Model Reductions with Special CSP Data

S. H. Lam*
Princeton University, Princeton, N.J. 08544, U.S.A.

June 11, 2013

Abstract

The model reduction methodology of *computational singular perturbation* (CSP) is enhanced for chemical kinetics problems. The leading order *converged* CSP-refined fast basis vectors are obtained analytically by exploiting the special format of the given reaction-specific chemical kinetics data. The leading order *net* reaction rates of the fast reactions (after they are exhausted) are also obtained analytically. Taking advantage of these leading order analytical results, a *new* model reduction strategy is advocated. The implementation of this new strategy is very straightforward, and extraction of general insights on the reaction system is also easier.

Keywords: Chemical kinetics, model reductions, stiff equations, CSP.

1 Statement of the problem

Consider the following general initial-value problem:

$$\frac{d\mathbf{y}}{dt} = \mathbf{g}(\mathbf{y};\epsilon), \quad \mathbf{y}(t=0) = \widehat{\mathbf{y}}, \tag{1}$$

where both $\mathbf{y}$ and $\mathbf{g}$ are N-dimensional column vectors, all components of $\mathbf{g}(\mathbf{y};\epsilon)$ are *given* algebraic functions of $\mathbf{y}$, and ϵ is a small dimensionless

*Professor emeritus, Department of Mechanical and Aerospace Engineering. Email: lam@princeton.edu

parameter. The arbitrary $\mathbf{y}$ initial value is denoted by $\widehat{\mathbf{y}}$. We are interested in model reductions when Eq.(1) is known to be *stiff* in the small ϵ limit.

For chemical kinetics problems, the right hand side of Eq.(1) is usually *given* in the following form:

$$\mathbf{g}(\mathbf{y};\epsilon) = \sum_{r=1}^{R} \boldsymbol{\alpha}_r \Omega^r(\mathbf{y};k_r), \tag{2}$$

where R is the number of reactions in the reaction system, and each term in the summation represents a *chemically/physically meaningful* chemical reaction. The $\boldsymbol{\alpha}_r$'s are N-dimensional column vectors; their elements are proportional to the *stoichiometric coefficients* of the r-th reaction, and their physical dimensions are identical to that of the corresponding elements of $\mathbf{y}$. For chemical kinetics, the $\boldsymbol{\alpha}_r$'s do not depend on $\mathbf{y}$. However, the subsequent developments remain valid so long as $\partial\boldsymbol{\alpha}_r/\partial\mathbf{y}$ is non-singular in the small ϵ limit. Usually, R is much bigger than N—so many $\boldsymbol{\alpha}_r$'s are linearly dependent. The $\Omega^r(\mathbf{y};k_r)$'s and the k_r's are the *net reaction rate* and the kinetic rate parameter of the r-th reaction, respectively. The physical dimensions of all the $\Omega^r(\mathbf{y};k_r)$'s are reciprocal time. Usually, $\Omega^r(\mathbf{y};k_r)$ is proportional to k_r, depends on $\mathbf{y}$ *nonlinearly*, and contains some thermodynamics parameters (e.g. equilibrium constants). For the developments in this paper, the $\boldsymbol{\alpha}_r$'s and the $\Omega^r(\mathbf{y};\epsilon)$'s do not need to honor any other restrictions such as *law of mass action*, conservation of atomic elements, etc..

Eq.(1) is stiff when $M > 0$ reactions are much faster than all others. The stiffness of Eq.(1) is symbolically represented by the dimensionless parameter ϵ in $\mathbf{g}(\mathbf{y};\epsilon)$. The classical methodology to deal with the singular nature of the problem in the small ϵ limit is to apply the *quasi-steady approximation* (QSA) and/or the *partial equilibrium approximation* (PEA) [1]. However, such pencil and paper exercises cannot handle very large reaction systems. The goal of *computational singular perturbation* (CSP) [2, 3, 4, 5] is to translate the classical singular perturbation procedures into programmable algorithms so that model reductions can be computationally done for very large systems. The present paper executes the main CSP algorithms analytically in terms of the *chemically/physically meaningful* $\boldsymbol{\alpha}_r$'s and the $\Omega^r(\mathbf{y};k_r)$'s, and obtains insightful model reduction formulas which are programmable. A new strategy of computational model reductions will be advocated.

