RESEARCH ARTICLE

A note on joint diagonalizers of shear matrices

Francesco G. Russo

DIEETCAM, Universitá Degli Studi di Palermo, Viale Delle Scienze, Edificio 8, 90128, Palermo, Italy

e-mail: francescog.russo@yahoo.com

ABSTRACT: The matrix joint diagonalization is a fundamental notion for studying the blind signal processing and the independent component analysis. We reformulate some classical results on the topic in terms of shear matrices. These matrices appear naturally in differential geometry and group theory and the present paper deals with certain perturbations of joint diagonalizers for which the knowledge of shear matrices is crucial.

KEYWORDS: indipendent component analysis, transvections, cost functions, linear groups

INTRODUCTION

Digital signal processing involves linear algebra in several areas of investigation and, for instance, a series of recent contributions show applications of digital signal processing to mathematical models for music[1–3]. Ref. 4 is a classic source for the study of digital signal processing, but more specifically Refs. 5–7 contribute to the blind signal processing, which is an interesting branch of research. The present paper deals with some ideas of Cardoso treated in Refs. 8–10 and is related to indipendent component analysis. These ideas have been recently generalized by Afsari in Refs. 11–14 and Yeredor in Refs. 15, 16. Most of these references contain such fundamental notions as exact joint diagonalization (EJD), orthogonal joint diagonalization and non-orthogonal joint diagonalization (NOJD). The motivation of the present paper is the fact that the original ideas of Cardoso[8,9] have been recently improved in Refs. 7, 14, 15 in terms of problems of optimization and calculus over smooth manifolds (see Refs. 17, 18 for relations to Riemannian geometry, control theory, topology, theory of compact Lie groups and numerical analysis).

Our main results generalize Proposition 1 of Ref. 9 in terms of certain families of matrices which are obtained by shear matrices. This allows us to describe the joint diagonalizers of many other matrices and find new criteria of classification.

We first recall some notions from Refs. 6, 17–19 and then deal with the main results. Terminology and notation are standard and follow Refs. 2, 6, 18, 19.

FEEDBACK OF DIFFERENTIAL GEOMETRY

Following Refs. 2, 17–19, the Hilbert space $\mathbb{C}^n$ of dimension $n \geqslant 1$ over the field $\mathbb{C}$ of the complex numbers allows us to consider the following groups

$$\mathrm{GL}(n, \mathbb{C}) = \{B \in \mathbb{C}^{n \times n} \mid \det(B) \neq 0\},$$

$$\mathrm{GL}(n, \mathbb{R}) = \{B \in \mathbb{R}^{n \times n} \mid \det(B) \neq 0\},$$

$$Z(\mathrm{GL}(n, \mathbb{C})) = \{A, B \in \mathrm{GL}(n, \mathbb{C}) \mid AB = BA\},$$

$$Z(\mathrm{GL}(n, \mathbb{C})) = \{aI \in \mathrm{GL}(n, \mathbb{C}) \mid a \in \mathbb{C} - \{0\}\},$$

$$\mathrm{PGL}(n, \mathbb{C}) = \mathrm{GL}(n, \mathbb{C})/Z(\mathrm{GL}(n, \mathbb{C})),$$

$$\mathrm{PGL}(n, \mathbb{R}) = \mathrm{GL}(n, \mathbb{R})/Z(\mathrm{GL}(n, \mathbb{R})),$$

$$\mathrm{SL}(n, \mathbb{C}) = \{B \in \mathbb{C}^{n \times n} \mid \det(B) = 1\},$$

$$\mathrm{SL}(n, \mathbb{R}) = \{B \in \mathbb{R}^{n \times n} \mid \det(B) = 1\},$$

$$\mathrm{PSL}(n, \mathbb{R}) = \mathrm{SL}(n, \mathbb{C})/Z(\mathrm{SL}(n, \mathbb{C})),$$

$$\mathrm{PSL}(n, \mathbb{R}) = \mathrm{SL}(n, \mathbb{R})/Z(\mathrm{SL}(n, \mathbb{R})).$$

Assume that $B \in \mathrm{GL}(n, \mathbb{C})$. $B^{\mathrm{H}} = \overline{B}^*$, where $*$ denotes the transpose and the bar denotes complex conjugation, denotes the conjugate transpose of B. We have

$$\mathrm{U}(n) = \{B \in \mathrm{GL}(n, \mathbb{C}) \mid (Bx, Bx) = (x, x), x \in \mathbb{C}^n\}$$
$$= \{B \in \mathrm{GL}(n, \mathbb{C}) \mid B^{-1} = B^{\mathrm{H}}\},$$

$$\mathrm{O}(n) = \{B \in \mathrm{GL}(n, \mathbb{R}) \mid (Bx, Bx) = (x, x), x \in \mathbb{R}^n\}$$
$$= \{B \in \mathrm{GL}(n, \mathbb{R}) \mid B^{-1} = B^*\},$$

$$\mathrm{SO}(n) = \mathrm{O}(n) \cap \mathrm{SL}(n, \mathbb{R}).$$

Furthermore, $B \in \mathrm{GL}(n, \mathbb{C})$ is called *hermitian*, if $B^{\mathrm{H}} = B$, and *anti-hermitian*, if $B^{\mathrm{H}} = -B$.

