

The CSP Method for Simplifying Kinetics*

S. H. Lam† and D.A. Goussis‡

Department of Mechanical and Aerospace Engineering

†Princeton University, Princeton, N.J. 08544, U.S.A.

‡University of Patras, Patra, Greece

Published in the
International Journal of Chemical Kinetics
26, pp. 461-486, 1994

Key Words:

Chemical kinetics, steady-state approximation, partial-equilibrium approximation, numerical computation.

Abstract

The *Computational Singular Perturbation* (CSP) method of simplified kinetics modeling is presented with emphasis on its comparative merits versus conventional methodologies. A new “refinement” procedure for the basis vectors spanning the fast reaction subspace is presented. A simple example is first worked through using the conventional partial-equilibrium and quasi-steady approximations, and is then treated in some detail using CSP.

1 Introduction

When an investigator is confronted with an unfamiliar problem in chemical kinetics, the traditional first step is to identify the relevant chemical species

*This work was supported by NASA Langley’s Aerothermodynamics Branch, Space Systems Division, and AFOSR URI Grant F49620-93-1-0427.

and the important elementary reactions which occur among them. After establishing a “complete model” of the reaction system, it is usually desirable to obtain a simplified model by taking advantage of available approximations. For sufficiently simple problems, conventional analytical methods can be used^{[1],[2],[3],[4],[5]}. In most situations, the success of the resulting simplified model is measured not only by its quantitative predictive capability, but also by its simplicity—the fewer superfluous terms the better. Generally speaking, approximate analytical results are highly valued because of the insights they can provide when inspected by a competent theoretician. But this approach is not a practical option when the reaction system is massively complex.

Recently, databases containing extensive, reliable and up-to-date data for certain reaction systems are available. Computations using complete models from such databases can now be routinely carried out. In this new computational era, it is no longer necessary to pick out only the relevant chemical species and the important elementary reactions—because inclusion of benign superfluous terms in the formulation is not a problem. An option increasingly available to modern theoreticians is to first generate a complete model numerical solution, examine the resulting data to discern significant and interesting causes and effects—making additional diagnostic runs if necessary—and then try to propose simplifications and approximations. Why do modern theoreticians still care about simplified approximate models when double-precision numerical solutions to the complete model are easily available? The reasons are: physical understanding and “stiffness.” Most theoreticians like to be able to make general statements about the problem in addition to showing color slides of the numerical solutions. The observed behaviors of the computed solutions need to be described in terms of familiar concepts, such as chain-branching, chain-termination, ignition delay, building up of the radical pool, heat release etc. In particular, theoreticians would also like to be able to identify the rate-controlling reactions (for the chemical species of interest in the time interval of interest), fast reactions for which rate coefficients do not need to be known accurately, and superfluous reactants and reactions which need not be included at all. In addition, the vast disparity of time scales which is responsible for the simplifications and approximations is also responsible for *stiffness*^[6], a generally undesirable attribute of the governing differential equations from the computation point of view.

The theory of *computational singular perturbation*^{[7],[8],[9],[10],[11],[12]} (to

be referred to as CSP here) exploits the power of the computer to do simplified kinetics modeling. In essence, CSP is a systematic mathematical procedure to do boundary-layer type singular perturbation analysis. While it can be used to obtain analytical results for simple problems, it is designed to be used for massively complex problems using computations. A CSP computation not only generates the numerical solution of the given problem, but also the *simplified equations* in terms of the given information. Most interesting questions about the reaction systems can be answered merely by inspection of the numerical CSP data.

The present paper presents the latest theoretical developments of CSP, including a simple example to illustrate the essential features.

2 The Mathematical Problem

We consider a spatially homogeneous reaction system consisting of R elementary chemical reactions. The total number of unknowns, which include concentrations of the chemical species and other state variables such as temperature and total pressure, is N . We represent the N unknowns by a N -dimensional column vector, $\mathbf{y} = [y^1, y^2, \dots, y^N]^T$. The governing equation for \mathbf{y} is a system of quasi-linear ODE:

$$\frac{d\mathbf{y}}{dt} = \mathbf{g}(\mathbf{y}), \quad (2.1)$$

where \mathbf{g} is the sum of contributions from the R elementary reactions:

$$\mathbf{g} \equiv \sum_{r=1}^R \mathbf{s}_r F^r(\mathbf{y}), \quad (2.2)$$

and \mathbf{s}_r and $F^r(\mathbf{y})$ are the (generalized) *stoichiometric vector* and the *reaction rate* of the r -th elementary reaction, respectively. The value of R may be greater, less than, or equal to N . We shall call (2.2) the *physical representation* of \mathbf{g} , because each additive term can be satisfactorily explained by the investigator who formulated the problem.