2 The CSP reduced model

CSP utilizes the concept of basis vectors to do model reductions. When there are M linearly independent fast reactions in the reaction system, CSP formally introduces a dimensionless $N \times N$ *fast subspace projection matrix* $\mathcal{Q}^{fast}(M)$ as follows [4, 5]:

$$\mathcal{Q}^{fast}(M) \equiv \sum_{m=1}^{M} \mathbf{a}_m \mathbf{b}^m, \tag{3}$$

where the $\mathbf{a}_m$'s and the $\mathbf{b}^m$'s are *linearly independent* (column and row) *basis vectors* which can completely span the M-dimensional fast subspace. They are—by definition—orthonormal to each other:

$$\mathbf{b}^m \odot \mathbf{a}_{m'} = \delta^m_{m'}, \quad m, m' = 1, \ldots, M, \tag{4}$$

where $\odot$ is the inner product operator and $\delta^m_{m'}$ is the Kronecker Delta.

CSP uses this $\mathcal{Q}^{fast}(M)$ to partition the right-hand side of Eq.(1):

$$\begin{aligned} \mathbf{g}(\mathbf{y};\epsilon) &= \mathbf{g}^{fast}(\mathbf{y};\epsilon) + \mathbf{g}^{slow}(\mathbf{y};\epsilon), & \text{(5a)} \\ \mathbf{g}^{fast}(\mathbf{y};\epsilon) &\equiv \mathcal{Q}^{fast}(M) \odot \mathbf{g}(\mathbf{y};\epsilon), & \text{(5b)} \\ \mathbf{g}^{slow}(\mathbf{y};\epsilon) &= \mathbf{g}(\mathbf{y};\epsilon) - \mathbf{g}^{fast}(\mathbf{y};\epsilon). & \text{(5c)} \end{aligned}$$

It is duly noted that Eqs.(5a,b,c) are exact for *any* chosen $\mathcal{Q}^{fast}(M)$. The model reduction *quality* of any chosen $\mathcal{Q}^{fast}(M)$ is eventually to be adjudicated by how well it *decouples* the fast and slow reaction groups. During the initial fast transients, $\mathbf{g}^{fast}(\mathbf{y};\epsilon)$ is expected to be dominant in Eq.(5a) when initial condition $\widehat{\mathbf{y}}$ is arbitrary. After the fast reactions are "exhausted," the $\mathbf{g}^{fast}(\mathbf{y};\epsilon)$ of a good quality $\mathcal{Q}^{fast}(M)$'s is expected to rapidly decay to $O(\epsilon)$ in comparison to $\mathbf{g}^{slow}(\mathbf{y};\epsilon)$ so that its neglect can be justified. The CSP-derived reduced model is then:

$$\frac{d\mathbf{y}}{dt} = \mathbf{g}^{slow}(\mathbf{y};\epsilon) + O(\epsilon), \quad \mathbf{g}^{fast}(\mathbf{y}(t = O(\epsilon);\epsilon) = O(\epsilon). \tag{6}$$

The notation $O(\epsilon)$ is used to denote any entity which is known to be negligible in the $\epsilon \to 0$ limit. A good quality $\mathcal{Q}^{fast}(M)$ should also significantly reduce the stiffness of the original problem. In addition, the solution $\mathbf{y}(t)$ of Eq.(6) is expected to *automatically* stay inside the *slow subspace* defined by:

$$\mathbf{g}^{fast}(\mathbf{y}(t);\epsilon) = O(\epsilon), \quad t > O(\epsilon). \tag{7}$$

CSP provides iterative *refinement procedures* to find good quality $\mathcal{Q}^{fast}(M)$ for completely general $\mathbf{g}(\mathbf{y};\epsilon)$. The present paper focuses on chemical kinetics problems, and exploits the fact that each additive terms in $\mathbf{g}(\mathbf{y};\epsilon)$ as displayed in Eq.(2) are *physically/chemically meaningful* to knowledgeable investigators.

In early versions of CSP [2, 3, 4], basis vectors of both the fast and slow subspaces were involved. The above summary follows the 1994 version [5]—in which only the fast subspace basis vectors are involved. See §6.5 of [5], and see [6] for some interesting applications of CSP on biochemical systems.

A new CSP-based model reduction strategy shall be advocated in §7 later.

3 Reaction-specific CSP data

Eq.(2) clearly associates the column vector $\boldsymbol{\alpha}_r$ with the r-th reaction. We now associate the following row vector $\boldsymbol{\beta}^r$ with the r-th reaction:

$$\boldsymbol{\beta}^r \equiv \tau^r \frac{\partial \Omega^r(\mathbf{y};k_r)}{\partial \mathbf{y}} = \boldsymbol{\beta}^r(\mathbf{y}), \quad r = 1,\ldots,R, \tag{8a}$$

where τ^r, a most interesting **reaction-specific** finite scalar with *time* as its physical dimension, is *defined* by:

$$\tau^r \equiv \frac{1}{\frac{\partial \Omega^r}{\partial \mathbf{y}} \odot \boldsymbol{\alpha}_r} = \tau^r(\mathbf{y};k_r), \quad r = 1,\ldots,R. \tag{8b}$$