$\mathrm{GL}(n, \mathbb{C})$ has the topology induced by $\| \| : A \in \mathrm{GL}(n, \mathbb{C}) \mapsto \|A\| = \sup\{\|Ax\| \mid \|x\| \leqslant 1\} \in [0, +\infty[$ for all $x \in \mathbb{C}^n$ and is a topological group. The

same is true for $\mathrm{GL}(n, \mathbb{R})$. Notice that $\mathrm{O}(n)$, $\mathrm{SO}(n)$, $\mathrm{U}(n)$, $\mathrm{SU}(n)$ are compact groups for $n = 1, 2, \ldots$.

If $A \in \mathrm{GL}(n, \mathbb{C})$, *the off of A* is the map

$$\mathrm{off} : A \in \mathrm{GL}(n, \mathbb{C}) \longmapsto \tag{1}$$

$$\mathrm{off}(A) = \sum_{\substack{i,j \in \{1,2,\ldots,n\} \\ i \neq j}} |a_{ij}|^2 \in [0, +\infty[$$

where a_{ij} is the (i, j)-th entry of A (see Ref. 9). In general if $V \in \mathrm{U}(n)$ and

$$\mathcal{M} = (M_1, M_2, \ldots, M_k, \ldots, M_m) \tag{2}$$

is an m-tuple of matrices in $\mathrm{GL}(n, \mathbb{C})^m$, then

$$\mathcal{Y} : (V, \mathcal{M}) \in \mathrm{U}(n) \times \mathrm{GL}(n, \mathbb{C})^m \longmapsto \tag{3}$$

$$\mathcal{Y}(V, \mathcal{M}) = \sum_{k=1}^{m} \mathrm{off}(V^{\mathrm{H}} M_k V) \in [0, +\infty[$$

and V is called *unitary minimizer* (or *joint diagonalizer*) of $\mathcal{Y}$ of $\mathcal{M}$, if $\mathcal{Y}(V, \mathcal{M}) = 0$. In the case of connected compact groups, existence and uniqueness of non-trivial solutions for $\mathcal{Y}(V, \mathcal{M}) = 0$ are ensured by Weierstrass theorems (see Ref. 17), but we cannot say whether an arbitrary $(V, \mathcal{M})$ satisfies $\mathcal{Y}(V, \mathcal{M}) = 0$ or not (this is the EJD problem in Refs. 13, 14).

Now, instead of $\mathcal{M}$, consider the m-tuple of $\mathrm{GL}(n, \mathbb{C})^m$

$$\mathcal{M}_0 = (UD_1 U^{\mathrm{H}}, \ldots, UD_k U^{\mathrm{H}}, \ldots, UD_m U^{\mathrm{H}}) \tag{4}$$

where $U \in \mathrm{U}(n)$ and D_k is the diagonal matrix of $\mathbb{C}^{n \times n}$ with diagonal entries $d_1(k), \ldots, d_n(k)$. We have

$$\mathcal{Y}(U, \mathcal{M}_0) = \sum_{k=1}^{m} \mathrm{off}(U^{\mathrm{H}}(UD_k U^{\mathrm{H}})U)$$

$$= \sum_{k=1}^{m} \mathrm{off}(U^{-1} U D_k U^{-1} U) = \sum_{k=1}^{m} \mathrm{off}(D_k) = 0, \tag{5}$$

and $\mathcal{Y}$ is minimized. Here it is clear that U is a joint diagonalizer for $\mathcal{Y}$ with respect to $\mathcal{M}_0$.

A more general situation is the following:

$$\mathcal{M}_\lambda = (UD_1 U^{\mathrm{H}} + \lambda R_1, \ldots, UD_k U^{\mathrm{H}} + \lambda R_k, \ldots \tag{6}$$

$$\ldots, UD_m U^{\mathrm{H}} + \lambda R_m),$$

where $\lambda \in \mathbb{R}$ and $R_k \in \mathrm{GL}(n, \mathbb{C})$ (for $k = 1, 2, \ldots, m$). When $\lambda = 0$, (6) reduces to (4). Therefore it is very interesting to look for $\lambda \neq 0$

for which $(U, \mathcal{M}_\lambda) \in \mathrm{U}(n) \times \mathrm{GL}(n, \mathbb{C})^m$ satisfies $\mathcal{Y}(U, \mathcal{M}_\lambda) = 0$. The answer is positive in the sense of Proposition 1 of Ref. 9. On the other hand, small values of λ, close to 0, allow us to study the perturbations in a neighbourhood of the minimum value of $\mathcal{Y}$ on $(U, \mathcal{M}_0)$. Some discussions and more details on this point can be found in Refs. 5, 7, 8, 11–14.

Now let e_i be the $n \times 1$ row vector with 1 in the ith position and 0 elsewhere, that is, $e_1 = (1, 0, 0, \ldots, 0)$, $e_2 = (0, 1, 0, \ldots, 0)$, …, $e_i = (0, \ldots, 0, 1, 0, \ldots, 0)$, …, $e_n = (0, 0, \ldots, 1)$. e_i^* denotes the $1 \times n$ column vector with 1 in the ith position and 0 elsewhere. Then we have the matrix

$$\begin{pmatrix} 0 & 0 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \ldots & 0 \end{pmatrix} \tag{7}$$

whose entries e_{ij} are 1 in the (i, j)th position and 0 elsewhere. More generally, a_i is the ith row and a_i^* the ith column of $A \in \mathrm{GL}(n, \mathbb{C})$, whose entries are a_{ij}.

