Chemical Engineering Science, Vol. 49, No. 3, pp. 343–361, 1994.
Printed in Great Britain.

0009–2509/94 $6.00 + 0.00

A GENERAL ANALYSIS OF EXACT NONLINEAR LUMPING IN CHEMICAL KINETICS

GENYUAN LI and HERSCHEL RABITZ†
Department of Chemistry, Princeton University, Princeton, NJ 08540, U.S.A.

and

JÁNOS TÓTH
Department of Computer Science, Agricultural Engineering Faculty, University of Agricultural Science, H-2103 Gödöllő, Páter K. u. 1, Hungary

(*First received* 1 *December* 1992; *accepted in revised form* 15 *July* 1993)

Abstract—A general analysis of exact nonlinear lumping is presented. This analysis can be applied to the kinetics of any reaction system with n species described by a set of first-order ordinary differential equations $d\mathbf{y}/dt = \mathbf{f}(\mathbf{y})$, where $\mathbf{y}$ is an n-dimensional vector and $\mathbf{f}(\mathbf{y})$ is an arbitrary n-dimensional function vector. We consider lumping by means of an $\hat{n}$ ($\hat{n} \leqslant n$)-dimensional arbitrary transformation $\hat{\mathbf{y}} = \mathbf{h}(\mathbf{y})$. The lumped differential equation system is $d\hat{\mathbf{y}}/dt = \mathbf{h}_{\mathbf{y}}(\bar{\mathbf{h}}(\hat{\mathbf{y}}))\mathbf{f}(\bar{\mathbf{h}}(\hat{\mathbf{y}}))$, where $\mathbf{h}_{\mathbf{y}}(\mathbf{y})$ is the Jacobian matrix of $\mathbf{h}(\mathbf{y})$, $\bar{\mathbf{h}}$ is a generalized inverse transformation of $\mathbf{h}$ satisfying the relation $\mathbf{h}(\bar{\mathbf{h}}) = I_{\hat{n}}$. Three necessary and sufficient conditions of the existence of exact nonlinear lumping schemes have been determined. The geometric and algebraic interpretations of these conditions are discussed. It is found that a system is exactly lumpable by $\mathbf{h}$ only if $\mathbf{h}(\mathbf{y}) = \mathbf{0}$ is its invariant manifold. A linear partial differential operator $A = \sum_{i=1}^{n} f_i(\mathbf{y})\partial/\partial y_i$ corresponding to $d\mathbf{y}/dt = \mathbf{f}(\mathbf{y})$ is defined. Using the eigenfunctions and the generalized eigenfunctions of A, the operator can be transformed to Jordan or diagonal canonical forms which give the lumped differential equation systems without determination of $\bar{\mathbf{h}}$. These approaches are illustrated by a simple example. The results of this analysis serve as a theoretical basis for the development of approaches for approximate nonlinear lumping.

1. INTRODUCTION

Chemical kinetics is central to the development as well as to the subsequent design and operation of chemical reactors. When complex feedstocks, especially those related to petroleum, are involved, it is impossible to consider each species and we have to group, or lump, many of the individual species into pseudo-species and then develop kinetics based on this lumped reaction model. This situation arises in other areas of physico-chemical phenomena leading to coupled differential equation models including aeronomy, combustion processes, molecular dynamics, etc. In the modeling of combustion in a three-dimensional flow, e.g. as the number of chemical species and the geometric complexity of the computational domain increases, the modeling of such systems becomes computationally prohibitive on even the largest supercomputer (Smooke and Giovangigli, 1990). Rabinowitz (1990) has shown that current supercomputers are incapable of handling complete species descriptions in three-dimensional flow solvers. The dimension reduction in the modeling of complex reaction flow problems is inevitable. In all these cases dimension reduced models serve the dual purpose of being computationally more practical, especially for repeated design studies, and in a fundamental vein they render the system to its essence.

In 1969, Wei and Kuo (1969) gave a lumping analysis of unimolecular reaction systems. The lumped species were certain linear combinations of the original ones. The terms "exact lumping" and "approximate lumping" were used by them to distinguish two cases of lumping corresponding to the lumped model either having error or not. Their work was extended by the authors to linear lumping transformations of arbitrary nonlinear reaction–diffusion systems (Li and Rabitz, 1989, 1990, 1991a, b, c). In this earlier work the lumped variables were still considered as certain linear combinations of the original ones. Although this recently developed linear lumping can be used for any system with arbitrary nonlinearities, the need for the existence of fixed invariant subspaces of the transpose of the Jacobian matrix of the differential equation system is overly strict for many nonlinear systems. Using a linear lumping transformation one cannot always find suitable lumping schemes with a low enough dimension for them, especially under some imposed constraints. However, in practice, researchers have found lumped empirical kinetic equations to express some nonlinear systems to quite good accuracy. This suggests that we may be able to achieve more flexibility and further simplification by introducing nonlinear transformations for lumping.

A general analysis of dimension reduction by arbitrary transformations is the purpose of this paper. Actually, this is an old subject in many areas such as

†Author to whom correspondence should be addressed.

control theory, econometrics, biology and ecology. In these areas the words *perfect* and *approximate aggregation* were used instead of exact and approximate lumping. Some necessary and sufficient conditions for the existence of perfect aggregation have been given (Luckyanov *et al.*, 1983; Iwasa *et al.*, 1987, 1989). However, the main goal of lumping in chemical kinetics and other areas is the determination of the lumping schemes. Unfortunately, for nonlinear lumping, these conditions could not directly be used to reach this goal. Following our previous work of linear lumping, we deduce similar conditions for the existence of arbitrary lumping transformations from a different perspective and tried to find their geometric and algebraic interpretations so that they could be used to determine the nonlinear lumping schemes. In spite of the clear difference, there is a close relationship between linear and nonlinear lumping transformations. Corresponding to a fixed invariant subspace there exists an invariant manifold in the state space for the nonlinear transformation which should reduce the dimension of the original system. This manifold can be qualitatively viewed as a "curved subspace". Compared to the restriction of a linear subspace, obviously, the introduction of a manifold has more freedom to satisfy the imposed requirements. Therefore, for most nonlinear differential equation systems, the possibility of the existence for these manifolds should be higher than that of fixed invariant subspaces for the transpose of the Jacobian matrix. In order to make a distinction with linear lumping, we refer to this approach as *nonlinear lumping*. Actually, a linear transformation is included as a special case in it. Even though the possibility of the existence of exact nonlinear lumping schemes is higher than linear ones, the resultant exact lumping schemes may not meet practically desired goals. Approximate nonlinear lumping is more useful for realistic systems. In this paper we present a general analysis of exact nonlinear lumping which can serve as a starting point for developing approaches of approximate nonlinear lumping.

Like linear lumping, the lumped differential equation system is usually expressed by a formula containing the inverse of the lumping transformation. For a nonlinear transformation, the analytic expressions of the generalized inverse transformations may not exist. To avoid this difficulty, a linear partial differential operator $A = \sum_{i=1}^{n} f_i(\mathbf{y})\partial/\partial y_i$ corresponding to the unlumped n-dimensional ordinary differential equation system $\mathrm{d}\mathbf{y}/\mathrm{d}t = \mathbf{f}(\mathbf{y})$ and its eigenfunctions and generalized eigenfunctions are defined. After giving appropriate definitions of the canonical forms of A and by using its eigenfunctions and generalized eigenfunctions, the operator A can be transformed to Jordan or diagonal canonical forms that give lumped differential equation systems without using the generalized inverse of the lumping transformation.

The present paper is organized as follows. Three necessary and sufficient conditions for the existence of exact nonlinear lumping schemes and a discussion of constrained exact nonlinear lumping are presented in Section 2. The definition and determination of generalized eigenfunctions and their application in the transformation of A to a diagonal canonical form are given in Section 3. Similarly, the definition and determination of eigenfunctions and their application in the transformation of A to a Jordan canonical form are given in Section 4. An example is utilized in Section 5 to illustrate these approaches. Finally, Section 6 presents conclusions and a discussion.

As this paper is written for chemical engineers and chemists, some proofs may seem intuitive for a mathematician and contain tacit assumptions (e.g. most of the functions appeared are assumed to be differentiable or even analytic) to avoid complication.

2. NECESSARY AND SUFFICIENT CONDITIONS FOR THE EXISTENCE OF EXACT NONLINEAR LUMPING

We give three necessary and sufficient conditions for the existence of exact nonlinear lumping of a given ordinary differential equation system. Using these conditions we prove the existence of an invariant manifold if a system can be reduced by an arbitrary transformation. We also show the feasibility of developing an approach to determine constrained nonlinear lumping schemes.

2.1. *The necessary and sufficient conditions*

The kinematics of many dynamic systems with n dependent variables can often be described by an n-dimensional ordinary differential equation system.

$$\frac{\mathrm{d}\mathbf{y}}{\mathrm{d}t} = \mathbf{f}(\mathbf{y}), \quad \mathbf{y} \in \mathscr{R}^n \tag{1}$$

where $\mathbf{f}(\mathbf{y})$ is $\mathscr{C}^2$ with $\mathbf{f}(\mathbf{0}) = \mathbf{0}$. The assumption about the location of a zero of $\mathbf{f}(\mathbf{y})$ at the origin is either fulfilled by the system (e.g. a chemical reaction system) or realized by a simple transformation to bring a known stationary point to the origin (although the practical problem in calculating the stationary point may arise).

We also assume that all the solutions of eq. (1) can be defined for all possible value of t. This property directly follows from the law of mass conservation in the case of kinetic differential equations.

For a given $\hat{n}$ ($\hat{n} \leqslant n$)-dimensional nondegenerate $\mathscr{C}^2$ transformation,

$$\hat{\mathbf{y}} = \mathbf{h}(\mathbf{y}) \tag{2}$$

with $\mathbf{h}(\mathbf{0}) = \mathbf{0}$, if one can find an $\hat{n}$-dimensional differential equation system

$$\frac{\mathrm{d}\hat{\mathbf{y}}}{\mathrm{d}t} = \hat{\mathbf{f}}(\hat{\mathbf{y}}) \tag{3}$$

to describe the evolution of $\hat{\mathbf{y}}$, then we say that the system of eq. (1) is exactly lumpable by $\mathbf{h}$. (By nondegenerate we mean that $h_1, \ldots, h_{\hat{n}}$ are functionally independent.)

According to this definition, we give one of the necessary and sufficient conditions of the existence of

exact lumping for a system as eq. (1). From eq. (2), we have

$$\frac{d\hat{\mathbf{y}}}{dt} = \frac{d\mathbf{h}(\mathbf{y})}{dt} = \mathbf{h}_\mathbf{y}(\mathbf{y})\frac{d\mathbf{y}}{dt}$$
$$= \mathbf{h}_\mathbf{y}(\mathbf{y})\mathbf{f}(\mathbf{y}) \tag{4}$$

where $\mathbf{h}_\mathbf{y}(\mathbf{y})$ is the Jacobian matrix of $\mathbf{h}(\mathbf{y})$, i.e.

$$\mathbf{h}_\mathbf{y}(\mathbf{y}) = \frac{\partial \mathbf{h}(\mathbf{y})}{\partial \mathbf{y}} = \begin{pmatrix} \frac{\partial h_1}{\partial y_1} & \cdots & \frac{\partial h_1}{\partial y_n} \\ \cdots & \cdots & \cdots \\ \frac{\partial h_{\hat{n}}}{\partial y_1} & \cdots & \frac{\partial h_{\hat{n}}}{\partial y_n} \end{pmatrix}. \tag{5}$$

Comparing eqs (3) and (4) gives

$$\hat{\mathbf{f}}(\hat{\mathbf{y}}) = \mathbf{h}_\mathbf{y}(\mathbf{y})\mathbf{f}(\mathbf{y}). \tag{6}$$

Let $\bar{\mathbf{h}}$ be a mapping from $\mathscr{R}^{\hat{n}}$ to $\mathscr{R}^n$, i.e.

$$\mathbf{y} = \bar{\mathbf{h}}(\hat{\mathbf{y}}) \tag{7}$$

that satisfies

$$\mathbf{h}(\bar{\mathbf{h}}) = I_{\hat{n}}. \tag{8}$$

This mapping will be referred to as a generalized inverse of $\mathbf{h}$. If we consider the lumping in the whole n-dimensional real space $\mathscr{R}^n$, eq. (6) holds for any $\mathbf{y} \in \mathscr{R}^n$. Let $\mathbf{y}$ take the value $\bar{\mathbf{h}}(\hat{\mathbf{y}})$ and substitute it into eq. (6). Then we have

$$\hat{\mathbf{f}}\{\mathbf{h}[\bar{\mathbf{h}}(\hat{\mathbf{y}})]\} = \mathbf{h}_\mathbf{y}[\bar{\mathbf{h}}(\hat{\mathbf{y}})]\mathbf{f}[\bar{\mathbf{h}}(\hat{\mathbf{y}})]$$
$$\hat{\mathbf{f}}(\hat{\mathbf{y}}) = \mathbf{h}_\mathbf{y}\{[\bar{\mathbf{h}}[\mathbf{h}(\mathbf{y})]\}\mathbf{f}\{[\bar{\mathbf{h}}[\mathbf{h}(\mathbf{y})]\}. \tag{9}$$

Comparing eqs (6) and (9) yields

$$\mathbf{h}_\mathbf{y}(\mathbf{y})\mathbf{f}(\mathbf{y}) = \mathbf{h}_\mathbf{y}\{\bar{\mathbf{h}}[\mathbf{h}(\mathbf{y})]\}\mathbf{f}\{\bar{\mathbf{h}}[\mathbf{h}(\mathbf{y})]\}. \tag{10}$$

This shows that eq. (10) is a necessary condition for the existence of exact lumping. It is easy to prove that this condition is also sufficient. Multiplying eq. (1) from the left by $\mathbf{h}_\mathbf{y}(\mathbf{y})$ and using eq. (10) yields

$$\mathbf{h}_\mathbf{y}(\mathbf{y})\frac{d\mathbf{y}}{dt} = \mathbf{h}_\mathbf{y}(\mathbf{y})\mathbf{f}(\mathbf{y})$$
$$= \mathbf{h}_\mathbf{y}\{\bar{\mathbf{h}}[\mathbf{h}(\mathbf{y})]\}\mathbf{f}\{\bar{\mathbf{h}}[\mathbf{h}(\mathbf{y})]\}.$$

Let

$$\hat{\mathbf{y}} = \mathbf{h}(\mathbf{y})$$

then

$$\mathbf{h}_\mathbf{y}(\mathbf{y})\frac{d\mathbf{y}}{dt} = \mathbf{h}_\mathbf{y}[\bar{\mathbf{h}}(\hat{\mathbf{y}})]\mathbf{f}[\bar{\mathbf{h}}(\hat{\mathbf{y}})]. \tag{11}$$

Note that the left-hand side of eq. (11) is just $d\hat{\mathbf{y}}/dt$. Thus we have

$$\frac{d\hat{\mathbf{y}}}{dt} = \mathbf{h}_\mathbf{y}[\bar{\mathbf{h}}(\hat{\mathbf{y}})]\mathbf{f}[\bar{\mathbf{h}}(\hat{\mathbf{y}})]$$
$$= \hat{\mathbf{f}}(\hat{\mathbf{y}}). \tag{12}$$

The condition has been proved to be sufficient and eq. (12) also provides the form of the lumped differential equations.