For chemical kinetics problems, the $\tau^r(\mathbf{y};k_r)$'s are usually negative. We next *define* $\Gamma^r_{r'}(\mathbf{y})$, a $R \times R$ matrix with $O(1)$ *dimensionless* elements, by:

$$\Gamma^r_{r'}(\mathbf{y}) \equiv \boldsymbol{\beta}^r \odot \boldsymbol{\alpha}_{r'}, \quad r, r' = 1,\ldots R. \tag{9}$$

Note that in general $\Gamma^r_{r'}(\mathbf{y}) \neq 0$ when $r \neq r'$, while $\Gamma^r_r(\mathbf{y}) = 1$ is always honored by Eq.(8b), the definition of $\tau^r(\mathbf{y};k_r)$.

Note that whenever $\Omega^r(\mathbf{y};k_r) \propto k_r$ is true (and $\boldsymbol{\alpha}_r$ is independent of k_r), then $\boldsymbol{\beta}^r(\mathbf{y})$ is independent of k_r, and $\tau^r(\mathbf{y};k_r) \propto 1/k_r$.

4 Exploitation of the CSP data

The $\tau^r(\mathbf{y};k_r)$ data is most interesting. If the r-th reaction is the only reaction in the reaction system, then it is easy to show that $d\Omega^r/dt = \Omega^r/\tau^r$,

and that this solo reaction will be exhausted when $t >> |\tau^r(\mathbf{y}; k_r)|$ provided $\tau^r(\mathbf{y}; k_r)$ is negative. Thus $|\tau^r(\mathbf{y}; k_r)|$, which has time as its physical dimension, is an *intrinsic time scale* of the r-th reaction. The smaller $|\tau^r(\mathbf{y}; k_r)|$ is, the faster the intrinsic speed of the r-th reaction. Thus, at any time t, we can order the R reactions in ascending order of their $|\tau^r(\mathbf{y}(t); k_r)|$'s so that $r = 1$ is intrinsically the fastest reaction of them all at that time. We assume this is done on the fly. Ordering reactions this way is much easier (but admittedly less theoretically definitive) than ordering by eigenvalues—which physical dimensions are reciprocal time—of the Jacobian matrix $\partial\mathbf{g}/\partial\mathbf{y}$ [7, 8].

Mathematically, the stiffness of Eq.(1) is associated with a large gap in the values of the $\tau^r(\mathbf{y}; k_r)$'s. Such gap separates the fast and slow reactions. When the ratio of two successive values of $\tau^r(\mathbf{y}; k_r)$ is a small number, this ratio is a tentative estimate of the small parameter ϵ in $\mathbf{g}(\mathbf{y}; \epsilon)$.

When some of the given $\boldsymbol{\alpha}_r$'s are linearly dependent, tampering with the linear dependencies will affect the formulas for the $\Omega^r(\mathbf{y}; k_r)$'s, the $\beta^r(\mathbf{y})$'s, the $\tau^r(\mathbf{y}; k_r)$'s, and therefore also the ranked order of the reactions. Thus, the needed set of linearly independent fast reactions is most likely non-unique when some of the given $\boldsymbol{\alpha}_r$'s are linearly dependent. This non-uniqueness issue will be discussed later.

For the sake of simplicity, we assume that the chosen set of M fast $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_M$ are linearly independent, and that all the fast $\tau^m(\mathbf{y}; k_m)$'s are negative. In addition, we also assume that the linear dependency number for each $\boldsymbol{\alpha}_r$ in the reaction system is not large (see last paragraph of §6.2 later). The fast subspace is then M-dimensional (with $M < N$) and is stable. Thus the original $\mathbf{g}(\mathbf{y}; \epsilon)$ can be rewritten exactly as follows:

$$\mathbf{g}(\mathbf{y}; \epsilon) = \sum_{m"=1}^{M} \boldsymbol{\alpha}_{m"} \Omega^{m"}(\mathbf{y}; k_{m"}) + \sum_{n=M+1}^{R} \boldsymbol{\alpha}_n \Omega^n(\mathbf{y}; k_n). \tag{10}$$

It is important to *emphatically* point out here that this is *not* a good quality fast-slow partition for $\mathbf{g}(\mathbf{y}; \epsilon)$—the mere fact that $|\tau^M| << |\tau^{M+1}|$ is no guarantee that the first summation term on the right hand side of Eq.(10) would become negligible after the fast transients are over.