Proposition 1 (See Proposition 1 of Ref. 9) *Assume that $\mathcal{M}_\lambda$ in (6) satisfies*

$$\forall i, j \in \{1, 2, \ldots, n\} \text{ such that } i \neq j \tag{8}$$

$$\exists k \in \{1, 2, \ldots, m\} \text{ such that } d_i(k) \neq d_j(k).$$

Then $\mathcal{Y}(U(I + \lambda G), \mathcal{M}_\lambda) = 0$ for $\lambda \neq 0$ small enough, where G is an anti-hermitian matrix whose diagonal is null. Furthermore, its off-diagonal entries are

$$g_{ij} = \frac{1}{2} \sum_{k=1}^{m} f_{ij}^*(k), u_i^* R_k u_j + f_{ij}(k) u_i^* R_k^{\mathrm{H}} u_j, \tag{9}$$

where

$$f_{ij}(k) = \frac{d_j(k) - d_i(k)}{\sum_{l=1}^{m} |d_j(l) - d_i(l)|^2} \in \mathbb{C}. \tag{10}$$

Notice that (8) is essential for expressing (10).

Remark 1 In Proposition 1, $J = I$ and we consider only small enough $\lambda \neq 0$ so that $o(\lambda)$ is omitted. As noted in Ref. 9, if $U(I + \lambda G + o(\lambda))$ minimizes $\mathcal{Y}$, then $U(I + \lambda G + o(\lambda))J$ minimizes $\mathcal{Y}$. In particular, this is true when λ is small enough and $o(\lambda)$ is omitted.

MAIN RESULTS

From Chapter 3 of Ref. 19, each invertible matrix with coefficients in $\mathbb{C}$ (resp. $\mathbb{R}$) and determinant equal to 1 can be generated by matrices of the form

$$K(1, a) = I + ae_{ij}$$

called *shear matrices* (or *transvections*), where $a \in \mathbb{C}$ (resp. $a \in \mathbb{R}$). Now $K(1, a)$ differs from I only in that there is an a in the (i, j)th position, and, if $a = 0$, then $K(1, 0) = I$. These matrices are important because they generate $\mathrm{SL}(n, \mathbb{C})$ (resp. $\mathrm{SL}(n, \mathbb{R})$) and each $B \in \mathrm{SL}(n, \mathbb{C})$ (resp. $\mathrm{SL}(n, \mathbb{R})$) can be written uniquely as product of finitely many $K(1, a)$, for suitable a. This is the so-called *rational canonical form of a special linear matrix* (pp. 73–76 of Ref. 19).

Lemma 1 (See 3.2.10 of Ref. 19) *For all $n \geqslant 2$,*

$$\mathrm{SL}(n, \mathbb{C}) = \langle K(1, a) \mid a \in \mathbb{C} \rangle$$

and

$$\mathrm{SL}(n, \mathbb{R}) = \langle K(1, a) \mid a \in \mathbb{R} \rangle.$$

Now we make the following observation.

Remark 2 We note that $\mathrm{off}(K(1, a)) = |a|^2$ and

$$a = 0 \implies \mathrm{off}(K(1, 0)) = \mathrm{off}(I) = 0.$$

The matrix $K(d_i(k), a) = d_i(k) + ae_{ij}$ is obtained from $I + ae_{ij}$ by replacing the 1's on the principal diagonal with the $d_1(k), \ldots, d_n(k)$, corresponding to D_k. Now

$$d_i(k) = 1 \, \forall i = 1, 2, \ldots, n \implies K(d_i(k), a) = K(1, a).$$

Then we can consider the family of matrices

$$\mathcal{M}_{a,\lambda} = (U \cdot K(d_i(1), a) \cdot U^{\mathrm{H}} + \lambda R_1, \ldots$$
$$\ldots, U \cdot K(d_i(2), a) \cdot U^{\mathrm{H}} + \lambda R_2, \ldots$$
$$\ldots, U \cdot K(d_i(m), a) \cdot U^{\mathrm{H}} + \lambda R_m), \quad (11)$$

so that $\mathcal{M}_{0,\lambda} = \mathcal{M}_\lambda$ and $\mathcal{M}_{0,0} = \mathcal{M}_0$. Consequently, we may generalize the results in Refs. 8, 9, 14, replacing (6) with (11).

Now we consider the linear maps

$$\gamma_{ijk} : V \in \mathrm{U}(n) \mapsto \gamma_{ijk}(V) = e_i^* V^{\mathrm{H}} M_k^{\mathrm{H}} V e_j \in \mathbb{C} \tag{12}$$

and

$$T_{ijk} : V \in \mathrm{U}(n) \longmapsto T_{ijk}(V) \tag{13}$$
$$= e_i e_j^* V^{\mathrm{H}} M_k^{\mathrm{H}} V - V^{\mathrm{H}} M_k^{\mathrm{H}} V e_i e_j^* \in \mathrm{GL}(n, \mathbb{C}).$$

Immediately, we deduce

$$\mathcal{Y}(V, \mathcal{M}) = \sum_{k=1}^m \sum_{\substack{i,j \in \{1, \ldots, n\} \\ i \neq j}} |\gamma_{ijk}|^2. \tag{14}$$