For a given $\mathbf{h}$, one may find an infinite number of $\bar{\mathbf{h}}$ to satisfy eq. (8) when $\hat{n} < n$. We can prove that the form of the exactly lumped differential equations is independent of the choice of $\bar{\mathbf{h}}$. Suppose $\bar{\mathbf{h}}^*$ is another transformation satisfying eq. (8). Since eq. (10) is valid for any value of $\mathbf{y}$, let $\mathbf{y}$ take the value $\bar{\mathbf{h}}^*(\hat{\mathbf{y}})$ and substitute into eq. (10). Then we have

$$\mathbf{h}_\mathbf{y}[\bar{\mathbf{h}}^*(\hat{\mathbf{y}})]\mathbf{f}[\bar{\mathbf{h}}^*(\hat{\mathbf{y}})] = \mathbf{h}_\mathbf{y}(\bar{\mathbf{h}}\{\mathbf{h}[\bar{\mathbf{h}}^*(\hat{\mathbf{y}})]\})\mathbf{f}(\bar{\mathbf{h}}\{\mathbf{h}[\bar{\mathbf{h}}^*(\hat{\mathbf{y}})]\})$$
$$= \mathbf{h}_\mathbf{y}[\bar{\mathbf{h}}(\hat{\mathbf{y}})]\mathbf{f}[\bar{\mathbf{h}}(\hat{\mathbf{y}})]. \tag{13}$$

This shows that the exactly lumped differential equation system does not depend on the choice of $\bar{\mathbf{h}}$.

Equation (10) has a clear geometrical interpretation. We first recall the definition of an invariant manifold (Carr, 1981) for eq. (1). A set $S \subset \mathscr{R}^n$ is said to be a *local invariant manifold* for eq. (1) if for $\mathbf{y}_0 \in S$ the solution $\mathbf{y}(t)$ of eq. (1) with $\mathbf{y}(0) = \mathbf{y}_0$ remains in S for $|t| < T$ where $T > 0$. If we can always choose $T = \infty$, then we say that S is an *invariant manifold*. We can now prove that if, eq. (1) is exactly lumpable by the transformation $\hat{\mathbf{y}} = \mathbf{h}(\mathbf{y})$, then

$$\mathbf{h}(\mathbf{y}) = \mathbf{0} \tag{14}$$

is an invariant manifold of it. To prove this first note that $\mathbf{h}(\mathbf{y}) = \mathbf{0}$ represents an $(n - \hat{n})$-dimensional manifold because this equation gives $\hat{n}$ restrictions on the variation of $\mathbf{y}$ (Fleming, 1977, p. 153). When eq. (1) is exactly lumpable by $\mathbf{h}$, it satisfies eq. (10). If $\mathbf{y}$ takes a value from this manifold, then according to eq. (10) we have

$$\frac{d\mathbf{h}(\mathbf{y})}{dt} = \mathbf{h}_\mathbf{y}(\mathbf{y})\mathbf{f}(\mathbf{y})$$
$$= \mathbf{h}_\mathbf{y}\{\bar{\mathbf{h}}[\mathbf{h}(\mathbf{y})]\}\mathbf{f}\{\bar{\mathbf{h}}[\mathbf{h}(\mathbf{y})]\}$$
$$= \mathbf{h}_\mathbf{y}[\bar{\mathbf{h}}(\mathbf{0})]\mathbf{f}[\bar{\mathbf{h}}(\mathbf{0})], \quad \forall \mathbf{y} \in \mathbf{h}(\mathbf{y}) = \mathbf{0}. \tag{15}$$

For $\mathbf{y} \in \mathbf{h}(\mathbf{y}) = \mathbf{0}$, eq. (10) becomes

$$\mathbf{h}_\mathbf{y}(\mathbf{y})\mathbf{f}(\mathbf{y}) = \mathbf{h}_\mathbf{y}[\bar{\mathbf{h}}(\mathbf{0})]\mathbf{f}[\bar{\mathbf{h}}(\mathbf{0})] = \text{const.} \quad \forall \mathbf{y} \in \mathbf{h}(\mathbf{y}) = \mathbf{0}.$$

Considering that the origin $\mathbf{0}$ is in this manifold, i.e. $\mathbf{h}(\mathbf{0}) = \mathbf{0}$ and $\mathbf{f}(\mathbf{0}) = \mathbf{0}$ we have

$$\mathbf{h}_\mathbf{y}(\mathbf{0})\mathbf{f}(\mathbf{0}) = \mathbf{h}_\mathbf{y}[\bar{\mathbf{h}}(\mathbf{0})]\mathbf{f}[\bar{\mathbf{h}}(\mathbf{0})] = \mathbf{0}.$$

This implies that

$$\frac{d\mathbf{h}(\mathbf{y})}{dt} = \mathbf{0}, \quad \forall \mathbf{y} \in \mathbf{h}(\mathbf{y}) = \mathbf{0}. \tag{16}$$

Considering a trajectory of eq. (1) with the initial value $\mathbf{y}(0)$ located in the manifold $\mathbf{h}(\mathbf{y}) = \mathbf{0}$. Suppose this trajectory $\mathbf{y}(t)$ would cross the boundary of the manifold $\mathbf{h}(\mathbf{y}) = \mathbf{0}$ and enter another manifold $\mathbf{h}(\mathbf{y}) \neq \mathbf{0}$. Then there would be a point $\mathbf{y} \in \mathbf{h}(\mathbf{y}) = \mathbf{0}$ with $d\mathbf{h}(\mathbf{y})/dt \neq \mathbf{0}$. This is a contradiction to eq. (15). Thus any trajectory with initial point within the manifold $\mathbf{h}(\mathbf{y}) = \mathbf{0}$ will remain in it for any t. This shows that if $\mathbf{y}(0) \in \mathbf{h}(\mathbf{y}) = \mathbf{0}$, then $\mathbf{y}(t) \in \mathbf{h}(\mathbf{y}) = \mathbf{0}$ for any $|t| > 0$. Hence, $\mathbf{h}(\mathbf{y}) = \mathbf{0}$ is an invariant manifold of eq. (1). However, this is only a necessary condition for exact lumping. Equation (10) also requires that for any $\mathbf{y}$ in the manifold $\mathbf{h}(\mathbf{y}) = \mathbf{c}$ which is *parallel* to

$\mathbf{h}(\mathbf{y}) = \mathbf{0}$ for $\mathbf{c}$ being a constant vector,

$$\frac{d\mathbf{h}(\mathbf{y})}{dt} = \bar{\mathbf{c}}(\mathbf{c}), \quad \forall \mathbf{y} \in \mathbf{h}(\mathbf{y}) = \mathbf{c} \tag{17}$$

where $\bar{\mathbf{c}}$ is a constant vector and a function of $\mathbf{c}$.

We now give another necessary and sufficient condition of exact lumping. Differentiating eq. (10) with respect to $\mathbf{y}$ from both sides gives

$$(\mathbf{h}_y\mathbf{f})_y(\mathbf{y}) = (\mathbf{h}_y\mathbf{f})_y\{\bar{\mathbf{h}}[\mathbf{h}(\mathbf{y})]\}\bar{\mathbf{h}}_{\hat{\mathbf{y}}}[\mathbf{h}(\mathbf{y})]\mathbf{h}_y(\mathbf{y}) \tag{18}$$

where $(\mathbf{h}_y\mathbf{f})_y(\mathbf{y})$ and $\bar{\mathbf{h}}_{\hat{\mathbf{y}}}(\hat{\mathbf{y}})$ represent the Jacobian matrix of $\mathbf{h}_y(\mathbf{y})\mathbf{f}(\mathbf{y})$ and $\bar{\mathbf{h}}(\hat{\mathbf{y}})$, respectively. Equation (18) is another necessary and sufficient condition of exact lumping because integrating it with respect to $\mathbf{y}$ and taking into account that $\mathbf{h}(\mathbf{0}) = \mathbf{0}$ and $\mathbf{f}(\mathbf{0}) = \mathbf{0}$ will give eq. (10). Since $\mathbf{h}$ is an $\hat{n}$-dimensional nondegenerate transformation and thus $\mathbf{h}_y(\mathbf{y})$ is an $\hat{n} \times n$ matrix with rank $\hat{n}$, there exists an $n \times l$ $(l \geqslant n - \hat{n})$-matrix $X(\mathbf{y})$ with rank $n - \hat{n}$ such that

$$\mathbf{h}_y(\mathbf{y})X(\mathbf{y}) = 0. \tag{19}$$

Multiplying both sides of eq. (18) from the right by $X(\mathbf{y})$ yields

$$(\mathbf{h}_y\mathbf{f})_y(\mathbf{y})X(\mathbf{y}) = 0. \tag{20}$$

As eqs (19) and (20) are obtained from eq. (18), the combination of them is a necessary condition of exact nolinear lumping.

We now prove that the combination of eqs (19) and (20) is also a sufficient condition for exact nonlinear lumping. Take the following nonsingular transformation

$$\mathbf{z} = T(\mathbf{y}) = \begin{pmatrix} \mathbf{h}(\mathbf{y}) \\ \mathbf{u}(\mathbf{y}) \end{pmatrix}. \tag{21}$$

The Jacobian matrix of $T(\mathbf{y})$ is

$$\begin{pmatrix} \mathbf{h}_y(\mathbf{y}) \\ \mathbf{u}_y(\mathbf{y}) \end{pmatrix}. \tag{22}$$

As the transformation $T(\mathbf{y})$ is nonsingular, the inverse transformation

$$\mathbf{y} = T^{-1}(\mathbf{z}) \tag{23}$$

exists. Let the Jacobian matrix of $T^{-1}(\mathbf{z})$ be represented as

$$T_z^{-1}(\mathbf{z}) = T_z^{-1}[T(\mathbf{y})] = Y[(\mathbf{y})|X(\mathbf{y})] \tag{24}$$

where $Y(\mathbf{y})$ and $X(\mathbf{y})$ are $n \times \hat{n}$ and $n \times (n - \hat{n})$ matrices, respectively. We also know that the product of the Jacobian matrices for T and T^{-1} is the identity matrix. This product gives that

$$\begin{aligned} I_n &= T_y(\mathbf{y})T_z^{-1}(\mathbf{z}) \\ &= T_y(\mathbf{y})T_z^{-1}[T(\mathbf{y})] \\ &= \begin{pmatrix} \mathbf{h}_y(\mathbf{y}) \\ \mathbf{u}_y(\mathbf{y}) \end{pmatrix}[Y(\mathbf{y})|X(\mathbf{y})] \\ &= \begin{pmatrix} \mathbf{h}_y(\mathbf{y})Y(\mathbf{y}) & \mathbf{h}_y(\mathbf{y})X(\mathbf{y}) \\ \mathbf{u}_y(\mathbf{y})Y(\mathbf{y}) & \mathbf{u}_y(\mathbf{y})X(\mathbf{y}) \end{pmatrix} \\ &= \begin{pmatrix} I_{\hat{n}} & 0 \\ 0 & I_{n-\hat{n}} \end{pmatrix} \end{aligned} \tag{25}$$

implying that

$$\mathbf{h}_y(\mathbf{y})X(\mathbf{y}) = 0. \tag{26}$$

We consider the differential equation system for $\mathbf{z}$ and its corresponding Jacobian matrix.

$$\begin{aligned} \frac{d\mathbf{z}}{dt} &= \begin{pmatrix} \mathbf{h}_y(\mathbf{y}) \\ \mathbf{u}_y(\mathbf{y}) \end{pmatrix}\mathbf{f}(\mathbf{y}) \\ &= \begin{pmatrix} \mathbf{h}_y[T^{-1}(\mathbf{z})] \\ \mathbf{u}_y[T^{-1}(\mathbf{z})] \end{pmatrix}\mathbf{f}[T^{-1}(\mathbf{z})] \\ &= \mathbf{g}(\mathbf{z}). \end{aligned} \tag{27}$$

The corresponding Jacobian matrix of $\mathbf{g}(\mathbf{z})$ is

$$\begin{aligned} \mathbf{g}_z(\mathbf{z}) &= \frac{\partial \mathbf{g}(\mathbf{z})}{\partial \mathbf{z}} \\ &= \frac{\partial}{\partial \mathbf{y}}\left[\begin{pmatrix} \mathbf{h}_y(\mathbf{y}) \\ \mathbf{u}_y(\mathbf{y}) \end{pmatrix}\mathbf{f}(\mathbf{y})\right]\frac{\partial \mathbf{y}}{\partial \mathbf{z}} \\ &= \begin{pmatrix} (\mathbf{h}_y\mathbf{f})_y(\mathbf{y}) \\ (\mathbf{u}_y\mathbf{f})_y(\mathbf{y}) \end{pmatrix}T_z^{-1}[T(\mathbf{y})] \\ &= \begin{pmatrix} (\mathbf{h}_y\mathbf{f})_y(\mathbf{y}) \\ (\mathbf{u}_y\mathbf{f})_y(\mathbf{y}) \end{pmatrix}[Y(\mathbf{y})|X(\mathbf{y})] \\ &= \begin{pmatrix} (\mathbf{h}_y\mathbf{f})_y(\mathbf{y})Y(\mathbf{y}) & (\mathbf{h}_y\mathbf{f})_y(\mathbf{y})X(\mathbf{y}) \\ (\mathbf{u}_y\mathbf{f})_y(\mathbf{y})Y(\mathbf{y}) & (\mathbf{u}_y\mathbf{f})_y(\mathbf{y})X(\mathbf{y}) \end{pmatrix}. \end{aligned} \tag{28}$$

When eqs (19) and (20) hold, we have

$$\mathbf{g}_z(\mathbf{z}) = \begin{pmatrix} (\mathbf{h}_y\mathbf{f})_y(\mathbf{y})Y(\mathbf{y}) & 0 \\ (\mathbf{u}_y\mathbf{f})_y(\mathbf{y})Y(\mathbf{y}) & (\mathbf{u}_y\mathbf{f})_y(\mathbf{y})X(\mathbf{y}) \end{pmatrix}. \tag{29}$$

Since the transformation T is nonsingular and applicable for all values of $\mathbf{y} \in \mathscr{R}^n$, this implies that its image consists of all values of $\mathbf{z} \in \mathscr{R}^n$. Thus from eq. (29), we have

$$\frac{\partial g_i(\mathbf{z})}{\partial z_j} = 0 \tag{30}$$

$(i = 1, 2, \ldots, \hat{n}, j = \hat{n} + 1, \hat{n} + 2, \ldots, n) \quad \forall \mathbf{z} \in \mathscr{R}^n.$

Equation (30) shows that the function $g_i(\mathbf{z})$ $(i = 1, 2, \ldots, \hat{n})$ do not contain the last $n - \hat{n}$ variables z_j $(j = \hat{n} + 1, \hat{n} + 2, \ldots, n)$. Hence the first $\hat{n}$ equations in eq. (27) compose an exactly lumped model. Therefore, the combination of eqs (19) and (20) is a sufficient condition for exact nonlinear lumping.

In order to determine all exact lumping schemes, we first need to determine all invariant manifolds of a given differential equation system and then examine which of them satisfy eqs (10) or (18) or (19) and (20). The expressions $\mathbf{h}(\mathbf{y})$ of these invariant manifolds can be used as transformations to exactly lump the system.