Usually, an investigator is interested only in certain special species, each with a different accuracy requirement and perhaps in a different time interval. In many cases, not all the initial conditions needed to compute the solutions

are known, and many rate constants needed in the computations are only educated estimates. The mathematical problem is to derive the simplest model of the reaction system tailored to the interest of the investigator, consistent with the user-specified accuracy requirements.

3 Some Definitions

When the forward and reverse reaction rates of a single or a group of fast reactions are in approximate balance, we say the reaction or the reaction group is in *partial-equilibrium*. When the production and destruction rates of a particular species are in approximate balance, we say the species is in *quasi-steady* state. In either case, an approximate algebraic relation is obtained between components of the \mathbf{y} vector. Such relations, which do *not* contain free parameters, shall be called *equations of state*. In conventional simplified kinetics modeling, the fast reactions and the quasi-steady species are identified based on experience and intuition. In many cases, species in quasi-steady state are chemical radicals, but not always.

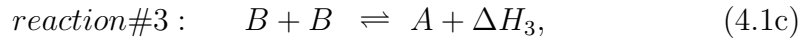
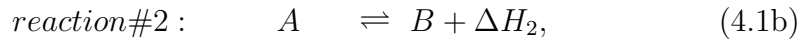
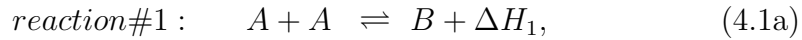
When M equations of state are available from applying these approximations, they can be used to selectively solve for M unknowns in terms of the others—in so doing the need for M of the ODE's is eliminated. Which M unknowns should be solved for in terms of the $(N - M)$ others, and which M ODE's should be discarded? It is not true that the choice of these M unknowns can be made arbitrarily. Lam^[12] provides explicit theoretical guidance to make these choices, and calls the M special unknowns *CSP radicals*, or simply radicals for short, when they can be accurately solved for from the M equations of state. When used in this CSP context, the term radical carries a special meaning distinct from its usual chemical structure context. In most situations, a CSP radical is also a chemical radical, and vice versa. But it is not always true. The identification of CSP radicals is done via a “radical pointer” which shall be demonstrated later.

Frequently, additional (exact or approximate) algebraic relations between the components of the \mathbf{y} vector exist. The conservation law of atomic species is one such example. Such relations are distinguished from the above-defined equations of state by the presence of free parameters which are determined by initial or boundary conditions. Following Lam^[12], we shall call all such relations containing free parameters *conservation laws*.

4 A Simple Example

We shall use a simple hypothetical reaction system^[12] to illustrate the issues involved.

Let the state vector be $\mathbf{y} = [A, B, C]^T$ where A and B are chemical concentrations and C is temperature. The reaction system consists of three elementary reactions:



where ΔH_1 , ΔH_2 and ΔH_3 are the heats of reaction (with the proper units) of the respectively reactions. The (generalized) stoichiometric vectors and the reaction rates are:

$$\mathbf{s}_1 = [-2, 1, \Delta H_1]^T, \quad F^1 = k_1(A^2 - K_1B), \quad (4.2a)$$

$$\mathbf{s}_2 = [-1, 1, \Delta H_2]^T, \quad F^2 = k_2(A - K_2B), \quad (4.2b)$$

$$\mathbf{s}_3 = [1, -2, \Delta H_3]^T, \quad F^3 = k_3(B^2 - K_3A). \quad (4.2c)$$

where the reaction rate coefficients k_1, k_2, k_3 and the equilibrium constants K_1, K_2, K_3 are known and—for the sake of simplicity—their dependence on C is assumed negligible. We shall find it useful to separately identify the forward and reverse reaction rates as follows:

$$F^r = F_+^r - F_-^r, \quad r = 1, 2, \dots, R, \quad (4.3)$$

where F_+^r and F_-^r are both positive.