Now the linear map

$$\mathcal{S} : (V, \mathcal{M}) \in \mathrm{U}(n) \times \mathrm{GL}(n, \mathbb{C})^m \longmapsto \mathcal{S}(V, \mathcal{M})$$
$$= \sum_{k=1}^m \sum_{\substack{i,j \in \{1, \ldots, n\} \\ i \neq j}} \gamma_{ijk}^* \, T_{ijk} \in \mathrm{GL}(n, \mathbb{C}) \tag{15}$$

(recall that $(A^*)^* = A$ and $(AB)^* = B^* A^*$ in $\mathrm{GL}(n, \mathbb{C})$ and the equality $V^{\mathrm{H}} = \overline{V}^*$) implies

$$\gamma_{ijk}^*(V) = (e_i^* V^{\mathrm{H}} M_k^{\mathrm{H}} V e_j)^*$$
$$= e_j^* V^* (M_k^{\mathrm{H}})^* (V^{\mathrm{H}})^* e_i$$
$$= e_j^* V^* (M_k^{\mathrm{H}})^* \overline{V} e_i. \tag{16}$$

In particular, if $(V, \mathcal{M}) \in (\mathrm{U}(n) \cap \mathrm{GL}(n, \mathbb{R})) \times \mathrm{GL}(n, \mathbb{R})^m$, then $\overline{V} = V$, $V^{\mathrm{H}} = V^*$, $(M_k^{\mathrm{H}})^* = \overline{M_k} = M_k$, hence

$$\gamma_{ijk}^*(V) = e_j^* V^{\mathrm{H}} M_k V e_i. \tag{17}$$

Once we substitute (17) and (13) in (15), we find Equation 12 of Ref. 9. We may do the same for the linear map

$$\mathcal{S}^{\mathrm{H}} : (V, \mathcal{M}) \in \mathrm{U}(n) \times \mathrm{GL}(n, \mathbb{C})^m \mapsto \mathcal{S}^{\mathrm{H}}(V, \mathcal{M})$$
$$= \left(\sum_{k=1}^m \sum_{\substack{i,j \in \{1, \ldots, n\} \\ i \neq j}} \gamma_{ijk}^* \, T_{ijk} \right)^{\mathrm{H}}$$
$$= \sum_{k=1}^m \sum_{\substack{i,j \in \{1, \ldots, n\} \\ i \neq j}} (\gamma_{ijk}^*)^{\mathrm{H}} \, T_{ijk}^{\mathrm{H}} \in \mathrm{GL}(n, \mathbb{C}). \tag{18}$$

Here is a short proof of Lemma 1 of Ref. 9.

Lemma 2 *If $(V, \mathcal{M}) \in (\mathrm{U}(n) \cap \mathrm{GL}(n, \mathbb{R})) \times \mathrm{GL}(n, \mathbb{R})^m$, then $\mathcal{S}$ is hermitian.*

Proof: Since (17) is satisfied, $(\gamma_{ijk}^*)^{\mathrm{H}} = \gamma_{ijk}^*$ and $T_{ijk}(V)^{\mathrm{H}} = T_{ijk}(V)$, then

$$\mathcal{S}(V, \mathcal{M}) = \sum_{k=1}^m \sum_{\substack{i,j \in \{1, \ldots, n\} \\ i \neq j}} \gamma_{ijk}^* \, T_{ijk}$$
$$= \sum_{k=1}^m \sum_{\substack{i,j \in \{1, \ldots, n\} \\ i \neq j}} \gamma_{ijk} \, T_{ijk}^{\mathrm{H}} = \mathcal{S}^{\mathrm{H}}(V, \mathcal{M}). \tag{19}$$

Now we generalize Lemmas 2 and 3 of Ref. 9 by using $\mathcal{M}_{a,\lambda}$ instead of $\mathcal{M}_\lambda$. The condition $\mathcal{S}(V,\mathcal{M}_{a,\lambda}) = \mathcal{S}^H(V,\mathcal{M}_{a,\lambda})$, described by Lemma 2, has a physical meaning. It is a stationary equation which allows us to study the perturbations around the solution $(U,\mathcal{M}_{a,\lambda}) \in (\mathrm{U}(n) \cap \mathrm{GL}(n,\mathbb{R})) \times \mathrm{GL}(n,\mathbb{R})^m$, minimizing $\mathcal{Y}$. A more detailed analysis can be done if we stop at the terms of the first order in $\lambda \neq 0$ and look for an anti-hermitian $L \in \mathrm{GL}(n,\mathbb{C})$ such that, for small enough $\lambda \neq 0$ and $a \neq 0$,

$$\mathcal{S}(U(I+\lambda L),\mathcal{M}_{a,\lambda}) = \mathcal{S}^H(U(I+\lambda L),\mathcal{M}_{a,\lambda}). \tag{20}$$

The problem can be more conveniently centred at U, noting that

$$\mathcal{S}(U(I+\lambda L),\mathcal{M}_{a,\lambda}) = \mathcal{S}(I+\lambda L,\mathcal{N}_{a,\lambda}), \tag{21}$$

where

$$\begin{aligned}\mathcal{N}_{a,\lambda} &= (K(d_i(1),a) + \lambda U^H R_1 U, \ K(d_i(2),a) \\ &\quad + \lambda U^H R_2 U, \dots, K(d_i(k),a) + \lambda U^H R_k U, \\ &\qquad \dots, \ K(d_i(m),a) + \lambda U^H R_m U). \end{aligned} \tag{22}$$

Eq. (21) is exactly Eq. 14 of Ref. 9 for $a = 0$ and we stop at the terms of first order in λ.

