The transformation of second-order linear systems into independent equations

Permalink
https://escholarship.org/uc/item/637104qb

Journal
SIAM Journal on Applied Mathematics, 71(4)

ISSN
0036-1399

Authors
Morzfeld, M
Ma, F
Parlett, BN

Publication Date
2011-09-19

DOI
10.1137/100818637

Peer reviewed
THE TRANSFORMATION OF SECOND-ORDER LINEAR SYSTEMS INTO INDEPENDENT EQUATIONS∗

MATTHIAS MORZFELD†, FAI MA‡, AND BERESFORD N. PARLETT§

Abstract. The class of second-order linear dynamical systems is considered. A method and algorithm are presented to transform any system with \( n \) degrees of freedom into \( n \) independent second-order equations. The conversion utilizes a real, invertible but nonlinear mapping and is applicable to practically every linear system. Two examples from earthquake engineering are provided to indicate the utility of this approach.

Key words. second-order linear differential equations, quadratic eigenvalue problem, modal analysis

AMS subject classifications. 15A22, 34A30, 70J10, 93A10

DOI. 10.1137/100818637

1. Introduction. We consider the second-order linear dynamical system

\[
M\ddot{q}(t) + C\dot{q}(t) + Kq(t) = f(t),
\]

where all quantities are real and where dots denote derivatives with respect to the independent variable \( t \geq 0 \). The coefficients \( M, C, \) and \( K \) are \( n \times n \) matrices; \( q(t) \) and \( f(t) \) are \( n \)-dimensional column vectors. The initial conditions \( q(0) \) and \( \dot{q}(0) \) as well as \( f(t) \) are given. For simplicity, we assume \( f(t) \) continuously differentiable and restrict our attention to invertible \( M \), thus avoiding differential algebraic equations. Equation (1.1) is a cornerstone in vibration theory and, for example, models small oscillations of particles [1, 2, 3, 4, 5]. In vibration terminology, (1.1) determines a nonconservative linear system [1, 2, 3, 4, 5].

Two symmetric positive definite (SPD) matrices \( M \) and \( K \) can be simultaneously diagonalized by a congruence transformation [6, 7]. The same congruence transformation that diagonalizes \( M \) and \( K \) also diagonalizes a symmetric \( C \) if and only if (see [8])

\[
CM^{-1}K = KM^{-1}C.
\]

Similar restrictive conditions on simultaneous diagonalization apply if \( M, C, \) and \( K \) are arbitrary square matrices (see [9, 10] for an application-based discussion). It follows that system (1.1) generally cannot be uncoupled (decoupled in modern terminology) into a set of mutually independent, real, scalar, second-order equations by a linear mapping \( q(t) \rightarrow \mathcal{L}q(t) \), with \( \mathcal{L} \) independent of \( t \).

Nonetheless, the decoupling of (1.1) is desirable from a practical as well as a theoretical perspective. We present a new approach to this old problem and consider

∗Received by the editors December 20, 2010; accepted for publication (in revised form) April 4, 2011; published electronically July 7, 2011. This research was supported in part by U.S. Department of Energy under contract DEAC02-05CH11231.
†Department of Mathematics, Computational Research Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720-8142 and UC Berkeley Department of Mechanical Engineering, Berkeley, CA (mmo@math.lbl.gov).
‡Department of Mechanical Engineering, University of California, Berkeley, CA 94720-1740 (fma@me.berkeley.edu).
§Department of Mathematics and Department of Electrical Engineering and Computer Science, University of California, Berkeley, CA 94720-3840 (parlett@math.berkeley.edu).

1026
a time-dependent, nonlinear mapping $q(t) \rightarrow N(q,t)$ to obtain a set of mutually independent, real, second-order equations. We then compute the solution $q(t)$ of the coupled equation from the solutions of the independent equations. With apologies for using a terminology that may cause confusion, any methodology that uses a linear or nonlinear mapping to obtain a diagonal second-order system is herein referred to as decoupling.

The possibility of decoupling any system possessing SPD coefficients by a nonlinear mapping was recently discovered [11, 12]. The decoupling procedure, termed phase synchronization, can be briefly described as follows. First, the quadratic eigenvalue problem associated with (1.1) is solved. A real and diagonal (i.e., decoupled) system is then constructed using the eigenvalues. Let the solution of the decoupled system be $p(t)$. If system solution is an objective, we can recover $q(t)$ by evaluating each component of $p(t)$ with a different time lag. It is here that we exploit the special parameter $t$, which is characteristic of a dynamical system, making it possible to decouple system (1.1) without restrictions for the first time.

In this paper, we extend the developments in [11, 12] to systems possessing nonsymmetric coefficients. We also clarify the decoupling procedure under real eigenvalues and provide a broader perspective of the decoupling operations. The few restrictions placed on $M$, $C$, and $K$ are an indication of the extensive scope of our method. We emphasize the theory of decoupling, rather than its numerous applications such as model reduction, stability analysis, optimal control, earthquake, or rotor design. Nevertheless, two examples from earthquake engineering are employed to demonstrate the utility of the decoupling method. Additional examples showing phase synchronization in action can be found in [11, 12].

Based upon the notion of structure-preserving transformations (SPT), Garvey and others [13, 14, 15, 16, 17] introduced a method for decoupling homogeneous systems. There, algorithms employing linear coordinate transformations in a higher dimensional space (the state space) are utilized to compute a real and second-order diagonal system. If available, this diagonal system is identical to the system we obtain by phase synchronization without recourse to state space or SPTs (modulo a normalization $M_d \rightarrow I$; see (3.3) and (4.6)). A more detailed discussion of our work in relation to SPTs is given in section 4.