The geometric meaning of eqs (19) and (20) is that the normal space (the $\hat{n}$-dimensional orthogonal compliment to the tangent space) of the manifold $\mathbf{h}_y(\mathbf{y})\mathbf{f}(\mathbf{y}) = \bar{\mathbf{c}}$ for $\mathbf{y} \in \mathbf{h}(\mathbf{y}) = \mathbf{c}$ is contained in the normal space of the manifold $\mathbf{h}(\mathbf{y}) = \mathbf{c}$. This will give the

conclusion of Luckyanov *et al.* (1983), i.e. the matrices $\mathbf{h}_{\mathbf{y}}(\mathbf{y})$ and

$$\begin{pmatrix} \mathbf{h}_{\mathbf{y}}(\mathbf{y}) \\ [(h_i)_{\mathbf{y}}\mathbf{f}]_{\mathbf{y}}(\mathbf{y}) \end{pmatrix}, \quad i = 1, 2, \ldots, \hat{n} \tag{31}$$

where h_i is the ith element of $\mathbf{h}$, have the same rank. Iwasa *et al.* (1987) gave a necessary and sufficient condition for exact lumping as

$$AB^{+}B = A \tag{32}$$

where B^{+} is the generalized inverse of B. According to our notation

$$A = (\mathbf{h}_{\mathbf{y}}\mathbf{f})_{\mathbf{y}}(\mathbf{y}) \tag{33}$$

$$B = \mathbf{h}_{\mathbf{y}}(\mathbf{y}). \tag{34}$$

Iwasa's condition can be readily obtained from eqs (19) and (20) by setting

$$X = I_n - \bar{\mathbf{h}}_{\hat{\mathbf{y}}}[\mathbf{h}(\mathbf{y})]\mathbf{h}_{\mathbf{y}}(\mathbf{y}). \tag{35}$$

Noting that

$$\mathbf{h}_{\mathbf{y}}(\mathbf{y})\bar{\mathbf{h}}_{\hat{\mathbf{y}}}[\mathbf{h}(\mathbf{y})] = I_{\hat{n}} \tag{36}$$

one can prove that this X satisfies eq. (19). Substituting this X into eq. (20) gives the condition

$$A(I_n - B^{+}B) = 0 \tag{37}$$

i.e.

$$A - AB^{+}B = 0$$

or

$$AB^{+}B = A \tag{38}$$

where

$$B^{+} = \bar{\mathbf{h}}_{\hat{\mathbf{y}}}[\bar{\mathbf{h}}(\mathbf{y})]. \tag{39}$$

This is Iwasa's condition.

Although Iwasa's and our results are similar, there still exists some difference. First, they did not give the form of the lumped system. Iwasa *et al.* (1987) gave the form of the Jacobian matrix of the right-hand side for the lumped differential equation system as

$$C = AB^{+} \tag{40}$$

but, since in Iwasa's paper A and B^{+} are functions of the original variables, so is C. It is not a function of the lumped variables. Only for a linear transformation and a linear differential equation system can they give the form of the lumped system. This is why the examples for the determination of the lumped models in Iwasa's paper (1989) are only for linear transformations and linear differential equation systems. For nonlinear systems, Iwasa *et al.* can only consider the linearized dynamics around some point, which is still a linear system.

There is another important difference between our two works. They gave some necessary and sufficient conditions so that one can examine whether a given lumping transformation is exact for a given differential equation system. In addition, supposing it is an exact transformation, they can check what condition the original system must satisfy. Our goal is to determine the exact lumping schemes for a given system. The latter goal is more difficult because the necessary and sufficient conditions alone can be readily obtained by simply differentiation. For determination of lumping schemes, we first need to find the geometric or algebraic interpretations of these conditions and then use these results to establish some general approaches to determine exact lumping schemes.

We now prove that exact linear lumping is only a special case of this analysis and its necessary and sufficient conditions can be readily obtained from those given above. For the linear transformation,

$$\mathbf{h}(\mathbf{y}) = M\mathbf{y} \tag{41}$$

where M is an $\hat{n} \times n$ constant matrix with rank $\hat{n}$: its $\bar{\mathbf{h}}(\hat{\mathbf{y}})$ is $\bar{M}\hat{\mathbf{y}}$, where $\bar{M}$ is a generalized inverse of M satisfying

$$M\bar{M} = I_{\hat{n}}. \tag{42}$$

In this case eqs (10) and (18) become

$$M\mathbf{f}(\mathbf{y}) = M\mathbf{f}(\bar{M}M\mathbf{y}) \tag{43}$$

$$M\mathbf{f}_{\mathbf{y}}(\mathbf{y}) = M\mathbf{f}_{\mathbf{y}}(\bar{M}M\mathbf{y})\bar{M}M. \tag{44}$$

Similarly, eqs (19) and (20) become

$$MX = 0 \tag{45}$$

$$M\mathbf{f}_{\mathbf{y}}(\mathbf{y})X = 0 \tag{46}$$

where X is a constant matrix. These relations have been obtained in our previous work on exact linear lumping. The only difference is that we used the notation $J(\mathbf{y})$ instead of $\mathbf{f}_{\mathbf{y}}(\mathbf{y})$ to represent the Jacobian matrix of $\mathbf{f}(\mathbf{y})$ (Li and Rabitz, 1989).

2.2. *Constrained nonlinear lumping*

In practice, most lumping schemes must fulfill some requirements. For example, some dependent variables may be desired to be left unlumped or lumped by a given transformation for physical reasons. This refers to finding a nonlinear lumping transformation which contains a set of given transformations. Using eqs (19) and (20), we can determine the constrained exact lumping transformation with the smallest dimension.

Suppose $\mathbf{h}_1(\mathbf{y})$ is a given k-dimensional transformation with $k < n$. We need to find an exact lumping transformation $\mathbf{h}(\mathbf{y})$ which contains $\mathbf{h}_1(\mathbf{y})$ as its first k coordinates. Let $(\mathbf{h}_1)_{\mathbf{y}}(\mathbf{y})$ be the Jacobian matrix of $\mathbf{h}_1(\mathbf{y})$. We can construct an $n \times (n-k)$ matrix $X_1(\mathbf{y})$ with rank $n-k$ satisfying

$$(\mathbf{h}_1)_{\mathbf{y}}(\mathbf{y})X_1(\mathbf{y}) = 0. \tag{47}$$

Note that

$$\frac{\mathrm{d}\mathbf{h}_1(\mathbf{y})}{\mathrm{d}t} = (\mathbf{h}_1)_{\mathbf{y}}(\mathbf{y})\mathbf{f}(\mathbf{y}). \tag{48}$$

If $\mathbf{h}(\mathbf{y})$ is an exact lumping transformation and contains $\mathbf{h}_1(\mathbf{y})$, the flux vector $(\mathbf{h}_1)_{\mathbf{y}}(\mathbf{y})\mathbf{f}(\mathbf{y})$ should be located within the range of $\mathbf{h}$, otherwise $\mathbf{h}(\mathbf{y})$ is not an exact lumping transformation. There are two cases:

$(\mathbf{h}_1)_{\mathbf{y}}(\mathbf{y})\mathbf{f}(\mathbf{y})$ is or is not located in the range of $\mathbf{h}_1(\mathbf{y})$. This can be judged by examining whether the following equation:

$$(\mathbf{h}_1\mathbf{f})_{\mathbf{y}}(\mathbf{y})X_1(\mathbf{y}) = 0 \tag{49}$$

holds. If eq. (49) holds, then $\mathbf{h}_1(\mathbf{y})$ is an exact transformation. Then we can use $\mathbf{h}(\mathbf{y}) = \mathbf{h}_1(\mathbf{y})$ as a transformation for exact lumping. If eq. (49) does not hold, let $\mathbf{h}_2(\mathbf{y})$ be the set of the elements of $(\mathbf{h}_1)_{\mathbf{y}}(\mathbf{y})\mathbf{f}(\mathbf{y})$ which are functionally independent from $\mathbf{h}_1(\mathbf{y})$. The new transformation $\mathbf{h}(\mathbf{y})$ is composed of $\mathbf{h}_1(\mathbf{y})$ and $\mathbf{h}_2(\mathbf{y})$. The Jacobian matrix $\mathbf{h}_{\mathbf{y}}(\mathbf{y})$ consists of $(\mathbf{h}_1)_{\mathbf{y}}(\mathbf{y})$ and $(\mathbf{h}_2)_{\mathbf{y}}(\mathbf{y})$. We now construct a new matrix $X_2(\mathbf{y})$ satisfying

$$\mathbf{h}_{\mathbf{y}}(\mathbf{y})X_2(\mathbf{y}) = 0 \tag{50}$$

and examine whether the following equation

$$(\mathbf{h}_{\mathbf{y}}\mathbf{f})_{\mathbf{y}}(\mathbf{y})X_2(\mathbf{y}) = 0 \tag{51}$$

holds. If it holds, the new $\mathbf{h}(\mathbf{y})$ can be used as an exact lumping transformation. Otherwise let $\mathbf{h}_3(\mathbf{y})$ be the set of the elements of $(\mathbf{h}_2)_{\mathbf{y}}(\mathbf{y})\mathbf{f}(\mathbf{y})$ which are functionally independent from $\mathbf{h}_1(\mathbf{y})$ and $\mathbf{h}_2(\mathbf{y})$. The new transformation $\mathbf{h}(\mathbf{y})$ is composed of $\mathbf{h}_i(\mathbf{y})$ $(i = 1, 2, 3)$. We follow the same procedure until we find a transformation $\mathbf{h}(\mathbf{y})$ satisfying eqs (19) and (20). According to the results of Section 2.1, the transformation $\mathbf{h}(\mathbf{y})$ is an exact one [the largest dimension of $\mathbf{h}(\mathbf{y})$ is n, which corresponds to no reduction]. By this procedure, we have shown that any other transformation, whose expression contains $\mathbf{h}_1(\mathbf{y})$, with a dimension lower than that of $\mathbf{h}(\mathbf{y})$ does not satisfy eqs (19) and (20) and then is not exact. Therefore, the resultant $\mathbf{h}(\mathbf{y})$ is the exact lumping transformation containing $\mathbf{h}_1(\mathbf{y})$ with the smallest dimension.

3. THE RELATIONSHIP BETWEEN EXACT LUMPING SCHEMES AND GENERALIZED EIGENFUNCTIONS

In this section we consider the algebraic explanation of exact nonlinear lumping. The algebraic explanation will present a general analysis of lumping by the discussion of the relationship between exact lumping schemes and a set of functions named *generalized eigenfunctions* for an ordinary differential equation system or its corresponding linear partial differential operator. An n-dimensional ordinary differential equation system has an infinite number of generalized eigenfunctions, but at most only n of them are functionally independent. It can be proven that any $\hat{n}$-dimensional lumped system has $\hat{n}$ functionally independent generalized eigenfunctions of the original n-dimensional differential equation system, and any $\hat{n}$-dimensional exact lumping transformation $\mathbf{h} = (h_1, h_2, \ldots, h_{\hat{n}})$ consists of only $\hat{n}$ functions of $\hat{n}$ functionally independent generalized eigenfunctions. Thus, the determination of exact lumping schemes of a differential equation system reduces to the determination of its generalized eigenfunctions. We first give the definition of the generalized eigenfunction and then discuss its relation with exact lumping schemes. At last, we present some approaches for the determination of generalized eigenfunctions of a given differential equation system.

3.1. *Generalized eigenfunctions*

We first give the definition of a generalized eigenfunction for an n-dimensional ordinary differential equation system as eq. (1). Using this equation system, we define a linear partial differential operator

$$A = \sum_{i=1}^{n} f_i(\mathbf{y})\frac{\partial}{\partial y_i}. \tag{52}$$

A is linear since for any two functions F_1 and F_2 and any two constants a and b we have

$$A(aF_1 + bF_2) = aAF_1 + bAF_2. \tag{53}$$

If eq. (1) can be exactly lumped by transformation $\hat{\mathbf{y}} = \mathbf{h}(\mathbf{y})$, eqs (10)–(12) hold. From these equations, we see that $\hat{f}_i$ results from the action of A upon h_i. For example, for $\hat{y}_i$ these equations give

$$\frac{\mathrm{d}\hat{y}_i}{\mathrm{d}t} = \sum_{j=1}^{n} f_j(\mathbf{y})\frac{\partial h_i(\mathbf{y})}{\partial y_j} = Ah_i = \hat{f}_i(h_1, h_2, \ldots, h_{\hat{n}}). \tag{54}$$

Thus if eq. (1) is exactly lumped by $\mathbf{h}$, then there exist $\hat{n}$ functions $h_i(\mathbf{y})$ $(i = 1, 2, \ldots, \hat{n})$ such that

$$Ah_i = \hat{f}_i(h_1, h_2, \ldots, h_{\hat{n}}), \quad i = 1, 2, \ldots, \hat{n}. \tag{55}$$

Equation (55) implies the existence of an invariant property of the functions h_i with respect to operator A. This property may give us a way to determine the functions h_i. Consider a special case of a one-dimensional lumped system

$$A\psi = \Omega(\psi) \tag{56}$$

where Ω is an arbitrary function. We refer to ψ as a *generalized eigenfunction* of A or eq. (1). Note that this definition is more general than the common definition of an eigenfunction of a differential operator $A\psi = \lambda\psi$, where the right-hand side is a linear function of ψ (we discuss eigenfunctions of A in Section 4). This is why we call ψ a generalized eigenfunction. If eq. (56) holds with Ω not identically zero, we can always find a function $\Psi(\psi)$ such that

$$A\Psi(\psi) = 1 \tag{57}$$

because using the solution of the following equation

$$\frac{\mathrm{d}\Psi(\psi)}{\mathrm{d}\psi} = 1/\Omega(\psi) \tag{58}$$

we have

$$A\Psi(\psi) = \frac{\mathrm{d}\Psi(\psi)}{\mathrm{d}\psi}A\psi = \frac{\Omega(\psi)}{\Omega(\psi)} = 1.$$

We call Ψ a *normal generalized eigenfunction* of A or eq. (1). Obviously, a normal generalized eigenfunction is a generalized eigenfunction if we refer to 1 in eq. (57) as Ψ^0.

Equation (57) is easier to analyze than eq. (56) because it does not contain an arbitrary function Ω. We first determine how many solutions eq. (57) has.

Certainly, there are an infinite number of solutions, but at most only n solutions are functionally independent because Ψ is only a function of n variables y_i ($i = 1, 2, \ldots, n$). Using the theory of partial differential equations, the system of equations

$$A\Psi = 1, \quad A\Phi_i = 0, \quad i = 1, 2, \ldots, n-1 \tag{59}$$

always has nontrivial solutions $\{\Psi, \Phi_i\}$, where Φ_i ($i = 1, 2, \ldots, n-1$) are functionally independent (Stephani, 1989). Ψ is not unique because $\Psi + \Phi_i$ also satisfy

$$A(\Psi + \Phi_i) = A\Psi + A\Phi_i = 1. \tag{60}$$

Since Ψ and $\Psi + \Phi_i$ as well as $\Psi + \Phi_i$ and $\Psi + \Phi_j$ are functionally independent, they compose n functionally independent normal generalized eigenfunctions of A or n functionally independent solutions of eq. (57).