Corollary 1 *From* (21) *and Lemma 2, the equations* (20) *and* $\mathcal{S}(I+\lambda L,\mathcal{N}_{a,\lambda}) = \mathcal{S}^H(I+\lambda L,\mathcal{N}_{a,\lambda})$ *have the same solutions for small enough values of* $\lambda \neq 0$ *and* $a \neq 0$.

We can say more.

Lemma 3 *Assume L is an anti-hermitian matrix and* $(I+\lambda L,\mathcal{N}_{a,\lambda}) \in (\mathrm{U}(n) \cap \mathrm{GL}(n,\mathbb{R}) \times \mathrm{GL}(n,\mathbb{R})^m$. *If we stop at the terms of the first order in* $\lambda \neq 0$, *then the entries of* $\mathcal{S}(I+\lambda L,\mathcal{N}_{a,\lambda})$ *are*

$$\begin{aligned}&s_{ij}(I+\lambda L,\mathcal{N}_{a,\lambda}) \\ &= \sum_{k=1}^m \delta_{ijk}\alpha_{ijk} + (\delta_{ijk}\beta_{ijk} + \epsilon_{ijk}\alpha_{ijk})\lambda, \end{aligned} \tag{23}$$

where α_{ijk}, β_{ijk}, δ_{ijk}, ϵ_{ijk} are linear maps depending only on $K(d_i(k),a)$ and $\lambda \neq 0$ and $a \neq 0$ are small enough.

Proof: From (13),

$$\begin{aligned}T_{ijk}(I+\lambda L) &= e_i e_j^*(I+\lambda L)^H \cdot K(d_i(k),a)^H \\ &\cdot (I+\lambda L) - (I+\lambda L)^H \cdot K(d_i(k),a)^H \cdot (I+\lambda L)e_i e_j^* \\ &= e_i e_j^*(I+\lambda L^H) \cdot K(d_i(k),a)^H \cdot (I+\lambda L) \\ &\quad - (I+\lambda L^H) \cdot K(d_i(k),a)^H \cdot (I+\lambda L)e_i e_j^* \\ &= (e_i e_j^* I + e_i e_j^* \lambda L^H) \cdot (K(d_i(k),a)^H \cdot I \\ &\quad + K(d_i(k),a)^H \cdot \lambda L) - (I \cdot K(d_i(k),a)^H \\ &\quad + \lambda L^H \cdot K(d_i(k),a)^H) \cdot (I \cdot e_i e_j^* + \lambda L \cdot e_i e_j^*) \\ &= e_i e_j^* K(d_i(k),a)^H + \lambda e_i e_j^* K(d_i(k),a)^H L \\ &\quad + \lambda e_i e_j^* L^H K(d_i(k),a)^H + \lambda^2 e_i e_j^* L^H K(d_i(k),a)^H L \\ &\quad - K(d_i(k),a)^H e_i e_j^* - \lambda K(d_i(k),a)^H L e_i e_j^* \\ &\quad - \lambda L^H K(d_i(k),a)^H e_i e_j^* - \lambda^2 L^H K(d_i(k),a)^H L e_i e_j^* \end{aligned} \tag{24}$$

Stopping at the terms of first order in λ, we delete $\lambda^2 e_i e_j^* L^H K(d_i(k),a)^H L$ and $\lambda^2 L^H K(d_i(k),a)^H L e_i e_j^*$, then

$$\begin{aligned}T_{ijk}(I+\lambda L) &= e_i e_j^* K(d_i(k),a)^H \\ &\quad + \lambda e_i e_j^* K(d_i(k),a)^H L + \lambda e_i e_j^* L^H K(d_i(k),a)^H \\ &\quad - K(d_i(k),a)^H e_i e_j^* - \lambda K(d_i(k),a)^H L e_i e_j^* \\ &\qquad - \lambda L^H K(d_i(k),a)^H e_i e_j^*. \end{aligned}$$

Expanding in λ, we put

$$\alpha_{ijk} = e_i e_j^* K(d_i(k),a)^H - K(d_i(k),a)^H e_i e_j^* \tag{25}$$

and

$$\begin{aligned}\beta_{ijk} &= (e_i e_j^* K(d_i(k),a)^H L - L^H K(d_i(k),a)^H e_i e_j^*) \\ &\quad + (e_i e_j^* L^H K(d_i(k),a)^H - K(d_i(k),a)^H L e_i e_j^*). \end{aligned} \tag{26}$$

Hence

$$T_{ijk}(I+\lambda L) = \alpha_{ijk} + \beta_{ijk}\lambda. \tag{27}$$

On the other hand, (17) becomes

$$\begin{aligned}\gamma_{ijk}^*(I+\lambda L) &= e_i^*(I+\lambda L)^H K(d_i(k),a)(I+\lambda L)e_j \\ &= e_i^*(I+\lambda L^H)K(d_i(k),a)(I+\lambda L)e_j \\ &= (e_i^* K(d_i(k),a) + \lambda e_i^* L^H K(d_i(k),a)) \cdot (Ie_j + \lambda Le_j) \\ &= e_i^* K(d_i(k),a)e_j + \lambda e_i^* K(d_i(k),a)Le_j \\ &\quad + \lambda e_i^* L^H K(d_i(k),a)e_j + \lambda^2 e_i^* L^H K(d_i(k),a)Le_j. \end{aligned} \tag{28}$$