It is *intuitively* obvious that the fast $\boldsymbol{\alpha}_m(\mathbf{y})$'s are excellent choices for the $\mathbf{a}_m(\mathbf{y})$'s:

$$\mathbf{a}_m(\mathbf{y}) = \boldsymbol{\alpha}_m(\mathbf{y}), \quad m = 1, \ldots, M. \tag{11a}$$

This chosen set of $\boldsymbol{\alpha}_m(\mathbf{y})$'s must be linearly independent. We next choose

the $\mathbf{b}^m(\mathbf{y})$'s to be some linear combinations of the fast $\boldsymbol{\beta}^m(\mathbf{y})$'s:

$$\mathbf{b}^m(\mathbf{y}) = \sum_{m'=1}^{M} \Theta^m_{m'}(\mathbf{y})\boldsymbol{\beta}^{m'}(\mathbf{y}), \quad m = 1, ..., M, \tag{11b}$$

where $\Theta^m_{m"}(\mathbf{y})$, a $M \times M$ dimensionless matrix, is uniquely determined by imposing $\mathbf{b}^m \odot \mathbf{a}_{m"} = \delta^m_{m"}$:

$$\Theta^m_{m"}(\mathbf{y}) = \left(\Gamma^{m"}_m(\mathbf{y})\right)^{-1}, \quad m, m" = 1, \ldots, M. \tag{11c}$$

This chosen set of $\boldsymbol{\beta}^m$'s must also be linearly independent. Hence, all elements of the $\Theta^m_{m'}(\mathbf{y})$ matrix are $O(1)$ dimensionless numbers.

Using Eqs.(11a,b) in Eq.(3), we obtain the following analytical formula:

$$\mathcal{Q}^{fast}(M) = \sum_{m,m'=1}^{M} \boldsymbol{\alpha}_m(\mathbf{y})\Theta^m_{m'}(\mathbf{y})\boldsymbol{\beta}^{m'}(\mathbf{y}). \tag{12}$$

This recommended $\mathcal{Q}^{fast}(M)$ is programmable. When M is a large number, efficient numerical inversion algorithms are needed to update $\Theta^m_{m'}(\mathbf{y}(t))$ on the fly.

5 CSP credentials of this $\mathcal{Q}^{fast}(M)$

The CSP step #1 $\mathbf{b}^m$-refinement procedure was given by Eq.(17a) of [5]:

$$\mathbf{b}^m_o(\mathbf{y}) \equiv \sum_{m'=1}^{M} \tau^m_{m'} \left(\frac{d\mathbf{b}^{m'}}{dt} + \mathbf{b}^{m'} \odot \frac{\partial \mathbf{g}}{\partial \mathbf{y}} \right), \quad m = 1, \ldots, M. \tag{13}$$

where $\tau^m_{m'}(\mathbf{y})$ is a $M \times M$ matrix. Since the time dependence of $\mathbf{b}^{m'}(t)$ arises only through its dependence on $\mathbf{y}(t)$, we have:

$$\frac{d\mathbf{b}^{m'}}{dt} = \frac{\partial \mathbf{b}^{m'}}{\partial \mathbf{y}} \odot \frac{d\mathbf{y}}{dt} = \frac{\partial \mathbf{b}^{m'}}{\partial \mathbf{y}} \odot \mathbf{g}, \quad m' = 1, \ldots, M. \tag{14}$$

With the help of Eq.(14), we can rewrite Eq.(13) as follows:

$$\mathbf{b}^m_o(\mathbf{y}) \equiv \sum_{m'=1}^{M} \tau^m_{m'} \frac{\partial(\mathbf{b}^{m'} \odot \mathbf{g})}{\partial \mathbf{y}}, \quad m = 1, \ldots, M. \tag{15}$$

Using Eq.(11b) for $\mathbf{b}^m(\mathbf{y})$, Eq.(10) for $\mathbf{g}(\mathbf{y};\epsilon)$, and choosing

$$\tau^m_{m'}(\mathbf{y}) = \Theta^m_{m'}(\mathbf{y})\tau^{m'}(\mathbf{y};k_{m'}), \quad m,m' = 1,\ldots,M, \tag{16}$$

Eq.(15) straightforwardly yields:

$$\mathbf{b}^m_o(\mathbf{y}) = \mathbf{b}^m(\mathbf{y}) + O(\epsilon), \quad m = 1,\ldots,M. \tag{17}$$

In other words, the $\mathbf{b}^m(\mathbf{y})$'s as given by Eq.(11b) are the *converged* CSP iterants—to leading order in the small ϵ limit—of the fast row basis vectors. A nominal estimate of ϵ in the $O(\epsilon)$ above is $|\tau^M/\tau^{M+1}|$. Since all the reaction-specific $\tau^{m'}(\mathbf{y};k_{m'})$'s are negative and small, the real part of all the eigenvalues of $\tau^m_{m'}(\mathbf{y})$ are also expected to be negative and small.

