Affine Lumping Formalism for Comparison of Model Reduction Techniques

Geoffrey M. Oxberry*, William H. Green*, Paul I. Barton*
*Massachusetts Institute of Technology/Department of Chemical Engineering, Cambridge, MA, USA

***Abstract*— Numerous methods exist for generating smaller, reduced chemical models from large, detailed chemical models. These methods arise from a variety of different theoretical backgrounds, yet few comparisons have been made between different model reduction methods. In order to assess the relative quality of these different model reduction techniques, these methods must be compared with each other using a common framework. As one element of such a framework, we propose a formalism called affine lumping. This formalism defines two affine mappings. The first affine mapping is used to lump the state variables from a detailed model to a reduced model with reduced state variables. The second affine mapping is used to unlump the reduced state variables and lift the reduced model back into the original state space. Conditions are stated under which the application of these two affine mappings in succession yields a solution of the original model. Finally, the techniques of species lumping by Li *et al*., computational singular perturbation and reaction invariants are all cast as special cases of affine lumping, to illustrate the potential usefulness of the affine lumping formulation. Given that three different model reduction techniques can be recast using the affine lumping formalism, it is possible that other model reduction techniques may be also be special cases of affine lumping. The affine lumping formalism could then be used as a common standard against which different model reduction techniques can be compared in order to assess their relative quality.**

I. Introduction

Many practical problems involve combustion under inhomogeneous, transient conditions, and therefore require the use of numerical methods that solve large systems of coupled, nonlinear partial differential equations of the kind typically found in reacting flow solvers. These practical problems typically require large, detailed chemical models in order to simulate faithfully the physics involved. However, reacting flow simulations of large chemical models have prohibitively large computational costs. Consequently, smaller reduced models are used in place of large chemical models in order to obtain approximate numerical solutions to the large chemical models at decreased computational cost.

Many methods are available for generating reduced models from detailed chemical models (see [2], [5], [3] and [4] for examples). However, these different methods originate from different theoretical backgrounds, including methods based on singular perturbation, methods based on a graph-theoretic interpretation of chemistry, and others. Given this variation in the theoretical development of model reduction methods, it is difficult to compare two given model reduction techniques. Previous work has made progress in this area by showing that the intrinsic low-dimensional manifold (ILDM) technique is a special case of computational singular perturbation (CSP) [2], facilitating the comparison of the two techniques in subsequent work. To make further progress in comparing model reduction techniques, we propose a formalism called "affine lumping."

In this work, we provide a definition of affine lumping. Affine lumping defines two affine mappings. The first affine mapping transforms a detailed model to a reduced-dimension representation called the reduced model; this process is called lumping. The second affine mapping lifts the reduced state variables into the space of the original state variables, recovering a representation of the reduced model in the original state variables that approximates the detailed model; this process is called unlumping. It can be shown that applying these two affine mappings in sequence is equivalent to projecting the original model onto an affine subspace. This projection is the reduced model, lifted into the space of the original state variables. It can be shown that under certain conditions, lifting the solution of the reduced model into the space of the original state variables yields a solution of the detailed model. In other words, under certain conditions, the solution of the detailed model lies in an affine subspace so that the solution of the reduced model can be used to construct the exact solution of the detailed model, and the reduced model can be described with fewer differential equations and state variables than the detailed model.

After presenting a definition of affine lumping, we present examples of existing model reduction techniques that are special cases of affine lumping. We show that the species lumping technique of Li *et al*. [3], CSP and the technique of reaction invariants reviewed by Waller and Mäkilä [6] are all special cases of affine lumping by translating the nomenclature and mathematics used in each of these techniques to the nomenclature and mathematics of affine lumping. These results suggest that other methods could also be expressed using the affine lumping formalism, and indicate that apparently different techniques could have similar properties despite different theoretical backgrounds. These similarities could then be used in future work to assess the relative merits of each model reduction technique that fits the formalism.