We can do again an approximation at the terms of the first order in λ, deleting $\lambda^2 e_i^* L^{\mathrm{H}} K(d_i(k), a) L e_j$, and we find

$$= e_i^* K(d_i(k), a) e_j + \lambda e_i^* K(d_i(k), a) L e_j$$
$$+ \lambda e_i^* L^{\mathrm{H}} K(d_i(k), a) e_j.$$

Expanding in λ, we put

$$\delta_{ijk} = e_i^* K(d_i(k), a) e_j \qquad (29)$$

and

$$\epsilon_{ijk} = e_i^* K(d_i(k), a) L e_j + e_i^* L^{\mathrm{H}} K(d_i(k), a) e_j \qquad (30)$$

and hence

$$\gamma_{ijk}^* = \delta_{ijk} + \lambda \epsilon_{ijk}. \qquad (31)$$

We may conclude that

$$s_{ij}(I + \lambda L, \mathcal{N}_{a,\lambda}) = \sum_{k=1}^{m} \sum_{\substack{i,j\{1,\ldots,n\} \\ i \neq j}} \gamma_{ijk}^* \, T_{ijk}$$
$$= \sum_{k=1}^{m} \sum_{\substack{i,j\{1,\ldots,n\} \\ i \neq j}} (\delta_{ijk} + \epsilon_{ijk}\lambda)(\alpha_{ijk} + \beta_{ijk}\lambda). \qquad (32)$$

Now

$$(\delta_{ijk} + \epsilon_{ijk}\lambda)(\alpha_{ijk} + \beta_{ijk}\lambda)$$
$$= \delta_{ijk}\alpha_{ijk} + (\delta_{ijk}\beta_{ijk} + \epsilon_{ijk}\alpha_{ijk})\lambda + (\epsilon_{ijk}\beta_{ijk})\lambda^2$$
$$= \delta_{ijk}\alpha_{ijk} + (\delta_{ijk}\beta_{ijk} + \epsilon_{ijk}\alpha_{ijk})\lambda$$

and the result follows. $\qquad \square$

We are ready to prove the main result.

Theorem 1 *Assume that L is an anti-hermitian matrix and $(I + \lambda L, \mathcal{N}_{a,\lambda}) \in (\mathrm{U}(n) \cap \mathrm{GL}(n, \mathbb{R}) \times \mathrm{GL}(n, \mathbb{R})^m$. Then*

$$\mathcal{Y}(I + \lambda L, \mathcal{N}_{a,\lambda}) = \sum_{k=1}^{m} \sum_{\substack{i,j \in \{1,\ldots,n\} \\ i \neq j}} |\gamma_{ijk}(I + \lambda L)|^2, \qquad (33)$$

for small enough $\lambda \neq 0$ and $a \neq 0$ around $U \in \mathrm{U}(n) \cap \mathrm{GL}(n, \mathbb{R})$. Furthermore, the stationary condition $\mathcal{S}(I + \lambda L, \mathcal{N}_{a,\lambda}) = \mathcal{S}^{\mathrm{H}}(I + \lambda L, \mathcal{N}_{a,\lambda})$ is satisfied with entries $s_{ij}(I + \lambda L, \mathcal{N}_{a,\lambda})$ as in (23).

Proof: Apply Corollary 1 and Lemma 3. $\qquad \square$

The cases $\lambda = a = 0$ and $\lambda \neq 0$ with $a = 0$ can be found in Refs. 8, 9, 14. Now notice that

$$\mathcal{N}_{a,\lambda} = (K(d_i(1), a), \ldots, K(d_i(m), a))$$
$$+ \lambda(U^{\mathrm{H}} R_1 U, \ldots, U^{\mathrm{H}} R_m U) = \mathcal{C}_a + \lambda \mathcal{D}, \quad (34)$$

where

$$\mathcal{C}_a = (K(d_i(1), a), \ldots, K(d_i(m), a)) \in \mathrm{SL}(n, \mathbb{C})^m$$

and

$$\mathcal{D} = (U^{\mathrm{H}} R_1 U, \ldots, U^{\mathrm{H}} R_m U) \in \mathrm{GL}(n, \mathbb{C})^m.$$

Then

$$\mathcal{N}_\lambda = \{(K(d_i(1), a), \ldots, K(d_i(m), a)) + \lambda \mathcal{D} \mid a \in \mathbb{R}\}$$
$$= \{(K(d_i(1), a), \ldots, K(d_i(m), a)) \mid a \in \mathbb{R}\} + \lambda \mathcal{D} \qquad (35)$$

and, in particular, for $d_i(1) = \ldots = d_i(m) = 1$,

$$\mathcal{N} = \{(K(1, a), \ldots, K(1, a)) \mid a \in \mathbb{R}\} + \lambda \mathcal{D} \quad (36)$$

generate $\mathrm{SL}(n, \mathbb{C})^m + \lambda \mathcal{D}$ (see Lemma 2).