Please note our notational conventions. We try to reserve capital letters for matrices, lowercase Roman letters for column vectors, and lowercase Greek letters for scalars. The main exceptions are $t$ (time), $e$ (Euler’s number), and $i = \sqrt{-1}$, for obvious reasons. We also reserve $n$ for the number of degrees of freedom (the length of the vector $q$), and the letters $j$ and $k$ for indexing scalars or vectors. Thus, $v_j$, $j = 1, \ldots, n$, denotes the sequence of vectors $v_1, v_2, \ldots, v_n$. We sometimes use $v_j^k$ to denote the $k$th entry of $v_j$. In addition, we make use of the Hadamard (or Schur, or pointwise) product of two vectors of the same dimension, i.e., $(v \odot w)^k = v^k w^k$. In this context, the vector $u = (1, \ldots, 1)^T$ is useful, where $T$ denotes the transpose ($u$ stands for unit and we would have preferred to use $e$ here, as is more common, but decided to reserve $e$ for Euler’s number). We apply scalar functions like $\sin(\cdot)$ or $\cos(\cdot)$ to vectors in a componentwise fashion. Thus, $\exp(v)$ is a vector with $k$th component given by $(\exp(v))^k = \exp(v^k)$. Likewise, $\Re(v)$ and $\Im(v)$ denote, respectively, the real and imaginary parts of the vector $v$. We construct diagonal matrices using the notation $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_n)$ so that $\Lambda \in \mathbb{C}^{n \times n}$ is diagonal with elements $\lambda_1, \ldots, \lambda_n$. Given a sequence of vectors $v_j$, $j = 1, \ldots, n$, we construct a matrix $V = (v_1, \ldots, v_n) \in \mathbb{C}^{n \times n}$, whose columns are the vectors $v_j$. The identity matrix is denoted by $I$, and a square matrix with all elements equal to zero is denoted by $0$. Based upon compatibility, the
dimensions of $I$ and $0$ can be inferred from the context. Last, we express complex vectors in a quasi-polar form, so that $v = r \exp(-if)$ with $f = (\varphi_1, \ldots, \varphi_n)^T$. In strictly polar form, $r^j \geq 0$ and $|\varphi_j| \leq \pi$ for $j = 1, \ldots, n$. For our purposes it appears more sensible to allow $r^j$ to be negative (if necessary) so that $|\varphi_j| \leq \pi/2$. There is no advantage in the problem we consider to have $r$ nonnegative. The choice of sign in $v = r \exp(-if)$ is for convenience only. We call this representation the quasi-polar form.

2. Phase synchronization. Associated with (1.1) is the regular quadratic eigenvalue problem [2, 18, 19, 20]

\begin{equation}
(M\lambda^2 + C\lambda + K)v = 0,
\end{equation}

where $\lambda$ is termed an eigenvalue and $v$ is the corresponding eigenvector. There are $2n$ eigenpairs $(\lambda_j, v_j)$, $j = 1, \ldots, 2n$, complex in general. All eigenvalues are finite because $M$ is invertible. Since $M$, $C$, and $K$ are real, complex eigenvalues and eigenvectors occur in complex conjugate pairs. Throughout this paper, we will associate real eigenvectors with real eigenvalues. For simplicity, we assume that the quadratic eigenvalue problem is nondefective. This assumption is neither unduly restrictive (almost all systems are nondefective) nor essential. We refer to the examples in [11, 12] for a glimpse on how to handle defective situations and to [21] for the details.

The solution of (1.1) with $f(t) = 0$ can be written as

\begin{equation}
q(t) = \sum_{j=1}^{2n} v_je^{\lambda_j t} \gamma_j,
\end{equation}

where $\gamma_j, j = 1, \ldots, 2n$, are constants determined by the initial conditions [2, 18]. Real initial conditions force those constants $\gamma_j$, associated with complex conjugate eigensolutions, to likewise occur in complex conjugate pairs. If all eigenvalues are complex, the solution becomes

\begin{equation}
q(t) = \sum_{j=1}^{n} \left(v_je^{\lambda_j t} \gamma_j + \bar{v}_je^{\bar{\lambda}_j t} \bar{\gamma}_j\right).
\end{equation}

Every summand is real and represents a vibration pattern, observable in experiments. Following [11, 12], we refer to a summand

\begin{equation}
s_j(t) = ve^{\lambda_j t} \gamma_j + \bar{v}_je^{\bar{\lambda}_j t} \bar{\gamma}_j = 2\Re(v_je^{\lambda_j t} \gamma_j)
\end{equation}

as a mode. We hope this terminology will not cause additional confusion, since the word “mode” has been used in different contexts before. For example $v_je^{\lambda_j t}$ is sometimes called a complex mode.

The real modes $s_j(t)$ play a key role in decoupling (1.1). When vibrating in a mode, each component performs exponentially decaying or growing harmonic motion with the same frequency and the same exponential decay. However, there is a constant phase difference between any two components. The key, but radical idea is to synchronize all modes $s_j(t)$ by evaluating each component $s_j^K(t)$ at a different, but fixed time lag. We call this process phase synchronization. Phase synchronization generates a nonlinear mapping to decouple the system. The main attraction in phase synchronization is thus not the formalism itself, but the decoupling algorithm we will extract from it.
Our terminology is not to be confused with the phenomenon of synchronization in nonlinear systems (we specifically mention the work of Blekhman; see, e.g., [22] and references therein). The main idea in phase synchronization is to enforce a synchronization by appropriate manipulation of the modes given by (2.4). Because the idea of phase synchronization is new and somewhat unconventional, the derivations of principal formulas will be presented in sufficient detail. We drop the index \( j \) on \( s(t) \) temporarily because all modes evolve independently of each other.

2.1. Phase synchronization of complex conjugate eigensolutions. A pair of complex conjugate eigenvalues

\[
\lambda, \bar{\lambda} = \alpha \pm i\omega, \quad \omega > 0,
\]

generates a real mode \( s(t) \) given by (2.4). We wish to physically interpret \( s(t) \) and, for this reason, we express the eigenvector \( v \) and the scalar \( \gamma \) in quasi-polar form,

\[
v = r \circ \exp(-il), \quad l = (\eta_1, \ldots, \eta_n)^T, \quad |\eta_j| \leq \pi/2,
\]

\[
\gamma = \frac{1}{2} \rho e^{-i\theta}, \quad |\theta| \leq \pi/2.
\]

Combining (2.5), (2.6), and (2.7) gives the intermediate result

\[
ve^{\lambda t} \gamma = (r \circ \exp(-il)) \exp(\alpha t + i\omega t) \frac{1}{2} \rho \exp(-i\theta).
\]

Since all parameters on the right-hand side of the above equation are real, we can rewrite (2.4) with the aid of the vector \( u = (1, \ldots, 1)^T \) in the desired form

\[
s(t) = r \circ \cos(u(\omega t - \theta) - l)\rho e^{\alpha t}.
\]