II. Definition and properties of affine lumping

In order to define affine lumping, we first motivate the definition by exploring the concept of lumping in

the specific case of chemical reaction in an adiabatic-isobaric batch reactor. Since the chemical source term in this model is used within reacting flow solvers that use operator splitting, it is the primary model under consideration; we postpone discussion of any other models to future work, since they will be extensions of this case. After briefly discussing the model and informally discussing ideas behind affine lumping, a definition of affine lumping is provided. Finally, we state without proof the main result of affine lumping: the solution of the reduced model obtained from affine lumping can be used to reconstruct the solution of the detailed model with no approximation error, provided that the affine lumping and reduced model have certain properties.

Consider the following model for a well-mixed, isolated chemically reacting system (that is, an adiabatic-isobaric batch reactor):

$$\dot{\mathbf{y}}(t) = \mathbf{\Gamma}(\mathbf{y}(t)), \tag{1}$$

where $\mathbf{y}(t) \in \mathbb{R}^{N_S}$ represents the original state variables (such as species compositions, temperatures, and pressures), and the source function $\mathbf{\Gamma} : \mathbb{R}^{N_S} \rightarrow \mathbb{R}^{N_S}$ describes changes in the state variables due to chemistry.

Consider the "affine lumping":

$$\boldsymbol{\phi}(t) = \mathbf{A}(\mathbf{y}(t) - \mathbf{y}_0), \tag{2}$$

where $N_S \geq N_L$, $\boldsymbol{\phi}(t) \in \mathbb{R}^{N_L}$ is a set of reduced state variables, $\mathbf{A} \in \mathbb{R}^{N_L \times N_S}$ is the assembly (or lumping) matrix, and $\mathbf{y}_0 \in \mathbb{R}^{N_S}$ is a point in the (original) state variable space corresponding to the origin of the new reduced-dimension space corresponding to the reduced state variables. We treat the matrix $\mathbf{A}$ as a function of the original state variables, and therefore, as a function of the solution $\mathbf{y}(t)$, but assume that the functional form of $\mathbf{A}(\mathbf{z})$ is piecewise constant over $\mathbb{R}^{N_S}$ (here, $\mathbf{z}$ is a dummy variable). Therefore, instead of using the notation $\mathbf{A}(\mathbf{z})$, we will use the notation $\mathbf{A}$ under the assumption that we have restricted the use of $\mathbf{A}$ to a subset of $\mathbb{R}^{N_S}$ over which its elements have constant values; this subset must be specified when defining a specific instance of lumping.

Having given an informal description of affine lumping, additional properties are needed to make a description of the lumping process more concrete. Suppose that we can find a generalized inverse $\mathbf{D}$ corresponding to $\mathbf{A}$ (for the properties of generalized inverses, see [1]). The matrix $\mathbf{D} \in \mathbb{R}^{N_S \times N_L}$ will be called the disassembly (or unlumping) matrix, since its purpose will be to recover approximately the original state variables from their reduced state variable counterparts. Like $\mathbf{A}$, we treat the matrix $\mathbf{D}$ as a piecewise constant function over state space. Therefore, instead of using the notation $\mathbf{D}(\mathbf{z})$, we will use the notation $\mathbf{D}$ under the assumption that we have restricted the use of $\mathbf{D}$ to a subset of $\mathbb{R}^{N_S}$ over which its elements have constant values; this subset must be specified when defining a specific instance of lumping, and it must be equal to the subset used in defining $\mathbf{A}$.

For any $\mathbf{y}^* \in \mathbb{R}^{N_S}$, the values of the reduced state variables resulting from lumping once the values of the original state variables should be equivalent to the values of the reduced state variables resulting from lumping the values of the state variables, unlumping the resulting values of the reduced state variables, and then lumping again the values of the state variables obtained from unlumping. Put another way, after lumping once, repeated unlumping and lumping should yield the same values for the lumped variables. Mathematically, this implies that for any $\mathbf{y}^* \in \mathbb{R}^{N_S}$, $\mathbf{ADAy}^* = \mathbf{Ay}^*$, and thus

$$\mathbf{ADA} = \mathbf{A}. \tag{3}$$

Equation (3) is Penrose's First Equation, indicating that $\mathbf{D}$ is a $\{1\}$-inverse of $\mathbf{A}$.

Also, assume that $\mathbf{A}$ has full rank; if it does not, we could describe the subspace $\mathcal{R}(\mathbf{A})$ with a smaller basis and obtain an equivalent lumped representation using a smaller assembly matrix and fewer variables. Since the goal of this work is to reduce as much as possible the number of state variables used to represent a chemically reacting system, a full rank assumption is not restrictive.