The system of ODE is:

$$\frac{d\mathbf{y}}{dt} = \mathbf{s}_1 F^1 + \mathbf{s}_2 F^2 + \mathbf{s}_3 F^3, \quad (4.4)$$

which can be written out as follows:

$$\frac{dA}{dt} = -2F^1 - F^2 + F^3, \quad (4.5a)$$

$$\frac{dB}{dt} = F^1 + F^2 - 2F^3, \quad (4.5b)$$

$$\frac{dC}{dt} = \Delta H_1 F^1 + \Delta H_2 F^2 + \Delta H_3 F^3. \quad (4.5c)$$

To make things concrete, the rate coefficients are given numerical values:

$$k_1 \approx 10^4 \text{cc/mole-second}, \quad K_1 \approx 1.1 \times 10^{-2} \text{mole/cc}, \quad (4.6a)$$

$$k_2 \approx 10^{-1} \text{/second}, \quad K_2 \approx 1.1 \times 10^2, \quad (4.6b)$$

$$k_3 \approx 10^4 \text{cc/mole-second}, \quad K_3 \approx 0.8 \times 10^{-8} \text{mole/cc}, \quad (4.6c)$$

and

$$\Delta H_1 \approx +1.1 \times 10^4 \text{cc-}^\circ K \text{/mole}, \quad (4.7a)$$

$$\Delta H_2 \approx +1.0 \times 10^5 \text{cc-}^\circ K \text{/mole}, \quad (4.7b)$$

$$\Delta H_3 \approx -2.9 \times 10^5 \text{cc-}^\circ K \text{/mole}. \quad (4.7c)$$

The initial conditions are also given numerical values:

$$A(0) \approx 1.5 \times 10^{-4} \text{mole/cc}, \quad B(0) \approx 0.1 \times 10^{-6} \text{mole/cc}, \quad C(0) \approx 300^\circ K. \quad (4.8)$$

The investigator is interested in $A(t)$, accurate to two significant figures, after the first few seconds for the next few minutes.

Experience and intuition can play no role here because the problem is hypothetical, and indeed may not even make “chemical sense.” Note that *detailed balance* would require $K_3 = K_1/(K_2)^3$, thermodynamics would require $\Delta H_3 = \Delta H_1 - 3\Delta H_2$, and the *Law of Mass Action* would require \mathbf{s}_r and F^r to be consistent. The numerical data provided above do not satisfy some of these relations exactly—they were chosen to demonstrate that the techniques under discussion are completely mathematical, and do not require consistency with constraints “external” to the given mathematical problem.

Because this example is sufficiently simple, conventional asymptotic analyses can be performed—provided that the dependent and independent variables can be intelligently non-dimensionalized, and a small dimensionless parameter can be identified. The reader can readily confirm that even for this simple problem the task of non-dimensionalization of variables is not straightforward or trivial. As a consequence, highly complex problems are usually dealt without using non-dimensionalization. Instead of formal asymptotics, progress is made via intuitive *ad hoc* judgment on the speed rankings of the elementary reactions. In addition, the use of partial-equilibrium or quasi-steady approximations must follow certain special procedures. We shall illustrate the conventional methodologies in the following section.

5 The Conventional Methodologies

In dealing with the example problem, one must first decide whether to apply the partial-equilibrium or the quasi-steady state approximation. No clear cut guideline exist in the literature at the present time. We shall work the example problem using both: the partial-equilibrium approximation first, followed by the quasi-steady approximation. We shall show that different results are obtained, and that the quasi-steady results are more restrictive than the partial-equilibrium results.

5.1 The Partial-Equilibrium Approximation

For the example problem, a competent investigator will (correctly) conclude that, for the problem under investigation, reaction #1 is fastest, reaction #2 is next, and reaction #3 is slowest. Hence, in the time period of interest, it is expected that reaction #1 will have exhausted itself, reaction #2 is the rate-controlling reaction, and reaction #3 is essentially dormant. Experience and intuition of the investigator play a major role here.

Reaction #1 is exhausted: The partial-equilibrium approximation for reaction #1 consists of setting the net reaction rate of reaction #1 to zero. The following equation of state is obtained:

$$F^1 = k_1(A^2 - K_1B) \approx 0, \quad (5.1)$$

Note that (5.1) contains no free parameters. One may use it to solve for A in terms of B , or B in terms of A .

Equation (5.1) must be handled with care—it must *never* be substituted directly into the original system of equations. If this advice is not heeded and (5.1) is substituted into (4.5a), (4.5b) and (4.5c), the resulting equations are simply wrong. This is because the small net reaction rate of an exhausted fast reaction is in general competitive with the currently active slower reactions. In other words, (5.1), which is an approximation, is only adequate to be used as an equation of state, but is inadequate to be used in the original equations. See (5.10) later.

Reaction #3 is dormant: The contribution by reaction #3 to (4.4) can be neglected and need not be included in the time period of interest.