If we can find $\hat{n}$ functions h_i satisfying eq. (55), we can prove that there always exist $\hat{n}$ functionally independent normal generalized eigenfunctions $\Psi_i(\mathbf{h})$ of eq. (1). As

$$\begin{aligned} A\Psi(\mathbf{h}) &= \sum_{j=1}^{\hat{n}} \frac{\partial \Psi(\mathbf{h})}{\partial h_j} Ah_j \\ &= \sum_{j=1}^{\hat{n}} \frac{\partial \Psi(\mathbf{h})}{\partial h_j} \hat{f}_j(\mathbf{h}) \end{aligned} \tag{61}$$

and

$$\begin{gathered} \sum_{j=1}^{\hat{n}} \frac{\partial \Psi(\mathbf{h})}{\partial h_j} \hat{f}_j(\mathbf{h}) = 1 \\ \sum_{j=1}^{\hat{n}} \frac{\partial \Phi_i(\mathbf{h})}{\partial h_j} \hat{f}_j(\mathbf{h}) = 0, \quad i = 1, 2, \ldots, \hat{n}-1 \end{gathered} \tag{62}$$

always have nontrivial solutions, we can certainly find a normal generalized eigenfunction $\Psi(\mathbf{h})$ of A. Since

$$\sum_{j=1}^{\hat{n}} \frac{\partial [\Psi(\mathbf{h}) + \Phi_i(\mathbf{h})]}{\partial h_j} \hat{f}_j(\mathbf{h}) = 1 \tag{63}$$

and $\Psi(\mathbf{h})$ and $\Psi(\mathbf{h}) + \Phi_i(\mathbf{h})$ ($i = 1, 2, \ldots, \hat{n}-1$) are functionally independent, we may find $\hat{n}$ functionally independent normal generalized eigenfunctions of A. As a normal generalized eigenfunction is a generalized eigenfunction, an $\hat{n}$-dimensional lumped system has $\hat{n}$ functionally independent generalized eigenfunctions of eq. (1). Note that any function of a normal generalized eigenfunction $\psi(\Psi)$ is a generalized eigenfunction provided that the inverse $\Psi(\psi)$ exists. Indeed,

$$\begin{aligned} A\psi(\Psi) &= \frac{\mathrm{d}\psi}{\mathrm{d}\Psi}(\Psi) A\Psi \\ &= \frac{\mathrm{d}\psi}{\mathrm{d}\Psi}[\Psi(\psi)] = \Omega(\psi). \end{aligned} \tag{64}$$

For a given set of $\hat{n}$ functionally independent generalized eigenfunctions $\psi_1(\mathbf{y}), \psi_2(\mathbf{y}), \ldots, \psi_{\hat{n}}(\mathbf{y})$, one can use them to construct $\hat{n}$ functionally independent functions $h_1(\psi), h_2(\psi), \ldots, h_{\hat{n}}(\psi)$, where $\psi = (\psi_1\ \psi_2\ \ldots\ \psi_{\hat{n}})^{\mathrm{T}}$. Since this transformation is nonsingular, the inverse transformation $\psi_1(\mathbf{h}), \psi_2(\mathbf{h}), \ldots, \psi_{\hat{n}}(\mathbf{h})$ exists. Then

$$\begin{aligned} \frac{\mathrm{d}h_i}{\mathrm{d}t} &= \sum_{j=1}^{\hat{n}} \frac{\partial h_i(\psi)}{\partial \psi_j} \frac{\mathrm{d}\psi_j}{\mathrm{d}t} \\ &= \sum_{j=1}^{\hat{n}} \frac{\partial h_i(\psi)}{\partial \psi_j} \Omega_j(\psi_j) \\ &= \sum_{j=1}^{\hat{n}} \frac{\partial h_i[\psi(\mathbf{h})]}{\partial \psi_j} \Omega_j[\psi_j(\mathbf{h})] \\ &= \hat{f}_i(h_1, h_2, \ldots, h_{\hat{n}}), \quad i = 1, 2, \ldots, \hat{n}. \end{aligned} \tag{65}$$

These equations compose an $\hat{n}$-dimensional lumped system. Thus $\mathbf{h}$ is an $\hat{n}$-dimensional exact lumping transformation if and only if h_i is a function of $\hat{n}$ functionally independent generalized eigenfunctions of a given differential equation system. If one can find all generalized eigenfunctions, then all exact lumping transformations with different dimensions can be constructed by them. Hence, the determination of exact lumping transformations reduces to the determination of generalized eigenfunctions or normal generalized eigenfunctions of A.

If one has determined n functionally independent normal generalized eigenfunctions Ψ_i ($i = 1, 2, \ldots, n$) of a given differential equation system, then for an arbitrary function ϕ and with the first integrals defined as

$$\Phi_i = \Psi_{i+1} - \Psi_1, \quad i = 1, 2, \ldots, n-1 \tag{66}$$

the function

$$\Psi = \Psi_1 + \phi(\Phi_1, \Phi_2, \ldots, \Phi_{n-1}) \tag{67}$$

is again a normal generalized eigenfunction.

Moreover, we can prove that any normal generalized eigenfunction Ψ can be represented in the form of eq. (67) with some function ϕ. Indeed, since Ψ_i ($i = 1, 2, \ldots, n$) are functionally independent, we can choose them as a base of n variable functions. Then any normal generalized eigenfunction can be expressed as a function of them, i.e. $\Psi(\Psi_1, \Psi_2, \ldots, \Psi_n)$ and satisfies the equation

$$\begin{aligned} A\Psi &= \sum_{i=1}^{n} \frac{\partial \Psi}{\partial \Psi_i} A\Psi_i \\ &= \sum_{i=1}^{n} \frac{\partial \Psi}{\partial \Psi_i} = 1. \end{aligned} \tag{68}$$

The general solution of eq. (68) is eq. (67). Therefore, all normal generalized eigenfunctions can be obtained provided n functionally independent normal generalized eigenfunctions have been determined. This is equivalent to knowing a normal generalized eigenfunction and all $n-1$ functionally independent global first integrals of an n-dimensional ordinary differential equation system. Unfortunately, there is presently no general approach to determine all global first integrals. This means that in practice we may not find all normal generalized eigenfunctions as well as all other higher dimensional exact lumping transformations for an arbitrary first-order ordinary differential

equation system. Indeed, some generalized eigenfunctions may even not be represented by elementary or other simple functions. Hence, for an arbitrary system, one will generally not be able to determine all exact lumping schemes by the above general approach. However, it may be quite satisfactory to determine all generalized eigenfunctions and other higher dimensional exact lumping transformations which can be described by elementary or other simple functions.

We consider a special case where the generalized eigenfunction ψ is a linear function of y_i

$$\psi = \sum_{i=1}^{n} c_i y_i \tag{69}$$

which satisfies eq. (56). Then

$$A\psi = \sum_{i=1}^{n} c_i f_i(\mathbf{y}) = \Omega(\psi). \tag{70}$$

Differentiating both sides of the above equation with respect to $\mathbf{y}$ gives

$$(c_1 \ c_2 \ \ldots \ c_n)\mathbf{f}_{\mathbf{y}}(\mathbf{y}) = \frac{d\Omega}{d\psi}(c_1 \ c_1 \ \ldots \ c_n). \tag{71}$$

This shows that vector $\mathbf{c} = (c_1 \ c_2 \ \ldots \ c_n)^{\mathrm{T}}$ is a left fixed (i.e. independent from $\mathbf{y}$) eigenvector of $\mathbf{f}_{\mathbf{y}}(\mathbf{y})$ or a fixed eigenvector of $\mathbf{f}_{\mathbf{y}}^{\mathrm{T}}(\mathbf{y})$ with eigenvalue $d\Omega/d\psi$.

If there exist $\hat{n}$ linear generalized eigenfunctions

$$\psi_j = \sum_{i=1}^{n} c_i^j y_i, \quad j = 1, \ldots, \hat{n} \tag{72}$$

i.e.

$$A\psi_j = \Omega_j(\psi_j) \tag{73}$$

then from eq. (71) we have

$$C\mathbf{f}_{\mathbf{y}}(\mathbf{y}) = Q(\psi)C \tag{74}$$

where

$$Q = \begin{pmatrix} \dfrac{d\Omega_1}{d\psi_1} & & \\ & \ddots & \\ & & \dfrac{d\Omega_{\hat{n}}}{d\psi_{\hat{n}}} \end{pmatrix} \tag{75}$$

$$C = \begin{pmatrix} c_1^1 & \ldots & c_n^1 \\ \vdots & & \vdots \\ c_1^{\hat{n}} & \ldots & c_n^{\hat{n}} \end{pmatrix}. \tag{76}$$

Equation (74) shows that the subspace $\mathscr{C}$ spanned by the row vectors of C is a left fixed invariant subspace of the Jacobian matrix $\mathbf{f}_{\mathbf{y}}(\mathbf{y})$ of $\mathbf{f}(\mathbf{y})$ or a fixed invariant subspace of $\mathbf{f}_{\mathbf{y}}^{\mathrm{T}}(\mathbf{y})$. This is the conclusion obtained in linear lumping (Li and Rabitz, 1989). If the original differential equation system is linear, $\mathbf{f}_{\mathbf{y}}(\mathbf{y})$ is a constant matrix. When this constant matrix is diagonalizable, it has n left linearly independent eigenvectors $\mathbf{c}^j$ and

$$\psi_j = \sum_{i=1}^{n} c_i^j y_i, \quad j = 1, \ldots, n \tag{77}$$

give n functionally independent generalized eigenfunctions. Then all normal generalized eigenfunctions and consequently all exact lumping schemes can be determined from them, at least in principle. However, it may happen that some of the functions cannot be expressed explicitly.

For nonlinear systems, the determination of generalized eigenfunctions and other higher dimensional exact lumping transformations is problem dependent. For different problems, we may have to employ different approaches to determine as many generalized eigenfunctions and other higher dimensional exact lumping transformtions as possible. The exact linear lumping schemes may be used as a starting point. The one-dimensional linear exact lumping schemes have given some generalized eigenfunctions with the form of eq. (77). If the original system has been reduced to a lower dimensional system by linear lumping, then the determination of generalized eigenfunctions and other higher dimensional exact lumping transformations of an n-dimensional system becomes the determination of generalized eigenfunctions and other higher dimensional exact lumping transformations of a lower dimensional system. The resultant problem may become simpler, especially when some generalized eigenfunctions contained in this lumped system are known.

Last, we show that using the resultant generalized eigenfunctions one can transform the lumped differential equation system, eq. (12), to some canonical forms. If we can determine $\Psi_1 = \Psi(\mathbf{h})$, $\Psi_{i+1} = \Psi(\mathbf{h}) + \Phi_i(\mathbf{h})$ $(i = 1, 2, \ldots, \hat{n} - 1)$ from eq. (62), then eq. (12) can be transformed into

$$\frac{d\Psi_i}{dt} = 1, \quad i = 1, 2, \ldots, \hat{n}.$$

Let

$$\psi_i = e^{\Psi_i}, \quad i = 1, 2, \ldots, \hat{n}. \tag{78}$$

Then

$$\frac{d\psi_i}{dt} = \frac{d\psi_i}{d\Psi_i}\frac{d\Psi_i}{dt} = e^{\Psi_i} = \psi_i, \quad i = 1, 2, \ldots, \hat{n}. \tag{79}$$

Generally, $d\psi_i/dt = \Omega_i(\psi_i)$ $(i = 1, 2, \ldots, \hat{n})$ is referred to as a diagonal form. Equation (79) is the simplest one of this form.

We can also transform eq. (12) into a triangular form. Let

$$\psi_i = \psi_i(\Psi_1, \Psi_2, \ldots, \Psi_i), \quad i = 1, 2, \ldots, \hat{n} \tag{80}$$

be an $\hat{n}$-dimensional nonsingular transformation. Its inverse transformation also has the form

$$\Psi_i = \Psi_i(\psi_1, \psi_2, \ldots, \psi_i), \quad i = 1, 2, \ldots, \hat{n}. \tag{81}$$

Then

$$\frac{d\psi_i}{dt} = \sum_{j=1}^{i} \frac{\partial \psi_i(\Psi_1, \Psi_2, \ldots, \Psi_i)}{\partial \Psi_j}\frac{d\Psi_j}{dt} = \sum_{j=1}^{i} \frac{\partial \psi_i[\Psi_1(\psi_1), \Psi_2(\psi_1, \psi_2), \ldots, \Psi_i(\psi_1, \ldots, \psi_i)]}{\partial \Psi_j} = \Omega_i(\psi_1, \psi_2, \ldots, \psi_i), \quad i = 1, 2, \ldots, \hat{n} \tag{82}$$

is a triangular form. Equations (79) and (82) are equivalent to the existence of functions $\psi_1(\mathbf{h}), \psi_2(\mathbf{h}), \ldots, \psi_{\hat{n}}(\mathbf{h})$ such that

$$\begin{aligned} A\psi_1 &= \psi_1 \\ A\psi_2 &= \psi_2 \\ &\cdots\cdots \\ A\psi_{\hat{n}} &= \psi_{\hat{n}} \end{aligned} \tag{83}$$

or

$$\begin{aligned} A\psi_1 &= \Omega_1(\psi_1) \\ A\psi_2 &= \Omega_2(\psi_1, \psi_2) \\ &\cdots\cdots \\ A\psi_{\hat{n}} &= \Omega_{\hat{n}}(\psi_1, \ldots, \psi_{\hat{n}}). \end{aligned} \tag{84}$$

If $\hat{n} = n$, the original system eq. (1) can be transformed to a diagonal or a triangular form, and the operator A in these coordinates will be called diagonal or triangular even though operator A does not possess a matrix form. We discuss these canonical forms of A further in Section 4.

3.2. *Determination of the generalized eigenfunctions*

As mentioned above one may not be able to find out all generalized eigenfunctions of an arbitrary ordinary differential equation system, but we wish to determine as many of them as possible. There are very likely several ways to determine generalized eigenfunctions. For example, Lie transformation groups may serve for this purpose (Wulfman and Rabitz, 1989). We present two approaches. One is using Lie algebra which can be applied to an arbitrary differential equation system under some conditions. If the right-hand side of a differential equation system is polynomial functions, then we can prove that any generalized eigenfunction of it is a function of the generalized eigenfunctions of the differential equation system whose right-hand side only consists of the linear terms of the original one. As the n functionally independent generalized eigenfunctions of a linear differential equation system can be completely determined, one may use them to find some generalized eigenfunctions of the original system.

3.2.1. *Determination of the generalized eigenfunctions by Lie algebra.* If the general solution of an n-dimensional ordinary differential equation system can be given in terms of some special solutions and a set of freely chosen constants, we say that the general solution of the system follows a nonlinear superposition principle (Stephani, 1989). When this principle holds, the right-hand side of the differential equation system can be represented as a linear combination of some n-dimensional function vectors whose combination coefficients are only functions of the independent variable or constants. These function vectors define a Lie algebra. Using a special representation, this Lie algebra is related to a linear differential equation system and its generalized eigenfunctions can be readily determined. Actually, these generalized eigenfunctions are just the generalized eigenfunctions of the original differential equation system, but represented in new dependent variables. Transforming the new variables to original ones will give the desired generalized eigenfunctions. We discuss this approach below.

