No. 10.

10.1098/rspa.2001.0820

THE ROYAL SOCIETY

Taussky's theorem, symmetrizability and modal analysis revisited

BY ATUL BHASKAR

School of Engineering Sciences, Aeronautics and Astronautics, University of Southampton, Highfield, Southampton SO17 1BJ, UK

Received 24 February 2000; revised 9 March 2001; accepted 19 March 2001

This paper is concerned with symmetrization and diagonalization of real matrices and their implications for the dynamics of linear, second-order systems governed by equations of motion having *asymmetric* coefficient matrices. Results in the light of Taussky's theorem are presented. The connection of the symmetrizers with the eigenvalue problem is brought out. An alternative proof of Taussky's theorem for real matrices is presented. Diagonalization of two real symmetric (but not necessarily positive-definite) matrices is discussed in the context of undamped non-gyroscopic systems. A commutator of two matrices with respect to a given third matrix is defined; this commutator is found to play an interesting role in deciding simultaneous diagonalizability of two or three matrices. Errors in a few previously known results are brought out. Pseudo-conservative systems are studied and their connection with the so-called 'symmetrizable systems' is critically examined. Results for modal analysis of general non-conservative systems are presented. Illustrative examples are given.

Keywords: eigenvalue; normal nodes; diagonalization

1. Introduction

Taussky (1959, 1968) and Taussky & Zassenhaus (1959) proved a remarkable result that every square matrix (real or complex) is related to its transpose via a similarity transform brought about by a symmetric matrix. In this paper, we establish the connection of this similarity transform with the eigensolutions of the original matrix for the real case, extend a few results to a matrix pencil, and examine the implications of these results in the context of modal analysis of non-conservative systems. Consider a non-conservative system governed by

$$\boldsymbol{A}\ddot{\boldsymbol{q}} + \boldsymbol{B}\dot{\boldsymbol{q}} + \boldsymbol{C}\boldsymbol{q} = \boldsymbol{f}(t), \tag{1.1}$$

where $\boldsymbol{A}$, $\boldsymbol{B}$ and $\boldsymbol{C}$ are assumed to be non-defective $n \times n$ real square arrays (i.e. a sufficient number of eigenvectors exist) and they need not be symmetric. They may not have the usual interpretation of being the mass, the damping and the stiffness matrices, respectively. Whenever $\boldsymbol{A}^{-1}$ is used, it will be assumed that it exists. Asymmetric coefficient matrices appear in problems involving follower forces (see, for example, Bolotin 1963), gyroscopy, aero-/hydro-elasticity and control effects, etc.

While carrying out modal analysis of system (1.1) using a suitable *real transform*, we are interested in two questions.

(i) How can equation (1.1) be decoupled?

(ii) When can equation (1.1) be decoupled?

Concerning the first, there are two main results.

(1) Diagonalization using orthogonality with respect to the two real symmetric matrices (i.e. using a congruence transform) when $\boldsymbol{B} = \boldsymbol{0}$ (Thomson & Tait 1867, article 337; Rayleigh 1894, article 87), when either $\boldsymbol{A}$ or $\boldsymbol{C}$ is positive semidefinite. When $\boldsymbol{B} \neq \boldsymbol{0}$, simultaneous diagonalization of all three matrices may not be guaranteed (Caughey & O'Kelley 1965).

(2) Diagonalization using biorthogonality (i.e. using an equivalence transform) if the matrices are not symmetric (Lancaster 1966; Fawzy & Bishop 1976).

Regarding the conditions for simultaneous diagonalization, the following are the main contributions.

(1) Rayleigh (1894) realized that '... in terms of normal coordinates, T and V (the *kinetic* and the *potential* energies) are reduced to sums of squares', implying that two real symmetric matrices, of which one is positive semidefinite, can *always* be diagonalized simultaneously.

(2) Caughey & O'Kelley (1965) stated that commutativity of $\boldsymbol{A}^{-1}\boldsymbol{B}$ and $\boldsymbol{A}^{-1}\boldsymbol{C}$ is the required condition for simultaneous diagonalization when a symmetric system of matrices is considered and $\boldsymbol{A}$ is positive definite. This condition will be referred to as Caughey's condition hereafter. The corresponding normal modes are known as the classical normal modes.

(3) Liu & Wilson (1992) and Ma & Caughey (1995) obtained Caughey's condition when matrix $\boldsymbol{B}$ is assumed to be a general asymmetric real matrix, thus extending Caughey & O'Kelley's (1965) result. We shall observe later in this paper that this extension is not valid as a *necessary as well as sufficient condition* if we seek diagonalization by a *real* transformation.

(4) Rayleigh considered a special case of Caughey's condition: when one of the matrices is a linear combination of the other two (see Rayleigh (1894, article 97), 'T, V, F simultaneously reducible'). The corresponding damping model is known as the 'Rayleigh damping model'.

Vector $\boldsymbol{q}$ of the generalized coordinates, in equation (1.1), belongs to the so-called n-dimensional configuration space. Equation (1.1) can also be written as $\dot{\boldsymbol{s}} = \mathcal{A}\boldsymbol{s} + \boldsymbol{g}$ by defining the $2n$-dimensional state-vector $\boldsymbol{s} = [\boldsymbol{q}^{\mathrm{T}} \mid \dot{\boldsymbol{q}}^{\mathrm{T}}]^{\mathrm{T}}$. Eigenvalues of the $2n \times 2n$ state matrix $\mathcal{A}$ are called eigenvalues of the system. In the configuration space the λ-matrix (see, for example, Lancaster 1966) associated with equation (1.1) is $(\lambda^2\boldsymbol{A} + \lambda\boldsymbol{B} + \boldsymbol{C})$, whose non-trivial solutions are called the latent roots and the latent vectors, respectively. Working with the state-matrix has many advantages and this approach has been very popular with control theorists. However, we shall not pursue this matter here.

In the next section, Taussky's theorem is considered and its connection with the left and right eigenvectors is established. Extensions of this theorem to a pair of matrices is considered in § 3. The dynamical system $\boldsymbol{A}\ddot{\boldsymbol{q}}+\boldsymbol{C}\boldsymbol{q} = \boldsymbol{f}$ is then studied in the light of these results. A point to note is that the symmetric *generalized* eigenproblem $\boldsymbol{Au} = \beta\boldsymbol{Cu}$ does not necessarily admit real solutions, whereas the symmetric *standard* eigenproblem $\boldsymbol{Au} = \beta\boldsymbol{u}$ does. The general dynamical system $\boldsymbol{A}\ddot{\boldsymbol{q}} + \boldsymbol{B}\dot{\boldsymbol{q}} + \boldsymbol{Cq} = \boldsymbol{f}$ is taken up in § 4.

2. Taussky's theorem and eigenvectors of a real matrix

Taussky & Zassenhaus (1959) showed that for every square matrix $\boldsymbol{A}$, there exists a non-singular symmetric matrix $\boldsymbol{S}$ such that $\boldsymbol{A}^{\mathrm{T}} = \boldsymbol{S}^{-1}\boldsymbol{AS}$. If $\boldsymbol{A}$ is real, then $\boldsymbol{S}$ is real-symmetric and if $\boldsymbol{A}$ is complex, then $\boldsymbol{S}$ is complex-symmetric. The result, for real $\boldsymbol{A}$, can be presented in another way: for a real square matrix $\boldsymbol{A}$, there exists a non-singular real *symmetric* matrix $\boldsymbol{R} = \boldsymbol{S}^{-1}$ such that $\boldsymbol{RA}$ is *symmetric*. It will be shown here that the equivalence transform $\boldsymbol{V}^{\mathrm{T}}\boldsymbol{AU}$ over $\boldsymbol{A}$, $\boldsymbol{V}$ and $\boldsymbol{U}$, being the right and left eigenvectors arranged column-wise in a matrix, plays an interesting role in this context.

A proof of Taussky's theorem for real matrices based on eigenvectors is presented next. It is assumed throughout that the system is not defective and a full set of eigenvectors is available. The case of multiple eigenvalues (i.e. the case of *degenerate* systems) is not ruled out, since it is always possible to obtain a set of linearly independent eigenvectors for the case of non-defective but degenerate matrices. This means that the set of left as well as right eigenvectors spans $\mathbb{C}^n$ and, therefore, it is possible to introduce an invertible linear transform between them:

$$\boldsymbol{V} = \boldsymbol{X}^{\mathrm{T}}\boldsymbol{U} \quad \text{or} \quad \boldsymbol{X} = \boldsymbol{U}^{-\mathrm{T}}\boldsymbol{V}^{\mathrm{T}}. \tag{2.1}$$

Although $\boldsymbol{U}$ and $\boldsymbol{V}$ are non-singular, their real and imaginary parts are rank deficient. If $\boldsymbol{U}$ possesses p number of real eigenvectors and $2q$ number of complex conjugate eigenvectors, then the nullity of $\mathrm{Re}(\boldsymbol{U})$, the real part of $\boldsymbol{U}$, is q, due to q repeated columns for $2q$ complex conjugate eigenvectors. On the other hand, $\mathrm{Im}(\boldsymbol{U})$, the imaginary part of $\boldsymbol{U}$, has p columns of zeros (since there are p real eigenvectors) and $2q$ non-zero columns, of which half are repeated (after changing the signs), since they correspond to complex-conjugate pairs. Hence the rank of $\mathrm{Im}(\boldsymbol{U})$ is q. Since the eigenvalues of a matrix and its transpose are the same, the numbers of the real left eigenvectors and the complex-conjugate left eigenvectors are the same as those for the right eigenvectors. Therefore, similar statements follow for the real and the imaginary parts of $\boldsymbol{V}$. We can summarize these observations as

$$\left.\begin{aligned} \mathrm{rank}\{\mathrm{Re}(\boldsymbol{U})\} &= \mathrm{null}\{\mathrm{Im}(\boldsymbol{U})\} = \mathrm{rank}\{\mathrm{Re}(\boldsymbol{V})\} = \mathrm{null}\{\mathrm{Im}(\boldsymbol{V})\}, \\ \mathrm{rank}\{\mathrm{Re}(\boldsymbol{U})\} + \mathrm{rank}\{\mathrm{Im}(\boldsymbol{U})\} &= \mathrm{rank}\{\mathrm{Re}(\boldsymbol{V})\} + \mathrm{rank}\{\mathrm{Im}(\boldsymbol{V})\} \\ &= \mathrm{rank}(\boldsymbol{U}) = \mathrm{rank}(\boldsymbol{V}) = (p + 2q) = n. \end{aligned}\right\} \tag{2.2}$$

Let us construct two *real* $n \times n$ matrices $\tilde{\boldsymbol{U}}$ and $\tilde{\boldsymbol{V}}$ as follows

$$\left.\begin{aligned} \tilde{\boldsymbol{V}} &= [\boldsymbol{v}_1 \mid \boldsymbol{v}_2 \mid \cdots \boldsymbol{v}_p \mid \boldsymbol{v}_{p+1}^{\mathrm{R}} \mid \boldsymbol{v}_{p+1}^{\mathrm{I}} \mid \cdots \boldsymbol{v}_{p+q}^{\mathrm{R}} \mid \boldsymbol{v}_{p+q}^{\mathrm{I}}], \\ \tilde{\boldsymbol{U}} &= [\boldsymbol{u}_1 \mid \boldsymbol{u}_2 \mid \cdots \boldsymbol{u}_p \mid \boldsymbol{u}_{p+1}^{\mathrm{R}} \mid \boldsymbol{u}_{p+1}^{\mathrm{I}} \mid \cdots \boldsymbol{u}_{p+q}^{\mathrm{R}} \mid \boldsymbol{u}_{p+q}^{\mathrm{I}}], \end{aligned}\right\} \tag{2.3}$$

where $\boldsymbol{u}_i$ and $\boldsymbol{v}_i$ are the ith left and right eigenvectors and the superscripts refer to the real or the imaginary part. The first p eigenvectors of $\boldsymbol{A}$ are real and the remaining $2q = (n - p)$ are q pairs of complex-conjugates. $\tilde{\boldsymbol{U}}$ and $\tilde{\boldsymbol{V}}$ have their first

p columns as the real eigenvectors of $\boldsymbol{A}$, whereas the real and imaginary parts of the $2q$ complex-conjugate eigenvectors occupy the last $2q$ columns.