It is emphasized that reaction #3 being negligible from the original system of equations does *not* mean $F^3 \approx 0$ is a valid equation. In other words, when F^3 is neglected from (4.4), $B^2 \approx K_3 A$ is *not* valid and must *never* be used. After neglecting $s_3 F^3$ from (4.4), one can show that:

$$(\Delta H_1 - \Delta H_2)A + (\Delta H_1 - 2\Delta H_2)B + C \approx \text{Constant}. \quad (5.2)$$

which is a (temporary) conservation law. An astute investigator will be able to attribute the physical meaning of “conservation of total energy” to (5.2). If one were interested in the reaction system over a time period of the order of hours, reaction #3 would not be negligible and (5.2) would not be valid.

The Simplified Model: The procedures recommended by Williams^[1] to follow-up the partial-equilibrium approximation proceeds as follows. First, one of the ODE in the system of equations is chosen to eliminate F^1 from the rest of the equations. For example, (4.5b) can be used to eliminate F^1 from (4.5a) and (4.5c). These two equations are then supplemented by (5.1) differentiated with respect to time. The simplified model is obtained by solving for dA/dt , dB/dt and dC/dt from these three equations. We obtain:

$$\frac{dA}{dt} \approx \frac{K_1}{K_1 + 4A} F^2 - \frac{3K_1}{K_1 + 4A} F^3, \quad (5.3a)$$

$$\frac{dB}{dt} \approx \frac{2A}{K_1 + 4A} F^2 - \frac{6A}{K_1 + 4A} F^3, \quad (5.3b)$$

$$\begin{aligned} \frac{dC}{dt} \approx & [\Delta H_2 - \Delta H_1 \left(\frac{K_1 + 2A}{K_1 + 4A} \right)] F^2 \\ & + [\Delta H_3 + \Delta H_1 \left(\frac{2(K_1 + A)}{K_1 + 4A} \right)] F^3. \end{aligned} \quad (5.3c)$$

Note that either (5.3a) or (5.3b) can be replaced by the equation of state, (5.1). Since F^1 does not appear in the above simplified model, this system of ODE is no longer stiff (regardless of how large k_1 is). It can easily be verified that its solution satisfies (5.2) when either F^3 or $(\Delta H_1 - 3\Delta H_2 - \Delta H_3)F^3$ is neglected.

The New Initial Conditions: The initial conditions at $t \rightarrow 0^+$ are no longer given by (4.8), since they must satisfy (5.1). A detailed analysis will show that reaction #2 is dormant in a brief initial transient period when reaction #1 alone dominates. The dormancy of reaction #2 yields a temporary conservation law, $A + 2B = \text{constant}$, which together with (5.1) can be used to find the modified initial conditions for A and B:

$$A(0^+) \approx A(0) + 2(B(0) - B(0^+)) = 1.46 \times 10^{-4}, \quad (5.4a)$$

$$B(0^+) \approx \frac{A^2(0^+)}{K_1} = 1.94 \times 10^{-6}. \quad (5.4b)$$

The value $C(0^+) \approx 301.04$ can be found from the conservation law (5.2) applied between $t = 0$ and $t = 0^+$.

5.2 The Quasi-Steady Approximation

Alternatively, an investigator may prefer to proceed with the quasi-steady approximation instead of the partial-equilibrium approximation.

The quasi-steady state approximation requires that certain species be chosen as “radicals.” If B is chosen to be the radical, we neglect dB/dt from (4.5b) to yield the following equation of state:

$$F^1 \approx -F^2 + 2F^3. \quad (5.5)$$

Unlike (5.1) in the partial-equilibrium development, this equation is substituted directly into (4.5a) and (4.5c) to eliminate F^1 . We obtain:

$$\frac{dA}{dt} \approx F^2 - 3F^3, \quad (5.6a)$$

$$\begin{aligned} \frac{dC}{dt} \approx & [\Delta H_2 - \Delta H_1]F^2 \\ & + [\Delta H_3 + 2\Delta H_1]F^3, \end{aligned} \quad (5.6b)$$

The radical, B , is to be solved for algebraically from (5.5) in terms of A . These equations differ from (5.3a), (5.3b) and (5.3c), but agree with them in the $K_1 \gg A$ limit.

If A is chosen as the radical instead, we neglect dA/dt from (4.5a) to yield:

$$F^1 \approx (-F^2 + F^3)/2. \quad (5.7)$$

Substituting this equation into (4.5b) and (4.5c) to eliminate F^1 , we obtain:

$$\frac{dB}{dt} \approx \frac{1}{2}F^2 - \frac{3}{2}F^3, \quad (5.8a)$$

$$\begin{aligned} \frac{dC}{dt} \approx & [\Delta H_2 - \frac{1}{2}\Delta H_1]F^2 \\ & + [\Delta H_3 + \frac{1}{2}\Delta H_1]F^3, \end{aligned} \quad (5.8b)$$

The radical, A , is to be solved for algebraically from (5.7) in terms of B . These equations also differ from (5.3a), (5.3b) and (5.3c), but agree with them in the $K_1 \ll A$ limit.