The CSP step # 2 $\mathbf{a}_m$-refinement procedure was given by Eq.(18b) of [5]. It can similarly be demonstrated that the $\mathbf{a}_m$'s as chosen by Eq.(11a) are also the *converged* CSP iterants—to leading order in the small ϵ limit—of the fast column basis vectors.

Thus the $\mathcal{Q}^{fast}(M)$ as given by Eq.(12) has excellent CSP credentials in the small ϵ limit.

6 The quality of this $\mathcal{Q}^{fast}(M)$

Using $\mathcal{Q}^{fast}(M)$ as given by Eq.(3) in Eq.(5b), the CSP-partitioned $\mathbf{g}^{fast}(\mathbf{y};\epsilon)$ is:

$$\mathbf{g}^{fast}(\mathbf{y};\epsilon) = \mathcal{Q}^{fast}(M) \odot \mathbf{g}(\mathbf{y};\epsilon) = \sum_{m=1}^{M} \boldsymbol{\alpha}_m f^m, \tag{18a}$$

$$f^m(\mathbf{y}) \equiv \mathbf{b}^m \odot \mathbf{g}(\mathbf{y};\epsilon), \quad m = 1,\ldots,M. \tag{18b}$$

In order for $\mathbf{g}^{fast}(\mathbf{y};\epsilon)$ to become small in the slow epoch, the chosen $\mathcal{Q}^{fast}(M)$ must guarantee that the $f^m(\mathbf{y})$'s will become small when the fast reactions are exhausted.

Using Eq.(11b) for $\mathbf{b}^m$ and Eq.(10) for $\mathbf{g}(\mathbf{y};\epsilon)$ in Eq.(18b) and taking advantage of Eq.(11c), we obtain the following exact formula for $f^m(\mathbf{y})$:

$$\begin{aligned} f^m(\mathbf{y}) &= \sum_{m'=1}^{M} \Theta^m_{m'}\boldsymbol{\beta}^{m'} \odot \left(\sum_{m"=1}^{M} \boldsymbol{\alpha}_{m"}\Omega^{m"}(\mathbf{y};k_m) + \sum_{n=M+1}^{R} \boldsymbol{\alpha}_n \Omega^n(\mathbf{y};k_n) \right) \\ &= \Omega^m(\mathbf{y};k_m) - \Omega^m_\infty(\mathbf{y};k_{slow}), \quad m = 1,\ldots,M, \end{aligned} \tag{18c}$$

where k_{slow} denotes the slow reaction rates $k_{M+1}, \ldots, k_R$, and

$$\Omega^m_\infty(\mathbf{y}; k_{slow}) \equiv -\sum_{n=M+1}^{R} \left(\sum_{m'=1}^{M} \Theta^m_{m'} \Gamma^{m'}_n \right) \Omega^n(\mathbf{y}; k_n), \quad m = 1, \ldots, M. \tag{18d}$$

Note that Eq.(18d) is clearly programmable, and that only slow reactions contribute to $\Omega^m_\infty(\mathbf{y}; k_{slow})$.

How does Ω^m evolve with time? To find out, we formally take the time derivative of $\Omega^{m'}$ to obtain:

$$\frac{d\Omega^{m'}}{dt} = \frac{\partial \Omega^{m'}}{\partial \mathbf{y}} \odot \frac{d\mathbf{y}}{dt} = \frac{\beta^{m'}}{\tau^{m'}} \odot \mathbf{g}, \quad m' = 1, \ldots, M. \tag{19}$$

Multiplying Eq.(19) by $\tau^m_{m'}$ as given by Eq.(16) and summing over m', we obtain another exact equation for f^m with the help of Eq.(11b) and Eq.(18b):

$$\sum_{m'=1}^{M} \tau^m_{m'} \frac{d\Omega^{m'}}{dt} = f^m, \quad m = 1, \ldots, M. \tag{20}$$

Eliminating f^m from Eq.(20) using Eq.(18c), we obtain:

$$\sum_{m'=1}^{M} \tau^m_{m'} \frac{d\Omega^{m'}}{dt} = \Omega^m - \Omega^m_\infty(\mathbf{y}; k_{slow}), \quad m = 1, \ldots, M. \tag{21}$$

This is an exact differential equation for Ω^m, and it is the heart of this paper. When the real part of all eigenvalues of $\tau^m_{m'}$ are negative, Eq.(21) says $\Omega^m \to \Omega^m_\infty(\mathbf{y}; k_{slow})$ as time marches on. The quality of this approximation depends on the smallness of $\tau^m_{m'}$—which is dependent on the quality of the chosen $\mathcal{Q}^{fast}(M)$.