First, we briefly recall the definition of Lie algebra and its adjoint representation. Suppose $\mathscr{G} = \{X_1, X_2, \ldots\}$ is a vector space of operators X_i. If any commutator of X_i and X_j still belongs to $\mathscr{G}$, i.e.

$$[X_i, X_j] = c_{ij}^k X_k \tag{85}$$

where X_k are basis operators of $\mathscr{G}$ and summation is taken over repeated index k, then $\mathscr{G}$ is a Lie algebra. The coefficients c_{ij}^k are called *structure constants*. When the commutator relations, eq. (85), of Lie algebra have been given, we then seek representations of X_i in some coordinates satisfying these relations. This process is called the realization of an abstract Lie algebra. A relevant realization is called the *adjoint representation*. Each operator is defined as a linear partial differential operator through the structure constants and variables z^i as follows:

$$X_i = -c_{ij}^k z^j \frac{\partial}{\partial z^k}. \tag{86}$$

It is easy to verify that the operators defined by eq. (86) obey the commutator relations given by eq. (85).

We now recall the relation between the nonlinear superposition of solutions and the existence of a certain Lie algebra composed by the operators corresponding to the differential equation system. Consider a first-order ordinary differential equation system

$$\frac{\mathrm{d}y^a}{\mathrm{d}x} = \omega^a(x, y^i), \quad a, i = 1, 2, \ldots, n. \tag{87}$$

For the sake of discussion, we do not use the vector representation and summation is taken over repeated indices. If the general solution of this system can be written as

$$\begin{aligned} y^a(x) &= F^a[y_L^b(x), c^k], \quad a, b, k = 1, 2, \ldots, n \\ & L = 1, 2, \ldots, m \end{aligned} \tag{88}$$

where $y_L^b(x)$ are special solutions of eq. (87), c^k are arbitrary constants, then we say that the general solution satisfies a *nonlinear superposition principle*. The necessary and sufficient condition for the existence of a nonlinear superposition solution has been obtained (Stephani, 1989): there exist a set of r functions $v^i(x)$ and nr functions $\eta_i^a(y^b)$ such that

$$\omega^a = v^i(x)\eta_i^a(y^b), \quad a, b = 1, 2, \ldots, n, \quad i = 1, 2, \ldots, r \tag{89}$$

holds and the operators

$$X_i = \eta_i^a \frac{\partial}{\partial y^a} \tag{90}$$

form a Lie algebra of dimension $r \leqslant mn$.

Equations (86) and (90) are different representations for the same X_i in the same Lie algebra. Using the same coefficients $v^i(x)$ in eq. (89) and using eq. (86) we can construct a linear differential equation system which is equivalent to eq. (87). The generalized eigenfunctions of the resultant linear differential equation system can then be easily determined. Equations (86) and (90) also supply the relations between z^j and y^i. Substituting these relations into the resultant generalized eigenfunctions of the linear system will give the desired generalized eigenfunctions of eq. (87). We now discuss the nonlinear superposition principle a little further. Suppose that there exists a nonsingular transformation

$$\mathbf{z} = T\mathbf{y} \tag{91}$$

such that the differential equation system for new variable vector $\mathbf{z}$ obtained by eqs (1) and (91) is of the form

$$\frac{d\mathbf{z}_1}{dt} = \mathbf{g}_1(\mathbf{z}_1)$$
$$\frac{d\mathbf{z}_2}{dt} = \mathbf{g}_2(\mathbf{z}_2) \tag{92}$$

where

$$\mathbf{z} = \begin{pmatrix} \mathbf{z}_1 \\ \mathbf{z}_2 \end{pmatrix}. \tag{93}$$

Since $\mathbf{z}_1$ and $\mathbf{z}_2$ are uncoupled, the general solution of eq. (92) is a superposition of solutions $\mathbf{z}_1$ and $\mathbf{z}_2$. As transformation T is nonsingular, the inverse T^{-1} exists and the general solution of the original differential equation system, eq. (1), also possesses the superposition property. This implies that $\mathbf{f}(\mathbf{y})$ can be divided into parts and the operators coresponding to these parts form a Lie algebra. The inverse negative proposition is also true, i.e. if the operators corresponding to the terms of $\mathbf{f}(\mathbf{y})$ in any decomposition cannot form a Lie algebra, then eq. (1) cannot be transformed to an uncoupled form as eq. (92). This analysis may be further applied to the lumped systems of $\mathbf{z}_1$ and $\mathbf{z}_2$. If $\mathbf{g}_i(\mathbf{z}_i)$ forms a Lie algebra, then the differential equation system of $\mathbf{z}_i$ can be further lumped. Note that the differential equation systems for $\mathbf{z}_1$ and $\mathbf{z}_2$ are lumped systems of eq. (1) and each system has a part of its generalized eigenfunctions. Using Lie algebra, we can separate the generalized eigenfunctions. At some stage when a lumped system does not form a Lie algebra, it implies that the generalized eigenfunctions contained in the lumped system cannot be further separated. It is likely that at least one of these generalized eigenfunctions cannot be represented by elementary or other simple functions. Then without further insight we have to stop there. However, quite adequate further approximate representations of the generalized eigenfunctions may exist, but the latter topic is beyond the scope of the present paper.

3.2.2. *Determination of the generalized eigenfunctions through the corresponding linear differential equations.* We now consider a special case of differential equation system as

$$d\mathbf{y}/dt = A\mathbf{y} + \mathbf{g}(\mathbf{y}) \tag{94}$$

where A is an $n \times n$ nonzero constant matrix and $\mathbf{g}(\mathbf{y})$ is a nonzero arbitrary function vector with the property

$$\mathbf{g}(\mathbf{0}) = \mathbf{0}, \quad \mathbf{g}_\mathbf{y}(\mathbf{0}) = \mathbf{0}. \tag{95}$$

This form is common in chemical kinetics, where $A\mathbf{y}$ represents the unimolecular reactions and $\mathbf{g}(\mathbf{y})$ describes higher-order reactions. [Note that the assumption of the form eq. (94) does not mean a restriction of generality as all functions can be expanded into this form with a nonzero A and a nonzero $\mathbf{g}(\mathbf{y})$.] We prove that any exact lumping transformation $\hat{\mathbf{y}} = \mathbf{h}(\mathbf{y})$ of eq. (94) is also an exact one for the linear differential equation system

$$d\mathbf{y}/dt = A\mathbf{y} \tag{96}$$

and the differential equation system

$$d\mathbf{y}/dt = \mathbf{g}(\mathbf{y}). \tag{97}$$

Obviously, if $\hat{\mathbf{y}} = \mathbf{h}(\mathbf{y})$ is an exact lumping transformation for both eqs (96) and (97), it is an exact lumping transformation of eq. (94) as well. As the generalized eigenfunctions and other higher dimensional exact lumping transformations of a linear system can be completely determined, then one can examine which of them are exact lumping transformations for eq. (94). It may possibly be simpler than the direct determination of some exact lumping schemes of eq. (94).

The proof is simple. If $\hat{\mathbf{y}} = \mathbf{h}(\mathbf{y})$ is an exact lumping transformation of eq. (1), then eq. (10) is satisfied. For eq. (94), eq. (10) is of the form

$$\mathbf{h}_\mathbf{y}(\mathbf{y})A\mathbf{y} + \mathbf{h}_\mathbf{y}(\mathbf{y})\mathbf{g}(\mathbf{y}) = \mathbf{h}_\mathbf{y}\{\bar{\mathbf{h}}[\mathbf{h}(\mathbf{y})]\}A\bar{\mathbf{h}}[\mathbf{h}(\mathbf{y})] + \mathbf{h}_\mathbf{y}\{\bar{\mathbf{h}}[\mathbf{h}(\mathbf{y})]\}\mathbf{g}\{\bar{\mathbf{h}}[\mathbf{h}(\mathbf{y})]\}. \tag{98}$$

Note that $A\mathbf{y}$ and $\mathbf{g}(\mathbf{y})$ are functionally independent. Equation (98) holds if and only if the following equations hold:

$$\mathbf{h}_\mathbf{y}(\mathbf{y})A\mathbf{y} = \mathbf{h}_\mathbf{y}\{\bar{\mathbf{h}}[\mathbf{h}(\mathbf{y})]\}A\bar{\mathbf{h}}[\mathbf{h}(\mathbf{y})] \tag{99}$$

$$\mathbf{h}_\mathbf{y}(\mathbf{y})\mathbf{g}(\mathbf{y}) = \mathbf{h}_\mathbf{y}\{\bar{\mathbf{h}}[\mathbf{h}(\mathbf{y})]\}\mathbf{g}\{\bar{\mathbf{h}}[\mathbf{h}(\mathbf{y})]\}. \tag{100}$$

These equations show that $\hat{\mathbf{y}} = \mathbf{h}(\mathbf{y})$ is an exact lumping transformation of eqs (96) and (97). Therefore, after the determination of all generalized eigenfunctions and other higher dimensional exact lumping transformations of eq. (96), we examine which of them are also ones of eq. (94) or simply of eq. (97). Those are the generalized eigenfunctions and other higher dimensional exact lumping transformations of eq. (94).

4. THE RELATIONSHIP BETWEEN EXACT LUMPING SCHEMES AND EIGENFUNCTIONS

From the above discussions, one knows that when n functionally independent generalized eigenfunctions ψ_i of A or the original ordinary differential equation system can be completely determined, A can be transformed into a diagonal form and all exact lumping schemes will be given by some functions of these

generalized eigenfunctions. For most systems, these generalized eigenfunctions cannot be completely determined. This conclusion follows from the fact that some generalized eigenfunctions may not be represented by elementary and other simple functions, and then they cannot be separated from other generalized eigenfunctions. However, they can be contained in some higher dimensional lumped coupled systems. Finding these lumped coupled systems is a task for future study. If all generalized eigenfunctions are separable, A possess a diagonal form. If they are not, can we transform A into some canonical form? In the following we prove that, if we give a proper definition, A can be transformed to a Jordan canonical form which will give some lumped coupled systems.

In order to do so, we first need to examine the eigenvalues and eigenfunctions of A or the original ordinary differential equation system. For a linear partial differential operator A if there exist a function $\lambda(\mathbf{y})$ and a function $\varphi(\mathbf{y})$, which is not identically zero, such that

$$A\varphi(\mathbf{y}) = \lambda(\mathbf{y})\varphi(\mathbf{y}) \tag{101}$$

where $\lambda(\mathbf{y})$ satisfies

$$A\lambda(\mathbf{y}) = 0 \tag{102}$$

we call $\lambda(\mathbf{y})$ and $\varphi(\mathbf{y})$ an eigenvalue and an eigenfunction of A. A function $\lambda(\mathbf{y})$ obeying eq. (102) is referred to as an invariant of A. It may be a function of $\mathbf{y}$ or a constant.

We now prove that for operator $A = \sum_{i=1}^{n} f_i(\mathbf{y})\,\partial/\partial y_i$ such eigenvalues and n functionally independent eigenfunctions always exist. For the operator A, we have proved that there exist n functionally independent normal generalized eigenfunctions Ψ_i and $n-1$ functionally independent first integrals Φ_i. Suppose $\lambda(\Phi_1, \Phi_2, \ldots, \Phi_{n-1})$ is an invariant and Ψ is a normal generalized eigenfunction of A. Then

$$\varphi = e^{\lambda\Psi} \tag{103}$$

is an eigenfunction of A corresponding to eigenvalue λ because

$$A\varphi = Ae^{\lambda\Psi} = \lambda e^{\lambda\Psi} = \lambda\varphi. \tag{104}$$

As there are n functionally independent Ψ_i, we can also find n functionally independent eigenfunctions. Even though there are an infinite number of eigenfunctions, at most only n of them are functionally independent.

If φ_1, φ_2 are two eigenfunctions of A corresponding eigenvalues λ_1 and λ_2, then $\varphi_1\varphi_2$ is also an eigenfunction of A with eigenvalue $\lambda_1 + \lambda_2$.

$$\begin{aligned} A\varphi_1\varphi_2 &= \varphi_2 A\varphi_1 + \varphi_1 A\varphi_2 \\ &= \lambda_1\varphi_1\varphi_2 + \lambda_2\varphi_1\varphi_2 \\ &= (\lambda_1 + \lambda_2)\varphi_1\varphi_2. \end{aligned} \tag{105}$$

We now use eigenfunctions to transform A to a Jordan canonical form. Since A is not of matrix form, we need a definition of Jordan form for operator A as follows. If there exist l eigenvalues $\lambda_i(y)$ $(i = 1, 2, \ldots, l)$ and for each $\lambda_i(\mathbf{y})$ there are k_i basis functions $\varphi_j^i(y)$ $(j = 1, 2, \ldots, k_i$ and $n = \sum_i k_i)$ such that

$$\begin{aligned} A\varphi_1^i &= \lambda_i\varphi_1^i \\ A\varphi_2^i &= \lambda_i\varphi_2^i + \varphi_1^i \\ &\cdots \\ A\varphi_{k_i}^i &= \lambda_i\varphi_{k_i}^i + \varphi_{k_i-1}^i \end{aligned} \tag{106}$$

we say that A possesses a Jordan canonical form under the coordinates φ_j^i.

We prove that these coordinates exist. For a given eigenvalue λ and $\hat{n}$ normal generalized eigenfunctions Ψ_i $(i = 1, 2, \ldots, \hat{n})$, we construct $\hat{n}$ eigenfunctions φ_i corresponding to the same eigenvalue λ by eq. (103). They satisfy the relation

$$A\varphi_i = \lambda\varphi_i, \quad i = 1, 2, \ldots, \hat{n}. \tag{107}$$

Let

$$\tilde{\varphi}_1 = \varphi_1 \tag{108}$$

$$\tilde{\varphi}_2 = \varphi_2 + \frac{\varphi_1}{\lambda}\ln\varphi_1 \tag{109}$$

$$\tilde{\varphi}_3 = (\varphi_2 + \varphi_3) + \frac{\varphi_2}{\lambda}\ln\varphi_1 + \frac{\varphi_1}{2!\lambda^2}(\ln\varphi_1)^2. \tag{110}$$

One can readily prove that

$$\begin{aligned} A\tilde{\varphi}_1 &= \lambda\tilde{\varphi}_1 \\ A\tilde{\varphi}_2 &= \lambda\tilde{\varphi}_2 + \tilde{\varphi}_1 \\ A\tilde{\varphi}_3 &= \lambda\tilde{\varphi}_3 + \tilde{\varphi}_2. \end{aligned} \tag{111}$$

Generally, let

$$\begin{aligned} \tilde{\varphi}_k = &\sum_{i=2}^{k}\varphi_i + \frac{1}{\lambda}\sum_{i=2}^{k-1}\varphi_i\ln\varphi_1 \\ &+ \frac{1}{2!\lambda^2}\sum_{i=2}^{k-2}\varphi_i(\ln\varphi_1)^2 + \cdots \\ &+ \frac{1}{l!\lambda^l}\sum_{i=2}^{k-l}\varphi_i(\ln\varphi_1)^l + \cdots \\ &+ \frac{\varphi_2}{(k-2)!\lambda^{k-2}}(\ln\varphi_1)^{k-2} \\ &+ \frac{\varphi_1}{(k-1)!\lambda^{k-1}}(\ln\varphi_1)^{k-1} \\ &k = 3, 4, \ldots, \hat{n}. \end{aligned} \tag{112}$$

Then we have

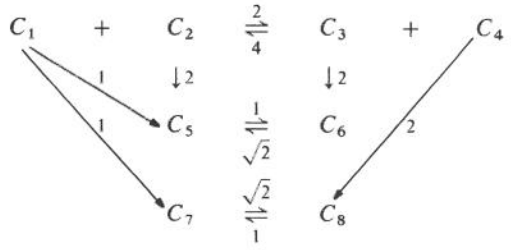
$$\begin{aligned} A\tilde{\varphi}_k = \lambda\Bigg[&\sum_{i=2}^{k}\varphi_i + \frac{1}{\lambda}\sum_{i=2}^{k-1}\varphi_i\ln\varphi_1 \\ &+ \frac{1}{2!\lambda^2}\sum_{i=2}^{k-2}\varphi_i(\ln\varphi_1)^2 + \cdots \\ &+ \frac{1}{l!\lambda^l}\sum_{i=2}^{k-l}\varphi_i(\ln\varphi_1)^l + \cdots \\ &+ \frac{\varphi_2}{(k-2)!\lambda^{k-2}}(\ln\varphi_1)^{k-2} \end{aligned}$$

$$+\frac{\varphi_1}{(k-1)!\,\lambda^{k-1}}(\ln\varphi_1)^{k-1}\Bigg]$$
$$+\Bigg\{\sum_{i=2}^{k-1}\varphi_i+\frac{1}{\lambda}\sum_{i=2}^{(k-1)-1}\varphi_i\ln\varphi_1+\cdots$$
$$+\frac{1}{(l-1)!\,\lambda^{l-1}}\sum_{i=2}^{(k-1)-(l-1)}\varphi_i(\ln\varphi_1)^{l-1}+\cdots$$
$$+\frac{\varphi_2}{[(k-1)-2]!\,\lambda^{(k-1)-2}}(\ln\varphi_1)^{(k-1)-2}$$
$$+\frac{\varphi_1}{[(k-1)-1]!\,\lambda^{(k-1)-1}}(\ln\varphi_1)^{(k-1)-1}\Bigg\}$$
$$=\lambda\tilde{\varphi}_k+\tilde{\varphi}_{k-1}. \tag{113}$$

The Jordan canonical forms will give us some lumped coupled differential equation systems. The Jordan form of A may also serve as a basis for developing an approach for approximate lumping without using the inverse transformation $\bar{\mathbf{h}}$. It will be presented in another paper.