Since $\mathbf{A}$ has full rank, by Lemma 1.2 of [1], it follows that

$$\mathbf{AD} = \mathbf{I}_{N_L}, \tag{4}$$

where $\mathbf{I}_{N_L}$ is an $N_L \times N_L$ identity matrix. From this equation,

$$\mathbf{DAD} = \mathbf{D} \tag{5}$$

also holds. Equation (5) is Penrose's Second Equation, indicating that $\mathbf{D}$ and $\mathbf{A}$ are $\{1,2\}$-inverses of each other.

These arguments motivate the following definition of an affine lumping:

Definition 2.1 (Affine Lumping): An **affine lumping** is a 3-tuple $(\mathbf{A}, \mathbf{D}, \mathbf{y}_0)$, where

1) $\mathbf{A}$ is a full rank $N_L \times N_S$ matrix called the *assembly matrix*, with $N_L \leq N_S$,
2) $\mathbf{D}$ is a $\{1,2\}$-inverse of $\mathbf{A}$ called the *disassembly matrix*,
3) $\mathbf{y}_0 \in \mathbb{R}^{N_S}$ is the *origin* of the lumping.

Having defined affine lumping, we next explore the properties of the definition. Suppose we have an affine lumping $(\mathbf{A}, \mathbf{D}, \mathbf{y}_0)$ associated with (1), and suppose that $\mathbf{y}(t)$ is a solution to (1). From (2), the chain rule, and (1), it follows that

$$\dot{\boldsymbol{\phi}}(t) = \mathbf{A}\dot{\mathbf{y}}(t) = \mathbf{A}\mathbf{\Gamma}(\mathbf{y}(t)). \tag{6}$$

Therefore, given a solution $\mathbf{y}^*(t)$ of (1) with $\mathbf{y}^*(0) = \mathbf{y}_0^*$, it is possible to prescribe initial conditions $\boldsymbol{\phi}_0$ and solve (6). Actually carrying out this process, however, would defeat the purpose of lumping, since the idea of species lumping is to cast the original state space into a reduced-dimensional state space for the purposes of reducing the computational effort needed to solve the ODE system in the reduced space; solving the original

ODE would yield all of the desired results. Consequently, it is necessary to determine a closure relation that allows us to express (6) as an autonomous ODE system.

The closure relation for constructing the solution of the original model from a suitably defined reduced model can be found under certain conditions using our main result, stated here without proof:

Theorem 2.1: Given an affine lumping $(\mathbf{A}, \mathbf{D}, \mathbf{y}_0)$, let

1) $Z = \{\mathbf{z} \in \mathbb{R}^{N_S} : \mathbf{z} = \mathbf{D}\mathbf{w} + \mathbf{y}_0, \mathbf{w} \in \mathbb{R}^{N_L}\}$,
2) $\boldsymbol{\phi}(t) \in \mathbb{R}^{N_L}$ be a solution of the reduced model

$$\dot{\boldsymbol{\phi}}(t) = \mathbf{A}\boldsymbol{\Gamma}(\mathbf{D}\boldsymbol{\phi}(t) + \mathbf{y}_0) \tag{7}$$

with $\boldsymbol{\phi}(0) = \mathbf{0}$,

3) $\mathbf{x}(t) \equiv \mathbf{D}\boldsymbol{\phi}(t) + \mathbf{y}_0$.

Suppose that $\boldsymbol{\Gamma}(\mathbf{z}) \in \mathcal{R}(\mathbf{D})$, $\forall \mathbf{z}$ in Z. It follows that $\mathbf{x}(t)$ is a solution of (1) with $\mathbf{x}(0) = \mathbf{y}_0$ and that $\boldsymbol{\phi}(t) = \mathbf{A}(\mathbf{x}(t) - \mathbf{y}_0)$.

