.Y. Wu).

[1] Deceased.

0024-3795/00/$ - see front matter
PII: S 0 0 2 4 - 3 7 9 5 (9 9) 0 0 2 3 9 - 6

where $\alpha \neq \beta$ *and* $m, r \geqslant 1$. *Then A is a sum of three square-zero matrices if and only if r is a divisor of* $2m$.

Lemma 1. *Let*

$$A = \begin{bmatrix} \alpha & 0 \\ 0 & \beta \end{bmatrix}^{(2)} \in \mathbb{C}^{4\times 4} \quad \textit{and} \quad \alpha \neq \beta.$$

Then, for any γ *and* δ *such that* $\gamma \neq \delta$ *and* $\gamma + \delta = \alpha + \beta$, *there is a square-zero matrix N such that* $A + N$ *is similar to*

$$\begin{bmatrix} \gamma & 1 \\ 0 & \gamma \end{bmatrix} \oplus \begin{bmatrix} \delta & 1 \\ 0 & \delta \end{bmatrix}.$$

Proof. Clearly, A is similar to

$$\begin{bmatrix} \gamma I_2 & I_2 \\ c I_2 & \delta I_2 \end{bmatrix},$$

where $c = (\alpha - \gamma)(\gamma - \beta)$. By considering the matrix

$$S^{-1} \begin{bmatrix} \gamma I_2 & I_2 \\ c I_2 & \delta I_2 \end{bmatrix} S \quad \text{for} \quad S = I_1 \oplus \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \oplus I_1,$$

we see that there is a square-zero matrix N such that $A + N$ is similar to

$$\begin{bmatrix} \gamma & 1 & 1 & 0 \\ 0 & \gamma & 0 & 1 \\ 0 & 0 & \delta & -1 \\ 0 & 0 & 0 & \delta \end{bmatrix},$$

which is similar to

$$\begin{bmatrix} \gamma & 1 \\ 0 & \gamma \end{bmatrix} \oplus \begin{bmatrix} \delta & 1 \\ 0 & \delta \end{bmatrix}$$

because $\gamma \neq \delta$. □

The following lemma is a special case of Proposition 1.

Lemma 2. *Let A be an* $n \times n$ *matrix with* tr $A = 0$ *and suppose that A is similar to*

$$\begin{bmatrix} \beta & 0 \\ 0 & \alpha \end{bmatrix}^{(m)} \oplus \alpha I_r,$$

where α *and* β *are scalars with* $\alpha \neq \beta$. *If* $r \leqslant 2$, *then A is a sum of three square-zero matrices.*

Proof. Since tr $A = 0$, the condition $\alpha \neq \beta$ is equivalent to $\alpha \neq 0$. The case when $r = 0$ or 1 follows from [1, Proposition 3.3] and its proof. (Indeed, since $\alpha \neq \beta$, the proof of [1, Proposition 3.3] with $c = -\alpha$ shows the case of $r = 1$.) Thus we consider the case when $r = 2$. Suppose that m is even. Then A is similar to $A_1 \oplus A_1$, where

$$A_1 = \begin{bmatrix} \beta & 0 \\ 0 & \alpha \end{bmatrix}^{(m/2)} \oplus \alpha I_1 \quad \text{and} \quad \operatorname{tr} A_1 = 0.$$

As remarked above, A_1 is a sum of three square-zero matrices, and hence so is A. Next, suppose that m is odd and $m = 2k + 1$. The condition $\operatorname{tr} A = 0$ implies $\beta + \alpha = -2\alpha/m$. By Lemma 1, for $1 \leqslant i \leqslant k$, there is a square-zero matrix N_i such that

$$\begin{bmatrix} \beta & 0 \\ 0 & \alpha \end{bmatrix}^{(2)} + N_i$$

is similar to

$$\begin{bmatrix} -(2\mathrm{i}+1)\alpha/m & 1 \\ 0 & -(2\mathrm{i}+1)\alpha/m \end{bmatrix} \oplus \begin{bmatrix} (2\mathrm{i}-1)\alpha/m & 1 \\ 0 & (2\mathrm{i}-1)\alpha/m \end{bmatrix}.$$

Also, there is a square-zero matrix M such that the matrix

$$\begin{bmatrix} \beta & 0 \\ 0 & \alpha \end{bmatrix} + M$$

is similar to

$$\begin{bmatrix} -\alpha/m & 1 \\ 0 & -\alpha/m \end{bmatrix}$$

(see [1] or the proof of Lemma 1). Let

$$N = N_1 \oplus \cdots \oplus N_k \oplus M \oplus \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