5. AN EXAMPLE

A uni- and bimolecular reaction system with eight species used in our previous papers will be utilized here again:



where the C_is are species; the numbers are unitless rate constants.

Letting y_i represent the concentration of C_i, it is easy to write out the kinetic equations

$$\begin{aligned}
dy_1/dt &= -2y_1 - 2y_1y_2 + 4y_3y_4\\
dy_2/dt &= -2y_2 - 2y_1y_2 + 4y_3y_4\\
dy_3/dt &= -2y_3 - 4y_3y_4 + 2y_1y_2\\
dy_4/dt &= -2y_4 - 4y_3y_4 + 2y_1y_2\\
dy_5/dt &= -y_5 + y_1 + 2y_2 + \sqrt{2}y_6\\
dy_6/dt &= -\sqrt{2}y_6 + 2y_3 + y_5\\
dy_7/dt &= -\sqrt{2}y_7 + y_1 + y_8\\
dy_8/dt &= -y_8 + 2y_4 + \sqrt{2}y_7.
\end{aligned} \tag{114}$$

All exact linear lumping matrices have been obtained. Some of them are given below (Li and Rabitz, 1989):

$$M_1=(\alpha_1+\alpha_2\ \ \alpha_3\ \ \alpha_2\ \ \alpha_1+\alpha_3\ \ 0\ \ 0\ \ 0\ \ 0)$$

$$M_2=(2\ \ -2\sqrt{2}\ \ 4\ \ -2-2\sqrt{2}\ \ 2-\sqrt{2}\ \ 2-2\sqrt{2}\ \ -\sqrt{2}\ \ 1),$$

$$M_3=(1\ \ 1\ \ 1\ \ 1\ \ 1\ \ 1\ \ 1\ \ 1)$$

$$M_8=\begin{pmatrix}1&-1&0&0&0&0&0&0\\1&0&1&0&0&0&0&0\\1&0&0&1&0&0&0&0\end{pmatrix}$$

$$M_{12}=\begin{pmatrix}1&&&&\\&1&&&\\&&1&&0\\&&&1&\end{pmatrix}$$

$$M_{14}=\begin{pmatrix}1&0&0&1&&&&\\1&0&1&0&&0&&\\0&1&0&1&&&&\\0&0&0&0&2-\sqrt{2}&2-2\sqrt{2}&-\sqrt{2}&1\end{pmatrix}$$

$$M_{17}=\begin{pmatrix}1&&&&&&&\\&1&&&&0&&\\&&1&&&&&\\&&&1&&&&\\&&&&\gamma&\gamma&\delta&\delta\end{pmatrix}$$

where $\alpha_i, \gamma, \delta\in\mathscr{R}$ are arbitrary numbers. Note that matrix M_8 here has been chosen in a different way from that in our previous paper (Li and Rabitz, 1989).

The exact lumping matrices with dimension one yield some linear generalized eigenfunctions. Since M_1 has three linearly independent representations as the three rows of M_8, we know five generalized eigenfunctions of eq. (114). They are

$$\psi_1=y_1-y_2 \tag{115}$$
$$\psi_2=y_1+y_3 \tag{116}$$
$$\psi_3=y_1+y_4 \tag{117}$$
$$\psi_4=\sum_{i=1}^{8}y_i \tag{118}$$
$$\begin{aligned}\psi_5=&\,2y_1-2\sqrt{2}y_2+4y_3\\&-(2+2\sqrt{2})y_4+(2-\sqrt{2})y_5\\&+2(1-\sqrt{2})y_6-\sqrt{2}y_7+y_8.\end{aligned} \tag{119}$$

They are also the eigenfunctions corresponding to the eigenvalues $-2, -2, -2, 0$ and $-1-\sqrt{2}$.

5.1. *Lie algebra approach to determine generalized eigenfunctions*

We now want to determine the generalized eigenfunctions of eq. (114) by the Lie algebra approach. As discussed before, an $\hat{n}$-dimensional lumped system contains $\hat{n}$ functionally independent generalized eigenfunctions of the original system. For simplicity, we can determine generalized eigenfunctions from a lower dimensional lumped system. We start, e.g., from the lumped model given by M_{12}. The corresponding differential equation system is just the first

four equations of eq. (114):

$$\begin{aligned} dy_1/dt &= -2y_1 - 2y_1y_2 + 4y_3y_4 \\ dy_2/dt &= -2y_2 - 2y_1y_2 + 4y_3y_4 \\ dy_3/dt &= -2y_3 - 4y_3y_4 + 2y_1y_2 \\ dy_4/dt &= -2y_4 - 4y_3y_4 + 2y_1y_2 \end{aligned} \tag{120}$$

whose three linear generalized eigenfunctions given by eqs (115), (116) and (117) have been determined by linear lumping. There is another, yet undetermined nonlinear generalized eigenfunction.

There is no systematic way to determine the Lie algebra for a given system. However, if we fortunately know one operator X_1, then $X - X_1$, $[X_1, X]$, $[X_1, X - X_1]$ will give other operators. Here

$$X = \sum_{i=1}^{4} f_i(\mathbf{y}) \frac{\partial}{\partial y_i}. \tag{121}$$

We choose

$$X_1 = \sum_{i=1}^{4} y_i \frac{\partial}{\partial y_i}. \tag{122}$$

Then the remaining part of the right-hand side of eq. (120) multiplied by any constant will be another operator of the Lie algebra. Thus let

$$X_2 = (2y_1y_2 - 4y_3y_4)\left(\frac{\partial}{\partial y_1} + \frac{\partial}{\partial y_2} - \frac{\partial}{\partial y_3} - \frac{\partial}{\partial y_4}\right). \tag{123}$$

It is easy to check that X_1 and X_2 form a Lie algebra because

$$\begin{aligned} [X_1, X_2] &= X_2 \\ X &= -2X_1 - X_2. \end{aligned} \tag{124}$$

All structure constants of the Lie algebra are equal to zero except that

$$c_{12}^2 = -c_{21}^2 = 1. \tag{125}$$

Using eq. (86), we have the adjoint representations of the basis operators for this Lie algebra

$$X_1 = -z_2 \frac{\partial}{\partial z_2} \tag{126}$$

$$X_2 = z_1 \frac{\partial}{\partial z_2}. \tag{127}$$

Equations (122), (123) and (126), (127) are only different representations of the basis operators for the same Lie algebra. The operators of the Lie algebra obey eq. (124), and then the right-hand side $\mathbf{g}(\mathbf{z})$ of the differential equation system for $\mathbf{z}$ can be obtained by

$$\begin{pmatrix} 0 & 0 \\ -z_2 & z_1 \end{pmatrix}\begin{pmatrix} -2 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 2z_2 - z_1 \end{pmatrix} \tag{128}$$

$$\begin{aligned} \frac{dz_1}{dt} &= 0 \\ \frac{dz_2}{dt} &= 2z_2 - z_1. \end{aligned} \tag{129}$$

This is a linear system. Its generalized eigenfunctions can be readily determined. This system is so simple that by observation one can find a generalized eigenfunction as

$$z_3 = z_2 - \frac{1}{2}z_1 \tag{130}$$

we then have

$$\frac{dz_3}{dt} = 2z_3. \tag{131}$$

We now have a new differential equation system

$$\begin{aligned} \frac{dz_1}{dt} &= 0 \\ \frac{dz_3}{dt} &= 2z_3. \end{aligned} \tag{132}$$

Equation (132) shows that z_1 and z_3 are a first integral and a generalized eigenfunction of eq. (129), respectively. If one can find the expressions of z_i in terms of the coordinates y_i, then the first integrals and generalized eigenfunctions of eq. (120) are obtained. To find the connection between z_i and y_i, we apply eqs (122) and (123) to the left-hand side and eqs (126) and (127) to the right-hand side of z_i. For z_1, this yields

$$\sum_{i=1}^{4} y_i \frac{\partial z_1}{\partial y_i} = 0 \tag{133}$$

$$(2y_1y_2 - 4y_3y_4)\left(\frac{\partial z_1}{\partial y_1} + \frac{\partial z_1}{\partial y_2} - \frac{\partial z_1}{\partial y_3} - \frac{\partial z_1}{\partial y_4}\right) = 0. \tag{134}$$

The general solution of eqs (133) and (134) is

$$z_1 = \Omega(\Phi_1, \Phi_2) \tag{135}$$

where Ω is an arbitrary function and

$$\Phi_1 = \frac{y_1 - y_2}{y_1 + y_3} \tag{136}$$

$$\Phi_2 = \frac{y_1 - y_2}{y_1 + y_4}. \tag{137}$$

Equations (133) and (134) have four independent variables, and then there are two functionally independent solutions of z_1. Since Ω can be arbitrary, we simply choose

$$z_1^1 = \Phi_1 = \frac{y_1 - y_2}{y_1 + y_3} \tag{138}$$

$$z_1^2 = \Phi_2 = \frac{y_1 - y_2}{y_1 + y_4}. \tag{139}$$

They are the first integrals of eq. (120). Since any function of a first integral is still a first integral and the first integral may be a difference of two normal generalized eigenfunctions, we choose

$$z_1^1 = \frac{-1}{2}\ln\frac{y_1 - y_2}{y_1 + y_3} = \frac{-1}{2}\ln(y_1 - y_2) - \left(\frac{-1}{2}\right)\ln(y_1 + y_3) \tag{140}$$

$$z_1^2 = \frac{-1}{2}\ln\frac{y_1 - y_2}{y_1 + y_4} = \frac{-1}{2}\ln(y_1 - y_2) - \left(\frac{-1}{2}\right)\ln(y_1 + y_4). \quad (141)$$

They are all first integrals of eq. (120). These forms are especially convenient because

$$\Psi_1 = \frac{-1}{2}\ln(y_1 - y_2) \quad (142)$$

$$\Psi_2 = \frac{-1}{2}\ln(y_1 + y_3) \quad (143)$$

$$\Psi_3 = \frac{-1}{2}\ln(y_1 + y_4) \quad (144)$$

are normal generalized eigenfunctions of eq. (120). As $e^{-2\Psi_i}$ are generalized eigenfunctions, we obtain three generalized eigenfunctions

$$\psi_1 = y_1 - y_2 \quad (145)$$

$$\psi_2 = y_1 + y_3 \quad (146)$$

$$\psi_3 = y_1 + y_4 \quad (147)$$

which have been obtained by linear exact lumping.

Similarly for z_2 we have the following equations

$$\sum_{i=1}^{4} y_i \frac{\partial z_2}{\partial y_i} = -z_2 \quad (148)$$

$$(2y_1y_2 - 4y_3y_4)\left(\frac{\partial z_2}{\partial y_1} + \frac{\partial z_2}{\partial y_2} - \frac{\partial z_2}{\partial y_3} - \frac{\partial z_2}{\partial y_4}\right) = z_1. \quad (149)$$

The general solution of the above equations is

$$z_2 = \frac{1}{y_1 - y_2}\left\{\Omega(\Phi_1, \Phi_2) + \frac{\Phi_1\Phi_2 z_1}{4(\gamma_1 - \gamma_2)}\ln\frac{(y_1 - y_2)/y_1 - \gamma_1}{(y_1 - y_2)/y_1 - \gamma_2}\right\} \quad (150)$$

where Ω, Φ_1 and Φ_2 are the same as before, γ_1 and γ_2 are the roots of the following quadratic equation:

$$x^2 - \left(\Phi_1 + \Phi_2 - \frac{\Phi_1\Phi_2}{2}\right)x + \frac{\Phi_1\Phi_2}{2} = 0. \quad (151)$$

There are also two functionally independent solutions of z_2 which can be obtained by substituting functionally independent z_1 into eq. (150). We choose $\Omega = 0$ and $z_1 = 4/\Phi_2$, $4/\Phi_1$, respectively. Then

$$z_2^1 = \frac{\Phi_1}{(\gamma_1 - \gamma_2)(y_1 - y_2)}\ln\frac{(y_1 - y_2)/y_1 - \gamma_1}{(y_1 - y_2)/y_1 - \gamma_2} \quad (152)$$

$$z_2^2 = \frac{\Phi_2}{(\gamma_1 - \gamma_2)(y_1 - y_2)}\ln\frac{(y_1 - y_2)/y_1 - \gamma_1}{(y_1 - y_2)/y_1 - \gamma_2}. \quad (153)$$

Using eq. (130), we have

$$z_3^1 = \frac{\Phi_1}{(\gamma_1 - \gamma_2)(y_1 - y_2)}\ln\frac{(y_1 - y_2)/y_1 - \gamma_1}{(y_1 - y_2)/y_1 - \gamma_2} - \frac{2}{\Phi_2} \quad (154)$$

$$z_3^2 = \frac{\Phi_2}{(\gamma_1 - \gamma_2)(y_1 - y_2)}\ln\frac{(y_1 - y_2)/y_1 - \gamma_1}{(y_1 - y_2)/y_1 - \gamma_2} - \frac{2}{\Phi_1}. \quad (155)$$

Note that γ_1 and γ_2 are functions of the invariant first integrals Φ_1 and Φ_2. Therefore, Φ_1, Φ_2 and γ_1, γ_2 are all constant with respect to time t.