The constant phase difference between any two system components is now obvious. Phase synchronization eliminates these phase differences by evaluating each component \( s^k(t) \) at a different, but fixed time lag. Upon phase synchronization, we obtain a synchronized vector \( y(t) \) with components

\[
y^k(t) = s^k(t + \eta_k/\omega)
\]

\[
= r_k e^{\alpha \eta_k/\omega} \rho e^{\alpha t} \cos(\omega(t + \eta_k/\omega) - \theta + \eta_k)
\]

\[
= r_k e^{\alpha \eta_k/\omega} \rho e^{\alpha t} \cos(\omega t - \theta).
\]

More simply, we write

\[
y(t) = z\psi(t),
\]

with

\[
\psi(t) = \rho e^{\alpha t} \cos(\omega t - \theta), \quad z = \exp(\alpha^{-1} t) \circ r.
\]

The above two equations highlight the synchronization. The argument of the cosine in \( y(t) \) is scalar, so that all components perform exponentially decaying (or growing) harmonic motion with the same frequency, passing through their equilibrium positions at the same time. In other words, the mode \( y(t) \) physically represents synchronous motion of all system components.
To invert the synchronization, we apply the time-shifting operation

\[ s^k(t) = y^k(t - \eta_j/\omega). \]  

Combination of (2.13) and (2.12) gives a formula for the mode \( s(t) \) in terms of \( \psi(t) \),

\[ s(t) = (\psi(t - \eta_1/\omega), \ldots, \psi(t - \eta_n/\omega))^T \circ z. \]

The above equation plays an important role in the decoupling of system (1.1).

### 2.2. Phase synchronization of two real and distinct eigensolutions.

Before we can proceed to decouple (1.1), we must consider modes generated by real eigenvalues. It may appear natural to think of a real eigensolution \( ve^{\lambda t} \) as a mode (see section 2.4). However, to decouple (1.1) into a set of independent second-order equations, two linearly independent eigensolutions have to be paired up to generate a mode of the form

\[ s(t) = v_a e^{\lambda_a t} \gamma_a + v_b e^{\lambda_b t} \gamma_b, \quad \lambda_a \neq \lambda_b. \]

The goal is to derive a synchronized vector \( y(t) \) of the functional form (2.11). The phase shifts necessary to enforce the synchronization are, however, not obvious. A series of algebraic manipulations, summarized below, permits the necessary phase shifts to manifest themselves. The results in (2.23)–(2.26) should be noted.

For the generic case (all \( n \) components of \( v_a, v_b \) being nonzero), consider the \( k \)th component of (2.15),

\[ s^k(t) = \exp \left( \frac{\lambda_a + \lambda_b}{2} t \right) \times \left( v_a^k \exp \left( \frac{\lambda_a - \lambda_b}{2} t \right) \gamma_a + v_b^k \exp \left( -\frac{\lambda_a - \lambda_b}{2} t \right) \gamma_b \right). \]

Whenever \( v_a^k \) or \( v_b^k \) equals zero, the corresponding component \( s^k(t) \) is either \( v_a^k e^{\lambda_a t} \gamma_a \) or \( v_b^k e^{\lambda_b t} \gamma_b \) and no extra work is required. In analogy to section 2.1, we define a vector \( y(t) \) by evaluating each component of \( s(t) \) with a constant, but yet unknown time lag \( \tau_k \) so that

\[ y^k(t) = s^k(t + \tau_k) \]

\[ = \exp \left( \frac{-\lambda_a + \lambda_b}{2} (t + \tau_k) \right) \left( v_a^k \exp \left( \frac{\lambda_a - \lambda_b}{2} \tau_k \right) \exp \left( \frac{-\lambda_a + \lambda_b}{2} t \right) \gamma_a \right. 
\]

\[ + v_b^k \exp \left( -\frac{\lambda_a - \lambda_b}{2} \tau_k \right) \exp \left( -\frac{\lambda_a - \lambda_b}{2} t \right) \left. \gamma_b \right). \]

To achieve a synchronizing effect in \( y(t) \) without disturbing the coefficients \( \gamma_a \) and \( \gamma_b \) in (2.15), we require that

\[ v_a^k \exp \left( \frac{\lambda_a - \lambda_b}{2} \tau_k \right) = v_b^k \exp \left( -\frac{\lambda_a - \lambda_b}{2} \tau_k \right). \]

We solve the above equation for \( \tau_k \) using the principal value of the logarithm of a complex number, \( \ln z = \ln |z| + i \text{Arg}(z) \), and obtain

\[ \tau_k = \frac{\ln |v_b^k/v_a^k|}{\lambda_a - \lambda_b} \times \begin{cases} 1, & \text{for } v_a^k v_b^k > 0, \\ i\pi, & \text{for } v_a^k v_b^k < 0. \end{cases} \]
The above \( \tau_k \) yields the intermediate results

\[
(2.20) \quad v_a^k \exp \left( \frac{\lambda_a - \lambda_b}{2} \tau_k \right) = \text{sign}(v_a^k) \sqrt{|v_a^k v_b^k|} \times \begin{cases} 1, & \text{for } v_a^k v_b^k > 0, \\ i, & \text{for } v_a^k v_b^k < 0, \end{cases}
\]

and

\[
(2.21) \quad \exp \left( \frac{\lambda_a + \lambda_b}{2} \tau_k \right) = \sqrt{|v_a^k v_b^k|} \times \begin{cases} 1, & \text{for } v_a^k v_b^k < 0, \\ \exp \left( \frac{i \pi (\lambda_a + \lambda_b)}{\lambda_a - \lambda_b} \right), & \text{for } v_a^k v_b^k > 0, \end{cases}
\]