Then $N^2 = 0$ and the matrix

$$\left(\begin{bmatrix} \beta & 0 \\ 0 & \alpha \end{bmatrix}^{(m)} \oplus \alpha I_2 \right) + N$$

is similar to

$$\begin{aligned} B = \bigoplus_{i=1}^{k} & \left(\begin{bmatrix} -(2\mathrm{i}+1)\alpha/m & 1 \\ 0 & -(2\mathrm{i}+1)\alpha/m \end{bmatrix} \right. \\ & \left. \oplus \begin{bmatrix} (2\mathrm{i}-1)\alpha/m & 1 \\ 0 & (2\mathrm{i}-1)\alpha/m \end{bmatrix} \right) \\ & \oplus \begin{bmatrix} -\alpha/m & 1 \\ 0 & -\alpha/m \end{bmatrix} \oplus \begin{bmatrix} \alpha & 1 \\ 0 & \alpha \end{bmatrix}. \end{aligned}$$

Clearly, B is similar to $-B$, and so by [1, Theorem 2.11] B is a sum of two square-zero matrices. Therefore it follows that A is a sum of three square-zero matrices. □

Lemma 3. *Let A be an $n \times n$ matrix whose minimal polynomial is $m(\lambda) = (\lambda - \alpha)(\lambda - \beta)$, and let N be an $n \times n$ square-zero matrix. If γ is the eigenvalue of $A + N$ and $\gamma \neq \alpha, \beta$, then $\alpha + \beta - \gamma$ is also the eigenvalue of $A + N$.*

Proof. Since $N^2 = 0$, we can take an invertible matrix P such that

$$P^{-1}NP = \begin{bmatrix} 0 & 0 \\ N_1 & 0 \end{bmatrix},$$

where $N_1 \in \mathbb{C}^{r\times(n-r)}$. Let

$$P^{-1}AP = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},$$

where $A_{21} \in \mathbb{C}^{r\times(n-r)}$. Then, since $(A-\alpha I)(A-\beta I) = 0$, we have

$$A_{11}A_{12} + A_{12}A_{22} = (\alpha+\beta)A_{12}$$

and the invariant polynomials of the matrix polynomials

$$[A_{11} - \lambda I, A_{12}] \quad \text{and} \quad \begin{bmatrix} A_{12} \\ A_{22} - \lambda I \end{bmatrix}$$

are divisors of $(\lambda-\alpha)(\lambda-\beta)$.

Hence the lemma follows from [2, Theorem 6(b)]. □

Proof of Proposition 1. Suppose that $2m = rs$ for some integer s. Then, according as r is odd or even, A is similar to $C^{(r)}$ or $D^{(r/2)}$, where

$$C = \begin{bmatrix} \beta & 0 \\ 0 & \alpha \end{bmatrix}^{(s/2)} \oplus \alpha I_1 \quad \text{and} \quad D = \begin{bmatrix} \beta & 0 \\ 0 & \alpha \end{bmatrix}^{(s)} \oplus \alpha I_2.$$

In each case, the condition $\operatorname{tr} A = 0$ implies $\operatorname{tr} C = 0$ or $\operatorname{tr} D = 0$. By Lemma 2, the matrices C and D are sums of three square-zero matrices and so A is a sum of three square-zero matrices.

Conversely, assume that A is a sum of three square-zero matrices. Then there is a square-zero matrix N such that $A + N$ is a sum of two square-zero matrices. Since $\operatorname{rank}(A - \alpha I) < n/2$ and $\operatorname{rank} N \leqslant n/2$ because N is square-zero, we have $\alpha \in \sigma(A+N)$, so it follows from [1, Theorem 2.11] that $-\alpha \in \sigma(A+N)$. Since $\operatorname{tr} A = 0$, the conditions $\alpha \neq \beta$ and $r \geqslant 1$ imply that $-(k\alpha + (k-1)\beta) \neq \beta$ for every integer $k \geqslant 1$ (and $-\alpha \neq \alpha$). Therefore, if $-(k\alpha + (k-1)\beta) \neq \alpha$ for all integers $k \geqslant 2$, then it follows from Lemma 3 and [1, Theorem 2.11] that $k\alpha + (k-1)\beta \in \sigma(A+N)$ for all k, which is impossible. Thus we have $-(k\alpha + (k-1)\beta) = \alpha$ for some integer $k \geqslant 2$ and therefore $2m = (k-1)r$ because $\operatorname{tr} A = 0$. □

For a matrix A, let $\mu_A = \max\{\dim\ker(A - \alpha I) : \alpha \in \mathbb{C}\}$. If $\operatorname{tr} A = 0$ and $\mu_A = \dim\ker A$, then the rational form of A shows that A is similar to $A_1 \oplus \cdots \oplus A_m \oplus 0$, where each A_i is a cyclic matrix of size at least 2. By [1, Proposition 3.3], $A_1 \oplus \cdots \oplus A_m$ is a sum of three square-zero matrices, hence so is A. Thus we consider matrices A with $\mu_A = \dim\ker(A - \alpha I)$ for some $\alpha \neq 0$.