Let

$$\psi_6 = z_3. \quad (156)$$

Then we obtain the whole set of four generalized eigenfunctions for eq. (120). They can be composed of

$$y_1 - y_2, \quad y_1 + y_3, \quad y_1 + y_4, \quad z_3^1$$

or

$$y_1 - y_2, \quad y_1 + y_3, \quad y_1 + y_4, \quad z_3^2.$$

We now want to determine other functionally independent generalized eigenfunctions of eq. (114). In order to do so, we use the lumped model given by the lumping matrix M_{17}, for the sake of simplicity with $\gamma = \delta = 1$,

$$M_{17} = \begin{pmatrix} 1 & & & & & & & \\ & 1 & & & & 0 & & \\ & & 1 & & & & & \\ & & & 1 & & & & \\ & & & & 1 & 1 & 1 & 1 \end{pmatrix}.$$

The lumped differential equation system is as follows:

$$\begin{aligned} dy_1/dt &= -2y_1 - 2y_1y_2 + 4y_3y_4 \\ dy_2/dt &= -2y_2 - 2y_1y_2 + 4y_3y_4 \\ dy_3/dt &= -2y_3 - 4y_3y_4 + 2y_1y_2 \\ dy_4/dt &= -2y_4 - 4y_3y_4 + 2y_1y_2 \\ d\hat{y}_5/dt &= 2y_1 + 2y_2 + 2y_3 + 2y_4 \end{aligned} \quad (157)$$

where

$$\hat{y}_5 = \sum_{i=5}^{8} y_i. \quad (158)$$

Similarly, we can find a Lie algebra composed by

$$X_1 = \sum_{i=1}^{4} y_i \frac{\partial}{\partial y_i} \quad (159)$$

$$X_2 = (2y_1y_2 - 4y_3y_4)\left(\frac{\partial}{\partial y_1} + \frac{\partial}{\partial y_2} - \frac{\partial}{\partial y_3} - \frac{\partial}{\partial y_4}\right) \quad (160)$$

$$X_3 = \left(\sum_{i=1}^{4} y_i\right)\frac{\partial}{\partial \hat{y}_5}. \quad (161)$$

It is easy to check that X_i $(i = 1, 2, 3)$ form a Lie algebra because

$$[X_1, X_2] = X_2 \quad (162)$$

$$[X_1, X_3] = X_3 \quad (163)$$

$$[X_2, X_3] = 0 \quad (164)$$

and

$$X = -2X_1 - X_2 + 2X_3. \tag{165}$$

The nonzero structure constants of the Lie algebra are

$$c_{12}^2 = -c_{21}^2 = c_{13}^3 = -c_{31}^3 = 1. \tag{166}$$

On using eq. (86), we obtain the adjoint representation of the basis operators for this Lie algebra

$$X_1 = -z_2\frac{\partial}{\partial z_2} - z_3\frac{\partial}{\partial z_3} \tag{167}$$

$$X_2 = z_1\frac{\partial}{\partial z_2} \tag{168}$$

$$X_3 = z_1\frac{\partial}{\partial z_3}. \tag{169}$$

The operators of the Lie algebra obey eq. (165), and then the right-hand side $\mathbf{g}(\mathbf{z})$ of the differential equation system for $\mathbf{z}$ can be obtained by

$$\begin{pmatrix} 0 & 0 & 0 \\ -z_2 & z_1 & 0 \\ -z_3 & 0 & z_1 \end{pmatrix}\begin{pmatrix} -2 \\ -1 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 2z_2 - z_1 \\ 2z_3 + 2z_1 \end{pmatrix} \tag{170}$$

$$\frac{dz_1}{dt} = 0, \quad \frac{dz_2}{dt} = 2z_2 - z_1, \quad \frac{dz_3}{dt} = 2z_3 + 2z_1. \tag{171}$$

If we construct two new variables

$$z_4 = z_2 - \tfrac{1}{2}z_1 \tag{172}$$

$$z_5 = z_3 + z_1 \tag{173}$$

we then have a new differential equation system

$$\frac{dz_1}{dt} = 0, \quad \frac{dz_4}{dt} = 2z_4, \quad \frac{dz_5}{dt} = 2z_5. \tag{174}$$

Note that z_1 and z_4 are the same as z_1 and z_3 for the lumped model given by M_{12}. Therefore, we only discuss the generalized eigenfunction z_5 here.

Comparing X_i for different representations gives

$$\sum_{i=1}^{4} y_i\frac{\partial z_5}{\partial y_i} = -z_3 \tag{175}$$

$$(2y_1y_2 - 4y_3y_4)\left(\frac{\partial z_5}{\partial y_1} + \frac{\partial z_5}{\partial y_2} - \frac{\partial z_5}{\partial y_3} - \frac{\partial z_5}{\partial y_4}\right) = 0 \tag{176}$$

$$\left(\sum_{i=1}^{4} y_i\right)\frac{\partial z_5}{\partial \hat{y}_5} = z_1. \tag{177}$$

Using eq. (173) the solution of eqs (175), (176) and (177) is obtained as

$$z_5 = \frac{\sum_{i=1}^{8} y_i}{\sum_{i=1}^{4} y_i}z_1. \tag{178}$$

As z_1 has two functionally independent solutions, so does z_5.

Observing that $\sum_{i=1}^{4} y_i$ and z_1 are all functions of ψ_i $(i = 1, 2, 3)$, the new term is $\sum_{i=1}^{8} y_i$ and we can readily prove that

$$\psi_4 = \sum_{i=1}^{8} y_i \tag{179}$$

is another generalized eigenfunction which has already been given by linear lumping.

Similarly, ψ_5 can be determined from the lumped model given by M_{14}. Then for eq. (114) there are only two nonlinear generalized eigenfunctions left as unknown. Unfortunately, we have not found other Lie algebras to determine them. Since eq. (114) possesses the form of eq. (94), we determine the generalized eigenfunctions by the second approach, i.e. the determination of generalized eigenfunctions through the generalized eigenfunctions of the corresponding linear differential equation system.

5.2. *Approach through the corresponding linear differential equation system to determine generalized eigenfunctions*

As demonstrated in Section 3.2.2, any generalized eigenfunction of eq. (94) is also one of the corresponding linear systems [eq. (96)]. Suppose the n functionally independent generalized eigenfunctions x_i $(i = 1, 2, \ldots, n)$ of eq. (96) are known. Then any generalized eigenfunction of eq. (94) can be represented as

$$\psi = \psi(x_1, x_2, \ldots, x_n). \tag{180}$$

A normal generalized eigenfunction is a generalized eigenfunction, then we also have

$$\Psi = \Psi(x_1, x_2, \ldots, x_n). \tag{181}$$

After the determination of the generalized eigenfunctions of the linear system, we can determine the normal generalized eigenfunctions Ψ of eq. (94). The procedure is the same as the determination of the normal generalized eigenfunctions by solving eq. (57) except that the dependent variables y_i are replaced by x_i. We show the procedure by this example.

The corresponding linear differential equation system of eq. (114) is

$$\begin{aligned} dy_1/dt &= -2y_1 \\ dy_2/dt &= -2y_2 \\ dy_3/dt &= -2y_3 \\ dy_4/dt &= -2y_4 \\ dy_5/dt &= -y_5 + y_1 + 2y_2 + \sqrt{2}y_6 \\ dy_6/dt &= -\sqrt{2}y_6 + 2y_3 + y_5 \\ dy_7/dt &= -\sqrt{2}y_7 + y_1 + y_8 \\ dy_8/dt &= -y_8 + 2y_4 + \sqrt{2}y_7. \end{aligned} \tag{182}$$

Eight functionally independent generalized eigenfunctions of eq. (182) can be found by the determination of the left eigenvectors of the coefficient matrix of

eq. (182) as follows:

$$\begin{aligned}
x_1 &= y_1\\
x_2 &= y_2\\
x_3 &= y_3\\
x_4 &= y_4\\
x_5 &= 2y_1 - 2\sqrt{2}y_2 + 4y_3 - (2+2\sqrt{2})y_4\\
&\quad + (2-\sqrt{2})y_5 + 2(1-\sqrt{2})y_6\\
&\quad - \sqrt{2}y_7 + y_8\\
x_6 &= y_1 - \sqrt{2}y_4 + (1-\sqrt{2})y_7 + \frac{(2-\sqrt{2})}{2}y_8\\
x_7 &= \sum_{i=1}^{8} y_i\\
x_8 &= y_1 + 2y_4 + 2y_7 + 2y_8.
\end{aligned} \tag{183}$$

For these generalized eigenfunctions, Ω of eq. (56) is linear and they are eigenfunctions with the form $\lambda_i x_i$. The eigenvalues for x_1, x_2, x_3, x_4; x_5, x_6; x_7, x_8 are -2, $-(1+\sqrt{2})$ and 0, respectively. Therefore, any linear combination of x_1, x_2, x_3, x_4 or x_5, x_6 or x_7, x_8 is still a generalized eigenfunction of eq. (182). Considering the results of linear exact lumping we choose

$$\begin{aligned}
x_1 &= y_1\\
x_2 &= y_1 - y_2\\
x_3 &= y_1 + y_3\\
x_4 &= y_1 + y_4\\
x_5 &= 2y_1 - 2\sqrt{2}y_2 + 4y_3 - (2+2\sqrt{2})y_4\\
&\quad + (2-\sqrt{2})y_5 + 2(1-\sqrt{2})y_6 - \sqrt{2}y_7 + y_8\\
x_6 &= y_1 - \sqrt{2}y_4 + (1-\sqrt{2})y_7 + \frac{(2-\sqrt{2})}{2}y_8\\
x_7 &= \sum_{i=1}^{8} y_i\\
x_8 &= y_1 + 2y_4 + 2y_7 + 2y_8.
\end{aligned} \tag{184}$$

Suppose $\Psi = \Psi(x_1, x_2, \ldots, x_n)$ is a normal generalized eigenfunction of eq. (114). Then it satisfies the following equation:

$$A\Psi = \sum_{i=1}^{n} \frac{\partial \Psi}{\partial x_i} A x_i = 1 \tag{185}$$

which gives

$$\begin{aligned}
&(-2x_1 - u)\frac{\partial \Psi}{\partial x_1} + \sum_{i=2}^{4} -2x_i \frac{\partial \Psi}{\partial x_i} - (1+\sqrt{2})x_5 \frac{\partial \Psi}{\partial x_5}\\
&-(1+\sqrt{2})(x_6+u)\frac{\partial \Psi}{\partial x_6} + 0\frac{\partial \Psi}{\partial x_7} + u\frac{\partial \Psi}{\partial x_8} = 1
\end{aligned} \tag{186}$$

where

$$u = 2y_1y_2 - 4y_3y_4. \tag{187}$$

The corresponding characteristic equation of eq. (186) is

$$\begin{aligned}
\frac{dx_1}{-2x_1-u} &= \frac{dx_2}{-2x_2} = \frac{dx_3}{-2x_3} = \frac{dx_4}{-2x_4}\\
&= \frac{dx_5}{-(1+\sqrt{2})x_5} = \frac{dx_6}{-(1+\sqrt{2})(x_6+u)}\\
&= \frac{dx_7}{0} = \frac{dx_8}{u} = \frac{d\Psi}{1}.
\end{aligned} \tag{188}$$

The general solution of eq. (188) is

$$\Psi = \Psi_1 + \phi(\Phi_1, \Phi_2, \ldots, \Phi_{n-1}) \tag{189}$$

where ϕ is an arbitrary function and

$$\begin{aligned}
\Psi_1 &= -\tfrac{1}{2}\ln(y_1 - y_2)\\
\Phi_1 &= -(y_1 - y_2) - \frac{1}{\mu_1 - \mu_2}\ln\frac{y_1 - \mu_1(y_1 - y_2)}{y_1 - \mu_2(y_1 - y_2)}\\
\Phi_2 &= \frac{y_1 + y_3}{y_1 - y_2}\\
\Phi_3 &= \frac{y_1 + y_4}{y_1 - y_2}\\
\Phi_4 &= x_5/(y_1 - y_2)^{(1+\sqrt{2})/2}\\
\Phi_5 &= x_2^{(-1+\sqrt{2})/2}x_6 - (1+\sqrt{2})(\mu_1 - \mu_2)^2\\
&\quad \times \int \frac{e^v x_2^{(1-\sqrt{2})/2}\,dx_2}{(1-e^v)^2}\\
\Phi_6 &= \sum_{i=1}^{8} y_i\\
\Phi_7 &= 2y_1 + 2y_4 + 2y_7 + 2y_8 - \mu_2(y_1 - y_2)\\
&\quad + \ln\left(\frac{y_1}{y_1 - y_2} - \mu_1\right)
\end{aligned} \tag{190}$$

where μ_1 and μ_2 are the roots of the following quadratic equation:

$$x^2 + (1 - 2(\Phi_1 + \Phi_2))x + 2\Phi_1\Phi_2 = 0 \tag{191}$$

and

$$v = (\mu_2 - \mu_1)(\Phi_1 + x_2). \tag{192}$$

If we choose

$$\Psi_1 = \Psi_1, \quad \Psi_{i+1} = \Psi_1 + \Phi_i, \quad i = 1, 2, \ldots, 7 \tag{193}$$

then n functionally independent normal generalized eigenfunctions of eq. (114) are obtained. Note that Φ_5 cannot probably be given an analytic form explicitly. Actually only seven normal generalized eigenfunctions have been determined. Since some functions of a normal eigenfunction can be generalized ones, for a chosen function we can obtain a generalized eigenfunction. For example,

$$\psi = \exp[-2(\Psi_1 - \tfrac{1}{2}\ln\Phi_2)] = y_1 + y_3. \tag{194}$$

In this way we can construct functionally independent generalized eigenfunctions. Using $\hat{n}$ generalized eigenfunctions of them, one can construct $\hat{n}$ functionally

independent new functions $h_i(\psi_1, \psi_2, \ldots, \psi_{\hat{n}})$ which are $\hat{n}$-dimensional exact lumping transformations.