Corollary 2 *Assume L is an anti-hermitian matrix, $\mathcal{N}$ as in* (36) *and $(I + \lambda L, \mathcal{N}) \in (\mathrm{U}(n) \cap \mathrm{GL}(n, \mathbb{R}) \times \mathrm{GL}(n, \mathbb{R})^m$. Then*

$$\mathcal{Y}(I + \lambda L, \mathcal{N}) = \sum_{k=1}^{m} \sum_{\substack{i,j \in \{1,\ldots,n\} \\ i \neq j}} |\gamma_{ijk}(I + \lambda L)|^2, \qquad (37)$$

for small enough values of $\lambda \neq 0$ and $a \neq 0$ around $U \in \mathrm{U}(n) \cap \mathrm{GL}(n, \mathbb{R})$. Furthermore, the stationary condition $\mathcal{S}(I + \lambda L, \mathcal{N}) = \mathcal{S}^{\mathrm{H}}(I + \lambda L, \mathcal{N})$ is satisfied with entries $s_{ij}(I + \lambda L, \mathcal{N})$ as in (23).

The importance of Corollary 2 is emphasized by the following observation.

Remark 3 The perturbations of joint diagonalizers of shear matrices are described by (37).

Corollary 2 has the following consequence.

Remark 4 The rational canonical form of a special linear matrix allows us to conclude that the perturbations of joint diagonalizers of special linear matrices can be obtained by linear combinations of perturbations of joint diagonalizers of shear matrices.

Non-singular matrices were involved in most the present paper and it might seem to be a strong restriction. The next observation justifies our choice, showing that it is not really restrictive.

Remark 5 Let A be a singular matrix with n rows and n columns over $\mathbb{R}$ (resp. $\mathbb{C}$) and $(a_{11}, a_{22}, \ldots, a_{nn})$ be the entries of its diagonal. If the product $a_{11} \cdot a_{22} \cdot \ldots \cdot a_{nn} \neq 0$, then we may write uniquely $A = T_1 + T_2$ as the sum of the upper triangular matrix T_1, whose diagonal is $\frac{1}{2}(a_{11}, a_{22}, \ldots, a_{nn})$, and the lower triangular matrix T_2, whose diagonal is the same of T_1. Both T_1 and T_2 are non-singular matrices.

We end with a concrete example, in which we can see that the two families $\mathcal{M}_\lambda$ and $\mathcal{M}_{a,\lambda}$ are very different when $a \neq 0$.

Remark 6 In order to simplify the notation, we may concentrate on the case $n = m = 3$ of (11), consider real matrices and denote the (real) numbers $d_i(k)$ only by $\alpha, \beta, \gamma, \delta, \varepsilon, \phi, \iota, \kappa, \mu$, avoiding indices, and the matrices $K(d_i(k), a)$ only by K_1, K_2, K_3. Assume that $\theta \in \mathbb{R}$ and

$$U = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

so that

$$U^{\mathrm{H}} = \overline{U}^* = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Now take $R_1 = R_2 = R_3 = I_3$, just for fixing the ideas (I_3 is the identity matrix of $GL(3, \mathbb{R})$), and

$$D_1 = \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 0 & \gamma \end{pmatrix}, \quad D_2 = \begin{pmatrix} \delta & 0 & 0 \\ 0 & \varepsilon & 0 \\ 0 & 0 & \phi \end{pmatrix}$$

$$D_3 = \begin{pmatrix} \iota & 0 & 0 \\ 0 & \kappa & 0 \\ 0 & 0 & \mu \end{pmatrix},$$

$$K_1 = \begin{pmatrix} \alpha & a & 0 \\ 0 & \beta & 0 \\ 0 & 0 & \gamma \end{pmatrix}, \quad K_2 = \begin{pmatrix} \delta & 0 & a \\ 0 & \varepsilon & 0 \\ 0 & 0 & \phi \end{pmatrix},$$

$$K_3 = \begin{pmatrix} \iota & 0 & 0 \\ 0 & \kappa & a \\ 0 & 0 & \mu \end{pmatrix}.$$

Making the products, we get that the rows of

$$A = UD_1U^{\mathrm{H}} + \lambda I_3$$

are

$$a_{1,j} = (\lambda + \alpha\cos^2\theta + \beta\sin^2\theta, (\beta - \alpha)(\cos\theta\sin\theta), 0)$$

$$a_{2,j} = ((\beta - \alpha)(\cos\theta\sin\theta), \lambda + \beta\cos^2\theta + \alpha\sin^2\theta, 0)$$

$$a_{3,j} = (0, 0, \lambda + \gamma)$$

the rows of

$$B = UD_2U^{\mathrm{H}} + \lambda I_3$$

are

$$b_{1,j} = (\lambda + \delta\cos^2\theta + \varepsilon\sin^2\theta, (\varepsilon - \delta)(\cos\theta\sin\theta), 0)$$

$$b_{2,j} = ((\varepsilon - \delta)(\cos\theta\sin\theta), \lambda + \varepsilon\cos^2\theta + \delta\sin^2\theta, 0)$$

$$b_{3,j} = (0, 0, \lambda + \phi)$$

and the rows of

$$C = UD_3U^{\mathrm{H}} + \lambda I_3$$

are

$$c_{1,j} = (\lambda + \iota\cos^2\theta + \kappa\sin^2\theta, (\kappa - \iota)(\cos\theta\sin\theta), 0)$$

$$c_{2,j} = ((\kappa - \iota)(\cos\theta\sin\theta), \lambda + \kappa\cos^2\theta + \iota\sin^2\theta, 0)$$