Lemma 4. *Let A be an $n \times n$ matrix with $\mu_A > n/2$ and N be an $n \times n$ square-zero matrix. Then there is an invertible matrix P such that*

$$P^{-1}AP = \begin{bmatrix} \alpha I_{k_1} & 0 & 0 & 0 \\ * & B_{11} & B_{12} & 0 \\ * & B_{21} & B_{22} & 0 \\ * & * & * & \alpha I_{k_2} \end{bmatrix}$$

and

$$P^{-1}(A+N)P = \begin{bmatrix} \alpha I_{k_1} & 0 & 0 & 0 \\ * & B_{11} & B_{12} & 0 \\ * & B_{21}+M & B_{22} & 0 \\ * & * & * & \alpha I_{k_2} \end{bmatrix},$$

where α *is a scalar such that* $\dim\ker(A-\alpha I) = \mu_A, k_1+k_2 = 2\mu_A - n, B_{ij}\,(i,j = 1,2)$ *and* M *are* $(n-\mu_A)\times(n-\mu_A)$ *matrices and* $*$ *are some matrices.*

Proof. Let $r = n/2$ or $r = (n-1)/2$ according as n is even or odd. Since N is square-zero, we may assume that

$$N = \begin{bmatrix} 0 & 0 \\ N_1 & 0 \end{bmatrix},$$

where $N_1 \in \mathbb{C}^{r\times(n-r)}$. We write

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},$$

where $A_{21} \in \mathbb{C}^{r\times(n-r)}$. Then

$$\operatorname{rank}[A_{11}-\alpha I,\ A_{12}] \leqslant \operatorname{rank}[A-\alpha I] = n-\mu_A$$

and

$$\operatorname{rank}\begin{bmatrix} A_{12} \\ A_{22}-\alpha I \end{bmatrix} \leqslant n-\mu_A.$$

Therefore there are invertible matrices $Q_1 \in \mathbb{C}^{(n-r)\times(n-r)}$ and $Q_2 \in \mathbb{C}^{r\times r}$ such that

$$\begin{bmatrix} Q_1 & 0 \\ 0 & Q_2 \end{bmatrix}^{-1} A \begin{bmatrix} Q_1 & 0 \\ 0 & Q_2 \end{bmatrix} = \begin{bmatrix} \alpha I_{\mu_A - r} & 0 & 0 & 0 \\ * & B_{11} & B_{12} & 0 \\ * & B_{21} & B_{22} & 0 \\ * & * & * & \alpha I_{\mu_A + r - n} \end{bmatrix},$$

where $B_{ij} \in \mathbb{C}^{(n-\mu_A)\times(n-\mu_A)}(i,j = 1,2)$, and we can write

$$\begin{bmatrix} Q_1 & 0 \\ 0 & Q_2 \end{bmatrix}^{-1} N \begin{bmatrix} Q_1 & 0 \\ 0 & Q_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ * & M & 0 & 0 \\ * & * & 0 & 0 \end{bmatrix}$$

in the same block form as the one of the matrix

$$\begin{bmatrix} Q_1 & 0 \\ 0 & Q_2 \end{bmatrix}^{-1} A \begin{bmatrix} Q_1 & 0 \\ 0 & Q_2 \end{bmatrix}.$$

This proves the lemma. □

Proposition 2. *Let A be an $n \times n$ matrix with* $\operatorname{tr} A = 0$, *and suppose that* $\mu_A = \dim\ker(A - \alpha I)$ *for some* $\alpha \neq 0$.

(1) *If $n = 4m$ and $\mu_A = 3m$, then A is a sum of three square-zero matrices if and only if A is similar to $A_1^{(m)}$, where* $A_1 = \operatorname{diag}(-3\alpha, \alpha, \alpha, \alpha)$.

(2) *If $n = 4m - 1$ and $\mu_A = 3m - 1$, then A is a sum of three square-zero matrices if and only if A is similar to $A_1^{(m-1)} \oplus A_2$, where* $A_1 = \operatorname{diag}(-3\alpha, \alpha, \alpha, \alpha)$ *and* $A_2 = \operatorname{diag}(-2\alpha, \alpha, \alpha)$.