6.1 The $\mathbf{g}^{fast}$ and $\mathbf{g}^{slow}$ partitions of this $\mathcal{Q}^{fast}(M)$

Eliminating f^m from Eq.(18a) using Eq.(18c) or Eq.(20), we have:

$$\mathbf{g}^{fast}(\mathbf{y}; \epsilon) = \sum_{m=1}^{M} \boldsymbol{\alpha}_m \left[\Omega^m(\mathbf{y}; k_m) - \Omega^m_\infty(\mathbf{y}; k_{slow}) \right], \tag{22a}$$

$$= \sum_{m,m'=1}^{M} \boldsymbol{\alpha}_m \tau^m_{m'} \frac{d\Omega^{m'}}{dt}. \tag{22b}$$

- Eq.(21) tells us that $\mathbf{g}^{fast}(\mathbf{y};\epsilon)$ is expected to become small as time marches on.

Using Eq.(22a) in Eq.(5c), we have:

$$\mathbf{g}^{slow}(\mathbf{y};\epsilon) = \sum_{m=1}^{M} \boldsymbol{\alpha}_m \Omega_\infty^m(\mathbf{y};k_{slow}) + \sum_{n=M+1}^{R} \boldsymbol{\alpha}_n \Omega^n(\mathbf{y};k_n), \tag{23a}$$

$$= \sum_{n=M+1}^{R} \boldsymbol{\alpha}_n^{slow} \Omega^n(\mathbf{y};k_n), \tag{23b}$$

where

$$\boldsymbol{\alpha}_n^{slow}(\mathbf{y}) \equiv \boldsymbol{\alpha}_n - \mathcal{Q}^{fast}(M) \odot \boldsymbol{\alpha}_n. \tag{23c}$$

- Eq.(23a) says $\mathbf{g}^{slow}(\mathbf{y};\epsilon)$ is simply the original $\mathbf{g}(\mathbf{y};\epsilon)$ as given by Eq.(2) or Eq.(10), except that $\Omega_\infty^m(\mathbf{y};k_{slow})$ is used whenever $\Omega^m(\mathbf{y};k_m)$ is needed,

- Eq.(23b) showcases the $\boldsymbol{\alpha}_n^{slow}(\mathbf{y})$'s—which describe the *effective stoichiometries* of the slow reactions in $\mathbf{g}^{slow}(\mathbf{y};\epsilon)$.

Eqs.(22a,b) and Eqs.(23a,b,c) are exact for any $\mathcal{Q}^{fast}(M)$. Whenever $\mathbf{g}^{fast}(\mathbf{y};\epsilon)$ and $\mathbf{g}^{slow}(\mathbf{y};\epsilon)$ are both kept, the formulation is exact. All M $f^m(\mathbf{y})$'s in $\mathbf{g}^{fast}(\mathbf{y};\epsilon)$ should be *separately* computed and monitored. The quality of the chosen $\mathcal{Q}^{fast}(M)$'s is manifested by the smallness of $\mathbf{g}^{fast}(\mathbf{y};\epsilon)$ calculated via Eq.(22a) or Eq.(22b) after the fast transients are over.

The time evolution of the first term on the right hand side of Eq.(10) is also governed by Eq.(21). But there is no reason to expect that this term would become small after the fast transients are over.

6.2 The ϵ of any $\mathcal{Q}^{fast}(M)$

We can formally define the dimensionless parameter ϵ by:

$$\epsilon \equiv \frac{|\tau^{fast}|}{|\tau^{slow}|}. \tag{24}$$

where $|\tau^{fast}|$ is the slowest relevant time scale of the fast reaction group, and $|\tau^{slow}|$ is the fastest relevant time scale of the slow reaction group. The ratio of the intrinsic time scales $|\tau^M/\tau^{M+1}|$ previously used to rank order the

reactions was only a tentative estimate of this ϵ. The magnitude of the real part of the largest eigenvalue of $\tau^m_{m'}$ is probably a better choice for τ^{fast} than τ^M. See §13 of [7]. An interesting question is: what would be a better choice for τ^{slow} than τ^{M+1}?

Two exact alternative representations of $\mathbf{g}^{slow}(\mathbf{y};\epsilon)$ were derived in §6.1. Eq.(23a) provides R time scales: M new τ^m_∞'s from the first summation and $R-M$ original τ^n's from the second summation. Eq.(23b) provides another set of $R-M$ new τ^n_{slow}'s—which are different from the original τ^n's because Eq.(23b) uses the new and different $\boldsymbol{\alpha}^{slow}_n$'s. These new slow time scales could be mathematically useful/meaningful in their own rights on some issues—e.g. the stiffness of $\mathbf{g}^{slow}(\mathbf{y};\epsilon)$. Should one of these new slow time scales be the correct choice for τ^{slow}?