5.3. *Determination of basis functions for a Jordan form of* A

From the linear lumping, we have obtained some eigenfunctions given in eqs (115)–(119). We can use one of them to determine a function φ which will satisfy the equation

$$A\varphi = \lambda_i\varphi + \varphi_i \tag{195}$$

where φ_i is an eigenfunction of A, i.e.

$$A\varphi_i = \lambda_i\varphi_i. \tag{196}$$

We choose ψ_5 with $\lambda_5 = -(1+\sqrt{2})$ given in eq. (119). In order to determine φ, we need to solve an equation which is similar to eq. (186):

$$\begin{aligned}(-2x_1-u)\frac{\partial\varphi}{\partial x_1} + \sum_{i=2}^{4} -2x_i\frac{\partial\varphi}{\partial x_i} - (1+\sqrt{2})x_5\frac{\partial\varphi}{\partial x_5} \\ - (1+\sqrt{2})(x_6+u)\frac{\partial\varphi}{\partial x_6} + 0\frac{\partial\varphi}{\partial x_7} + \mu\frac{\partial\varphi}{\partial x_8} \\ = -(1+\sqrt{2})\varphi + x_5\end{aligned} \tag{197}$$

where $x_5 = \psi_5$. The corresponding characteristic equation of eq. (197) is

$$\begin{aligned}\frac{dx_1}{-2x_1-u} = \frac{dx_2}{-2x_2} = \frac{dx_3}{-2x_3} = \frac{dx_4}{-2x_4} \\ = \frac{dx_5}{-(1+\sqrt{2})x_5} = \frac{dx_6}{-(1+\sqrt{2})(x_6+u)} \\ = \frac{dx_7}{0} = \frac{dx_8}{u} = \frac{d\varphi}{-(1+\sqrt{2})\varphi + x_5}.\end{aligned} \tag{198}$$

All the solutions of Φ_i $(i = 1, 2, \ldots, 7)$ are the same as those given in eq. (190). We only need to determine the last invariant Φ_8 by solving the equation

$$\frac{dx_5}{-(1+\sqrt{2})x_5} = \frac{d\varphi}{-(1+\sqrt{2})\varphi + x_5}. \tag{199}$$

The resultant solution is

$$\Phi_8 = \frac{1}{x_5}\left(\varphi + \frac{x_5}{1+\sqrt{2}}\ln x_5\right). \tag{200}$$

Then the general solution of φ is

$$\begin{aligned}\varphi &= \Phi_8 x_5 - \frac{x_5}{1+\sqrt{2}}\ln x_5 \\ &= \phi(\Phi_1, \Phi_2, \ldots, \Phi_7)x_5 - \frac{x_5}{1+\sqrt{2}}\ln x_5 \\ &= \phi(\Phi_1, \Phi_2, \ldots, \Phi_7)\psi_5 - \frac{\psi_5}{1+\sqrt{2}}\ln\psi_5\end{aligned} \tag{201}$$

where ϕ is an arbitrary function. If $\phi(\Phi_1, \Phi_2, \ldots, \Phi_7)$ is functionally independent from ψ_i $(i = 1, 2, \ldots, 5)$, then we obtain a two-dimensional lumped system which contains a new functionally independent eigenfunction. Similarly, we can determine other lumped systems with a Jordan form by using other eigenfunctions or the resultant φ.

6. CONCLUSIONS AND A DISCUSSION

In this paper a general analysis of exact nonlinear lumping has been presented. It can be used for any system described by a set of first-order ordinary differential equations with any degree of nonlinearity. A chemical kinetic system is only a special case of this analysis.

Three necessary and sufficient conditions for the existence of exact lumping schemes for a given system have been derived. If $\hat{\mathbf{y}} = \mathbf{h}(\mathbf{y})$ is an exact lumping transformation, then $\mathbf{h}(\mathbf{y}) = \mathbf{0}$ is an invariant manifold of the original differential equation system $d\mathbf{y}/dt = \mathbf{f}(\mathbf{y})$. The lumped differential equation system can be described as $d\hat{\mathbf{y}}/dt = \mathbf{h}_{\mathbf{y}}[\bar{\mathbf{h}}(\hat{\mathbf{y}})]\mathbf{f}[\bar{\mathbf{h}}(\hat{\mathbf{y}})]$ and is independent of the choice of the generalized inverse transformation $\bar{\mathbf{h}}$.

For $d\mathbf{y}/dt = \mathbf{f}(\mathbf{y})$, a corresponding linear partial differential operator $A = \sum_{i=1}^{n} f_i(\mathbf{y})\partial/\partial y_i$ is defined. A definition of a generalized eigenfunction for A has been given. It has been proved that for an n-dimensional ordinary differential equation system there exist n functionally independent generalized eigenfunctions. Any $\hat{n}$-dimensional lumped system must have $\hat{n}$ functionally independent generalized eigenfunctions. Any $\hat{n}$ functionally independent functions of $\hat{n}$ functionally independent generalized eigenfunctions compose an $\hat{n}$-dimensional lumped system. Therefore, after the full determination of the generalized eigenfunctions, all lumped systems are obtained. Normal generalized eigenfunctions are special generalized eigenfunctions. In the coordinates of n functionally independent generalized eigenfunctions ψ_i the operator A is referred to as being in a diagonal canonical form.

Two approaches to determine generalized eigenfunctions are presented. If the general solution of a differential equation system follows a nonlinear superposition principle, some generalized eigenfunctions of the system can be determined through a Lie algebra approach. When a differential equation system contains linear terms, any generalized eigenfunction of the system is a generalized eigenfunction of the differential equation system which is only composed of the linear terms. As the generalized eigenfunctions of a linear system can be readily determined, using them one may possibly find some generalized eigenfunctions of the original system. It may be easier than the direct determination of the generalized eigenfunction of the original system.

In many cases generalized eigenfunctions cannot be completely determined. This may come from the fact that some generalized eigenfunctions may not be described by elementary or other simple functions. Therefore, we may not be able to separate all generalized eigenfunctions for some systems. However, these undetermined generalized eigenfunctions may compose a lumped coupled differential equation system

with some other generalized eigenfunctions. In order to determine these lumped coupled systems, the eigenvalues and eigenfunctions of A are defined. It has been proved that A has an infinite number of eigenfunctions, but at most only n eigenfunctions are functionally independent. Using these eigenfunctions the operator A can be transformed to a Jordan canonical forms which give some lumped coupled differential equation systems.

As discussed above, the nonlinear lumping schemes are related to global invariant manifolds in the geometric respect, or generalized eigenfunctions and consequently canonical forms of a partial differential operator in the algebraic respect. However, many global invariant manifolds and generalized eigenfunctions of an ordinary differential equation system cannot be represented by elementary and other simple functions. Therefore, we do not expect to find a general approach to determine all exact nonlinear lumping schemes. Actually, the determination of all generalized eigenfunctions is equivalent to the determination of all first integrals for an ordinary differential equation system in mathematics, or the determination of all conserved quantities for a classical dynamic system in physics. For a long time this task has been known to be very difficult, if not impossible. It is not surprising then that the two approaches to determine the exact lumping schemes in the present paper are quite restrictive.

Fortunately, in practice some degree of error of a lumped model is acceptable. Therefore, approximate nonlinear lumping is more important for realistic systems. The results of exact nonlinear lumping can serve as a theoretical basis for the development of approaches for approximate nonlinear lumping. Based on this analysis we developed two approaches for approximate nonlinear lumping (Li *et al.*, 1993, 1994; Tomlin *et al.*, 1994).

Although the main purpose of the present paper is to establish a theoretical basis for approximate nonlinear lumping, the significance of the result in exact nonlinear lumping is also evident. It is well known that the eigenfunctions of a harmonic system described by a set of ordinary differential equations correspond to normal modes of the motion. They are fundamental in the understanding of the motion. What is the meaning of the eigenfunctions for a chemical reaction system? How do they give the fundamental description of the movement for a reaction system? At least we know that for a first-order reaction system the eigenfunctions correspond to straight line reaction paths in the composition space (Wei and Prater, 1962). Discovering the physical meaning of the eigenfunctions and other high-order invariant manifolds of a reaction system will yield a deep understanding of a reaction course, not only the concentration change but also similar properties, such as driving force, chemical potential, etc. This is very interesting subject in the study of chemical kinetics and will be a future topic of study. Further open questions include extension of the method to considering lumping under some imposed restrictions and revealing whether there is a change of the qualitative properties of the lumped system compared to the original system.

Acknowledgments—The authors acknowledge support from the Department of Energy and Exxon Corporation. One of the authors (J.T.) has also been supported by the Soros Foundation and by the National Scientific Research Fund, Hungary, No. 3286.

NOTATION

Scalars

a	positive integer
b	positive integer
c_i	constant
c_i^j	constant
c_{ij}^k	structure constant
c^k	constant
C_i	ith species of a reaction system
$\mathscr{C}$	$\hat{n}$-dimensional subspace of $\mathscr{R}^n$
$\mathscr{G}$	vector space of X_i
h_i	ith element of vector $\mathbf{h}$
i	positive integer
j	positive integer
k	positive integer
l	positive integer
m	positive integer
n	dimension of vector $\mathbf{y}$
$\hat{n}$	dimension of vector $\hat{\mathbf{y}}$
r	positive integer
$\mathscr{R}^n$	n-dimensional real space
$\mathscr{R}^{\hat{n}}$	$\hat{n}$-dimensional real space
S	invariant manifold
t	time
T	real number
x	independent variable
y_i	ith element of vector $\mathbf{y}$
y^i	ith element of vector $\mathbf{y}$
$y_L^b(x)$	bth element of Lth special solution of eq. (87)
z_i	ith element of vector $\mathbf{z}$
z^i	ith element of vector $\mathbf{z}$

Vectors and matrices

Capital letters represent matrices; bold-face lower case letters represent vectors.

A	linear differential operator or $n \times n$ constant matrix or defined as $(\mathbf{h}_\mathbf{y}\mathbf{f})_\mathbf{y}(\mathbf{y})$
B	defined as $\mathbf{h}_\mathbf{y}(\mathbf{y})$
B^+	generalized inverse of B
$\mathbf{c}$	$\hat{n}$- or n-dimensional constant vector
$\bar{\mathbf{c}}$	$\hat{n}$-dimensional constant vector
$\mathbf{c}^j$	n-dimensional constant vector
C	$\hat{n} \times n$ constant matrix or defined as AB^+
$\mathbf{f}(\mathbf{y})$	n-dimensional function vector
$\mathbf{f}_\mathbf{y}(\mathbf{y})$	Jacobian matrix of $\mathbf{f}(\mathbf{y})$
$\hat{\mathbf{f}}(\hat{\mathbf{y}})$	$\hat{n}$-dimensional function vector
$\mathbf{g}(\mathbf{y})$	n-dimensional nonlinear function vector
$\mathbf{g}_\mathbf{y}(\mathbf{y})$	Jacobian matrix of $\mathbf{g}(\mathbf{y})$
$\mathbf{g}_i(\mathbf{z}_i)$	function vector
$\mathbf{h}(\mathbf{y})$	$\hat{n}$-dimensional function vector
$\mathbf{h}_\mathbf{y}(\mathbf{y})$	Jacobian matrix of $\mathbf{h}(\mathbf{y})$

$\mathbf{h}_i(\mathbf{y})$	$\hat{n}$-dimensional function vector
$\bar{\mathbf{h}}(\hat{\mathbf{y}})$	n-dimensional function vector satisfying $\mathbf{h}(\bar{\mathbf{h}}) = I_{\hat{n}}$
$\bar{\mathbf{h}}^*(\hat{\mathbf{y}})$	n-dimensional function vector satisfying $\mathbf{h}(\bar{\mathbf{h}}^*) = I_{\hat{n}}$
I	identity matrix
M	lumping matrix
$\bar{M}$	generalized inverse of M satisfying $M\bar{M} = I_{\hat{n}}$
Q	$\hat{n} \times \hat{n}$ diagonal matrix
T	nonsingular transformation
T^{-1}	inverse of T
x_i	generalized eigenfunction of a linear system
X	linear operator or $n \times (n - \hat{n})$ matrix
X_i	linear operator
$X(\mathbf{y})$	$n \times (n - \hat{n})$ matrix
$\mathbf{y}$	n-dimensional variable vector
$\hat{\mathbf{y}}$	$\hat{n}$-dimensional variable vector
$\mathbf{z}$	n-dimensional variable vector
$\mathbf{z}_i$	variable vector

Greek letters

α_i	real number
γ	real number
γ_i	function of first integrals Φ_i
δ	real number
$\eta_i^a(y^b)$	function of $\mathbf{y}$
λ	eigenvalue
λ_i	eigenvalue
μ_i	function of first integrals Φ_i
$v^i(x)$	function of x
ϕ	arbitrary function
φ	eigenfunction
φ_i	eigenfunction
φ_j^i	Jordan basis function
$\tilde{\varphi}_i$	Jordan basis function
Φ	first integral
Φ_i	first integral
ψ	generalized eigenfunction
ψ_i	generalized eigenfunction
Ψ	normal generalized eigenfunction
Ψ_i	normal generalized eigenfunction
$\omega^a(x, y^i)$	function of x and $\mathbf{y}$
Ω	arbitrary function
Ω_i	arbitrary function

Symbols

$\hat{}$	any property related to the lumped system
$\mathbf{0}$	null vector

REFERENCES

Carr, J., 1981, *Applications of Centre Manifold Theory*. Springer, New York.

Fleming, W. H., 1977, *Functions of Several Variables*, 2nd edition. Springer, New York.

Iwasa, Y., Andreasen, V. and Levin, S., 1987, Aggregation in model ecosystems. I. Perfect aggregation. *Ecological Modelling* **37**, 287–302.

Iwasa, Y., Levin, S. A. and Andreasen, V., 1989, Aggregation in model ecosystems. II. Approximate aggregation. *IMA J. Math. appl. Med. Biol.* **6**, 1–23.

Li, G. and Rabitz, R., 1989, A general analysis of exact lumping in chemical kinetics. *Chem. Engng Sci.* **44**, 1413–1430.

Li, G. and Rabitz, R., 1990, A general analysis of approximate lumping in chemical kinetics. *Chem. Engng Sci.* **45**, 977–1002.

Li, G. and Rabitz, R., 1991a, New approaches to determination of constrained lumping schemes for a reaction system in the whole composition space. *Chem. Engng Sci.* **46**, 95–111.

Li, G. and Rabitz, R., 1991b, Determination of constrained lumping schemes for nonisothermal first-order reaction systems. *Chem. Engng Sci.* **46**, 583–596.

Li, G. and Rabitz, R., 1991c, A general lumping analysis of a reaction system coupled with diffusion. *Chem. Engng Sci.* **46**, 2041–2053.

Li, G., Tomlin, A. S., Rabitz, H. and Tóth, J., 1993, Determination of approximate schemes by a singular perturbation. *J. chem. Phys.* **99**, 3562–3574.

Li, G., Tomlin, A. S., Rabitz, H. and Tóth, J., 1994, A general analysis of approximate nonlinear lumping in chemical kinetics, I. Unconstrained lumping. Submitted to *J. chem. Phys.*

Luckyanov, N. K., Svirezhev, Yu. M. and Voronkova, O. V., 1983, Aggregation of variables in simulation models of water ecosystems. *Ecological Modelling* **18**, 235–240.

Rabinowitz, M., 1990, Chemical kinetics research program at Lewis. HSRP Emission Workshop, NASA Lewis Research Center.

Smooke, M. D. and Giovangigli, V., 1990, Reduced chemistry and simplified transport modeling of premixed and nonpremixed combustion. HSRP Emission Workshop, NASA Lewis Research Center.

Stephani, H., 1989, in *Differential Equations: Their Solution Using Symmetries* (Edited by M. Maccullum). Cambridge University Press, Cambridge.

Tomlin, A. S., Li, G., Rabitz, H. and Tóth, J., 1994, A general analysis of approximate nonlinear lumping in chemical kinetics, II. Constrained lumping. Manuscript in preparation.

Wei, J. and Kuo, J. C. W., 1969, A lumping analysis in monomolecular reaction systems. *Ind. Engng Chem. Fundam.* **8**, 114–133.

Wei, J. and Prater, C. D., 1962, The structure and analysis of complex reaction systems. *Adv. Catal.* **13**, 203–392.

Wulfman, C. E. and Rabitz, H., 1989, Global sensitivity analysis of nonlinear chemical kinetic equations using Lie groups: II. Some chemical and mathematical properties of the transformation groups. *J. Math. Chem.* **3**, 261–297.