Proof. The "if" parts of the assertions (1) and (2) follow from the fact that A_1 and A_2 are sums of three square-zero matrices (see [1, Corollary 3.5]). So suppose that A is a sum of three square-zero matrices, or equivalently, there is a square-zero matrix N such that $A + N$ is a sum of two square-zero matrices.

(1) By Lemma 4, we may assume that

$$A = \begin{bmatrix} \alpha I_{k_1} & 0 & 0 & 0 \\ B_{10} & B_{11} & B_{12} & 0 \\ B_{20} & B_{21} & B_{22} & 0 \\ B_{30} & B_{31} & B_{32} & \alpha I_{k_2} \end{bmatrix}$$

and

$$A + N = \begin{bmatrix} \alpha I_{k_1} & 0 & 0 & 0 \\ B_{10} & B_{11} & B_{12} & 0 \\ * & B_{21} + M_1 & B_{22} & 0 \\ * & * & B_{32} & \alpha I_{k_2} \end{bmatrix},$$

where B_{11}, B_{22} and M_1 are $m \times m$ matrices and $k_1 = k_2 = m$. Let

$$B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \quad \text{and} \quad M = \begin{bmatrix} 0 & 0 \\ M_1 & 0 \end{bmatrix},$$

which are $2m \times 2m$ matrices. Since $A + N$ is a sum of two square-zero matrices, it follows from [1, Theorem 2.11] that $A + N$ is similar to $-(A + N)$ and therefore $\sigma(B + M) = \{-\alpha\}$, which implies that $A + N$ is similar to

$$\begin{bmatrix} \alpha I_{k_1} & 0 \\ * & \alpha I_{k_2} \end{bmatrix} \oplus (B + M).$$

Then, since $A + N$ and $-(A + N)$ are similar, it follows that $(B + M + \alpha I)^2 = 0$. On the other hand, the invertibility of $B + M - \alpha I$ implies

$$\operatorname{rank}[B_{11} - \alpha I,\ B_{12}] = m = \operatorname{rank}(A - \alpha I).$$

Hence there are matrices F and G such that

$$F[B_{10}, B_{11} - \alpha I,\ B_{12}] = [B_{20},\ B_{21},\ B_{22} - \alpha I]$$

and

$$G[B_{10},\ B_{11} - \alpha I,\ B_{12}] = [B_{30},\ B_{31},\ B_{32}],$$

and therefore we have

$$\begin{bmatrix} I & 0 & 0 & 0 \\ 0 & I & 0 & 0 \\ 0 & F & I & 0 \\ 0 & G & 0 & I \end{bmatrix}^{-1} A \begin{bmatrix} I & 0 & 0 & 0 \\ 0 & I & 0 & 0 \\ 0 & F & I & 0 \\ 0 & G & 0 & I \end{bmatrix} = \begin{bmatrix} \alpha I & 0 & 0 & 0 \\ B_{10} & B_{11} + B_{12}F & B_{12} & 0 \\ 0 & 0 & \alpha I & 0 \\ 0 & 0 & 0 & \alpha I \end{bmatrix}$$

and

$$\begin{bmatrix} I & 0 \\ F & I \end{bmatrix}^{-1} (B + M) \begin{bmatrix} I & 0 \\ F & I \end{bmatrix} = \begin{bmatrix} B_{11} + B_{12}F & B_{12} \\ M_1 & \alpha I \end{bmatrix}.$$

Since $(B + M + \alpha I)^2 = 0$, it follows that the matrix B_{12} is invertible and $(B_{11} + B_{12}F + 3\alpha I)B_{12} = 0$, so that $B_{11} + B_{12}F = -3\alpha I$. Therefore we can conclude that A is similar to $A_1^{(m)}$ because $-3\alpha \neq \alpha$.