From the mathematical vantage point, the critical issue is the smallness of the left hand side of Eq.(21) after the fast reactions are exhausted. This depends on the estimated smallness of the elements of $\tau^m_{m'}$ relative to the time scale of Eq.(21)'s forcing term, $\Omega^m_\infty(\mathbf{y}(t);k_{slow})$. In other words, the mathematically correct choice for τ^{slow} should be the smallest of the new τ^m_∞'s of the $\Omega^m_\infty(\mathbf{y};k_{slow})$'s calculated by Eq.(8b) using Eq.(18d).

Whenever computing the inverse of $\Gamma^m_{m'}$ is problematic, some of the $\boldsymbol{\alpha}_m$'s and/or the β^m's above the tentative τ^r gap must be either linearly dependent or nearly so. The set of fast reactions to be chosen to use in $\mathcal{Q}^{fast}(M)$ is then not unique nor definitive. The competitive merits of the different non-unique reduced models can be adjudicated via their ϵ values as calculated by Eq.(24) above. For example, the best quality $\mathcal{Q}^{fast}(M)$ is most likely to be that provided by partitioning $\mathbf{g}(\mathbf{y};\epsilon)$ using the eigenmodes of $\partial\mathbf{g}/\partial\mathbf{y}$. The generality and simplicity of the recommended partitioning procedure here are achieved by conceding and carefully monitoring the fast-slow mode mixings.

7 A new model reduction strategy

The paper and pencil analysis in §6 used the values of the reaction-specific $\tau^r(\mathbf{y})$'s to identify the M linearly independent fast reactions on the fly. The $M\times M$ matrices $\Theta^m_{m'}(\mathbf{y})$ and $\tau^m_{m'}(\mathbf{y})$ are programmable and can be computed on the fly. When the eigenvalues of the latter are sufficiently small (and negative), the f^m's defined by Eq.(18b) are expected to become small after the fast transients are over.

At this point, we may simply neglect the f^m terms after verifying that

they are indeed negligible numerically. Neglecting negligible terms is the classical paper and pencil strategy to do singular perturbation, and the resulting reduced models are totally correct [7]. However, we advocate here a *new* CSP-based model reduction strategy which takes full advantage of the fact that numerical computations are involved:

1. Do *not* use the given algebraic formulas of the fast reactions' $\Omega^m(\mathbf{y};\epsilon)$ to compute the Ω^m's needed for the evaluation of $\mathbf{g}^{fast}(\mathbf{y};\epsilon)$. The information contained in the given algebraic $\Omega^m(\mathbf{y};\epsilon)$ formula have already been fully exploited in the derivation of Eq.(21).

2. Use some *implicit* integration algorithm for the numerical integration of Eq.(21) to obtain the $\Omega^m(t)$'s and the $d\Omega^m(t)/dt$'s. For example, use backward finite difference for $d\Omega^m/dt$. Use Eq.(18d) to compute the $\Omega^m_\infty(\mathbf{y};k_{slow})$'s needed by Eq.(21).

3. Use the $\Omega^m(t)$'s or the $d\Omega^m(t)/dt$'s from item #2 above to compute $\mathbf{g}^{fast}(\mathbf{y};\epsilon)$ from Eq.(22a) or Eq.(22b). Either route is fine. Include this $\mathbf{g}^{fast}(\mathbf{y};\epsilon)$ in Eq.(5a) in the numerical integration of Eq.(1) to obtain $\mathbf{y}(t)$. Explicit integration algorithm is fine here. Setting $\Omega^m(\mathbf{y};k_m) = \Omega^m(t)$ yields M algebraic "equations of state"—which could be used to reduce the number of dependent variables.

Item #2 above is the heart of this new strategy. It is easy to show that the use of implicit integration algorithm for Eq.(21) resolves the stiffness issue caused by $\mathbf{g}^{fast}(\mathbf{y};\epsilon)$ in the original problem. Note that item #3 advocates $\mathbf{g}^{fast}(\mathbf{y};\epsilon)$ be kept in $\mathbf{g}(\mathbf{y};\epsilon)$. Thus *nothing is neglected in this new strategy*. Essentially, the new strategy tacitly tolerates whatever inadequacies the implicit integration algorithms may have—knowing full well that the terms affected by their use are small and unimportant.

Output files for the $f^m(t)$'s and the $\Omega^m_\infty(t)$'s should be generated. Whenever $f^m(t)$ is not small in comparison to $\Omega^m_\infty(t)$, that m-th fast reaction is *currently active*, and thus must be playing a major role in the current reaction dynamics—particularly for those species identified by the m-th CSP *radical pointer*. See §6.4 of [4].