If the conditions of Theorem 2.1 hold, then lifting the reduced model into the space of the original state variables yields the equation

$$\dot{\mathbf{x}}(t) = \mathbf{D}\mathbf{A}\boldsymbol{\Gamma}(\mathbf{x}(t)). \tag{8}$$

Aside from the change of variables from $\mathbf{y}$ to $\mathbf{x}$ (chosen to distinguish between the solution of the detailed model and the lifted solution of the reduced model), note that the right hand side of the lifted reduced model (8) is the same as the right-hand side of the original model (1) premultiplied by the matrix $\mathbf{DA}$. From Corollary 2.7 of [1], the matrix $\mathbf{DA}$ is a projector onto $\mathcal{R}(\mathbf{D})$ along $\mathcal{N}(\mathbf{A})$, yielding the interpretation that the solution of the reduced model, lifted into the space of the original state variables, is the solution to the original model projected onto the affine subspace $\mathcal{R}(\mathbf{D}) + \mathbf{y}_0$ along $\mathcal{N}(\mathbf{A})$.

It is worth noting that a simple choice for $\mathbf{D}$ that satisfies Theorem 2.1 is a maximal set of linearly independent columns of the stoichiometry matrix $\mathbf{N}$. However, in order to reduce models more aggressively, the assumptions of Theorem 2.1 will have to be relaxed so that the right-hand side of the reduced model, lifted into the space of the original state variables differs from the right-hand side of the reduced model to within a known error bound. A method for error-controlled affine lumping will be the focus of future work.

III. Special cases of affine lumping

Having defined the affine lumping formalism and established conditions under which a reduced model can be used to determine exactly the solution of the detailed model, we now give some examples of model reduction techniques that are special cases of affine lumping. We discuss three techniques: the technique of species lumping by Li *et al*. [3], the technique of CSP by Lam [2] and the technique of reaction invariants reviewed by Waller and Mäkilä [6].

Li *et al*. [3] define their species lumping scheme using the matrices $\mathbf{M}$ and $\bar{\mathbf{M}}$ that take the roles of the matrices $\mathbf{A}$ and $\mathbf{D}$, respectively, in an affine lumping. Cast in terms of the notation for affine lumping by setting $\mathbf{A} = \mathbf{M}$, $\mathbf{D} = \bar{\mathbf{M}}$, and $\boldsymbol{\phi} = \hat{\mathbf{y}}$, the species lumping scheme in [3] is defined as

$$\boldsymbol{\phi}(t) = \mathbf{A}\mathbf{y}(t), \tag{9}$$

and the matrices $\mathbf{A}$ and $\mathbf{D}$ are related by $\mathbf{AD} = \mathbf{I}_{N_L}$. Li *et al*. do not place any restrictions on $\mathbf{A}$ and $\mathbf{D}$ other than to state that $\bar{\mathbf{M}}$ is one of the generalized inverses of $\mathbf{M}$. If we restrict $\mathbf{M}$ and $\bar{\mathbf{M}}$ to be $\{1,2\}$-inverses of each other, we have an affine lumping $(\mathbf{A}, \mathbf{D}, \mathbf{y}_0)$ where $\mathbf{A} = \mathbf{M}$, $\mathbf{D} = \bar{\mathbf{M}}$ and $\mathbf{y}_0 = \mathbf{0}$. We postpone a more detailed comparison of affine lumping with the species lumping of Li *et al*. to future work.

Computational singular perturbation defines affine lumping-like objects through the CSP basis vectors. Let $\mathbf{A}^{CSP}$ be the CSP basis matrix whose columns are the CSP basis vectors, and let $\mathbf{B}^{CSP}$ be the CSP reciprocal basis matrix whose rows are the CSP reciprocal vectors, such that

$$\mathbf{B}^{CSP} = (\mathbf{A}^{CSP})^{-1}. \tag{10}$$

We assume here that the matrix $\mathbf{A}^{CSP}$ is a constant, rather than treating it as time-varying, as in the general case of CSP.

To discuss CSP in the context of affine lumping, some additional notation is necessary. Using the notation of Ben Israel and Greville [1], denote by

$$Q_{k,n} = \{(i_1, i_2, \ldots, i_k) : 1 \le i_1 < i_2 < \ldots < i_k \le n\} \tag{11}$$

the set of increasing sequences of k elements from the set $\{1, \ldots, n\}$ for given integers $0 < k \le n$. Also, for a given matrix $\mathbf{M} \in \mathbb{R}^{m \times n}$, and index sets $I \in Q_{p,m}$ and $J \in Q_{q,n}$, let the $p \times n$ submatrix $\mathbf{M}_{I*}$ be the matrix whose elements are m_{ij} for $i \in I$ and $j \in \{1, \ldots, n\}$ and let the $m \times q$ submatrix $\mathbf{M}_{*J}$ be the matrix whose elements are m_{ij} for $i \in \{1, \ldots, n\}$ and $j \in J$.