Since $\tilde{\boldsymbol{U}}$ and $\tilde{\boldsymbol{V}}$ are *real* square matrices, a real linear transform (say $\boldsymbol{T}_{\rm R}$) relates them through $\tilde{\boldsymbol{V}} = \boldsymbol{T}_{\rm R}\tilde{\boldsymbol{U}}$, i.e. $\tilde{\boldsymbol{v}}_j = \boldsymbol{T}_{\rm R}\tilde{\boldsymbol{u}}_j$, $j = 1, 2, \dots, n$. Since the transformation matrix is real, vectors superscripted R and I can be combined, and left and right eigenvectors $\boldsymbol{v}_j$ and $\boldsymbol{u}_j$ can be constructed as $(\boldsymbol{v}_s^{\rm R} \pm {\rm i}\boldsymbol{v}_s^{\rm I}) = \boldsymbol{T}_{\rm R}(\boldsymbol{u}_s^{\rm R} \pm {\rm i}\boldsymbol{u}_s^{\rm I})$, $s = (p+1), \dots, (p+q)$. For the first p eigenvectors $\boldsymbol{v}_k = \tilde{\boldsymbol{v}}_k$, $\boldsymbol{u}_k = \tilde{\boldsymbol{u}}_k$; and therefore the left and right eigenvectors are related via $\boldsymbol{v}_k = \boldsymbol{T}_{\rm R}\boldsymbol{u}_k$, $k = 1, 2, \dots, p$. Therefore, the real and the *complex* left and right eigenvectors are related through a common *real* transform $\boldsymbol{T}_{\rm R}$, i.e. $\boldsymbol{v}_k = \boldsymbol{T}_{\rm R}\boldsymbol{u}_k$, $k = 1, 2, \dots, n$; or $\boldsymbol{V} = \boldsymbol{T}_{\rm R}\boldsymbol{U}$. Comparing this with (2.1) we realize that $\boldsymbol{X}$ must be real and that $\boldsymbol{X}^{\rm T} = \boldsymbol{T}_{\rm R}$. Therefore, the equivalence transform $\boldsymbol{V}^{\rm T}\boldsymbol{A}\boldsymbol{U}$ takes the form $\boldsymbol{V}^{\rm T}\boldsymbol{A}\boldsymbol{U} = \boldsymbol{U}^{\rm T}\boldsymbol{T}_{\rm R}^{\rm T}\boldsymbol{A}\boldsymbol{U} = \boldsymbol{U}^{\rm T}\boldsymbol{X}\boldsymbol{A}\boldsymbol{U}$. The well-known biorthogonality relations between the left and right eigenvectors are given by $\boldsymbol{V}^{\rm T}\boldsymbol{U} = \boldsymbol{I}$ and $\boldsymbol{V}^{\rm T}\boldsymbol{A}\boldsymbol{U} = \boldsymbol{\Lambda}$, where $\boldsymbol{\Lambda}$ is the diagonal matrix of eigenvalues. The biorthogonality relation can now be rewritten as $\boldsymbol{U}^{\rm T}\boldsymbol{X}\boldsymbol{U} = \boldsymbol{I}$ and $\boldsymbol{U}^{\rm T}\boldsymbol{X}\boldsymbol{A}\boldsymbol{U} = \boldsymbol{\Lambda}$. Pre-multiplication by $\boldsymbol{U}^{-\rm T}$ and post-multiplication by $\boldsymbol{U}^{-1}$ implies that both $\boldsymbol{X}$ and $\boldsymbol{X}\boldsymbol{A}$ are symmetric. Since $\boldsymbol{X}\boldsymbol{A}$ and $\boldsymbol{X}$ are symmetric matrices, transposition leaves them unchanged, so that $\boldsymbol{X}\boldsymbol{A} = \boldsymbol{A}^{\rm T}\boldsymbol{X}^{\rm T} = \boldsymbol{A}^{\rm T}\boldsymbol{X}$ and, therefore, $\boldsymbol{A}^{\rm T} = \boldsymbol{X}\boldsymbol{A}\boldsymbol{X}^{-1}$. This proves Taussky's theorem. Taussky & Zassenhaus's (1959) original arguments did not involve the left or right eigenvectors. In a similar manner, $\boldsymbol{A}$ can also be symmetrized by post-multiplication by another real symmetric matrix, say $\boldsymbol{Y}$, by expressing each right eigenvector in terms of the left eigenvectors (details are omitted).

Definition 2.1. A linear mapping $\boldsymbol{X} : \mathbb{R}^{n\times n} \mapsto \mathbb{R}^{n\times n}$ is called a left symmetrizer to $\boldsymbol{A}$ if (i) $\boldsymbol{X}$ is non-singular, (ii) $\boldsymbol{X} = \boldsymbol{X}^{\rm T}$, and (iii) $(\boldsymbol{X}\boldsymbol{A}) = (\boldsymbol{X}\boldsymbol{A})^{\rm T}$. A linear mapping $\boldsymbol{Y} : \mathbb{R}^{n\times n} \mapsto \mathbb{R}^{n\times n}$ is called a right symmetrizer to $\boldsymbol{A}$ if (i) $\boldsymbol{Y}$ is non-singular, (ii) $\boldsymbol{Y} = \boldsymbol{Y}^{\rm T}$, and (iii) $(\boldsymbol{A}\boldsymbol{Y}) = (\boldsymbol{A}\boldsymbol{Y})^{\rm T}$.

That the symmetrizer is not unique, has been recognized by Taussky (1968). However, the procedure (procedure s) of finding them does (do) not appear to have been explored in the existing works of various authors. One of the aims of the present paper is to achieve this: equation (2.1) is the required relationship. Given definition 2.1, Taussky's theorem (over the field of reals) can be succinctly rephrased now.

Theorem 2.2. *For every real square matrix, there exists a real left (and a right) symmetrizer.*

Lemma 2.3. *The inverse of a left symmetrizer of a real square matrix $\boldsymbol{A}$ is a left symmetrizer of $\boldsymbol{A}^{\rm T}$ and the inverse of a right symmetrizer of a real square matrix $\boldsymbol{A}$ is a right symmetrizer of $\boldsymbol{A}^{\rm T}$.*

Lemma 2.4. *A left symmetrizer of a real square matrix $\boldsymbol{A}$ is a right symmetrizer of $\boldsymbol{A}^{\rm T}$ and a right symmetrizer of a real square matrix $\boldsymbol{A}$ is a left symmetrizer of $\boldsymbol{A}^{\rm T}$.*

Proofs of lemmas 2.3 and 2.4 follow easily from $\boldsymbol{A}^{\rm T} = \boldsymbol{X}\boldsymbol{A}\boldsymbol{X}^{-1}$. Equation (2.1) implies that for a right eigenvector $\boldsymbol{u}_i$, *there exists a left eigenvector which is a product of a left symmetrizer and the right eigenvector*, i.e. $\boldsymbol{v}_i = \boldsymbol{X}\boldsymbol{u}_i$. Using lemma 2.3, it is further observed that $\boldsymbol{u}_i = \boldsymbol{Y}\boldsymbol{v}_i$, where $\boldsymbol{Y}$ is a right symmetrizer. An example follows.

Example 2.5 (symmetrizer of a matrix with real eigensolutions). Consider a real asymmetric matrix

$$\boldsymbol{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}.$$

Eigenvectors of $\boldsymbol{A}$ and $\boldsymbol{A}^{\mathrm{T}}$ are calculated and arranged in matrices $\boldsymbol{U}$ and $\boldsymbol{V}$ ensuring that the sequence of eigenvectors in both matrices is consistent. A left symmetrizer is calculated according to $\boldsymbol{X} = \boldsymbol{U}^{-\mathrm{T}}\boldsymbol{V}^{\mathrm{T}} = \boldsymbol{X}^{\mathrm{T}} = \boldsymbol{V}\boldsymbol{U}^{-1}$ as

$$\boldsymbol{X} = \begin{bmatrix} 1.2014 & 0.1622 & 0.1229 \\ 0.1622 & 1.0650 & -0.0321 \\ 0.1229 & -0.0321 & 0.8128 \end{bmatrix},$$

$$\boldsymbol{X}\boldsymbol{A} = \begin{bmatrix} 2.7108 & 4.1973 & 5.6839 \\ 4.1973 & 5.3924 & 6.5875 \\ 5.6839 & 6.5975 & 7.4911 \end{bmatrix} = (\boldsymbol{X}\boldsymbol{A})^{\mathrm{T}}.$$

The inverse of $\boldsymbol{X}$ is a *right symmetrizer* of $\boldsymbol{A}$ due to lemmas 2.3 and 2.4:

$$\boldsymbol{Y} = \boldsymbol{X}^{-1} = \begin{bmatrix} 0.8646 & -0.1358 & -0.1362 \\ -0.1358 & 0.9614 & 0.0585 \\ -0.1362 & 0.0585 & 1.2532 \end{bmatrix},$$

$$\boldsymbol{A}\boldsymbol{Y} = \begin{bmatrix} 0.1846 & 1.9626 & 3.7407 \\ 1.9626 & 4.6151 & 7.2676 \\ 3.7407 & 7.2676 & 10.7946 \end{bmatrix} = (\boldsymbol{A}\boldsymbol{Y})^{\mathrm{T}}.$$

Matrices of left and right eigenvectors are unique only up to independent scaling of the vectors and their permutations. What is interesting in Taussky's theorem is the fact that the complex left and right eigenvectors are related via a *real symmetric* transform.

Example 2.6 (symmetrizer of a matrix having complex eigensolutions). The right and the left eigenvectors of

$$\boldsymbol{A} = \begin{bmatrix} 1 & 2 & -1 \\ 3 & 2 & 1 \\ 4 & -2 & 2 \end{bmatrix}$$

are calculated as

$$\boldsymbol{U} = \begin{bmatrix} -0.2653 + 0.2337\mathrm{i} & -0.2653 - 0.2337\mathrm{i} & 0.5071 \\ 0.2653 - 0.2337\mathrm{i} & 0.2653 + 0.2337\mathrm{i} & 0.8452 \\ 0.8505 + 0.1632\mathrm{i} & 0.8505 - 0.1632\mathrm{i} & 0.1690 \end{bmatrix},$$

$$\boldsymbol{V} = \begin{bmatrix} 0.7162 + 0.0474\mathrm{i} & 0.7162 - 0.0474\mathrm{i} & 0.7071 \\ -0.4791 - 0.1136\mathrm{i} & -0.4791 + 0.1136\mathrm{i} & 0.7071 \\ 0.2468 + 0.4260\mathrm{i} & 0.2468 - 0.4260\mathrm{i} & 0 \end{bmatrix},$$

respectively. The two complex matrices are related via a real symmetric transform

$$\boldsymbol{X} = \boldsymbol{U}^{-\mathrm{T}}\boldsymbol{V}^{\mathrm{T}} = \begin{bmatrix} 0.2324 & 0.5486 & 0.7434 \\ 0.5486 & 0.6249 & -0.5871 \\ 0.7434 & -0.5871 & 0.7052 \end{bmatrix} = \boldsymbol{X}^{\mathrm{T}}.$$

This transformation matrix is a left symmetrizer of $\boldsymbol{A}$, since

$$\boldsymbol{XA} = \begin{bmatrix} 4.8517 & 0.0750 & 1.8030 \\ 0.0750 & 3.5212 & -1.0978 \\ 1.8030 & -1.0978 & 0.0799 \end{bmatrix} = (\boldsymbol{XA})^{\mathrm{T}}.$$

The question of uniqueness of symmetrizers is examined now. Suppose $\boldsymbol{X}$ is a left symmetrizer. If there exists another symmetrizer, say, $\tilde{\boldsymbol{X}}$, then the non-singularity of both $\boldsymbol{X}$ and $\tilde{\boldsymbol{X}}$ implies that the columns (or rows) of either of the two matrices span $\mathbb{R}^n$. Therefore, $\boldsymbol{X}$ and $\tilde{\boldsymbol{X}}$ must be related via a non-singular linear transform (say, $\boldsymbol{T}$), i.e. $\tilde{\boldsymbol{X}} = \boldsymbol{XT}$. Since $\tilde{\boldsymbol{X}}$ and $\boldsymbol{X}$ are symmetric matrices,

$$\tilde{\boldsymbol{X}} = \tilde{\boldsymbol{X}}^{\mathrm{T}} \Rightarrow (\boldsymbol{XT}) = \boldsymbol{T}^{\mathrm{T}}\boldsymbol{X} \Rightarrow \boldsymbol{XTA} = \boldsymbol{T}^{\mathrm{T}}\boldsymbol{XA}.$$

Since $\tilde{\boldsymbol{X}}\boldsymbol{A}$ is also symmetric, we have

$$(\tilde{\boldsymbol{X}}\boldsymbol{A}) = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{T}^{\mathrm{T}}\boldsymbol{X}, \quad \text{i.e. } \boldsymbol{XTA} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{T}^{\mathrm{T}}\boldsymbol{X}. \tag{2.4}$$

Therefore, $\boldsymbol{T}^{\mathrm{T}}\boldsymbol{XA} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{T}^{\mathrm{T}}\boldsymbol{X}$. Since $\boldsymbol{X}$ is non-singular, appropriate pre-multiplication and post-multiplication results in $\boldsymbol{XAX}^{-1} = \boldsymbol{T}^{-\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{T}^{\mathrm{T}}$. But $\boldsymbol{XAX}^{-1} = \boldsymbol{A}^{\mathrm{T}}$, and, therefore, $\boldsymbol{TA} = \boldsymbol{AT}$. On similar lines, the set of right symmetrizers can also be handled. In the following, we use the notation for the commutator of two matrices $\boldsymbol{A}_1$ and $\boldsymbol{A}_2$ as $[\boldsymbol{A}_1, \boldsymbol{A}_2]$, which is given by $\boldsymbol{A}_1\boldsymbol{A}_2 - \boldsymbol{A}_2\boldsymbol{A}_1$.

Theorem 2.7. *If $\boldsymbol{X}$ is a left symmetrizer to $\boldsymbol{A}$, then so is $\tilde{\boldsymbol{X}} = \boldsymbol{XT}$, where $\boldsymbol{T}$ is a linear non-singular mapping such that $[\boldsymbol{T}, \boldsymbol{A}] = \boldsymbol{TA} - \boldsymbol{AT} = \boldsymbol{0}$. If $\boldsymbol{Y}$ is a right symmetrizer to $\boldsymbol{A}$, then so is $\tilde{\boldsymbol{Y}} = \boldsymbol{TY}$, where $\boldsymbol{T}$ is a linear non-singular mapping such that $[\boldsymbol{T}, \boldsymbol{A}] = \boldsymbol{TA} - \boldsymbol{AT} = \boldsymbol{0}$.*

Thus, the set of *all* the symmetrizers $\boldsymbol{X}_i$, $\boldsymbol{Y}_i$ can be generated by solving the commutator equation: $[\boldsymbol{T}_i, \boldsymbol{A}] = \boldsymbol{0}$ for $\boldsymbol{T}_i$ and substituting into $\boldsymbol{X}_i = \boldsymbol{XT}_i$ and $\boldsymbol{Y}_i = \boldsymbol{T}_i\boldsymbol{Y}$.