(2) The argument similar to the proof of (1) with $k_1 = m$ and $k_2 = m - 1$ shows that the characteristic polynomial of $B + M$ is $p(\lambda) = \lambda(\lambda + \alpha)^{2m-1}$ and its minimal polynomial is a divisor of $\lambda(\lambda + \alpha)^2$. Thus $B + M - \alpha I$ is invertible and as in the proof of (1), we see that A and $B + M$ are similar to

$$\begin{bmatrix} \alpha I & 0 & 0 & 0 \\ * & C & B_{12} & 0 \\ 0 & 0 & \alpha I & 0 \\ 0 & 0 & 0 & \alpha I \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} C & B_{12} \\ M_1 & \alpha I \end{bmatrix},$$

where $C \in \mathbb{C}^{m \times m}$. We also have that $\operatorname{rank}(B + M + \alpha I)^2 = 1$. Hence $\operatorname{rank}(C + 3\alpha I)B_{12} \leqslant 1$ and, since the invertibility of $B + M - \alpha I$ implies that of B_{12}, $\dim \ker(C + 3\alpha I) \geqslant m - 1$. This, together with the identity $\operatorname{tr} B = -(2m - 1)\alpha$, shows that -2α is the eigenvalue of C and $\dim \ker(C + 3\alpha I) = m - 1 (\alpha \neq 0)$. Thus A is similar to $A_1^{(m-1)} \oplus A_2$. $\square$

Lemma 5. *Let $A = B \oplus \alpha I_r$, where B is an $m \times m$ cyclic matrix and $r \leqslant m - 2$. Then, for any $m + r$ scalars $\delta_1, \ldots, \delta_{m+r}$ with $\sum_{i=1}^{m+r} \delta_i = \operatorname{tr} A$, there is a square-zero matrix N such that*

$$\sigma(A + N) = \{\delta_1, \ldots, \sigma_{m+r}\}.$$

Proof. The case of $r = 0$ is shown in [1, Lemma 3.2], and if $\alpha \notin \sigma(B)$, then $B \oplus \alpha I_1$ is cyclic. Thus, by considering $B \oplus \alpha I_1$ instead of B in this case, we may assume that $\alpha \in \sigma(B)$. Let $p(\lambda) = \prod_{i=1}^{m}(\lambda - \beta_i)$ be the characteristic polynomial of B, where $\beta_1 = \alpha$, and let P be an $(m + r) \times (m + r)$ invertible matrix whose jth column p_j is

$$p_j = \begin{cases} \left(\prod_{i=1}^{j}(B-\beta_i I)x\right) \oplus 0 & \text{for } j \leqslant m-1, \\ x \oplus 0 & \text{for } j = m, \\ \left(\prod_{i=1}^{j-m}(B-\beta_i I)x\right) \oplus e_{j-m} & \text{for } m+1 \leqslant j \leqslant m+r, \end{cases}$$

where x is a cyclic vector of B and $\{e_1, e_2, \ldots, e_r\}$ is a basis for $\mathbb{C}^r$. Then we have

$$P^{-1}AP = \begin{bmatrix} A_{11} & A_{12} \\ 0 & \alpha I_{r+1} \end{bmatrix},$$

where

$$A_{11} = \begin{bmatrix} \beta_2 & 0 & \cdots & \cdots & 0 \\ 1 & \beta_3 & 0 & \cdots & \vdots \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 & \beta_m \end{bmatrix} \in \mathbb{C}^{(m-1)\times(m-1)}$$

and

$$A_{12} = \begin{bmatrix} 1 & \beta_2 & 0 & \cdots & 0 \\ 0 & 1 & \beta_3 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \beta_{r+1} \\ \vdots & \ddots & \ddots & 0 & 1 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & 0 \end{bmatrix} \in \mathbb{C}^{(m-1)\times(r+1)}.$$

Since the pair (A_{11}, A_{12}) is of full range, that is,

$$\operatorname{rank}[A_{12}, A_{11}A_{12}, \ldots, A_{11}^{m-2}A_{12}] = m-1,$$

and rank $A_{12} = r+1$, by [2, Theorem 1] there is a matrix X such that

$$\sigma\left(\begin{bmatrix} A_{11} & A_{12} \\ X & \alpha I \end{bmatrix}\right) = \{\delta_1, \ldots, \delta_{m+r}\}.$$

Therefore, if N is the square-zero matrix defined by

$$N = P\begin{bmatrix} 0 & 0 \\ X & 0 \end{bmatrix}P^{-1},$$

then we have $\sigma(A+N) = \{\delta_1, \ldots, \delta_{m+r}\}$, which proves the lemma. □