Returning to CSP, let N_L be the number of active, slow CSP modes, and let $S \in Q_{N_L,N_S}$ be the set of indices of CSP basis vectors corresponding to the slow CSP modes. The matrices $\mathbf{B}^{CSP}_{S*}$ and $\mathbf{A}^{CSP}_{*S}$ take the roles of $\mathbf{A}$ and $\mathbf{D}$ in affine lumping and have the property that

$$\mathbf{B}^{CSP}_{S*}\mathbf{A}^{CSP}_{*S} = \mathbf{I}_{N_L}, \tag{12}$$

implying that $\mathbf{A}^{CSP}_{*S}$ is a $\{1\}$-inverse of $\mathbf{B}^{CSP}_{S*}$ by rearranging (12) to yield Penrose's First Equation. Since $\mathbf{A}^{CSP}_{*S}$ and $\mathbf{B}^{CSP}_{S*}$ are both full rank matrices, they are $\{1,2\}$-inverses of each other, by Corollary 2.1 of [1]. If the approximate equations of state in [2] are treated as equalities, then CSP replaces the original model (1) with the approximate, reduced model

$$\dot{\mathbf{y}}(t) = \mathbf{A}^{CSP}_{*S}\mathbf{B}^{CSP}_{S*}\boldsymbol{\Gamma}(\mathbf{y}(t)), \tag{13}$$

which resembles the reduced model lifted into the space of the original state variables, shown in (8). This result suggests that CSP fits the affine lumping formalism

without defining a lumping operation explicitly. In order to complete the definition of affine lumping from the CSP basis matrices, an appropriate origin of the lumping must be defined; in this case, set $\mathbf{y}_0$ equal to the current point in state space at which these particular CSP basis matrices are used to approximate the original model. Again, we postpone a more detailed comparison for future work.

The reaction invariants methods use a basis transformation to change the state variables from the state variables of the original model in (1) to state variables that can be classified as one of two types: reaction variants and reaction invariants. Reaction invariants are state variables that are constant with time, and reaction variants are state variables that change over time. Using the notation of Waller and Mäkilä [6], suppose that $\mathbf{v}(t) \in \mathbb{R}^{N_L}$ are new state variables representing the reaction invariants, and $\mathbf{w}(t) \in \mathbb{R}^{N_S-N_L}$ are new state variables representing the reaction invariants. Suppose also that there exist matrices $\mathbf{P}^{RI} \in \mathbb{R}^{N_S\times(N_S-N_L)}$, $\mathbf{T}^{RI} \in \mathbb{R}^{N_S\times N_L}$, $\mathbf{D}^{RI} \in \mathbb{R}^{(N_S-N_L)\times N_S}$ and $\mathbf{L}^{RI} \in \mathbb{R}^{N_L\times N_S}$ such that the matrix $[\mathbf{P}^{RI}\ \mathbf{T}^{RI}]$ is nonsingular, and the equations

$$\mathbf{y}(t) = \begin{bmatrix} \mathbf{P}^{RI} & \mathbf{T}^{RI} \end{bmatrix} \begin{bmatrix} \mathbf{w}(t) \\ \mathbf{v}(t) \end{bmatrix}, \tag{14a}$$

$$\begin{bmatrix} \mathbf{D}^{RI} \\ \mathbf{L}^{RI} \end{bmatrix} = \begin{bmatrix} \mathbf{P}^{RI} & \mathbf{T}^{RI} \end{bmatrix}^{-1}, \tag{14b}$$

both hold. From (14b), it follows that $\mathbf{L}^{RI}\mathbf{T}^{RI} = \mathbf{I}_{N_L}$, and by Penrose's First Equation, $\mathbf{T}^{RI}$ is a $\{1\}$-inverse of $\mathbf{L}^{RI}$. Since the columns of the matrix $[\mathbf{P}^{RI}\ \mathbf{T}^{RI}]$ are linearly independent, $\mathbf{L}^{RI}$ and $\mathbf{T}^{RI}$ are both full rank matrices. Consequently, by Corollary 2.1 of [1], $\mathbf{L}^{RI}$ and $\mathbf{T}^{RI}$ are $\{1,2\}$-inverses, suggesting that $\mathbf{L}^{RI}$ and $\mathbf{T}^{RI}$ take the roles of the assembly matrix $\mathbf{A}$ and disassembly matrix $\mathbf{D}$ in an affine lumping scheme.