Substitution of (2.20) and (2.21) into (2.17) now provides the sought-after expression for the components

\[
(2.22) \quad y^k(t) = (e^{\lambda_a t} \gamma_a + e^{\lambda_b t} \gamma_b) \times \sqrt{|v_a^k v_b^k|} \times \begin{cases} 1, & \text{for } v_a^k v_b^k > 0, \\ \exp \left( \frac{i \pi (\lambda_a + \lambda_b)}{\lambda_a - \lambda_b} \right), & \text{for } v_a^k v_b^k < 0, \end{cases}
\]

of the synchronized mode. By time shifting every component, we again obtain equation (2.11) with

\[
(2.23) \quad \psi(t) = \gamma_a e^{\lambda_a t} + \gamma_b e^{\lambda_b t}, \quad z = (\zeta_1, \ldots, \zeta_n)^T,
\]

where

\[
(2.24) \quad \zeta_k = \text{sign}(v_a^k) \sqrt{|v_a^k v_b^k|} \times \begin{cases} 1, & \text{for } v_a^k v_b^k < 0, \\ \exp \left( \frac{i \pi (\lambda_a + \lambda_b)}{\lambda_a - \lambda_b} \right), & \text{for } v_a^k v_b^k > 0, \end{cases}
\]

Straightforward calculation shows that the inverse formulas (2.13) and (2.14) remain valid with the substitutions

\[
(2.25) \quad l = (\eta_1, \ldots, \eta_n)^T, \quad \eta_k = \begin{cases} \ln \frac{v_b^k}{v_a^k}, & \text{for } v_a^k v_b^k > 0, \\ \ln \frac{v_a^k}{v_b^k} + i \pi, & \text{for } v_a^k v_b^k < 0, \end{cases}
\]

\[
(2.26) \quad \omega = \lambda_a - \lambda_b \neq 0.
\]

Finally, we would like to point out that phase synchronization reduces to the identity transformation if two distinct eigenvalues share the same real eigenvector.

\section*{2.3. Computation of homogeneous solution by phase synchronization.}

We wish to express the homogeneous solution in terms of the modes \( s(t) \) we just defined. Suppose \( 2d \) of the \( 2n \) eigenvalues of (2.1) are complex and \( 2d = 2(n - o) \) are real (the unusual indices \( o \) and \( d \) are used to denote oscillatory and nonoscillatory decaying solutions). For simplicity, we index the eigenvalues such that \( \lambda_1, \ldots, \lambda_o \) are complex with increasing positive imaginary parts; \( \lambda_{o+1} \leq \cdots \leq \lambda_n \leq \lambda_{n+o+1} \leq \cdots \leq \lambda_{2n} \) are real and \( \lambda_{n+1}, \ldots, \lambda_{n+o} \) are the complex conjugates of \( \lambda_1, \ldots, \lambda_o \). The indexing scheme is graphically illustrated in Figure 2.1. For \( j = 1, \ldots, o \), define \( s_j(t) \) by substituting \( \lambda_j, v_j \) in (2.4). Expressions for \( \rho_j, \theta_j, \alpha_j, \omega_j, r_j \), and

\[
(2.27) \quad l_j = (\eta_j^1, \ldots, \eta_j^n)
\]
can be inferred from (2.5), (2.6), and (2.7) by simply reintroducing the index $j$. Using this notation, define $\psi_j(t)$ and $z_j$ for $j = 1, \ldots, o$ by (2.12). Similarly, for $j = o+1, \ldots, n$, define $s_j(t)$ by substituting $\lambda_j, v_j$ for $\lambda_o, v_o$ and $\lambda_{j+n}, v_{j+n}$ for $\lambda_b, v_b$ in (2.15). With the same substitution, define $\psi_j$ and $z_j$ by (2.23), $l_j$ by (2.25), and, finally, $\omega_j$ by (2.26). Assuming that system (2.1) is nondefective, this pairing of real eigenvalues guarantees that eigensolutions with distinct real eigenvalues generate a mode, as required by its definition. We have thus shown that the homogeneous solution $q(t)$ is the superposition of $n$ real modes such that

\begin{equation}
q(t) = \sum_{j=1}^{n} s_j(t).
\end{equation}

Note that the above equation is valid for systems with complex and real eigenvalues. By using a similar notation for both real and complex eigensolutions, we are able to write down streamlined formulas for the general case. In particular, with $\omega$ in (2.14) representing either the imaginary part of a complex eigenvalue or the difference between two real and distinct eigenvalues (see (2.26)), we obtain

\begin{equation}
q(t) = \sum_{j=1}^{n} (\psi_j(t - \eta_j^1/\omega_j), \ldots, \psi_j(t - \eta_j^n/\omega_j))^T \circ z_j,
\end{equation}

with $\eta_j^k$ as in (2.27). The above equation represents a mapping from the set of mutually independent functions $\psi_j(t)$ to the homogeneous solution $q(t)$ of (1.1) and is linear in the function space spanned by $\{v_j e^{\lambda_j t} \}, j = 1, \ldots, 2n$. The equation appeared for the first time in [11], however, in a more restrictive setting and in different notation.

2.4. Choice of pairing schemes. As indicated in section 2.2, any two distinct real eigensolutions may be paired up to generate a real mode $s_j(t)$ according to (2.15). In section 2.3, we have described only one way. If all real eigenvalues are distinct, there are, in fact, $(2d!)/(2^d d!)$ ways to pair the $2d$ distinct eigensolutions to generate $d$ real modes. Each pairing scheme may generate a different set of modes and, consequently,
a different set of functions \( \psi_j(t) \). Formally, any such set of functions may be used. In applications, the pairing of real eigenvalues warrants further consideration. For example, when the coefficients \( M, C, \) and \( K \) are SPD, the commutativity condition in (1.2) guarantees that the \( 2n \) eigenvalues occur in \( n \) pairs, each pair sharing a common real eigenvector (even if some or all eigenvalues are complex; see [8]). By using this pairing scheme, phase synchronization becomes the identity transformation for all \( n \) modes, rendering this specific pairing particularly easy and the scheme of choice in vibration books. A real, invertible, linear mapping (coinciding with classical modal analysis) to decouple the system can be derived from this pairing of eigensolutions. In the general case ((1.2) not satisfied), phase synchronization is far from being the identity transformation, regardless of the pairing scheme.

It is also reasonable to describe the nonoscillatory part of the solution as a linear combination of \( 2d \) real eigensolutions \( \psi_\lambda(t) \), i.e., avoid the somewhat artificial modes \( s(t) \) in section 2.2 altogether. The price for this simplification is to have \( o+2d \) terms in (2.28). With appropriate modifications, \( o+2d \) functions \( \psi_j(t) \) can be used to derive an equation similar to (2.29). In other words, phase synchronization may be confined to the oscillatory part of the solution. Following this approach we ultimately generate a set of \( o+2d \) mutually independent equations, \( o \) of which are of second order, \( 2d \) of first order. The second-order structure inherent to (1.1) is thus lost. We prefer to apply phase synchronization to real eigensolutions and preserve the second-order structure.