A more impressive statement of Taussky's theorem for real matrices is that every real matrix can be factorized into *two real symmetric factors* (Taussky 1968). One of these two factors can be identified with a symmetrizer (as defined via the left and the right eigenvectors) or its inverse, since

$$\boldsymbol{A} = \boldsymbol{SA}^{\mathrm{T}}\boldsymbol{S}^{-1} = (\boldsymbol{X}^{-1}\boldsymbol{A}^{\mathrm{T}})(\boldsymbol{X}) = (\boldsymbol{X}^{-1})(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{X}),$$

where each term inside the parentheses is symmetric. Symmetric factorization of a real matrix is, thus, a known fact; its relationship with eigenvectors, however, seems to have been missing in the literature. There have been attempts to obtain symmetric factors of a matrix (see, for example, Inman 1983, appendix). Inman (1983) constructs 'the symmetric factors' by solving simultaneous equations by imposing the condition of symmetry on the factors in the product. Ahmadian & Chou (1987) present a fairly involved procedure of calculating symmetric factors of a real matrix. The approach here reduces this problem to solving the eigenproblem of the matrix and that of its adjoint, and obtaining the linear transform between the left and the right eigenvectors. Since symmetrizers are not unique, the possible factors of a matrix into a product of symmetric matrices is not unique. Sen & Venkaiah (1988*a*, *b*) and Venkaiah & Sen (1988) have approached the problem of computing symmetrizers via algorithms that operate directly on the rows and columns of the matrix in question.

The set of mappings $\{\boldsymbol{T}\} = \{\boldsymbol{T} : [\boldsymbol{T}, \boldsymbol{A}] = \boldsymbol{0}\}$ and the set of symmetrizers $\{\boldsymbol{X}\} = \{\boldsymbol{X} : \boldsymbol{XA} = (\boldsymbol{XA})^{\mathrm{T}}\}$ can be shown to have the following mathematical structure:

(i) the set of all linear mappings $\{\boldsymbol{T}\}$ which map a given left symmetrizer $\boldsymbol{X}$ to another left symmetrizer $\tilde{\boldsymbol{X}} = \boldsymbol{XT}$, together with $\boldsymbol{0}$ constitute a ring;

(ii) the set of all non-singular linear mappings $\{\boldsymbol{T}\}$ constitutes a multiplicative group; and, finally,

(iii) the set of symmetrizers $\boldsymbol{X}$ constitutes a commutative semigroup with respect to the addition operation.

This is true because the set is non-empty due to Taussky's theorem; and any non-zero linear combination of two symmetrizers is also a symmetrizer, hence the set is closed under addition. Elementary group definitions may be found, for example, in Valenza (1993, ch. 2).

3. Undamped non-gyroscopic systems: two real square arrays or a matrix pencil

Ma & Caughey (1995) have presented a method of decoupling the governing equations of vibratory motion based on an equivalence transform when the coefficient matrices are taken as real square arrays without any assumption regarding symmetry or definiteness. Their concern appears to be 'decoupling' equations of motion in the configuration space by the use of *a real or a complex* transform. This method is essentially the same as that of Fawzy & Bishop (1976). While decoupling equations in the configuration space, it is often implicitly assumed that the transformation involves real matrices.

Consider the case when velocity terms vanish, i.e. $\boldsymbol{B}$ is a null matrix. The use of equivalence transforms (or biorthogonality) $\boldsymbol{V}^{\mathrm{T}}\boldsymbol{AU} = \boldsymbol{I}$ and $\boldsymbol{V}^{\mathrm{T}}\boldsymbol{CU} = \boldsymbol{\Lambda}$ has been well known in this context (see, for example, Lancaster (1966, theorem 2.1), Fawzy & Bishop (1976), Wahed & Bishop (1976), Newland (1987) and Meirovitch (1980, §§ 6.7, 6.8) for simultaneous diagonalization in the state space). On this basis Ma & Caughey (1995) obtained a result: 'an undamped non-gyroscopic system that is not degenerate or defective can always be decoupled by equivalence transformation'. *If* $\boldsymbol{U}$ and $\boldsymbol{V}$ used for the equivalence transforms are restricted to be in the real space, then the claim of the result regarding decoupling becomes invalid.

While achieving 'decoupling' of coordinates in a manner presented by Ma & Caughey (1995), one needs to transform the coordinates according to $\boldsymbol{q} = \boldsymbol{Up}$, and, according to them, the transformed equation $\ddot{\boldsymbol{p}} + \boldsymbol{D}\dot{\boldsymbol{p}} = \boldsymbol{V}^{\mathrm{T}}\boldsymbol{f}(t) = \boldsymbol{P}(t)$ 'represents a completely decoupled system'. Complete decoupling must be interpreted as possession of classical normal modes in the sense of Caughey (1960) and Caughey & O'Kelley (1965). This necessarily means decoupling by a *real* transformation.

Generalized coordinates $\boldsymbol{p}$ are in the *configuration space* and they remain coupled in terms of the real variables. A typical 'decoupled' complex equation takes the form: $\ddot{p}_j + d_j p_j = P_j$. Note that $\boldsymbol{q}$ is real but $\boldsymbol{p}$ and $\boldsymbol{D}$ are, in general, complex. Separating the real and the imaginary parts of the variables and the parameters as $p_j = \eta_j + \mathrm{i}\phi_j$, $d_j = \Sigma_j + \mathrm{i}\Omega_j$ and $P_j = r_j + \mathrm{i}s_j$, each 'decoupled' complex equation becomes

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{Bmatrix} \ddot{\eta}_j \\ \ddot{\phi}_j \end{Bmatrix} + \begin{bmatrix} \Sigma_j & -\Omega_j \\ \Omega_j & \Sigma_j \end{bmatrix} \begin{Bmatrix} \eta_j \\ \phi_j \end{Bmatrix} = \begin{Bmatrix} r_j \\ s_j \end{Bmatrix}. \tag{3.1}$$

This equation represents a set of *coupled* ordinary differential equations in real variables η_j and ϕ_j. Eigenvalues of this system are given by $\gamma_j^{(1,2)} = \Sigma_j \pm \mathrm{i}\Omega_j$. Therefore, decoupling achieved via $\ddot{p}_j + d_j p_j = P_j$ is only that of appearance. The requirement of non-degeneracy in the theorem of Ma & Caughey (1995) is unnecessary, since biorthogonality conditions exist for degenerate systems too. Of course, this is not true of a defective system. If the d_j are real, the solution of $\ddot{p}_j + d_j p_j = P_j$ is immediate. Otherwise, calculations are considerably involved.

Equations (3.1) can be cast in the first-order form in a four-dimensional state-space spanned by the variables η_j, $\dot{\eta}_j$, ϕ_j and $\dot{\phi}_j$. This means solving n problems of size 4×4 using standard state-space methods. Alternatively, the *complex* second-order differential equation can be solved directly by the modal summation:

$$\boldsymbol{q}(t) = \sum_{r=1}^{l} [\boldsymbol{u}_r p_r(t) + \bar{\boldsymbol{u}}_r \bar{p}_r(t)] + \sum_{r=2l+1}^{n} [\boldsymbol{u}_r p_r(t)]. \tag{3.2}$$

Here, the $\boldsymbol{u}_r$ represent columns of $\boldsymbol{U}$. It has been assumed in (3.2) that there are $2l$ number of complex conjugate eigenvalues and the remaining $(n-2l)$ are real (positive or negative). The modal solution is constructed as

$$\begin{aligned}\boldsymbol{q}(t) = 2\sum_{r=1}^{l} \Re\Bigg[\int_0^t P_r(t-\tau)\left\{\frac{\exp(\mathrm{i}\mu_r\tau) - \exp(-\mathrm{i}\mu_r\tau)}{2\mathrm{i}\mu_r}\right\}\mathrm{d}\tau \\ + \left\{\frac{\exp(\mathrm{i}\mu_r t) + \exp(-\mathrm{i}\mu_r t)}{2}\right\} p_r(0) \\ + \left\{\frac{\exp(\mathrm{i}\mu_r t) - \exp(-\mathrm{i}\mu_r t)}{2\mathrm{i}\mu_r}\right\}\dot{p}_r(0)\Bigg]\boldsymbol{u}_r + R,\end{aligned} \tag{3.3}$$

where the last term R accounts for contributions due to classical (real) uncoupled normal modes. Familiar solutions in terms of sines and cosines result when the d_r are all positive; whereas overdamped and divergence modes contribute to R if any d_r is negative.

(a) *Symmetrization and diagonalization of a real matrix pencil*

The statement of Taussky's theorem for a pair of matrices is generalized now in the following theorem.

Theorem 3.1. *Given two real non-defective square matrices* $\boldsymbol{A}$ *and* $\boldsymbol{B}$ *of the same size, there exists a real matrix* $\boldsymbol{X}$ *such that* $(\boldsymbol{XA}) = (\boldsymbol{XA})^{\mathrm{T}}$ *and* $(\boldsymbol{XB}) = (\boldsymbol{XB})^{\mathrm{T}}$, *and there exists a real matrix* $\boldsymbol{Y}$ *such that* $(\boldsymbol{AY}) = (\boldsymbol{AY})^{\mathrm{T}}$ *and* $(\boldsymbol{BY}) = (\boldsymbol{BY})^{\mathrm{T}}$.

Proof. Biorthogonality relations for the eigenproblems $\boldsymbol{A}\boldsymbol{u}_i = \gamma_i \boldsymbol{B}\boldsymbol{u}_i$ and $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{v}_j = \gamma_j \boldsymbol{B}\boldsymbol{v}_j$ become $\boldsymbol{V}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{U} = \boldsymbol{D}_A$ and $\boldsymbol{V}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{U} = \boldsymbol{D}_B$, where $\boldsymbol{D}_A$ and $\boldsymbol{D}_B$ are diagonal matrices. Since $\boldsymbol{A}$ and $\boldsymbol{B}$ are non-defective, a full set of linearly independent eigenvectors can always be found and, therefore, $\boldsymbol{U}$ and $\boldsymbol{V}$ are taken to be non-singular. Define a matrix $\boldsymbol{X} = \boldsymbol{U}^{-\mathrm{T}}\boldsymbol{V}^{\mathrm{T}}$ so that the products

$$\left.\begin{aligned}\boldsymbol{XA} &= \boldsymbol{U}^{-\mathrm{T}}\boldsymbol{V}^{\mathrm{T}}\boldsymbol{V}^{-\mathrm{T}}\boldsymbol{D}_A\boldsymbol{U}^{-1} = \boldsymbol{U}^{-\mathrm{T}}\boldsymbol{D}_A\boldsymbol{U}^{-1}, \\ \boldsymbol{XB} &= \boldsymbol{U}^{-\mathrm{T}}\boldsymbol{V}^{\mathrm{T}}\boldsymbol{V}^{-\mathrm{T}}\boldsymbol{D}_B\boldsymbol{U}^{-1} = \boldsymbol{U}^{-\mathrm{T}}\boldsymbol{D}_B\boldsymbol{U}^{-1}\end{aligned}\right\} \tag{3.4}$$

are symmetric, since $\boldsymbol{D}_A$ and $\boldsymbol{D}_B$ are diagonal. If $\boldsymbol{U}$ and $\boldsymbol{V}$ are real matrices, then the matrix $\boldsymbol{X} = \boldsymbol{U}^{-\mathrm{T}}\boldsymbol{V}^{\mathrm{T}}$ is real and the proof is complete. If, on the other hand, there exists at least one complex eigenvector, then $\boldsymbol{X}$ as defined here may be complex. In that case, matrices $\boldsymbol{XA}$ and $\boldsymbol{XB}$ are complex *symmetric* matrices: symmetry being guaranteed due to equation (3.4). Real and imaginary parts of $\boldsymbol{X}$ can be separated as $\boldsymbol{X} = \boldsymbol{X}_{\mathrm{R}} + \mathrm{i}\boldsymbol{X}_{\mathrm{I}}$ so that $\boldsymbol{XA} = \boldsymbol{X}_{\mathrm{R}}\boldsymbol{A} + \mathrm{i}\boldsymbol{X}_{\mathrm{I}}\boldsymbol{A}$ and $\boldsymbol{XB} = \boldsymbol{X}_{\mathrm{R}}\boldsymbol{B} + \mathrm{i}\boldsymbol{X}_{\mathrm{I}}\boldsymbol{B}$.

Since $\boldsymbol{XA}$ and $\boldsymbol{XB}$ are complex symmetric, the corresponding real and imaginary parts must also be symmetric, and therefore $\boldsymbol{X}_{\mathrm{R}}\boldsymbol{A} = (\boldsymbol{X}_{\mathrm{R}}\boldsymbol{A})^{\mathrm{T}}$, $\boldsymbol{X}_{\mathrm{R}}\boldsymbol{B} = (\boldsymbol{X}_{\mathrm{R}}\boldsymbol{B})^{\mathrm{T}}$, $\boldsymbol{X}_{\mathrm{I}}\boldsymbol{A} = (\boldsymbol{X}_{\mathrm{I}}\boldsymbol{A})^{\mathrm{T}}$ and $\boldsymbol{X}_{\mathrm{I}}\boldsymbol{B} = (\boldsymbol{X}_{\mathrm{I}}\boldsymbol{B})^{\mathrm{T}}$. Hence, either $\boldsymbol{X}_{\mathrm{R}}$ or $\boldsymbol{X}_{\mathrm{I}}$ can play the role of $\boldsymbol{X}$ in the statement of the theorem, which proves the proposition. ■

Based on this result, we propose the following definitions for symmetrizers of a matrix pencil: the family of matrices $(\boldsymbol{A} + a\boldsymbol{B})$ generated by changing the value of the parameter a. For convenience of notation, we shall denote a pencil by the ordered pair $(\boldsymbol{A}, \boldsymbol{B})$.