Lemma 6. *Let $A = B \oplus \alpha I_{m-2}$, where B is an $m \times m$ cyclic matrix such that $\alpha \in \sigma(B)$, and let $1 \leqslant s \leqslant m-1$. Then, for scalars γ and $\delta_1, \ldots, \delta_{2(m-1)-s}$ such that*

$$s\gamma + \sum_{i=1}^{2(m-1)-s} \delta_i = \operatorname{tr} A \quad \textit{and} \quad \delta_i \neq \delta \quad \textit{for } 1 \leqslant i \leqslant 2(m-1)-s,$$

there is a square-zero matrix N such that $A + N$ is similar to $C_1 \oplus C_2$, where $C_1 \in \mathbb{C}^{s\times s}$ and $C_2 \in \mathbb{C}^{(2(m-1)-s)\times(2(m-1)-s)}$ are matrices such that

$$(C_1 - \gamma I)^2 = 0 \quad \textit{and} \quad \sigma(C_2) = \{\delta_1, \ldots, \delta_{2(m-1)-s}\}.$$

Proof. The proof of Lemma 5 with $r = m-2$ shows that A_{12} is invertible and so A is similar to

$$\tilde{A} = \begin{bmatrix} A_1 & I \\ 0 & \alpha I \end{bmatrix},$$

where $A_1 \in \mathbb{C}^{(m-1)\times(m-1)}$ is cyclic. Since $D = A_1 + \alpha I$ is cyclic, for a cyclic vector x of D, the matrix Q whose jth column is $\prod_{i=1}^{j-1}(D - d_i)x$, where $d_i = \delta_i + \delta_{m-1+i}$ for $1 \leqslant i \leqslant m-1-s$ and $d_i = \delta_i + \gamma$ for $m-s \leqslant i \leqslant m-1$, is invertible, and

$$Q^{-1}DQ = \begin{bmatrix} d_1 & 0 & \cdots & \cdots & c_1 \\ 1 & d_2 & \ddots & \ddots & c_2 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & c_{m-2} \\ 0 & \cdots & \cdots & 1 & d_{m-1} \end{bmatrix},$$

for some scalars $c_1, c_2, \ldots, c_{m-2}$. Let

$$G = Q \begin{bmatrix} \delta_1 & 0 & \cdots & \cdots & 0 \\ 1 & \delta_2 & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 1 & \delta_{m-1} \end{bmatrix} Q^{-1}$$

and $H = D - G$. Then

$$\sigma(G) = \{\delta_1, \ldots, \delta_{m-1}\} \quad \text{and} \quad H = \begin{bmatrix} H_{11} & H_{12} \\ 0 & H_{22} \end{bmatrix},$$

where $H_{11} = \operatorname{diag}(\delta_m, \delta_{m+1}, \ldots, \delta_{2(m-1)-s})$,

$$H_{12} = \begin{bmatrix} 0 & \cdots & 0 & c_1 \\ \vdots & \cdots & \vdots & \vdots \\ 0 & \cdots & 0 & c_{m-1-s} \end{bmatrix} \in \mathbb{C}^{(m-1-s)\times s}$$

and

$$H_{22} = \begin{bmatrix} \gamma & 0 & \cdots & c_{m-s} \\ 0 & \gamma & \ddots & \vdots \\ \vdots & \ddots & \ddots & c_{m-2} \\ 0 & \cdots & 0 & \gamma \end{bmatrix} \in \mathbb{C}^{s\times s}.$$

Since $\gamma \neq \delta_i$ for all i, H is similar to $\operatorname{diag}(\delta_m, \ldots, \delta_{2(m-1)-s}) \oplus H_{22}$, and $(H_{22} - \gamma I)^2 = 0$. We also have

$$\begin{bmatrix} I & 0 \\ G - A_1 & I \end{bmatrix}^{-1} \tilde{A} \begin{bmatrix} I & 0 \\ G - A_1 & I \end{bmatrix} = \begin{bmatrix} G & I \\ HG - \alpha A_1 & H \end{bmatrix},$$

so that if N is the square-zero matrix defined by

$$N = \begin{bmatrix} I & 0 \\ G - A_1 & I \end{bmatrix} \begin{bmatrix} 0 & 0 \\ \alpha A_1 - HG & 0 \end{bmatrix} \begin{bmatrix} I & 0 \\ G - A_1 & I \end{bmatrix}^{-1},$$

then $\tilde{A} + N$ is similar to

$$\begin{bmatrix} G & I \\ 0 & H \end{bmatrix}.$$

But, since $\gamma \neq \delta_i$ for all i, the matrix

$$\begin{bmatrix} G & I \\ 0 & H \end{bmatrix}$$

is similar to

$$\begin{bmatrix} G & J \\ 0 & H_{11} \end{bmatrix} \oplus H_{22},$$

where

$$J = \begin{bmatrix} I_{m-1-s} \\ 0 \end{bmatrix} \in \mathbb{C}^{(m-1)\times(m-1-s)}.$$

This proves the lemma. □

Note that the proof of Lemma 6 shows that in Lemma 6, if $s < m - 2$, the matrix C_1 can be also taken to be $C_1 = \gamma I$.