3. The decoupling of linear dynamical systems. We now proceed to construct a real and diagonal system whose solution generates the solution of the original system (1.1). It is convenient to separate the homogeneous and the inhomogeneous cases.

3.1. Decoupling the homogeneous equation. In section 2.3, we found a mapping from the set of independent functions \( \psi_j(t) \) in (2.12) to the homogeneous solution of (1.1). Straightforward calculation reveals that the functions \( \psi_j(t) \) satisfy the second-order differential equations

\[
\ddot{\psi}_j(t) - (\lambda_j + \lambda_{j+n}) \dot{\psi}_j(t) + (\lambda_j \lambda_{j+n}) \psi_j(t) = 0, \quad j = 1, \ldots, n.
\]

All coefficients in the above equations are real. To streamline the notation, define

\[
\begin{aligned}
    p(t) &= (\psi_1(t), \ldots, \psi_n(t))^T, \\
    \Lambda_1 &= \text{diag}(\lambda_1, \ldots, \lambda_n), \\
    \Lambda_2 &= \text{diag}(\lambda_{n+1}, \ldots, \lambda_{2n}), \\
    D &= -(\Lambda_1 + \Lambda_2), \\
    \Omega &= \Lambda_1 \Lambda_2,
\end{aligned}
\]

and express the \( n \) mutually independent, scalar equations (3.1) in a compact matrix form,

\[
\ddot{p}(t) + D\dot{p}(t) + \Omega p(t) = 0.
\]

The above represents a decoupled system to which the homogeneous equation associated with (1.1) reduces.

The remaining task is to connect the initial conditions of (1.1) and (3.3). The initial conditions render \( q(t) \), and hence \( p(t) \), unique. By simply evaluating (2.29) and its derivative at \( t = 0 \), we can only connect \( q \) with \( p \) at different times. This difficulty can be avoided by observing that phase synchronization does not disturb the constants \( \gamma_j \) in (2.2). Specifically, from (2.2), we have

\[
\begin{pmatrix}
    q(0) \\
    \dot{q}(0)
\end{pmatrix} = \begin{pmatrix}
    V_1 & V_2 \\
    V_1 L_1 & V_2 L_2
\end{pmatrix} \begin{pmatrix}
    c_1 \\
    c_2
\end{pmatrix},
\]

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.
where the columns of $V_1, V_2 \in \mathbb{C}^{n \times n}$ consist of the eigenvectors of (2.1) such that
\begin{align}
V_1 &= (v_1, \ldots, v_n), \quad V_2 = (v_{n+1}, \ldots, v_{2n}),
\end{align}
and where $c_1^T = (\gamma_1, \ldots, \gamma_n)^T$, $c_2^T = (\gamma_{n+1}, \ldots, \gamma_{2n})^T$. On the other hand, (2.12) and (2.23) imply that
\begin{align}
\begin{pmatrix}
p(0) \\
\dot{p}(0)
\end{pmatrix} &=
\begin{pmatrix}
I & I \\
\Lambda_1 & \Lambda_2
\end{pmatrix}
\begin{pmatrix}
c_1 \\
c_2
\end{pmatrix}.
\end{align}
Combining (3.6) and (3.4) to eliminate $(c_1^T c_2^T)^T$ yields the desired real mapping of initial conditions:
\begin{align}
\begin{pmatrix}
p(0) \\
\dot{p}(0)
\end{pmatrix} &=
\begin{pmatrix}
I & I \\
\Lambda_1 & \Lambda_2
\end{pmatrix}
\begin{pmatrix}
V_1 & V_2 \\
V_1\Lambda_1 & V_2\Lambda_2
\end{pmatrix}^{-1}
\begin{pmatrix}
q(0) \\
\dot{q}(0)
\end{pmatrix}.
\end{align}
The decoupling is now complete. To solve the homogeneous equation associated with (1.1), solve the quadratic eigenvalue problem (2.1), construct and solve the real, diagonal system (3.3) with initial conditions (3.7), and map $p = (\psi_1, \ldots, \psi_n)^T$ back to $q$ using (2.29).

### 3.2. Decoupling the inhomogeneous equation.

The inhomogeneous part of (1.1) can always be decoupled by phase synchronization as described earlier. Thus the remaining task is to determine how the excitation $f(t)$ transforms under phase synchronization. Put differently, we postulate that the coupled system (1.1) can be decoupled into the form
\begin{align}
\begin{pmatrix}
p(t) \\
\dot{p}(t)
\end{pmatrix} + D\dot{p}(t) + \Omega p(t) = g(t),
\end{align}
where $D$ and $\Omega$ are given by (3.2) and where $g(t)$ is real. To find $g(t)$, recast (1.1) in a first-order form
\begin{align}
\begin{pmatrix}
\dot{q}(t) \\
\ddot{q}(t)
\end{pmatrix} &=
\begin{pmatrix}
0 & I \\
-M^{-1}K & -M^{-1}C
\end{pmatrix}
\begin{pmatrix}
q(t) \\
\dot{q}(t)
\end{pmatrix}
+ \begin{pmatrix}
0 \\
M^{-1}f(t)
\end{pmatrix}.
\end{align}
Should $M$ be ill-conditioned, other forms of first-order conversion may be used [23, 24]. Inspired by (3.7), we define a real and invertible mapping by
\begin{align}
\begin{pmatrix}
q(t) \\
\dot{q}(t)
\end{pmatrix} &=
\begin{pmatrix}
V_1 & V_2 \\
V_1\Lambda_1 & V_2\Lambda_2
\end{pmatrix}
\begin{pmatrix}
I & I \\
\Lambda_1 & \Lambda_2
\end{pmatrix}^{-1}
\begin{pmatrix}
p_1(t) \\
p_2(t)
\end{pmatrix}.
\end{align}
Substitution into (3.9) yields the equations
\begin{align}
p_2(t) &= \dot{p}_1(t) - g_1(t), \\
\ddot{p}_1(t) + Dp_1(t) + \Omega p_1(t) &= (D + Id/dt)g_1(t) + g_2(t),
\end{align}
where $g_1(t)$ and $g_2(t)$ are given by
\begin{align}
g_1(t) &= ((V_1\Lambda_1 - V_2\Lambda_2V_2^{-1}V_1)^{-1} + (V_2\Lambda_2 - V_1\Lambda_1V_1^{-1}V_2)^{-1})M^{-1}f(t), \\
g_2(t) &= (\Lambda_1(V_1\Lambda_1 - V_2\Lambda_2V_2^{-1}V_1)^{-1} + \Lambda_2(V_2\Lambda_2 - V_1\Lambda_1V_1^{-1}V_2)^{-1})M^{-1}f(t).
\end{align}
Note that $g_1(t)$ and $g_2(t)$ are real and depend continuously on $f(t)$. Specifically, (3.11) defines how $p_2(t)$ is connected to the displacements and velocities of the decoupled system (3.8), while (3.12) represents the dynamics of the decoupled system, i.e., $p_1(t) = p(t)$. We thus obtained (3.8) with
\begin{align}
g(t) &= (D + Id/dt)g_1(t) + g_2(t).
\end{align}
The mapping from $p$ to $q$, as inferred from (3.10), (3.11), and (3.12), is