$$c_{3,j} = (0, 0, \lambda + \mu).$$

Now the rows of

$$L = UK_1U^{\mathrm{H}} + \lambda I_3$$

are

$$l_{1,j} = (\lambda + \alpha\cos^2\theta + \beta\sin^2\theta + a\sin\theta\cos\theta,$$
$$(\beta - \alpha)(\cos\theta\sin\theta) + a\cos^2\theta, 0)$$
$$l_{2,j} = ((\beta - \alpha)(\cos\theta\sin\theta) - a\sin^2\theta,$$
$$\lambda + \beta\cos^2\theta + \alpha\sin^2\theta - a\sin\theta\cos\theta, 0)$$
$$l_{3,j} = (0, 0, \lambda + \gamma)$$

which clearly differ from the respective rows $a_{1,j}$ and $a_{2,j}$ by the presence of the new terms $\pm a\sin\theta\cos\theta$, $-a\sin^2\theta$, $a\cos^2\theta$. Similarly, this happens for

$$M = UK_2U^{\mathrm{H}} + \lambda I_3$$

where

$$m_{1,j} = (\lambda + \delta\cos^2\theta + \varepsilon\sin^2\theta + a\sin\theta\cos\theta,$$
$$(\varepsilon - \delta)(\cos\theta\sin\theta) + a\cos^2\theta, 0)$$
$$m_{2,j} = ((\varepsilon - \delta)(\cos\theta\sin\theta) - a\sin^2\theta,$$
$$\lambda + \varepsilon\cos^2\theta + \delta\sin^2\theta - a\sin\theta\cos\theta, 0)$$
$$m_{3,j} = (0, 0, \lambda + \phi)$$

and for

$$N = UK_3U^{\mathrm{H}} + \lambda I_3$$

where

$$n_{1,j} = (\lambda + \iota\cos^2\theta + \kappa\sin^2\theta + a\sin\theta\cos\theta,$$
$$(\kappa - \iota)(\cos\theta\sin\theta) + a\cos^2\theta, 0)$$
$$n_{2,j} = ((\kappa - \iota)(\cos\theta\sin\theta) - a\sin^2\theta,$$
$$\lambda + \kappa\cos^2\theta + \iota\sin^2\theta - a\sin\theta\cos\theta, 0)$$
$$n_{3,j} = (0, 0, \lambda + \mu)$$

We conclude that $\mathcal{M}_{a,\lambda} = \{L, M, N\}$ is different from $\mathcal{M}_\lambda = \{A, B, C\}$ whenever $a \neq 0$.

ACKNOWLEDGEMENTS

REFERENCES

1. Russo FG, Schettino C (2007) *Problemi d'acustica nelle sale da concerto*, Aracne, Rome.
2. Russo FG (2004) *La musica algoritmica e l'Offerta Musicale di J. S. Bach*, Delta 3, Grottaminarda.
3. Russ FG, Schettino C (2010) A simulation of the Madrigal n.1 of the 3rd book of madrigals of Gesualdo. *Far East J Math Sci* **47**, 51–61.
4. Monoakis DG, Proakis JG (1996) *Digital signal processing: principles, algorithms and applications*, Prentice–Hall, New York.
5. Beyer R, Bunse–Gerstner A, Mehrmann V (1993) Numerical methods for simultaneous diagonalization. *SIAM J Matrix Anal Appl* **14**, 927–49.
6. Heylen W, Lammens S, Sas P (2003) *Modal analysis theory and testing*, Catholic Univ of Leuven, Leuven.
7. Pham DT (2001) Joint approximate diagonalization of positive definite matrices. *SIAM J Matrix Anal Appl* **22**, 1136–52.
8. Cardoso JF, Souloumiac A (1993) Blind beamforming for non–gaussian signals. *IEEE Proceedings F* **140**, 362–70.
9. Cardoso JF (1994) *Perturbation of joint diagonalizers*, Tech. Rep. 94D023, Signal Department, Telecom Paris, Paris.
10. Cardoso JF, Pham DT (2001) Blind separation of instantaneous mixtures of non stationary sources. *IEEE Trans Signal Process* **49**, 1837–48.
11. Afsari B, Krishnaprasad P (2004) Some gradient based joint diagonalization methods for ICA. In: *Proceedings of the Fifth International Conference on Independent Component Analysis and Blind Source Separation*, Lecture Notes in Computer Science, Springer, Berlin, pp 437–44.
12. Afsari B (2006) Simple LU and QR based non–orthogonal matrix joint diagonalization. In: *Proceedings of ICA2006*, Lecture Notes in Computer Science, Springer, Berlin, pp 1–7.
13. Afsari B (2007) What can make joint diagonalization difficult? In: *Proceedings of ICASSP07*, Vol.3, Honolulu, pp 1377–80.
14. Afsari B (2008) Sensitivity analysis for the problem of matrix joint diagonalization. *SIAM J Matrix Anal Appl* **30**, 1148–71.
15. Yeredor A (2002) Non–orthogonal joint diagonalization in the least–squares sense with application in blind source separation. *IEEE Trans Signal Process* **50**, 1545–53.
16. Yeredor A (2005) On using exact joint diagonalization for non–iterative approximate joint diagonalization. *IEEE Signal Process Lett* **12**, 645–8.
17. Helmke U, More JB (1994) *Optimization and dynamical systems*, Springer, Berlin.
18. Stewart GW, Guang Sun J (1990) *Matrix perturbation theory*, Academic Press, New York.
19. Robinson DJS (1982) *A course in the theory of groups*, Springer, Berlin.