Definition 3.2. A linear mapping $\boldsymbol{X} : \mathbb{R}^{n\times n} \mapsto \mathbb{R}^{n\times n}$ is called a left symmetrizer of a real matrix pencil $(\boldsymbol{A}, \boldsymbol{B})$ if (i) $\boldsymbol{X}$ is non-singular, (ii) $(\boldsymbol{XA}) = (\boldsymbol{XA})^{\mathrm{T}}$, and (iii) $(\boldsymbol{XB}) = (\boldsymbol{XB})^{\mathrm{T}}$. A linear mapping $\boldsymbol{Y} : \mathbb{R}^{n\times n} \mapsto \mathbb{R}^{n\times n}$ is called a right symmetrizer of a real matrix pencil $(\boldsymbol{A}, \boldsymbol{B})$ if (i) $\boldsymbol{Y}$ is non-singular, (ii) $(\boldsymbol{AY}) = (\boldsymbol{AY})^{\mathrm{T}}$, and (iii) $(\boldsymbol{BY}) = (\boldsymbol{BY})^{\mathrm{T}}$.

Therefore, *for every real, non-defective matrix pencil* $(\boldsymbol{A} + \mu\boldsymbol{B})$, *there exists a real matrix* $\boldsymbol{X}$ *such that* $\boldsymbol{X}(\boldsymbol{A} + \mu\boldsymbol{B})\boldsymbol{X}^{-\mathrm{T}} = (\boldsymbol{A}^{\mathrm{T}} + \mu\boldsymbol{B}^{\mathrm{T}})$ *for all values of* μ. It means, then, that every generalized eigenvalue problem involving two real matrices can be rendered self-adjoint by pre-multiplication with a suitable matrix. An interesting case for complex matrices is obtained: *for every complex matrix* $\boldsymbol{Z} = \boldsymbol{A} + \mathrm{i}\boldsymbol{B}$; $\boldsymbol{A}$, $\boldsymbol{B}$ *non-defective, there exist* real *matrices* $\boldsymbol{X}$ *and* $\boldsymbol{Y}$ *such that* $\boldsymbol{XZ} = (\boldsymbol{XZ})^{\mathrm{T}}$ *and* $\boldsymbol{ZY} = (\boldsymbol{ZY})^{\mathrm{T}}$. Note that the pre-multiplier $\boldsymbol{X}$ or a post-multiplier $\boldsymbol{Y}$ required to symmetrize the complex matrix $\boldsymbol{Z}$ may not be symmetric and, therefore, this statement is different from Taussky's theorem for a single complex matrix. Taussky's theorem for complex matrices asserts that there exists a *complex symmetric matrix* such that its product with a given complex matrix is complex. On the other hand, $\boldsymbol{X}$ in definition 3.2 is *real*, but not necessarily symmetric. The origin of lack of symmetry can be traced to the orthogonality relations of the left and right eigenvector matrices. In the case of two matrices (or a matrix pencil), the biorthogonality relations are *with respect to* matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ and, therefore, $\boldsymbol{X} = (\boldsymbol{U}^{-\mathrm{T}}\boldsymbol{D}_A\boldsymbol{U}^{-1})\boldsymbol{A}^{-1}$. The expression inside the parentheses is indeed symmetric, but a post-multiplication by $\boldsymbol{A}^{-1}$ destroys this symmetry.

All of the above discussions hold true for a post-multiplier $\boldsymbol{Y}$ which renders both $\boldsymbol{A}$ and $\boldsymbol{B}$ symmetric. In that case, matrix $\boldsymbol{Y}$ is defined as $\boldsymbol{Y} = \boldsymbol{UV}^{-1}$. Other results similar to the ones presented here in terms of a left symmetrizer matrix $\boldsymbol{X}$ can be extended easily to the ones involving a right symmetrizer matrix $\boldsymbol{Y}$. It is interesting to note that results on the lines of lemmas 2.3 and 2.4 do *not* hold for symmetrizers of a matrix pencil. The left symmetrizer and the right symmetrizer of a matrix pencil are not related via an inverse. This contrasting situation, with respect to the case of symmetrizers of a single real matrix, is due to asymmetry of the symmetrizers of a matrix pencil. The expression for a right symmetrizer of a matrix pencil $\boldsymbol{Y} = \boldsymbol{UV}^{-1}$ when combined with $\boldsymbol{X} = \boldsymbol{U}^{-\mathrm{T}}\boldsymbol{V}^{\mathrm{T}}$ leads to $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{Y} = \boldsymbol{I}$.

Example 3.3 (symmetrizer of a real matrix pencil). A symmetrizer of the non-defective *matrix pencil* $(\boldsymbol{A}, \boldsymbol{C})$, where

$$\boldsymbol{A} = \begin{bmatrix} 1 & 3 & 2 \\ 4 & 2 & 3 \\ 3 & 1 & 4 \end{bmatrix} \quad \text{and} \quad \boldsymbol{C} = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 1 & 2 \\ 1 & 1 & 2 \end{bmatrix}$$

is calculated using the procedure outlined above as

$$\boldsymbol{X} = \begin{bmatrix} 0.8183 - 0.1573\mathrm{i} & 0.7573 - 0.1841\mathrm{i} & -0.8993 - 0.3953\mathrm{i} \\ -0.1116 - 0.5660\mathrm{i} & 0.8183 - 0.1573\mathrm{i} & -0.0305 - 0.0134\mathrm{i} \\ 0.7878 - 0.1707\mathrm{i} & -0.1421 - 0.5794\mathrm{i} & 0.0305 + 0.0134\mathrm{i} \end{bmatrix}.$$

Separating the real and the imaginary parts of $\boldsymbol{X}$ according to $\boldsymbol{X}_{\mathrm{R}} = \mathrm{Re}(\boldsymbol{X})$ and $\boldsymbol{X}_{\mathrm{I}} = \mathrm{Im}(\boldsymbol{X})$, we obtain a pair of real symmetrizers to the pencil, since

$$\left.\begin{aligned} \boldsymbol{X}_{\mathrm{R}}\boldsymbol{A} &= \begin{bmatrix} 1.1494 & 3.0701 & 0.3111 \\ 3.0701 & 1.2714 & 2.1098 \\ 0.3111 & 2.1098 & 1.2714 \end{bmatrix}, \\ \boldsymbol{X}_{\mathrm{R}}\boldsymbol{C} &= \begin{bmatrix} 1.4335 & 1.4945 & 0.5342 \\ 1.4945 & 0.5647 & 1.4640 \\ 0.5342 & 1.4640 & 0.5647 \end{bmatrix}, \\ \boldsymbol{X}_{\mathrm{I}}\boldsymbol{A} &= \begin{bmatrix} -2.0796 & -1.2354 & -2.4480 \\ -1.2354 & -2.0260 & -1.6575 \\ -2.4480 & -1.6575 & -2.0260 \end{bmatrix}, \\ \boldsymbol{X}_{\mathrm{I}}\boldsymbol{C} &= \begin{bmatrix} -0.9208 & -0.8940 & -1.3161 \\ -0.8940 & -1.3027 & -0.9074 \\ -1.3161 & -0.9074 & -1.3027 \end{bmatrix} \end{aligned}\right\} \tag{3.5}$$

are all symmetric. The right symmetrizer $\boldsymbol{Y} = \boldsymbol{U}\boldsymbol{V}^{-1}$ is calculated as:

$$\boldsymbol{Y} = \begin{bmatrix} 0.0305 - 0.0134\mathrm{i} & -0.0305 + 0.0134\mathrm{i} & -0.9298 + 0.4087\mathrm{i} \\ -0.1268 + 0.5727\mathrm{i} & 0.8030 + 0.1640\mathrm{i} & 0.7725 + 0.1774\mathrm{i} \\ 0.7725 + 0.1774\mathrm{i} & -0.0963 + 0.5593\mathrm{i} & 0.8335 + 0.1506\mathrm{i} \end{bmatrix}.$$

Real and imaginary parts of $\boldsymbol{Y}$ can be separated as $\boldsymbol{Y} = \boldsymbol{Y}_{\mathrm{R}} + \mathrm{i}\boldsymbol{Y}_{\mathrm{I}}$. Matrices $\boldsymbol{A}$ and $\boldsymbol{C}$ can now be simultaneously symmetrized by post-multiplication either by $\boldsymbol{Y}_{\mathrm{R}}$ or by $\boldsymbol{Y}_{\mathrm{I}}$:

$$\left.\begin{aligned} \boldsymbol{A}\boldsymbol{Y}_{\mathrm{R}} &= \begin{bmatrix} 1.1952 & 2.1860 & 3.0548 \\ 2.1860 & 1.1952 & 0.3263 \\ 3.0548 & 0.3263 & 1.3171 \end{bmatrix}, \\ \boldsymbol{C}\boldsymbol{Y}_{\mathrm{R}} &= \begin{bmatrix} 0.5494 & 1.4793 & 1.4488 \\ 1.4793 & 0.5494 & 0.5799 \\ 1.4488 & 0.5799 & 1.5098 \end{bmatrix}, \end{aligned}\right\} \tag{3.6}$$

$$\left.\begin{aligned} \boldsymbol{A}Y_{\mathrm{I}} &= \begin{bmatrix} 2.0595 & 1.6240 & 1.2421 \\ 1.6240 & 2.0595 & 2.4413 \\ 1.2421 & 2.4413 & 2.0058 \end{bmatrix}, \\ \boldsymbol{C}Y_{\mathrm{I}} &= \begin{bmatrix} 1.3094 & 0.9007 & 0.9141 \\ 0.9007 & 1.3094 & 1.2960 \\ 1.9141 & 1.2960 & 0.8873 \end{bmatrix}. \end{aligned}\right\} \tag{3.6 cont.}$$

While it is true as observed above that every non-defective pair of real matrices can be rendered symmetric by a suitable pre-multiplication (and post-multiplication), simultaneous diagonalization by a real transformation is not always guaranteed. It is a common misconception that Rayleigh (1894) proved that two real symmetric matrices can always be rendered diagonal by the use of an appropriate congruence transform. A simple counterexample is enough to show that this is not true. Consider two real, symmetric, non-singular matrices

$$\boldsymbol{A} = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}, \qquad \boldsymbol{C} = \begin{bmatrix} 2 & 5 \\ 5 & 3 \end{bmatrix}.$$

It can be checked that this pair of matrices can never be simultaneously diagonalized by a real congruence transform (the sort of matrices used in 'classical modal analysis'), because the eigensolutions of the problem $\boldsymbol{A}\boldsymbol{u} = \beta\boldsymbol{C}\boldsymbol{u}$ are complex. Rayleigh, of course, did not claim this often wrongly assumed result, because he clearly stated that the two matrices need to be positive semidefinite.

A celebrated result about simultaneous diagonalization is that (see, for example, Bellman 1960, pp. 56–57) two real symmetric matrices can be simultaneously diagonalized by a real *orthogonal* transform if, and only if, they commute in multiplication. This result is generalized in theorem 3.5 to transforms that need not be orthogonal. A definition is presented first.

Definition 3.4. The matrix commutator of $\boldsymbol{A}$ and $\boldsymbol{B}$ with respect to $\boldsymbol{P}$, denoted by $[\boldsymbol{A},\boldsymbol{B}]_{\boldsymbol{P}}$, is defined as $[\boldsymbol{A},\boldsymbol{B}]_{\boldsymbol{P}} = \boldsymbol{APB} - \boldsymbol{BPA}$.

The commutator $[\boldsymbol{A},\boldsymbol{B}]_{\boldsymbol{P}}$ as defined here is a Lie product of $\boldsymbol{A}$ and $\boldsymbol{B}$ (*with respect to* $\boldsymbol{P}$), since for a scalar p, the following identities hold for the operation defined here:

(i) $[p(\boldsymbol{A}+\boldsymbol{B}),\boldsymbol{C}]_{\boldsymbol{P}} = p[\boldsymbol{A},\boldsymbol{C}]_{\boldsymbol{P}} + p[\boldsymbol{B},\boldsymbol{C}]_{\boldsymbol{P}}$ (linearity);

(ii) $[\boldsymbol{A},\boldsymbol{B}]_{\boldsymbol{P}} = -[\boldsymbol{B},\boldsymbol{A}]_{\boldsymbol{P}}$ (skew-symmetry); and

(iii) $[\boldsymbol{A},[\boldsymbol{B},\boldsymbol{C}]_{\boldsymbol{P}}]_{\boldsymbol{P}} + [\boldsymbol{B},[\boldsymbol{C},\boldsymbol{A}]_{\boldsymbol{P}}]_{\boldsymbol{P}} + [\boldsymbol{C},[\boldsymbol{A},\boldsymbol{B}]_{\boldsymbol{P}}]_{\boldsymbol{P}} = \boldsymbol{0}$ (Jacobi identity).

Theorem 3.5. *Two non-defective symmetric matrices* $\boldsymbol{A}$ *and* $\boldsymbol{B}$ *can be simultaneously diagonalized by a congruence transform brought about by a real non-singular matrix if, and only if, there exists a real symmetric positive-definite matrix* $\boldsymbol{P}$ *such that* $[\boldsymbol{A},\boldsymbol{B}]_{\boldsymbol{P}} = \boldsymbol{0}$, *i.e.* $\boldsymbol{APB} = \boldsymbol{BPA}$.