(3.15) \[ q(t) = (T_1 + T_2 \frac{d}{dt})p(t) - T_2 g_1(t), \]

where the real matrices $T_1$ and $T_2$ are given by

(3.16) \[ T_1 = (V_1 \Lambda_2 - V_2 \Lambda_1)(\Lambda_2 - \Lambda_1)^{-1}, \]

(3.17) \[ T_2 = (V_2 - V_1)(\Lambda_2 - \Lambda_1)^{-1}. \]

As a mapping between $p$ and $q$, (3.15) is real, time dependent, and nonlinear. The initial conditions of (1.1) and (3.8) are connected by

(3.18) \[ \begin{pmatrix} p(0) \\ \dot{p}(0) \end{pmatrix} = \begin{pmatrix} I & I \\ \Lambda_1 & \Lambda_2 \end{pmatrix} \begin{pmatrix} V_1 & V_2 \\ V_1 \Lambda_1 & V_2 \Lambda_2 \end{pmatrix}^{-1} \begin{pmatrix} q(t) \\ \dot{q}(0) \end{pmatrix} + \begin{pmatrix} 0 \\ g_1(0) \end{pmatrix}, \]

thus completing the decoupling of (1.1). The formulas for decoupling presented in this section are direct generalizations of those given in [12], applicable only when $M$, $C$, and $K$ are SPD.

3.3. Nonlinearity and nonuniqueness in decoupling. Based upon physical intuition, the dependence of the nonlinear mapping (3.15) on the excitation $f(t)$ can be explained as follows. If $f(t) = 0$, viscous damping and gyroscopic forces induce constant time shifts between the components of a mode $s(t)$. These time shifts are accounted for by the linear mapping (2.29). For $f(t) \neq 0$, the mapping must account for additional and not necessarily constant time shifts caused by the external force $f(t)$. Thus, a mapping leading to the decoupling of the inhomogeneous equation must depend on $f(t)$. If $f(t) = 0$, the nonlinear mapping in (3.15) for the inhomogeneous equation reduces to the linear mapping in (2.29).

In decoupling a homogeneous system, we have observed a degree of nonuniqueness when generating the modes. This nonuniqueness is carried over to the decoupling of (1.1). Recall that two systems with identical eigenvalues and multiplicities are termed strictly isospectral. Since the property of being strictly isospectral is reflexive, transitive, and symmetric, strictly isospectral systems generate an equivalence class [15]. It is easy to verify that system (1.1) and the decoupled systems (3.3) are strictly isospectral regardless of the pairing of real eigensolutions during phase synchronization. Indeed, every system within the equivalence class can be generated by suitably pairing the real eigensolutions. Thus, phase synchronization generates the path to all real and diagonal systems within the equivalence class of systems strictly isospectral to (1.1).

Finally, it is important to point out that decoupling by phase synchronization reduces to a linear mapping if $M$, $C$, and $K$ are simultaneously diagonalizable to real diagonal matrices by a real equivalence transformation (because phase synchronization becomes the identity in this case).

3.4. Decoupling algorithm. The procedure for decoupling the linear system (1.1) by phase synchronization may be summarized as an algorithm.

1. Solve the quadratic eigenvalue problem (2.1), enumerate the eigenvalues as in Figure 2.1, and generate the real and diagonal system (3.3).

2. If $f(t) = 0$, decoupling is complete. The solution $p$ of (3.3), with initial conditions (3.7) can be easily obtained. We may recover the homogeneous solution $q$ of (1.1) from $p$ by simply evaluating (2.29).
If $f(t) \neq 0$, the decoupled system is (3.8), which can be obtained from (3.3) by incorporating an excitation $g(t)$ given by (3.14). The solution $p$ of the real and diagonal system (3.8), with initial conditions (3.18), can be readily computed. We may recover the solution $q$ of (1.1) from $p$ by using (3.15).

The algorithm is graphically illustrated in the flowchart in Figure 3.1. Although complex quantities appear in the algorithmic development, the entire process can be implemented in real arithmetic. Computing the solution by decoupling is particularly attractive if (1.1) is very stiff (eigenvalues vary over many orders of magnitude). In this case, different time stepping methods can be applied to each decoupled equation, resulting in substantial improvement in algorithms for response computation. In
addition, low-energy components of \( p \) may be neglected to obtain powerful model reduction schemes (see section 5).