Whenever a currently small $f^m(t)$ is no longer small, this change of status is correctly dealt with by the new strategy effortlessly without ceremony. When the identities of the fast reactions are time dependent, a more inclusive $\mathcal{Q}^{fast}(M)$ which includes several marginally fast reactions could be useful. It

is also possible to skip over the fast transients if the time-resolved details of the fast decays are not of interest. Note that Eq.(22b) can be used to perform the so-called CSP *radical correction* procedure [3, 4].

For reacting flow problems, the right hand side of Eq.(1) would contain diffusion terms. So long as the estimated time scales of the diffusion processes (laminar or turbulent) are much larger than $|\tau^{fast}|$, these terms would be inside the $\mathbf{g}^{slow}$ partition. See §7.5 of [4] on how these diffusion terms should be handled.

After the fast transients, the current dynamics is dictated by reactions with the smallest $|\tau^n_{slow}|$'s. Reactions with finite but very large $|\tau^n_{slow}|$'s—computed from Eq.(23b) using $\boldsymbol{\alpha}^{slow}_n$'s—should only be ignored after their numerical contributions to $\mathbf{g}^{slow}$ have been found to be negligible numerically.

8 Concluding remarks

For chemical kinetics problems, the reaction-specific stoichiometric coefficients and the kinetic rate data contain useful information for doing model reductions. Subroutines for $\boldsymbol{\alpha}_r(\mathbf{y})$, $\boldsymbol{\beta}^r(\mathbf{y})$, and $\tau^r(\mathbf{y}; k_r)$ should be included for every reaction in chemical kinetics databases so that tentative values of M and $|\tau^M/\tau^{M+1}|$—for any reaction system of interest—can easily be determined, and that CSP data such as $\Gamma^m_r(\mathbf{y})$, $\Theta^m_{m'}(\mathbf{y})$, $\mathcal{Q}^{fast}(M)$, $\tau^m_{m'}(\mathbf{y})$, and *most importantly*, $\Omega^m_\infty(\mathbf{y}; k_{slow})$ and $\boldsymbol{\alpha}^{slow}_n(\mathbf{y})$, can be routinely made available. The quality of the resulting CSP-derived reduced model can be and should be *directly* assessed numerically.

The main scientific value of model reductions is the *insights* it can provide on the system of interest. Knowledgeable investigators can inspect the CSP-derived analytical formulas and computed numbers to extract informative answers to questions such as *what roles do the r-th reaction play at time t, which k_r's must be known accurately* and which k_r's could tolerate sloppiness, etc.

It is worthwhile to point out again that numerical solutions computed from the new strategy's reduced models are formally exact—the only concession made is the use of implicit finite difference to deal with Eq.(21). Of course, the option to trade off some exactness for some reduction of the amount of needed computations is always available.

References

[1] F. A. Williams, *Combustion Theory, The fundamental theory of chemically reacting systems*, 2nd edition, Benjamin/Cummings Publ. Co., Menlo Park, CA., 1985.

[2] S. H. Lam, "Singular perturbation for stiff equations using numerical methods," *Recent Advances in the Aerospace Sciences*, Corrado Casci, Ed., Plenum Press, New York and London, 1985, pp.3-20.

[3] S. H. Lam and D. A. Goussis, "Conventional asymptotics and computational singular perturbation for simplified kinetics modeling," in *Reduced Kinetic Mechanisms and Asymptotic Approximations for Methane-Air Flames*, Chapter 10. *Lecture Notes in Physics 384*, M Smooke, Ed., Springer-Verlag, 1991.

[4] S. H. Lam, "Using CSP to understand complex chemical kinetics," *Combust. Sci. and Tech*, **89**, 1993, pp. 375-404.

[5] S. H. Lam and D. A. Goussis, "The CSP Method for Simplifying Kinetics," *International Journal of Chemical Kinetics*, **26**, 1994, pp. 461-486.

[6] P. D. Kourdis, R. Steuer and D. A. Goussis, "Physical understanding of complex multiscale biochemical models via algorithmic simplification: Glycolysis in Saccharomyces cerevisiae," *Physica D*, 239, 2010, pp. 1798-1817.

[7] D. A. Goussis, "Quasi steady state and partial equilibrium approximations: their relation and their validity." *Combustion Theory and Modelling*, **6**, 5, 2012, pp. 869-926.

[8] U. Maas and S. B. Pope, "Simplifying Chemical Kinetics Intrinsic Low-Dimensional Manifolds in Composition Space," *Combustion and Flame*, **88**, 1992, pp. 239-264.