Proof. (i) *The 'if' part.* Given $\boldsymbol{P} = \boldsymbol{P}^{\mathrm{T}} > \boldsymbol{0}$, there exists a $\boldsymbol{Q}$ such that $\boldsymbol{P} = \boldsymbol{Q}\boldsymbol{Q}^{\mathrm{T}}$, i.e. $\boldsymbol{A}\boldsymbol{Q}\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{B} = \boldsymbol{B}\boldsymbol{Q}\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{A}$ holds for some $\boldsymbol{Q}$. Pre-multiplying by $\boldsymbol{Q}^{\mathrm{T}}$ and post-multiplying by $\boldsymbol{Q}$, we have $\tilde{\boldsymbol{A}}\tilde{\boldsymbol{B}} = \tilde{\boldsymbol{B}}\tilde{\boldsymbol{A}}$, where $\tilde{\boldsymbol{A}} = \tilde{\boldsymbol{A}}^{\mathrm{T}} = \boldsymbol{Q}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{Q}$ and $\tilde{\boldsymbol{B}} = \tilde{\boldsymbol{B}}^{\mathrm{T}} = \boldsymbol{Q}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{Q}$. Since $\tilde{\boldsymbol{A}}$ and $\tilde{\boldsymbol{B}}$ are commuting real symmetric matrices, there exists a real orthogonal matrix $\boldsymbol{R}$ such that $\boldsymbol{R}^{\mathrm{T}}\tilde{\boldsymbol{A}}\boldsymbol{R}=\boldsymbol{D}_1$ and $\boldsymbol{R}^{\mathrm{T}}\tilde{\boldsymbol{B}}\boldsymbol{R}=\boldsymbol{D}_2$, i.e. there exists a $\tilde{\boldsymbol{R}}=\boldsymbol{Q}\boldsymbol{R}$ such that $\tilde{\boldsymbol{R}}^{\mathrm{T}}\boldsymbol{A}\tilde{\boldsymbol{R}}=\boldsymbol{D}_1$ and $\tilde{\boldsymbol{R}}^{\mathrm{T}}\boldsymbol{B}\tilde{\boldsymbol{R}}=\boldsymbol{D}_2$; $\boldsymbol{D}_1$ and $\boldsymbol{D}_2$ are real diagonal.

(ii) *The 'only if' part.* Whenever real symmetric $\boldsymbol{A}$ and $\boldsymbol{B}$ are real diagonalizable by a congruence transform $\boldsymbol{U}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{U}=\boldsymbol{D}_1$ and $\boldsymbol{U}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{U}=\boldsymbol{D}_2$, $\boldsymbol{U}$ non-singular; we have $\boldsymbol{U}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{U}\boldsymbol{U}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{U}=\boldsymbol{U}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{U}\boldsymbol{U}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{U}$, since diagonal matrices always commute. With $\boldsymbol{P}=\boldsymbol{U}\boldsymbol{U}^{\mathrm{T}}$ and carrying out appropriate inversions, the statement is proved. ■

The two well known cases of simultaneous diagonalization of symmetric matrices now become special cases of the 'if part' of theorem 3.5, as sufficient conditions.

(i) When either $\boldsymbol{A}$ or $\boldsymbol{B}$ is positive (or negative) definite, then its inverse (which is also definite) assumes the role of $\boldsymbol{P}$ in $\boldsymbol{APB}=\boldsymbol{BPA}$, and therefore two symmetric matrices, of which one is definite, can always be diagonalized by a real congruence transform.

(ii) Whenever $\boldsymbol{AB}=\boldsymbol{BA}$ is satisfied, the identity matrix is the positive-definite matrix $\boldsymbol{P}$ required in theorem 3.5. This is the case of simultaneous diagonalization by an orthogonal matrix.

The condition of theorem 3.5 restricts the admissible combinations of $\boldsymbol{A}$ and $\boldsymbol{B}$ that are simultaneously real-diagonalizable. All pairs of real symmetric matrices do not satisfy this condition. While diagonalization is sought by an *orthogonal* transform for the well-known result (Bellman 1960), here in theorem 3.5 it is effected by a congruence transform—if, at all, it is possible. A generalization of theorem 3.5 to real square matrices (that need not be symmetric) is presented now.

Theorem 3.6. *A pair of non-defective real square matrices* $\boldsymbol{A}$ *and* $\boldsymbol{B}$ *can be simultaneously diagonalized by the equivalence transforms* $\boldsymbol{VAU}$ *and* $\boldsymbol{VBU}$ *with* $\boldsymbol{V}$ *and* $\boldsymbol{U}$ *real and non-singular matrices if, and only if, there exists a real matrix* $\boldsymbol{R}=\boldsymbol{S}_1\boldsymbol{X}=\boldsymbol{Y}\boldsymbol{S}_2$ *such that* $[\boldsymbol{A},\boldsymbol{B}]_{\boldsymbol{R}}=\boldsymbol{0}$, *where* $\boldsymbol{X}$ *is a left symmetrizer and* $\boldsymbol{Y}$ *is a right symmetrizer of* $(\boldsymbol{A},\boldsymbol{B})$; *and where* $\boldsymbol{S}_1$ *and* $\boldsymbol{S}_2$ *are real symmetric positive-definite matrices.*

Proof. (i) *The 'if' part.* Given that there exists an $\boldsymbol{R}$ with the properties stated above, the requirement of the if part is the simultaneous diagonalizability of a pair of symmetric matrices $\boldsymbol{XA}$ and $\boldsymbol{XB}$ or $\boldsymbol{AY}$ and $\boldsymbol{BY}$. Arguments further to this are the same as those in the 'if part' of theorem 3.5.

(ii) *The 'only if' part.* Given non-singular $\boldsymbol{U}$ and $\boldsymbol{V}$, the equivalence transforms $\boldsymbol{VAU}$ and $\boldsymbol{VBU}$ can be expressed as congruence transforms by writing $\boldsymbol{V}=\boldsymbol{U}^{\mathrm{T}}\boldsymbol{X}$ and $\boldsymbol{U}=\boldsymbol{Y}\boldsymbol{V}^{\mathrm{T}}$. The condition of simultaneous diagonalizability of $\boldsymbol{A}$ and $\boldsymbol{B}$ then becomes simultaneous diagonalizability of $\boldsymbol{XA}$ and $\boldsymbol{XB}$ or that of $\boldsymbol{AY}$ and $\boldsymbol{BY}$. Given that this transform diagonalizes, it is clear that $\boldsymbol{X}$ and $\boldsymbol{Y}$ are the common left and right symmetrizers of $\boldsymbol{A}$ and $\boldsymbol{B}$. Employing theorem 3.5 for a symmetric pair of matrices again, the proof follows. ■

Example 3.7 (a pair of simultaneously diagonalizable real symmetric matrices). Consider

$$\boldsymbol{A}=\begin{bmatrix}1 & 2\\ 2 & 4\end{bmatrix} \quad \text{and} \quad \boldsymbol{B}=\begin{bmatrix}4 & 5\\ 5 & 6\end{bmatrix}.$$

They are diagonalized simultaneously by a real matrix

$$U = \begin{bmatrix} -2 & 4 \\ 1 & -3 \end{bmatrix},$$

i.e. $\boldsymbol{U}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{U}$ and $\boldsymbol{U}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{U}$ are real diagonal matrices, since there exists a positive definite

$$P = \begin{bmatrix} 10 & -7 \\ -7 & 5 \end{bmatrix}$$

such that $\boldsymbol{APB} = \boldsymbol{BPA}$. It can be checked that $\boldsymbol{AB} \neq \boldsymbol{BA}$ and that $\boldsymbol{A}$ is singular, whereas $\boldsymbol{B}$ is indefinite. Therefore, other known sufficient conditions of simultaneous diagonalizability are not applicable to this example. Construction of $\boldsymbol{P}$ has required steps involved in the proof of theorem 3.5.

For an undamped non-gyroscopic multi-degree-of-freedom system, $\boldsymbol{B} = \boldsymbol{0}$ and therefore we are interested in simultaneous diagonalization of $\boldsymbol{A}$ and $\boldsymbol{C}$. Due to theorem 3.1 the dynamics of $\boldsymbol{A}\ddot{\boldsymbol{q}} + \boldsymbol{C}\boldsymbol{q} = \boldsymbol{f}(t)$, whether or not $\boldsymbol{A}$ or/and $\boldsymbol{C}$ is/are symmetric, can always be described by the use of symmetric coefficient matrices such that

$$\boldsymbol{S}_A\ddot{\boldsymbol{q}} + \boldsymbol{S}_C\boldsymbol{q} = \boldsymbol{g}(t).$$

Here $\boldsymbol{S}_A = \boldsymbol{XA} = \boldsymbol{S}_A^{\mathrm{T}}$, $\boldsymbol{S}_C = \boldsymbol{XC} = \boldsymbol{S}_C^{\mathrm{T}}$ and $\boldsymbol{g}(t) = \boldsymbol{X}\boldsymbol{f}(t)$; $\boldsymbol{X}$ assumes the role of a left symmetrizer associated with the matrix pencil $(\boldsymbol{A}, \boldsymbol{C})$. Symmetrization can also be achieved by a coordinate transformation according to $\boldsymbol{q}(t) = \boldsymbol{Y}\boldsymbol{\eta}$, where $\boldsymbol{Y}$ is the right symmetrizer of $\boldsymbol{A}$ and $\boldsymbol{C}$, resulting in $\boldsymbol{AY}\ddot{\boldsymbol{\eta}} + \boldsymbol{CY}\boldsymbol{\eta} = \boldsymbol{f}(t)$. Due to previous discussions, the symmetric eigenproblem $\boldsymbol{S_A u} = \beta \boldsymbol{S_C u}$ does not always admit real eigensolutions. Such situations frequently arise in the dynamics of *non-conservative* systems governed by $\boldsymbol{A}\ddot{\boldsymbol{q}} + \boldsymbol{C}\boldsymbol{q} = \boldsymbol{f}(t)$ when the non-conservative effects are not due to passive dissipation, but due to circulatory forces, control effects, etc. The eigensolutions of the transformed problem $\boldsymbol{S_A u} = \beta \boldsymbol{S_C u}$ (achieved by pre-multiplication by a symmetrizer $\boldsymbol{X}$) and the original eigenproblem $\boldsymbol{Au} = \beta\boldsymbol{Cu}$ are the same. A variety of dynamical behaviour is now possible when β is negative or complex. This is a clear departure from the familiar case of conservative dynamics when $\boldsymbol{A}$ and $\boldsymbol{C}$ are symmetric and in addition at least semi-definite.

Equations of motion of an undamped non-gyroscopic system in the absence of circulatory forces are sometimes not derived by the use of Lagrange's equations, resulting in asymmetric coefficient matrices. Reorganizing equations of motion (see, for example, Newland 1989, problem 5.2(iii), pp. 510–511) restores symmetry. It is not always obvious what the sequence of reorganization should be. The present method of pre-multiplication by the left symmetrizer achieves this in a systematic manner. It also follows that conservative non-gyroscopic systems with *follower forces* too can be described using symmetric coefficient matrices.

An *equivalence transform* such as the one described by Ma & Caughey (1995), can be regarded as a procedure of *symmetrization* by $\boldsymbol{X}$, followed by the usual role of the transformation matrix in the congruence transform. The process of symmetrization is not unique, since one could choose to *post-multiply*, instead of pre-multiplying, by an appropriate matrix, and since symmetrizers themselves are not unique. In the case of symmetrization achieved by post-multiplication, the matrix of left latent vectors assumes the role of transformation matrix for congruence transform. Symmetrization of a matrix pencil $(\boldsymbol{A}\gamma + \boldsymbol{C})$ can be viewed as symmetrization of the pencil $(\boldsymbol{I}\gamma + \boldsymbol{A}^{-1}\boldsymbol{C})$ for a non-singular $\boldsymbol{A}$, which is always possible using a symmetric $\boldsymbol{X}$, the symmetrizer of $\tilde{\boldsymbol{C}} = \boldsymbol{A}^{-1}\boldsymbol{C}$, due to Taussky's original theorem.

When the eigenvectors of the pencil $(\boldsymbol{S}_A, \boldsymbol{S}_C)$ are complex, true decoupling is not achieved, as discussed earlier (although a rather unusual set of orthogonality relations for *complex* modes exist (see, for example, Meirovitch 1980, § 3.6)). Ma & Caughey (1995) distinguish systems with asymmetric $\boldsymbol{A}$ and $\boldsymbol{C}$ compared with those with symmetric matrices due to different orthogonality conditions in the two cases: biorthogonality with respect to the two coefficient matrices in the first case and orthogonality with respect to coefficient matrices in the case of the latter.

They state: 'from a strictly mathematical viewpoint, the modes $\boldsymbol{u}_i$ are still not the same as classical normal modes when the corresponding eigenvalues are all positive'. This line of thought reappears when they discuss their example 1. In another paper (Caughey & Ma 1993), they state 'systems governed by equations for which $\boldsymbol{M}$, $\boldsymbol{C}$, $\boldsymbol{K}$ lack any specific symmetry or definiteness will be termed non-classical systems'. In what follows, we show that this distinction is unnecessary and inappropriate. Distinction ought to be made on the basis of the character of eigensolutions.

Example 3.8 (non-classical normal modes of Caughey & Ma (1993), Ma & Caughey (1995) and Ma (1995)). The set of the right and the left eigenvectors for the system

$$\begin{bmatrix} 0 & 2 \\ 1 & 0 \end{bmatrix} \begin{Bmatrix} \ddot{q}_1 \\ \ddot{q}_2 \end{Bmatrix} + \begin{bmatrix} 1 & 3 \\ 2 & 1 \end{bmatrix} \begin{Bmatrix} q_1 \\ q_2 \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix} \tag{3.7}$$

are calculated as

$$\boldsymbol{U} = \begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix} \quad \text{and} \quad \boldsymbol{V} = \begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix},$$

respectively. $\boldsymbol{U}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{U}$, $\boldsymbol{U}^{\mathrm{T}}\boldsymbol{C}\boldsymbol{U}$, $\boldsymbol{V}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{V}$ and $\boldsymbol{V}^{\mathrm{T}}\boldsymbol{C}\boldsymbol{V}$ are all non-diagonal matrices. However, *synchronous* free vibration in the two modes is possible and *classical* normal modes exist. Ma & Caughey (1995) and Ma (1995), would label them as non-classical modes, because the eigenvectors are not orthogonal with respect to the coefficient matrices. This is misleading. Consider the following system:

$$\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{Bmatrix} \ddot{q}_1 \\ \ddot{q}_2 \end{Bmatrix} + \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix} \begin{Bmatrix} q_1 \\ q_2 \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix}. \tag{3.8}$$

This system is symmetric and positive definite and, therefore, must possess classical normal modes. Equations (3.7) and (3.8) describe the same physical system, only the order of writing the equations has changed. According to Ma & Caughey (1995) the symmetric system (3.8) possesses classical normal modes, whereas the asymmetric description (3.7) of the same system does not. This interpretation of 'classical normal modes' is, in my view, inappropriate.