4. Phase synchronization and SPT. Garvey et al. [13, 14] defined SPT and applied them to decouple certain linear dynamical systems of second order. In our notation, an equivalence transformation \( \{ U_L, U_R \} \) is structure preserving if and only if

\[
U_L^T \begin{pmatrix} 0 & K \\ K & C \end{pmatrix} U_R = \begin{pmatrix} 0 & K_D \\ K_D & C_D \end{pmatrix},
\]

(4.1)

\[
U_L^T \begin{pmatrix} K & 0 \\ 0 & -M \end{pmatrix} U_R = \begin{pmatrix} K_D & 0 \\ 0 & -M_D \end{pmatrix},
\]

(4.2)

\[
U_L^T \begin{pmatrix} C & M \\ M & 0 \end{pmatrix} U_R = \begin{pmatrix} C_D & M_D \\ M_D & 0 \end{pmatrix},
\]

(4.3)

where \( U_L \) and \( U_R \) are real, invertible \( 2n \times 2n \) matrices and \( M_D, C_D, \) and \( K_D \) are real \( n \times n \) matrices. It is easy to check that the eigenvalues (and their multiplicities) of (2.1) remain the same if \( M, C, \) and \( K \) are replaced, respectively, by \( M_D, C_D, \) and \( K_D \). Thus, SPTs are strictly isospectral. An SPT is termed diagonalizing if \( M_D, C_D, \) and \( K_D \) are diagonal. To illustrate how a diagonalizing SPT decouples (1.1) for \( f(t) = 0 \), consider the first-order formulation

\[
\begin{pmatrix} C & M \\ M & 0 \end{pmatrix} \begin{pmatrix} \dot{q}(t) \\ \ddot{q}(t) \end{pmatrix} \begin{pmatrix} K & 0 \\ 0 & -M \end{pmatrix} \begin{pmatrix} q(t) \\ \dot{q}(t) \end{pmatrix} = 0.
\]

(4.4)

A diagonalizing SPT leads to

\[
\begin{pmatrix} C_D & M_D \\ M_D & 0 \end{pmatrix} \begin{pmatrix} \dot{q}_D(t) \\ \ddot{q}_D(t) \end{pmatrix} \begin{pmatrix} K_D & 0 \\ 0 & -M_D \end{pmatrix} \begin{pmatrix} q_D(t) \\ \dot{q}_D(t) \end{pmatrix} = 0,
\]

(4.5)

from which the decoupled second-order equation,

\[
M_D \ddot{q}_D(t) + C_D \dot{q}_D(t) + K_D q_D(t) = 0,
\]

(4.6)

can be extracted. It was pointed out in [17] that current algorithms for constructing diagonalizing SPTs can be quite restrictive and the conditions for convergence are not fully understood.

We have argued that the decoupled system (3.3) is the unique real and diagonal system, isospectral to (1.1). Thus, a diagonalizing SPT must generate the same decoupled system as phase synchronization. It is indeed straightforward to show that \( C_D = D \) and \( K_D = \Omega \), provided we choose \( M_D = I \), which can be done without loss of generality. Thus, a diagonalizing SPT leads to the same decoupled system as phase synchronization.

This somewhat surprising observation can be explained as follows. A diagonalizing SPT lives in the state space, i.e., the space of dimension \( 2n \) spanned by displacements and velocities. Phase synchronization, on the other hand, is a method in the configuration space of dimension \( n \) spanned only by the displacements. Any procedure in configuration space can, of course, be executed in state space (but not vice versa). We thus may interpret diagonalizing SPTs to be the state space version of phase synchronization. Starting from phase synchronization, a diagonalizing SPT can be constructed, if it is available. The mapping in (3.7), for example, defines a diagonalizing structure-preserving congruence transformation for symmetric and positive definite \( M, C, \) and \( K \) and a clever normalization of eigenvectors. On the other hand, it is generally not possible to obtain phase synchronization from diagonalizing SPTs [25].
The decoupling in state space by SPTs or the decoupling by phase synchronization in configuration space are not to be confused with diagonalization of the first-order state equation by similarity transformation (see, e.g., [2, 3]). The decoupling of the state equation, which is also routinely used in control theory, does not necessarily imply decoupling in configuration space. For example, state decoupling produces complex equations without physical meaning, whereas phase synchronization or SPTs are real.

5. Applications in earthquake engineering. The response of a building to earthquake excitation is of great importance in structural engineering. Although the dynamic properties are nonlinear, experimental studies using shaking tables have shown that the system response can often be simulated to a satisfactory degree of accuracy by a linear viscoelastic model. The equation of motion (1.1) is thus widely used in earthquake engineering to model the dynamic behavior of, for example, multi-story buildings, nuclear power plants, or base-isolated structures [5, 26, 27, 28, 29, 30, 31]. In such applications, \( M \) and \( K \) represent, respectively, the inertia and elastic properties of the structure under investigation, and \( C \) describes the energy dissipation. All three system matrices are SPD. Practically speaking, the commutativity condition (1.2) is satisfied if energy dissipation is almost uniformly distributed throughout the system. This condition is violated for systems consisting of two or more subsystems with significantly different levels of damping. We shall consider two examples in the analysis of (1.1) when (1.2) is not satisfied.

5.1. Response of light equipment in a base-isolated structure. The use of base isolation is known to attenuate not only the response of the building (termed the primary structure), but also the response of a secondary system mounted on the primary structure, for example, internal equipment. The isolation system, primary structure, and secondary system are usually made of different materials with significantly different energy dissipation characteristics. Hence the commutativity condition (1.2) is not satisfied [28, 29].

The classical engineering approach to the analysis of base-isolated systems is as follows. Solve the symmetric eigenvalue problem \( M \omega^2 = Ku \) to obtain \( n \) real eigenvectors \( u_j \) and \( n \) natural frequencies \( \omega_j \), \( j = 1, \ldots, n \). It is well known that the eigenvectors are orthogonal with respect to \( M \) or \( K \). Normalize the eigenvectors with respect to \( M \) and use the normalized eigenvectors as columns to construct a square matrix \( U = (u_1, \ldots, u_n) \). Upon modal transformation \( q(t) = Ux(t) \), the equation of motion (1.1) becomes

\[
\ddot{x}(t) + \tilde{C}\dot{x}(t) + \tilde{K}x(t) = U^Tf(t),
\]

where \( \tilde{C} = U^TCU \) is known as the modal damping matrix and \( \tilde{K} = U^TKU \) is a real and diagonal spectral matrix. The modal damping matrix \( \tilde{C} \) is diagonal if and only if (1.2) is satisfied [8]. When \( \tilde{C} \) is diagonal, (5.1) is decoupled and methods for analysis and design are readily available (see [5]). When \( \tilde{C} \) is not diagonal, (5.1) is often decoupled by simply neglecting the off-diagonal elements of \( \tilde{C} \). This procedure, termed the decoupling approximation, is relatively routine in structural engineering. The decoupling approximation appears intuitive if the off-diagonal elements in \( \tilde{C} \) are small in magnitude when compared to the diagonal elements; i.e., \( \tilde{C} \) is diagonally dominant. However, even when \( \tilde{C} \) is diagonally dominant, the errors can still be large and exhibit rather surprising behaviors [32, 33, 34].