In addition, $\mathbf{w}(t)$ does not vary with time, so for given initial conditions of the detailed model (1), $\mathbf{w}(t)$ is fixed. Setting $\mathbf{A} = \mathbf{L}^{RI}$ and $\mathbf{y}_0 = \mathbf{P}^{RI}\mathbf{w}(0)$, we can use (2) to define an affine lumping from the matrices of the reaction invariants technique, and $(\mathbf{L}^{RI}, \mathbf{T}^{RI}, \mathbf{P}^{RI}\mathbf{w}(0))$ is an affine lumping. A more detailed comparison of affine lumping and reaction invariants will follow in future work.

IV. Conclusions and future work

A formalism called affine lumping has been defined to facilitate the comparison of different model reduction techniques that attempt to reduce the number of state variables used to describe a chemically reacting system. In addition to defining affine lumping, conditions have been stated under which an affine lumping can be used to reconstruct exactly the solution of a detailed model from a reduced model. The model reduction techniques of species lumping by Li *et al*. [3], computational singular perturbation by Lam [2] and reaction invariants [6] have been shown to be affine lumpings, under certain assumptions.

To further develop this work, more detailed comparisons of the theory of affine lumping with the theory of existing model reduction techniques should be conducted to get a better idea of the breadth of techniques that can be expressed using the ideas of projection and generalized inverses as a theoretical basis. In parallel, a technique to determine an error-controlled affine lumping could be developed. The idea behind error-controlled affine lumping would be to replace the condition under which the solution of a detailed model can be reconstructed from a reduced model exactly with conditions under which the solution of a detailed model can be reconstructed from a reduced model approximately with known error bounds. These error bounds could then be used in conjunction with additional information to determine the affine lumping that yields the reduced model with the fewest reduced state variables, subject to error bounds. Such a technique could then be compared with existing techniques for model reduction with (or without) error control in order to better assess the strengths and weaknesses of various model reduction methods.

Acknowledgments

GMO would like to thank Joseph K. Scott, Dr. Ajay Selot and Mehmet Yunt for helpful discussions. He would also like to gratefully acknowledge support under a DOE CSGF fellowship provided under grant DE-FG02-97ER25308. Finally, he would like to thank Drs. Bert Debusschere and Habib N. Najm for their hospitality and for discussions about computational singular perturbation during a practicum at Sandia National Laboratory in Livermore, CA over Summer 2008. Financial support from the U.S. Department of Energy's Office of Basic Sciences, Division of Chemical Sciences, Geosciences and Biosciences through Grant DE-FG02-98ER14914 is gratefully acknowledged.

References

[1] A. Ben-Israel and T. N. E. Greville. *Generalized Inverses: Theory and applications*. Springer-Verlag, New York, 2nd edition, 2003.

[2] S. H. Lam. Using CSP to understand complex chemical kinetics. *Combustion Science and Technology*, 89(5):375–404, 1993.

[3] G. Li and H. Rabitz. A general analysis of exact lumping in chemical kinetics. *Chemical Engineering Science*, 44(6):1413–1430, 1989.

[4] T. Lu and C. K. Law. Systematic approach to obtain analytic solutions of quasi steady state species in reduced mechanisms. *Journal of Physical Chemistry A*, 110:13202–13208, 2006.

[5] S. Singh, J. M. Powers, and S. Paolucci. On slow manifolds of chemical reactive systems. *Journal of Chemical Physics*, 117(4):1482–1496, 2002.

[6] K. V. Waller and P. M. Mäkilä. Chemical reaction invariants and variants and their use in reactor modeling, simulation, and control. *Industrial and Engineering Chemistry Process Design and Development*, 20:1–13, 1981.