The following definition of classical normal modes (independent of symmetry of the coefficient matrices) is proposed now.

Definition 3.9. A dynamical system is said to possess a classical normal mode if, and only if, the corresponding free vibratory motion is synchronous.

Therefore, existence of classical normal modes, real eigenvectors or synchronous motion are all synonymous. A normal mode (classical or otherwise) of a system governed by equation (1.1) with $\boldsymbol{B} = \boldsymbol{0}$ is always given by a *right* eigenvector. When it is complex, the asynchronous (or non-classical) mode is obtained by a linear combination of the two complex conjugate modes. Eigenvectors of the adjoint problem (or the left eigenvectors $\boldsymbol{v}_i$) do not enjoy any such interpretation readily. They do not represent normal modes of this system and their association with this system is only mathematical. In fact, they represent normal modes of a *different* system! Therefore, the 'adjoint modes' appear to have little dynamical importance.

A further simplification (when $\boldsymbol{B} = \boldsymbol{0}$) to the symmetrical representation of equation (1.1) as $\boldsymbol{S}_A\ddot{\boldsymbol{q}}+\boldsymbol{S}_C\boldsymbol{q} = \boldsymbol{g}(t)$ is possible by noting that both coefficient matrices are real symmetric. Either $\boldsymbol{S}_A$ or $\boldsymbol{S}_C$ can be diagonalized using a real transform leaving the other matrix symmetric, which gives us the simplest possible canonical representation of (1.1) as $\boldsymbol{D}_A\ddot{\boldsymbol{q}} + \tilde{\boldsymbol{S}}_C\boldsymbol{q} = \boldsymbol{h}(t)$, where $\boldsymbol{D}_A$ is diagonal and $\tilde{\boldsymbol{S}}_C$ is symmetric. Since definiteness of $\boldsymbol{S}_A$ or $\boldsymbol{S}_C$ has not been assumed, the entries on the diagonal of $\boldsymbol{D}_A$ will, in general, have mixed sign.

(*b*) *Pseudo-conservative systems*

Pseudo-conservative systems are non-conservative systems with $\boldsymbol{B} = \boldsymbol{0}$ if the free vibratory motion is synchronous. $\boldsymbol{A}$ and $\boldsymbol{C}$ are assumed to be general square arrays. Pseudo-conservative systems possess *synchronous* stable modes or 'synchronous' unstable *divergent* modes.

Huseyin & Leipholz (1973) studied this class of systems, which they called *symmetrizable systems* because the eigenstructure of this class of (*non-conservative*) problems is very similar to that of the symmetric (and conservative) systems. They called a system pseudo-conservative if the matrix $\boldsymbol{A}^{-1}\boldsymbol{C}$ is symmetrizable, there being no restriction on the symmetry of $\boldsymbol{A}$ or $\boldsymbol{C}$ (see, for example, Huseyin 1978, § 4.2). A real asymmetric matrix is called *symmetrizable* if it can be expressed as a product of two real symmetric matrices, one of which is positive definite (Huseyin 1978, § 1.4), the concept being familiar in the context of operators (Taussky 1968).

The use of the term 'symmetrizable systems' requires care. While Huseyin (1978) means, by this term, a system $\boldsymbol{A}\ddot{\boldsymbol{q}} + \boldsymbol{C}\boldsymbol{q} = \boldsymbol{f}$ with symmetrizable $\boldsymbol{A}^{-1}\boldsymbol{C}$, Ahmadian & Inman (1984) call a symmetrizable asymmetric system 'one which is similar to a symmetric system'. Now symmetric $\boldsymbol{A}$ and $\boldsymbol{C}$ do not necessarily mean a symmetrizable $\boldsymbol{A}^{-1}\boldsymbol{C}$, and worst, since every $\boldsymbol{A}\ddot{\boldsymbol{q}} + \boldsymbol{C}\boldsymbol{q} = \boldsymbol{f}$ can be cast in a symmetric form, we shall abandon the use of the term 'symmetrizable systems' in favour of the term 'pseudo-conservative' or 'real diagonalizable'. For matrices, we continue to call a 'symmetrizable matrix' one that is similar to a symmetric matrix.

If one of the two symmetric factors of a real matrix is positive definite, it has been shown (Taussky 1972) that the eigenvalues are real, i.e. symmetrizable matrices possess real eigenvalues. The converse (Sen & Venkaiah's (1988*b*) remark 'is not yet proved') follows easily from $\boldsymbol{U}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{U} = \boldsymbol{I}$ (see the paragraph before definition 2.1), since

$$\boldsymbol{X} = \boldsymbol{U}^{-\mathrm{T}}\boldsymbol{U}^{-1} \tag{3.9}$$

has the form $\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{Q}$, where $\boldsymbol{Q} = \boldsymbol{U}^{-1}$ is real if the eigenvalues have to be real and, therefore, $\boldsymbol{X}$ must be positive definite.

For symmetric $\boldsymbol{A}$ and $\boldsymbol{C}$, any of the conditions $\boldsymbol{A} > 0$, $\boldsymbol{C} > 0$, $\boldsymbol{A} < 0$, $\boldsymbol{C} < 0$, $\boldsymbol{A} \leqslant 0$, $\boldsymbol{C} \leqslant 0$, $\boldsymbol{A} \geqslant 0$, $\boldsymbol{C} \geqslant 0$ guarantee classical normal modes. The converse is not true, i.e. pseudo-conservative systems may possess indefinite symmetric factors of $\boldsymbol{A}^{-1}\boldsymbol{C}$ (see, example 3.10). Pseudo-conservative systems are also possible when neither $A^{-1}C$ nor $C^{-1}A$ may exist. In those cases, a criterion based on $A^{-1}C$ being a symmetrizable matrix will not be applicable, whereas theorem 3.5 or theorem 3.6 will be (e.g. example 3.7 had A singular).

Example 3.10 (a pseudo-conservative system with indefinite symmetric factors of $A^{-1}C$). Consider a system (1.1) with

$$B = 0, \qquad A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \text{and} \quad C = \begin{bmatrix} -1 & 3 \\ 3 & -1 \end{bmatrix}.$$

It can be checked that neither of the two matrices is definite and their eigenvalues have mixed sign.

Therefore, $A^{-1}C$ has two indefinite symmetric factors, A^{-1} and C. The two matrices are, however, simultaneously diagonalizable by a real matrix

$$U = \begin{bmatrix} \sqrt{2} & \sqrt{2} \\ -\sqrt{2} & \sqrt{2} \end{bmatrix},$$

since $U^T AU = \mathrm{diag}(-1, 1)$ and $U^T CU = \mathrm{diag}(-4, 2)$. Therefore, the system is pseudo-conservative having two synchronous stable modes with frequencies 2 and $\sqrt{2}$.

When the eigenvalues are positive, the normal modes are classical, *whether or not the coefficient matrices are symmetric.* The conclusion of Ma (1995) in this context,

> in free vibration, all components of an undamped non-gyroscopic structure can perform harmonic vibration with identical frequency if the associated eigenvalue problem... possesses positive eigenvalues. The natural frequencies are simply the square roots of these positive eigenvalues, and the mode shapes can be determined from the corresponding complex eigenvectors. Unlike classical modal vibration, the system components generally vibrate with different phase angles

is misleading. While it is true that natural frequencies are the square roots of positive eigenvalues, the deduction of the type of possible free vibratory motion is wrong. Corresponding to a positive eigenvalue (necessarily real), the eigenvector must be *real* (with the exception of degenerate systems) and the motion must be synchronous (i.e. in or out of phase).

Complex eigenvectors may be associated with real eigenvalues for a real eigenproblem when degeneracy exists. This is not due to asymmetry of matrices, rather it is associated with degeneracy of modes, and could be observed in otherwise perfectly well-behaved positive-definite systems. This case is discussed below.

(c) *Non-classical modes for conservative non-gyroscopic positive-definite systems*

Non-classical modes are usually attributed to the presence of (non-classical) damping and/or gyroscopy as well as to asymmetries induced in the coefficient matrices due to control, aerodynamic effects (such as those in aeroelasticity) or follower forces. They have also been traditionally attributed to the inability to decouple a set of coupled differential equations using a real transformation. We illustrate by means of example 3.11 that this need not necessarily be true.

Example 3.11 (non-classical modes for conservative non-gyroscopic systems: degeneracy). Consider (1.1) with $\boldsymbol{A} = \boldsymbol{I}_3$ and $\boldsymbol{B} = \boldsymbol{0}$. The stiffness matrix $\boldsymbol{C}$ and the set of eigenvectors of the eigenproblem $\boldsymbol{Cu} = \gamma \boldsymbol{Au}$ are

$$\boldsymbol{C} = \frac{1}{4}\begin{bmatrix} 5 & -\sqrt{2} & 1 \\ -\sqrt{2} & 6 & -\sqrt{2} \\ 1 & -\sqrt{2} & 5 \end{bmatrix} \quad \text{and} \quad \boldsymbol{U} = \begin{bmatrix} 0.8619 & -0.0839 & 0.5 \\ 0.3504 & -0.6142 & -0.7071 \\ -0.3664 & -0.7847 & 0.5 \end{bmatrix}.$$

The system is non-defective, since a full set of eigenvectors exists. Eigenvalues are all positive implying the existence of three classical normal modes. However, non-classical modes are not ruled out, since, for example,

$$\boldsymbol{u} = \{1 \quad 0.5484 + 0.2586\mathrm{i} \quad -0.2244 + 0.3658\mathrm{i}\}^{\mathrm{T}}$$

satisfies $\boldsymbol{Cu} = \gamma \boldsymbol{Au}$. Since the eigenvector is complex, asynchronous *harmonic* motion is possible in a *non-classical normal mode*. The origin of this is in the degeneracy of modes: eigenvalues are $(1, 1, 2)$. A complex eigenvector of the kind associated with a degenerate system is genuinely complex: no scaling (real or complex) will render it real.

(*d*) *Distribution of the latent roots of* $(\lambda^2 \boldsymbol{A} + \boldsymbol{C})$ *on the complex plane*

The latent roots of $(\lambda^2 \boldsymbol{A} + \lambda \boldsymbol{B} + \boldsymbol{C})$ always appear in complex conjugate pairs (unless they are purely real) whether or not $\boldsymbol{B} = \boldsymbol{0}$, since they are eigenvalues of a real matrix $\mathcal{A}$, the state matrix. This amounts to a mirror symmetry of the set of latent roots about the real axis. When $\boldsymbol{B} = \boldsymbol{0}$, the latent roots of $(\lambda^2 \boldsymbol{A} + \boldsymbol{C})$ are related to the eigenvalues of $\boldsymbol{Cu} = \gamma \boldsymbol{Au}$ via $\gamma = -\lambda^2$. Since $\gamma_j = -\lambda_j^2 = -r_j \exp(\pm \mathrm{i}\theta_j)$, we always have *four* values of λ_j for each pair $(\gamma_j, \bar{\gamma}_j)$. They are given by

$$\lambda_j^{1,2,3,4} = \pm\sqrt{r}_j \exp(\pm \mathrm{i}\theta_j/2) = \pm\sigma_j \pm \tilde{\omega}_j.$$

Therefore, the four values of λ possess a mirror symmetry about *both real and imaginary axes* and they lie in *all the four* quadrants if they do not lie on one of the axes.

The pair $-\sigma_j \pm \tilde{\omega}_j$ results in a stable damped oscillatory mode and the pair $+\sigma_j \pm \tilde{\omega}_j$ results in an unstable (oscillatory) *flutter mode*. When γ_j is real, there are two possibilities. When γ_j is positive, a pair $\lambda_j = \pm \mathrm{i}\omega_j$ is pure imaginary. This results in the well-known case of neutrally stable in- or out-of-phase oscillatory motion with ω_j cyclic frequency and constant amplitude. When γ_j is negative, two real values of λ_j are found: one positive and one negative. They correspond to an unstable *divergent* mode or a stable *overdamped* mode, respectively. Both of these are non-oscillatory.

The doubly symmetric distribution of eigenvalues about the real and the imaginary axes results in a special character of the normal modes of $\boldsymbol{A}\ddot{\boldsymbol{q}} + \boldsymbol{C}\boldsymbol{q} = \boldsymbol{0}$. These attributes are summarized here.

(1) The asynchronous flutter modes and asynchronous damped modes accompany each other; divergence modes and overdamped modes accompany each other.

(2) The accompanying modes (flutter, divergence, overdamped and underdamped oscillatory) are such that they possess identical frequencies and characteristic time-scales of decay/growth.

(3) It is impossible to observe classically damped oscillations or synchronous flutter.

Characteristic time-scales of decay/growth can be defined using the real part of λ. The half-life is given by $\tau_j^{1/2} = \ln 2/|\sigma_j|$ and the doubling time is given by $\tau_j^{(2)} = \ln 2/\sigma_j$. Alternatively, the relaxation time defined as $\tau_j = 1/|\sigma_j|$ could be used. A complex conjugate pair of eigenvalues $(\lambda_j, \bar{\lambda}_j) = \sigma_j \pm \mathrm{i}\tilde{\omega}_j$ can be expressed as

$$(\lambda_j, \bar{\lambda}_j) = -\zeta_j\omega_j \pm \mathrm{i}\Big(\sqrt{1-\zeta_j^2}\Big)\omega_j,$$

where the 'natural frequency' ω_j and the 'damping ratio' ζ_j are defined as

$$\omega_j^2 = \lambda_j\bar{\lambda}_j = \sigma_j^2 + \tilde{\omega}_j^2 \quad \text{and} \quad 2\zeta_j\omega_j = -(\lambda_j + \bar{\lambda}_j) = -2\sigma_j.$$

Given these definitions and that the *magnitude* of real and/or imaginary parts is the same for all the four λ, result (2) follows easily.