Tsai and Kelly [28] investigated the validity of the decoupling approximation in computation of the seismic response of attached equipment in a base-isolated building.
In [28], they derived a linear viscoelastic model for a five-story, base-isolated building with internal equipment. The model has three degrees of freedom, representing the displacement of the base, the primary structure, and the equipment, respectively (see [28] for the details of the modeling process). The response of the equipment to the 1940 El Centro earthquake is studied in detail. Here, we utilize their formulation and apply the decoupling algorithm described in section 3.4 to this problem.

Given $M$, $C$, and $K$ as in [28], we first solve the quadratic eigenvalue problem (2.1). Using the eigendata, we construct the matrices $\Lambda_1$ and $\Lambda_2$ according to (3.2) and $V_1$ and $V_2$ according to (3.5). With an excitation $g(t)$ derived from (3.13) and (3.14), the decoupled equation (3.8) is obtained, where $D = \text{diag}(0.1910, 0.1644, 0.8178)$ and $\Omega = \text{diag}(0.3505, 33.0195, 33.7899)$. We apply an explicit Runge–Kutta (4,5) formula [35] (implemented in the Matlab function ode45) to solve the decoupled equations (3.8), i.e., compute $p(t)$. The time step is adaptively chosen for each independent decoupled coordinate and can be as large as 0.08. Equation (3.15) is used to map $p(t)$ back to $q(t)$. The upper part of Figure 5.1 shows the response of the attached equipment as computed by the method of phase synchronization. The lower part shows the results obtained by employing the decoupling approximation.

As already observed by Tsai and Kelly [28], the results obtained by the decoupling approximation underestimate the response of the equipment because the coupling in (5.1) is significant. In contrast, the method of phase synchronization generates the decoupled equation (3.8) by capturing all coupling effects. Using (3.8) instead of decoupling approximation, an accurate system response is computed. Furthermore, system solution by phase synchronization can be substantially more efficient than direct simulation because a different time step may be chosen for each independent decoupled equation. For example, the time step in the present example could be chosen twice as big for the decoupled equations. However, the main advantage of phase synchronization lies beyond response calculation: All streamlined design and analysis methods applicable only to independent single-degree-of-freedom equations can be used once (3.8) is obtained.
5.2. A simplified linear viscoelastic model of a nuclear power plant.

We now consider the seismic response of a nuclear power plant. We use the 8 degree-of-freedom linear viscoelastic model presented in [26] and shown in Figure 5.2 for convenience. The model consists of four interconnected rigid structures representing the core, prestressed concrete pressure vessel (PCPV), basement, and building, respectively. Each structure has two degrees of freedom representing the sway ($U$) and rocking angle ($\Phi$). The coefficient matrices $M$, $C$, and $K$ are given in [26]. We are interested in the response of the core ($U_1$) when the structure is subject to the 1940 El Centro earthquake.

We compare three approaches to this problem. First, we simulate the response of the system using the Runge–Kutta formula already mentioned in the previous section. The results serve as the benchmark solution and are shown in the second uppermost part of Figure 5.3. Next, we apply the method of phase synchronization and decouple the system to obtain (3.8). Upon solution of the decoupled equations using the same Runge–Kutta formula as above, we compute the energy

\[
E_j = \frac{1}{2} \int_0^T \dot{p}_j(t)^2 + \Omega_j p_j(t)^2 dt
\]

in each decoupled coordinate $p_j(t)$. In the above equation, $T = 10$, and $\Omega_j$ is the $j$th diagonal entry of $\Omega$ in (3.2). The results are shown in the top part of Figure 5.3, and it is clear that the coordinates $p_1(t)$ and $p_5(t)$ carry most of the energy within the system. As a result, an approximation to the response can be obtained by solving only the first five decoupled equations (which are associated with the lowest frequencies), followed by transformation to $q(t)$ using (3.15). The results of this approximation, shown in the third uppermost part of Figure 5.3, are in very good agreement with the core response obtained by direct numerical simulation.

Finally, we can obtain another decoupled system using the decoupling approximation. When the first five coordinates of this decoupled system are used to generate the system response, the resulting approximation dramatically underestimates the core response, as is shown in the bottom part of Figure 5.3. Perhaps this is not surprising. The decoupling approximation produces a decoupled system by neglecting the coupling caused by nonzero off-diagonal elements of the modal damping matrix $\tilde{C}$.

Fig. 5.2. An 8 degree-of-freedom model of a nuclear power plant (taken from [26]).
In contrast, the method of phase synchronization generates a decoupled equation by incorporating all coupling effects properly.

To be sure, direct numerical simulation certainly yields the correct core response, but the method of phase synchronization offers a lot more than just efficient evaluation of response. Incidentally, we have outlined in this example a new method of model reduction, i.e., solving fewer differential equations while achieving good accuracy in response computations. The decoupling approximation, although in popular use, is not feasible for model reduction unless coupling is negligible.

6. Conclusions. We have developed a theory and an algorithm to decouple almost all second-order linear dynamical systems. While previous work on decoupling has focused on simultaneous diagonalization of the three system matrices $M$, $C$, and $K$, our theory exploits the parameter $t$ (time), characteristic of a dynamical system. The decoupling is achieved by a real, invertible, but nonlinear mapping. This mapping simplifies to a real, linear time-invariant transformation when the three system matrices can be simultaneously diagonalized. All parameters required for the decoupling of a second-order system are obtained from the solution of a quadratic eigenvalue problem.

Unlike any methods of decoupling approximation, phase synchronization is an exact decoupling technique that accounts for all effects of coupling in full. As demonstrated by examples, the main advantage of decoupling lies beyond system solution. It is the possibility, for example, of model reduction, damping characterization, or stability design that would make system decoupling worthwhile.

Acknowledgments. We would like to thank members of the Applied Mathematics Group at the University of California at Berkeley and at Lawrence Berkeley National Laboratory for valuable comments. Opinions, findings, and conclusions are those of the authors and do not reflect the views of the sponsor.