Proposition 3. *Let A be an $n \times n$ matrix with* $\operatorname{tr} A = 0$, *and let m be the number of its invariant polynomials of degree* 2. *If $\mu_A \leqslant (2n - m)/3$, then A is a sum of three square-zero matrices.*

Proof. Let α be a scalar such that $\dim \ker(A - \alpha I) = \mu_A$, and let ℓ and r be the numbers of the invariant polynomials of A of degreeat least 3 and of degree 1, respectively. Using the rational form of A, we may assume that $A = \left(\bigoplus_{i=1}^{\ell+m} B_i\right) \oplus \alpha I_r$, where each B_i is a cyclic matrix with $\alpha \in \sigma(B_i)$ whose size k_i is $k_i \geqslant 3$ for $1 \leqslant i \leqslant \ell$ and $k_i = 2$ for $\ell + 1 \leqslant i \leqslant \ell + m$. (Indeed, $B_{\ell+1} = \cdots = B_{\ell+m}$.) Note that since $n = \sum_{i=1}^{\ell} k_i + 2m + r$ and $\mu_A = \ell + m + r$, the condition $\mu_A \leqslant (2n - m)/3$ is equivalent to $r \leqslant \sum_{i=1}^{\ell}(2k_i - 3)$. First suppose that $r \leqslant \sum_{i=1}^{\ell}(k_i - 2)$, which is equivalent to $\mu_A \leqslant n/2$. Take ℓ nonnegative integers $r_1, \ldots, r_\ell$ such that $\sum_{i=1}^{\ell} r_i = r$ and $r_i \leqslant k_i - 2$ for all i, and let $A_i = B_i \oplus \alpha I_{r_i}$ for $1 \leqslant i \leqslant \ell$ and $A_i = B_i$ for $\ell + 1 \leqslant i \leqslant \ell + m$. Then A is similar to $\tilde{A} = \bigoplus_{i=1}^{\ell+m} A_i$. Let $t_i = \operatorname{tr} A_i$ for $1 \leqslant i \leqslant \ell + m$ and take a scalar c such that $c > \sum_{i=1}^{\ell+m} |t_i|$. As in the proof of [1, Proposition 3.3], we apply Lemma 5 to the matrices $A_1, \ldots, A_{\ell+m}$ to obtain square-zero matrices $N_1, \ldots, N_{\ell+m}$ such that

$$\sigma(A_i + N_i) = \left\{ c - \sum_{j=1}^{i-1} t_j, \sum_{j=1}^{i} t_j - c,\ 0, \ldots, 0 \right\}$$

for $1 \leqslant i \leqslant \ell + m$. Then, since $c - \sum_{j=1}^{i-1} t_j$ and $\sum_{j=1}^{i} t_j - c$ are different nonzero numbers, each $A_i + N_i$ is similar to

$$\operatorname{diag}\left(c - \sum_{j=1}^{i-1} t_j,\ \sum_{j=1}^{i} t_j - c,\ 0, \ldots, 0 \right),$$

and so the matrix $\tilde{A} + N$, where $N = \bigoplus_{i=1}^{\ell+m} N_i$, is similar to $-(\tilde{A} + N)$. Hence it follows from [1, Theorem 2.11] that $\tilde{A} + N$ is a sum of two square-zero matrices. Since $N^2 = 0$, we can conclude that A is a sum of three square-zero matrices.

Next suppose that $r > \sum_{i=1}^{\ell}(k_i - 2)$, and let $s = r - \sum_{i=1}^{\ell}(k_i - 2)$. Then A is similar to $\tilde{A} = (\bigoplus_{i=1}^{\ell+m} A_i) \oplus \alpha I_s$, where $A_i = B_i \oplus \alpha I_{k_i - 2}$ for $1 \leqslant i \leqslant \ell + m$. Since $r \leqslant \sum_{i=1}^{\ell}(2k_i - 3)$ by assumption, $0 < s \leqslant \sum_{i=1}^{\ell}(k_i - 1)$, so we can take q integers $s_1, \ldots, s_q$ $(q \leqslant \ell)$ such that $\sum_{i=1}^{q} s_i = s$ and $1 \leqslant s_i \leqslant k_i - 1$ for each i. Let $t_i = \operatorname{tr} A_i + \alpha s_i$ for $1 \leqslant i \leqslant q$ and $t_i = \operatorname{tr} A_i$ for $q + 1 \leqslant i \leqslant \ell + m$, and let c be a scalar with $c > \sum_{i=1}^{\ell+m} |t_i| + |\alpha|$. Then, for each i, the numbers $-\alpha$, $c - \sum_{j=1}^{i-1} t_j$ and $\sum_{j=1}^{i} t_j - c$ are nonzero and mutually different. Hence, for $1 \leqslant i \leqslant q$, by Lemma 6 there is a square-zero matrix N_i such that $A_i + N_i$ is similar to $C_i \oplus D_i$, where C_i is an $s_i \times s_i$ matrix with $(C_i + \alpha I)^2 = 0$ and