The third of the results above rules out synchronous flutter and classically damped motion when $\boldsymbol{B} = \boldsymbol{0}$. This is easily proved by realizing that both these cases require the eigenvector $\boldsymbol{u}$ in the eigenproblem $\boldsymbol{Cu} = \gamma\boldsymbol{Au}$ to be real, i.e. γ to be real. When this is so, the resulting behaviour is oscillatory with constant amplitude or non-oscillatory divergence-overdamping, as discussed above. When $\boldsymbol{B} \neq \boldsymbol{0}$, other possibilities exist (see, example 4.7 for synchronous flutter). $\zeta_j = 1$ defines the boundary between underdamping and overdamping. In an analogous manner, we find here that $\zeta_j = -1$ defines the boundary between the oscillatory (flutter) modes (when $0 > \zeta_j > -1$) and the non-oscillatory (divergent) modes (when $\zeta_j < -1$). This is because λ_j is real when $|\zeta_j| > 1$, since the imaginary part of $(\lambda_j, \bar{\lambda}_j)$ equals

$$\pm\Big(\sqrt{1-\zeta_j^2}\Big)\omega_j.$$

Example 3.12 (accompanying modes of an undamped non-gyroscopic system).

$$\boldsymbol{A} = \begin{bmatrix} 7361 & 7564 & 2470 \\ 3282 & 9910 & 9826 \\ 6326 & 3653 & 7227 \end{bmatrix} \quad \text{and} \quad \boldsymbol{C} = \begin{bmatrix} 5269 & 4160 & 7622 \\ 920 & 7012 & 2625 \\ 6539 & 9103 & 475 \end{bmatrix}.$$

Eigenvalues of $\boldsymbol{A}^{-1}\boldsymbol{C}$ are calculated as $\{\gamma_j\} = \{-1.4091, 0.7073 \pm 0.1924\mathrm{i}\}$. Six eigenvectors of the state matrix are

$$\boldsymbol{W} = \begin{bmatrix} -0.0229 & 0.0229 & -0.4375 \mp 0.4914\mathrm{i} & 0.3781 \pm 0.5384\mathrm{i} \\ 0.4014 & -0.4014 & 0.3139 \pm 0.1537\mathrm{i} & -0.0778 \mp 0.3407\mathrm{i} \\ -0.5035 & 0.5035 & 0.1470 \pm 0.0205\mathrm{i} & -0.0137 \mp 0.1478\mathrm{i} \\ 0.0272 & 0.0272 & 0.4666 \mp 0.3155\mathrm{i} & -0.4140 \pm 0.3819\mathrm{i} \\ -0.4764 & -0.4764 & -0.1661 \pm 0.2489\mathrm{i} & 0.2803 \mp 0.1046\mathrm{i} \\ 0.5976 & 0.5976 & -0.0341 \pm 0.1224\mathrm{i} & -0.0800 \mp 0.0052\mathrm{i} \end{bmatrix}$$

and the six λ_j are calculated as $\{\lambda_j\} = \{\mp 1.1871 \quad \mp 0.1134 \pm 0.8486\mathrm{i}\}$. Note the four complex values and the types of associated free vibratory motion. The two real eigenvalues correspond to two non-oscillatory modes: an overdamped mode and an unstable divergence mode. Rescaling columns of $\boldsymbol{W}$ reveals that the first three elements of the first column and the second column are the same, equal to $\{1, -17.5373, 21.9979\}^{\mathrm{T}}$ corresponding to the two non-oscillatory accompanying modes. The third and fourth columns can be rescaled similarly, and the first three entries turn out to be the same, equal to $\{1, -0.4918 \pm 0.2009\mathrm{i}, -0.1719 \pm 0.1461\mathrm{i}\}^{\mathrm{T}}$, and they correspond to the two accompanying oscillatory modes.

4. General non-conservative systems: three real square arrays

A result for n simultaneously diagonalizable real square matrices is proved first.

Theorem 4.1. *If a common equivalence transform*

$$\boldsymbol{V}^{\mathrm{T}}\boldsymbol{A}_i\boldsymbol{U} = \boldsymbol{D}_i, \quad \forall i = 0, 1, \dots, n,$$

$\boldsymbol{U}$ and $\boldsymbol{V}$ non-singular, diagonalizes a set of real, non-defective matrices $\boldsymbol{A}_0, \boldsymbol{A}_1, \dots,$ $\boldsymbol{A}_n$, then there exists a real matrix $\boldsymbol{X}$ such that $(\boldsymbol{X}\boldsymbol{A}_i) = (\boldsymbol{X}\boldsymbol{A}_i)^{\mathrm{T}}$, $\forall i = 0, 1, \dots, n$.

Proof. Define $\boldsymbol{V} = \boldsymbol{X}^{\mathrm{T}}\boldsymbol{U}$, i.e. $\boldsymbol{X} = \boldsymbol{U}^{-\mathrm{T}}\boldsymbol{V}^{\mathrm{T}}$. Thus it is given that $\boldsymbol{U}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{A}_i\boldsymbol{U} = \boldsymbol{D}_i$, $\forall i = 0, 1, \dots, n$. Pre-multiply and post-multiply both sides by appropriate matrices to get

$$\boldsymbol{X}\boldsymbol{A}_i = \boldsymbol{U}^{-\mathrm{T}}\boldsymbol{D}_i\boldsymbol{U}^{-1} = (\boldsymbol{X}\boldsymbol{A}_i)^{\mathrm{T}}, \quad i = 0, 1, \dots, n. \tag{4.1}$$

If the matrices $\boldsymbol{U}$, $\boldsymbol{V}$ used in the equivalence transform are real, the proof is immediate. If they are complex, $\boldsymbol{X}$ can be written as a sum of its real and imaginary parts on both sides of equation (4.1); separation of real and imaginary parts completes the proof. It can be further shown that, under the conditions of theorem 4.1, there also exists a real matrix $\boldsymbol{Y}$ such that $(\boldsymbol{A}_i\boldsymbol{Y}) = (\boldsymbol{A}_i\boldsymbol{Y})^{\mathrm{T}}$. ■

It is trivial that real matrices that can be simultaneously rendered real diagonal by a common transform can also be rendered real symmetric by a common transform, since the diagonal form is also a symmetric form. What is not obvious is that the simultaneously diagonalizable matrices (the transform used for diagonalization need not be real) possess a common *real* transform that renders all the matrices symmetric on pre-multiplication. An interesting feature of the above result is that even when $\boldsymbol{U}$ and $\boldsymbol{V}$ are complex, $\boldsymbol{X} = \boldsymbol{U}^{-\mathrm{T}}\boldsymbol{V}^{\mathrm{T}}$ is real!

Example 4.2. (simultaneous diagonalizability by a real or a complex equivalence transform implies simultaneous symmetrizability by real pre-multiplication). The complex matrices

$$\boldsymbol{V} = \begin{bmatrix} 1 & 1 \\ -0.2(2-\mathrm{i}) & -0.2(2+\mathrm{i}) \end{bmatrix}, \qquad \boldsymbol{U} = \begin{bmatrix} 1 & 1 \\ -0.5(1-\mathrm{i}) & -0.5(1+\mathrm{i}) \end{bmatrix}$$

simultaneously diagonalize the three real matrices

$$\boldsymbol{A}_1 = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \qquad \boldsymbol{A}_2 = \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix}, \qquad \boldsymbol{A}_3 = \begin{bmatrix} 1.5 & 1 \\ 2.5 & 0 \end{bmatrix},$$

since

$$\boldsymbol{V}^{\mathrm{T}}\boldsymbol{A}_1\boldsymbol{U} = 0.4\begin{bmatrix} -2+\mathrm{i} & 0 \\ 0 & -2-\mathrm{i} \end{bmatrix},$$

$$\boldsymbol{V}^{\mathrm{T}}\boldsymbol{A}_2\boldsymbol{U} = -0.2\begin{bmatrix} 1-3\mathrm{i} & 0 \\ 0 & 1+3\mathrm{i} \end{bmatrix},$$

$$\boldsymbol{V}^{\mathrm{T}}\boldsymbol{A}_3\boldsymbol{U} = \begin{bmatrix} \mathrm{i} & 0 \\ 0 & -\mathrm{i} \end{bmatrix}.$$

The common *real* symmetrizer

$$\boldsymbol{X} = \boldsymbol{U}^{-\mathrm{T}}\boldsymbol{V}^{\mathrm{T}} = \begin{bmatrix} 1 & -0.2 \\ 0 & 0.4 \end{bmatrix}$$

renders $\boldsymbol{A}_1$, $\boldsymbol{A}_2$ and $\boldsymbol{A}_3$ simultaneously symmetric:

$$\boldsymbol{X}\boldsymbol{A}_1 = 0.4\begin{bmatrix} 1 & 3 \\ 3 & 4 \end{bmatrix}, \quad \boldsymbol{X}\boldsymbol{A}_2 = 0.2\begin{bmatrix} 3 & 4 \\ 4 & 2 \end{bmatrix} \quad \text{and} \quad \boldsymbol{X}\boldsymbol{A}_3 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}.$$

It can be similarly shown that real matrices that are diagonalized by a common (real or complex) equivalence transform possess a common *real right symmetrizer*. Therefore, *if a non-defective system governed by equation (1.1) can be decoupled by the use of a real equivalence transform, then there exists a real matrix* $\boldsymbol{X}$ *such that the dynamics are described by* $\boldsymbol{S}_A\ddot{\boldsymbol{q}} + \boldsymbol{S}_B\dot{\boldsymbol{q}} + \boldsymbol{S}_C\boldsymbol{q} = \boldsymbol{g}(t)$, *where the coefficient matrices* $\boldsymbol{S}_A = \boldsymbol{X}\boldsymbol{A} = (\boldsymbol{S}_A)^{\mathrm{T}}$, $\boldsymbol{S}_B = \boldsymbol{X}\boldsymbol{B} = (\boldsymbol{S}_B)^{\mathrm{T}}$ *and* $\boldsymbol{S}_C = \boldsymbol{X}\boldsymbol{C} = (\boldsymbol{S}_C)^{\mathrm{T}}$ *are symmetric, the right side vector is given by* $\boldsymbol{g}(t) = \boldsymbol{X}\boldsymbol{f}(t)$. It is also readily seen that *if a symmetrizer of the matrix pencil* $(\boldsymbol{A}^{-1}\boldsymbol{B}, \boldsymbol{A}^{-1}\boldsymbol{C})$ *is symmetric, then the system governed by equation (1.1) can be described by one having symmetric coefficient matrices.*

For symmetric-definite systems, Caughey's commutativity condition is *necessary and sufficient* for the existence of classical normal modes. On the other hand, for an asymmetric system (1.1), Caughey's condition is only necessary. If three real matrices $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ are simultaneously diagonalized by a real equivalence transform, $\boldsymbol{V}^{\mathrm{T}}(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})\boldsymbol{U} = (\boldsymbol{D}_A, \boldsymbol{D}_B, \boldsymbol{D}_C)$ and if $\boldsymbol{A}$ is assumed to be non-singular, then

$$\boldsymbol{A}^{-1}\boldsymbol{B}\boldsymbol{A}^{-1}\boldsymbol{C} = \boldsymbol{U}\boldsymbol{D}_A^{-1}\boldsymbol{D}_B\boldsymbol{D}_A^{-1}\boldsymbol{D}_C\boldsymbol{U}^{-1},$$

and similarly

$$\boldsymbol{A}^{-1}\boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B} = \boldsymbol{U}\boldsymbol{D}_A^{-1}\boldsymbol{D}_C\boldsymbol{D}_A^{-1}\boldsymbol{D}_B\boldsymbol{U}^{-1}.$$

This proves the necessity of Caughey's condition.

Example 4.3 (Caughey's condition is not sufficient for the existence of classical normal modes if the coefficient matrices are not symmetric definite). Consider $\boldsymbol{A} = \boldsymbol{B} = \boldsymbol{I}_2$ and

$$\boldsymbol{C} = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

in equation (1.1). Clearly, $[\boldsymbol{A}^{-1}\boldsymbol{B}, \boldsymbol{A}^{-1}\boldsymbol{C}] = \boldsymbol{0}$ is satisfied. However, all the modes are complex, since

$$\boldsymbol{W}^{\mathrm{T}} = \begin{bmatrix} 1 & \pm\mathrm{i} & -1\pm\mathrm{i} & -1\mp\mathrm{i} \\ 1 & \mp\mathrm{i} & \pm\mathrm{i} & 1 \end{bmatrix},$$

where $\boldsymbol{W}$ is the 4×4 matrix of eigenvectors.

Equation (1.1) can be cast as $\boldsymbol{I}\ddot{\boldsymbol{q}} + \tilde{\boldsymbol{B}}\dot{\boldsymbol{q}} + \tilde{\boldsymbol{C}}\boldsymbol{q} = \boldsymbol{0}$ if $\boldsymbol{A}$ is non-singular. Ahmadian & Inman (1984) prove that a 'symmetrizable' system possesses classical normal modes if, and only if, Caughey's condition holds. With *their* definition of a symmetrizable system (quoted here in § 3 above), the result becomes invalid.