$$D_i = \operatorname{diag}\left(c - \sum_{j=1}^{i-1} t_j, \sum_{j=1}^{i} t_j - c,\ 0, \ldots, 0 \right) \in \mathbb{C}^{(2(k_i-1)-s_i)\times(2(k_i-1)-s_i)}.$$

Also, for $q + 1 \leqslant i \leqslant \ell + m$, we have a square-zero matrix N_i such that $A_i + N_i$ is similar to

$$D_i = \begin{bmatrix} c - \sum_{j=1}^{i-1} t_j & 0 \\ 0 & \sum_{j=1}^{i} t_j - c \end{bmatrix}$$

(see [1] or Lemma 5). By [1, Theorem 2.11], $\bigoplus_{i=1}^{\ell+m} D_i$ is a sum of two square-zero matrices, and since $\sigma(C_i) = \{-\alpha\}$, the matrices $C_i \oplus (C_i + 2\alpha I_{s_i})$, $i = 1, \ldots, q$, are sums of two square-zero matrices too. Now, let

$$N = \left(\bigoplus_{i=1}^{\ell+m} N_i \right) \oplus \left(\bigoplus_{i=1}^{q} (C_i + \alpha I) \right).$$

Then $N^2 = 0$ and $\tilde{A} + N$ is similar to

$$\left(\bigoplus_{i=1}^{\ell+m} D_i \right) \oplus \left(\bigoplus_{i=1}^{q} (C_i \oplus (C_i + 2\alpha I)) \right),$$

which is a sum of two square-zero matrices. Thus it follows that A is a sum of three square-zero matrices. □

Corollary 1. *Let A be an $n \times n$ matrix with* $\operatorname{tr} A = 0$. *If $\mu_A \leqslant n/2 + 1$, then A is a sum of three square-zero matrices.*

Proof. Let ℓ and r be the numbers of the invariant polynomials of A of degree at least 3 and of degree 1, respectively. The condition $\mu_A \leqslant n/2 + 1$ is equivalent to $r \leqslant \sum_{i=1}^{\ell}(k_i - 2) + 2$, where $k_1, \ldots, k_\ell$ are the degrees of the invariant polynomials of A with degree $\geqslant 3$. So, if $\ell = 0$, then $r \leqslant 2$ and it follows from Lemma 2 that A is a sum of three square-zero matrices. If $\ell \neq 0$, then the inequality $r \leqslant \sum_{i=1}^{\ell}(k_i - 2) + 2$ implies $r \leqslant \sum_{i=1}^{\ell}(2k_i - 3)$, so the assertion follows from Proposition 3. □

Corollary 2. *Let A be an $n \times n$ matrix with* $\operatorname{tr} A = 0$ *and suppose that* $\mu_A = \dim \ker(A - \alpha I)$ *for some* $\alpha \neq 0$.

(1) *When $n = 6$, A is a sum of three square-zero matrices if and only if $\mu_A \leqslant 4$.*
(2) *When $n = 7$, A is a sum of three square-zero matrices if and only if* (i) $\mu_A \leqslant 4$ *or* (ii) $\mu_A = 5$ *and A is similar to* $(-3\alpha I_1) \oplus (-2\alpha I_1) \oplus \alpha I_5$.
(3) *When $n = 8$, A is a sum of three square-zero matrices if and only if* (i) $\mu_A \leqslant 5$ *or* (ii) $\mu_A = 6$ *and A is similar to* $(-3\alpha I_2) \oplus \alpha I_6$.

Proof. The "only if" parts of (1)–(3) follow from [1, Theorem 3.1] and Proposition 2. On the other hand, the "if" parts follow from [1, Corollary 3.4] and Corollary 1. □

References

[1] J.-H. Wang, P.Y. Wu, Sums of square-zero operators, Studia Math. 99 (2) (1991) 115–127.
[2] K. Takahashi, Eigenvalues of matrices with given block upper triangular part, Linear Algebra Appl. 239 (1996) 175–184.