We now study the implication of Caughey's condition $[\tilde{\boldsymbol{B}}, \tilde{\boldsymbol{C}}] = \boldsymbol{0}$. Suppose $\boldsymbol{X}$ is a symmetrizer of $\tilde{\boldsymbol{B}}$, so that $\boldsymbol{X} = \boldsymbol{X}^{\mathrm{T}}$ and $\boldsymbol{X}\tilde{\boldsymbol{B}} = (\boldsymbol{X}\tilde{\boldsymbol{B}})^{\mathrm{T}}$. Given the commutativity of $\tilde{\boldsymbol{B}}$ and $\tilde{\boldsymbol{C}}$, we conclude that $\tilde{\boldsymbol{C}}$ must be a transform (see theorem 2.7) that generates another symmetrizer according to $\boldsymbol{X}' = \boldsymbol{X}\boldsymbol{C}$. But a symmetrizer, by definition, is itself symmetric, hence both $\boldsymbol{X}'$ and $\boldsymbol{X}$ are symmetric. Pre-multiplying by $\boldsymbol{X}$ throughout, we have $\boldsymbol{X}\ddot{\boldsymbol{q}} + \boldsymbol{X}\tilde{\boldsymbol{B}}\dot{\boldsymbol{q}} + \boldsymbol{X}\tilde{\boldsymbol{C}}\boldsymbol{q} = \boldsymbol{0}$. Therefore, *if Caughey's condition holds, then the dynamics of the system (1.1) with asymmetric coefficient matrices can always be described by means of three real symmetric coefficient matrices.*

Theorem 4.4. *A general dynamical system governed by equation (1.1) possesses classical normal modes if*

(a) $\boldsymbol{A}^{-1}\boldsymbol{B}$ *and* $\boldsymbol{A}^{-1}\boldsymbol{C}$ *commute in multiplication, and*

(b) *the symmetrizer matrix* $\boldsymbol{X}$ *that exists due to (a), i.e.*

$$\boldsymbol{X}(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}) = (\boldsymbol{A}^{\mathrm{T}}, \boldsymbol{B}^{\mathrm{T}}, \boldsymbol{C}^{\mathrm{T}})(\boldsymbol{X}^{\mathrm{T}}),$$

is definite, i.e. $\boldsymbol{X} > 0$ *or* $\boldsymbol{X} < 0$.

The two conditions together form a set of *sufficient* conditions for the existence of classical normal modes. The proof follows from the fact that the existence of a common definite symmetrizer renders the problem to one considered by Caughey & O'Kelley (1965) and that $\boldsymbol{S}_B\boldsymbol{S}_A^{-1}\boldsymbol{S}_C = \boldsymbol{S}_C\boldsymbol{S}_A^{-1}\boldsymbol{S}_B \Leftrightarrow \boldsymbol{B}\boldsymbol{A}^{-1}\boldsymbol{C} = \boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B}$. While condition (a) of theorem 4.4 is necessary, the second condition is *not*.

Liu & Wilson (1992) state '...a damped system, whether symmetric or not, can be transformed into an equivalent undamped one, if and only if the Caughey condition is satisfied. However, equation...[the equation that describes Caughey's condition in their paper]...cannot be simply used to determine if a non-symmetrical general damping matrix can be decoupled'. These two statements juxtaposed to each other appear to be in conflict: the 'only if' part of the first statement is incorrect.

Caughey & Ma (1993) state that 'the non-classical system...can be decoupled if and only if the coefficient matrices...are diagonalizable and pairwise commutative' (see Caughey & Ma 1993, theorem 1). This is not correct. To show this, choose two arbitrary non-singular matrices, $\boldsymbol{U}$, $\boldsymbol{V}$, and three arbitrary diagonal matrices, $\boldsymbol{D}_1$, $\boldsymbol{D}_2$, $\boldsymbol{D}_3$, and construct matrices $\boldsymbol{A}$, $\boldsymbol{B}$ and $\boldsymbol{C}$ such that

$$\boldsymbol{U}\boldsymbol{D}_1\boldsymbol{V} = \boldsymbol{A}, \quad \boldsymbol{U}\boldsymbol{D}_2\boldsymbol{V} = \boldsymbol{B} \quad \text{and} \quad \boldsymbol{U}\boldsymbol{D}_3\boldsymbol{V} = \boldsymbol{C}.$$

Clearly, $\boldsymbol{U}^{-1}$ and $\boldsymbol{V}^{-1}$ diagonalize all three coefficient matrices $\boldsymbol{A}$, $\boldsymbol{B}$ and $\boldsymbol{C}$ by the use of an equivalence transform. Theorem 1 of Caughey & Ma (1993) then requires that $\boldsymbol{A}\boldsymbol{B} = \boldsymbol{B}\boldsymbol{A}$. In other words, $\boldsymbol{U}\boldsymbol{D}_1\boldsymbol{V}\boldsymbol{U}\boldsymbol{D}_2\boldsymbol{V} = \boldsymbol{U}\boldsymbol{D}_2\boldsymbol{V}\boldsymbol{U}\boldsymbol{D}_1\boldsymbol{V}$ is required, i.e. $\boldsymbol{D}_1\boldsymbol{W}\boldsymbol{D}_2 = \boldsymbol{D}_2\boldsymbol{W}\boldsymbol{D}_1$, where $\boldsymbol{W} = \boldsymbol{V}\boldsymbol{U}$ is required for arbitrarily chosen $\boldsymbol{U}$, $\boldsymbol{V}$, $\boldsymbol{D}_1$, $\boldsymbol{D}_2$!

Example 4.5 (counter example to theorem 1 in Caughey & Ma (1993)).

$$\boldsymbol{V} = \begin{bmatrix} -1 & 0 & (\frac{2}{3}) \\ 2 & -3 & (\frac{4}{3}) \\ -1 & 2 & -1 \end{bmatrix} \quad \text{and} \quad \boldsymbol{U} = \begin{bmatrix} 2.2 & -2.8 & 1.2 \\ -3 & 3 & -1 \\ 1.2 & -0.8 & 0.2 \end{bmatrix}$$

Table 1. *Possible types of normal modes in a second-order system*

	oscillatory		
	synchronous	asynchronous	non-oscillatory
stable $\sigma_j < 0$	classically damped $0 < \zeta_j < 1$	non-classically damped $0 < \zeta_j < 1$	overdamped $\zeta_j > 1$
neutrally stable $\sigma_j = 0$	conservative non-gyroscopic $\zeta_j = 0$	conservative gyroscopic $\zeta_j = 0$	trivial static state
unstable $\sigma_j > 0$	classical flutter $-1 < \zeta_j < 0$	non-classical flutter $-1 < \zeta_j < 0$	divergence $\zeta_j < -1$

(i) that the first n elements of the $2n$-dimensional vector are real, implying the classical nature of the flutter mode, and

(ii) the last n elements of the $2n$-dimensional vector are $0.5 \pm 0.8660\mathrm{i}$ times the first n elements of this vector ($n = 2$ here).

The second observation, particularly in the context of damped passive systems, is not new (Newland 1987, 1989). Using an appropriate scaling (in this case by dividing the second column of $\boldsymbol{W}$ by $0.5 \pm 0.8660\mathrm{i}$), one could reduce the last n elements of the second column to be real. This shows that *by an appropriate scaling, either the first n or the last n elements of an eigenvector corresponding to a classical normal mode can be made real.* This holds for classically damped (i.e. stable) systems too. The whole of the $2n$-dimensional eigenvector will be real when, in addition, the corresponding eigenvalue is real; which is the case with classical undamped modes.

Analogous to the damped modes, I propose to call the synchronous flutter modes the 'classical flutter' modes and the asynchronous flutter modes the 'non-classical flutter' modes. A summary of the mathematical attributes of the possible normal mode motions is presented in table 1.

5. Conclusions

A proof of Taussky's theorem based on left and right eigenvectors was presented. It was shown that the matrix of left eigenvectors is related to the matrix of right eigenvectors via a symmetric matrix (also known as a symmetrizer) which appears in Taussky's theorem. The mathematical structure of the transforms that relate symmetrizers was explored and it was shown that they form a multiplicative group. A systematic procedure of finding symmetrizers (and, hence, symmetric factors) of a real matrix was presented.

It was shown that a real non-defective matrix pencil can always be rendered symmetric by means of pre-multiplication by a real (but not necessarily symmetric) matrix. On this basis we conclude that the governing equations of motion of an undamped non-gyroscopic system can always be cast in terms of two real symmetric matrices even in presence of circulatory/follower forces. The so-called symmetrizable matrices were critically examined. It was noted that pseudo-conservative systems do not necessarily need to possess positive-definite coefficient matrices: a fact not recognized in the literature.

The well-known result of simultaneous diagonalization of two real symmetric matrices by a real orthogonal transform was generalized to diagonalization by a general real transformation. It was noted that the required necessary and sufficient condition is the existence of a positive-definite matrix with respect to which the given matrices, whose diagonalization is in question, must commute; the classical result being a special case when this positive-definite matrix is the identity matrix. Finally, the result was generalized to simultaneous diagonalization of two real asymmetric matrices by a real transform.

Distribution of the latent roots of the quadratic eigenvalue problem associated with undamped non-gyroscopic systems was studied. It was observed that the latent roots fall symmetrically about both real and imaginary axes. As the damping ratio equal to 1 defines the boundary between overdamped and underdamped modes; it was shown that the damping ratio equal to -1 defines the boundary between divergence and flutter. It was proved that flutter modes accompany damped oscillatory modes, whereas divergence modes accompany overdamped modes for undamped non-gyroscopic systems. It was further shown that the accompanying modes have identical frequencies and time-scales of growth/decay.

General non-conservative systems governed by equations of motion having three coefficient matrices were studied. It was also shown that Caughey's condition (when the coefficient matrices are asymmetric) is a necessary condition for classical modes to exist; however, it is not sufficient. Several examples (and in some instances counterexamples, to point out an error) were presented.

I thank Dr Debashish Ghose, Department of Aerospace Engineering, Indian Institute of Science, Professor P. C. Dumir, Department of Applied Mechanics, and Dr V. Venkaiah, Department of Mathematics, Indian Institute of Technology, New Delhi, for very useful discussions. Suggestions by the unknown referees are gratefully acknowledged.

References

Ahmadian, M. & Chou, S.-H. 1987 A new method for finding symmetrical form of asymmetric finite-dimensional dynamic-systems. *ASME J. Appl. Mech.* **54**, 700–705.

Ahmadian, M. & Inman, D. J. 1984 Classical normal modes in asymmetric nonconservative dynamic systems. *AIAA J.* **22**, 1012–1015.

Bellman, R. 1960 *Introduction to matrix analysis.* New York: McGraw-Hill.

Bolotin, V. V. 1963 *Nonconservative problems of the theory of elastic stability.* New York: Pergamon.

Caughey, T. K. 1960 Classical normal modes in damped linear systems. *ASME J. Appl. Mech.* **27**, 269–271.

Caughey, T. K. & Ma, F. 1993 Complex modes and solvability of non-classical linear systems. *ASME J. Appl. Mech.* **60**, 26–28.

Caughey, T. K. & O'Kelley, M. E. J. 1965 Classical normal modes in damped linear dynamic systems. *ASME J. Appl. Mech.* **32**, 583–588.

Fawzy, I. & Bishop, R. E. D. 1976 On the dynamics of linear non-conservative systems. *Proc. R. Soc. Lond.* A **352**, 25–40.

Huseyin, K. 1978 *Vibrations and stability of multiple parameter systems.* Alphen aan den Rijn, The Netherlands: Noordhoff.

Huseyin, K. & Leipholz, H. H. E. 1973 Divergence instability of multiple parameter circulatory systems. *Q. Appl. Math.* **31**, 185–197.

Inman, D. J. 1983 Dynamics of asymmetric non-conservative systems. *ASME J. Appl. Mech.* **50**, 199–203.

Lancaster, P. 1966 *Lambda-matrices and vibrating systems.* New York: Pergamon.

Liu, M. & Wilson, J. M. 1992 Criterion for decoupling dynamic equations of motion of linear gyroscopic systems. *AIAA J.* **30**, 2989–2991.

Ma, F. 1995 Analysis of the equations of motion of linearised controlled structures, *Sādhanā* **20**, 709–719.

Ma, F. & Caughey, T. K. 1995 Analysis of linear non-conservative vibrations. *ASME J. Appl. Mech.* **62**, 685–691.

Meirovitch, L. 1980 *Computational methods in structural dynamics.* Alphen aan den Rijn, The Netherlands: Sijthoff & Noordhoff.

Newland, D. E. 1987 On the modal analysis of non-conservative linear systems. *J. Sound Vib.* **112**, 69–97.

Newland, D. E. 1989 *Mechanical vibration: analysis and computation.* Longman.

Rayleigh, J. W. S. 1894 *The theory of sound*, vol. 1. Reprinted 1945. New York: Dover.

Sen, S. K. & Venkaiah, V. Ch. 1988*a* On symmetrizing a matrix. *Indian J. Pure Appl. Math.* **19**, 554–561.

Sen, S. K. & Venkaiah, V. Ch. 1988*b* On computing an equivalent symmetric matrix for a non-symmetric matrix. *Int. J. Comp. Math.* **24**(2), 169–180.

Taussky, O. 1959 Problem. *Am. Math. Mon.* **66**, 427.

Taussky, O. 1968 Positive-definite matrices and their role in the study of the characteristic roots of general matrices. *Adv. Math.* **2**, 175–186. (Also in *Inequalities* (ed. O. Shisha), pp. 389–400.)

Taussky, O. 1972 The role of symmetric matrices in the study of general matrices. *Linear Algebra Appl.* **5**, 147–154.

Taussky, O. & Zassenhaus, H. 1959 On the similarity transformation between a matrix and its transpose. *Pac. J. Math.* **9**, 893–896.

Thomson, W. & Tait, P. G. 1867 *Treatise on natural philosophy*, 1st edn. Reprinted as *Principles of mechanics and dynamics* (1962). New York: Dover.

Valenza, R. J. 1993 *Linear algebra: an introduction to abstract mathematics.* Springer.

Venkaiah, V. Ch. & Sen, S. K. 1988 Computing a matrix symmetrizer exactly using modified multiple modulus residue arithmetic. *J. Comput. Appl. Math.* **21**, 27–40.

Wahed, I. F. A. & Bishop, R. E. D. 1976 On the equations governing the free and forced vibration of a general non-conservative system. *J. Mech. Engng Sci.* **18**, 6–10.