Whenever $K_1 = O(A)$, the quasi-steady approximation is simply wrong^{[12],[16]}. According to this presentation, the quasi-steady approximation is seen to be more restrictive and less robust than the partial-equilibrium approximation. See §5.4 later.

5.3 The Weaknesses of the Conventional Methods

The main weakness of the conventional method is, of course, the need to rely on experience and intuition to make educated estimates of the speed ranking of the elementary reactions, and to identify the radicals.

The analytical results obtained above are only the leading order term of an asymptotic theory valid in the limit of $k_1 \rightarrow \infty$. How good are they for the “real” problem at hand? For the example problem, a formal asymptotic analysis for large k_1 would use $\epsilon = k_2K_2/k_1K_1$ as the small parameter. For the given rate data, we have $\epsilon \approx 10^{-1}$, which is barely adequate to provide one significant figure in the answers. Because of the enormous amount of algebra involved, few if any would ever attempt to obtain the higher order terms.

More seriously, the educated guesses of speed rankings and the identification of quasi-steady radicals, *etc.* depend strongly on the initial conditions. For example, if in our example problem the initial data were $A(0) \approx O(10^{-4})$ and $B(0) \approx O(10^{-1})$ (while all the rate data remained the same), then nearly all the above derived results would no longer apply.

5.4 The Exhausted Fast Reaction Rate

We have shown that the equation of state obtained by the partial-equilibrium approximation, (5.1), must not be substituted directly into the original equations, while the equation of state obtained by the quasi-steady approximation, (5.5) or (5.7), can be more freely used—with the caveat that its validity is limited. It is thus obvious that neither (5.1) nor (5.5) or (5.7) are “sufficiently” accurate and general. The question is: can a “better” equation of state be found for large time than either (5.1), (5.5) or (5.7)? In other words, when reaction #1 is considered exhausted, what is the best way to evaluate F^1 ?

Differentiating F^1 with respect to time, we obtain without approximation:

$$\begin{aligned}\frac{dF^1}{dt} &= k_1(2A\frac{dA}{dt} - K_1\frac{dB}{dt}) \\ &= \frac{F^1 - F_\infty^1}{\tau(1)},\end{aligned}\tag{5.9a}$$

where

$$\tau(1) = -\frac{1}{k_1(K_1 + 4A)},\tag{5.9b}$$

$$F_\infty^1 = -\frac{K_1 + 2A}{K_1 + 4A}F^2 + \frac{2(K_1 + A)}{K_1 + 4A}F^3.\tag{5.9c}$$

It is now clear what F^1 does as time marches on. Initially, F^1 decays exponentially with characteristic time scale $O(|\tau(1)|)$, but eventually it follows F_∞^1 which evolves with a slower time scale. The long time asymptotic solution of (5.9a), for $t \gg O(|\tau(1)|)$, can be expanded in a Taylor Series of $\tau(1)$:

$$F^1 \approx F_\infty^1 + \tau(1)\frac{d}{dt}F_\infty^1 + \dots\tag{5.10}$$

In other words, when reaction #1 is exhausted, the leading approximation for F^1 when $\tau(1)$ is small is F_∞^1 . If $F^1 \approx F_\infty^1$ is substituted *directly* into (4.5a) and (4.5c), we will indeed recover (5.3a) and (5.3c), the partial-equilibrium results obtained earlier. Hence, an alternative way to arrive at the partial-equilibrium results is to use $F^1 \approx F_\infty^1$, which is clearly a superior equation of state than either (5.1), (5.5) or (5.7). It agrees with them in the appropriate limits.

5.5 Observations on the Conventional Methodologies

The most important information needed by the conventional methodologies is the speed ranking of the reactions and the identification of the CSP radicals. Once the fast reactions are somehow identified, and the radicals chosen, the partial-equilibrium and the quasi-steady approximations are available to make further progress—with appropriate caution on the unreliability of the latter. The derivation presented in §5.4 is a new and different way to make progress—it does not need to identify the radical and it clearly yields the most accurate equation of state, $F^1 \approx F_\infty^1$. All it needs is assurance that reaction #1 is the fastest.

The theory of CSP provides a systematic way to identify the fast reactions. In addition, it generalizes the procedures in §5.4 to find the exhausted fast reaction rates—and thereby derives the simplified approximate chemical kinetics models.

6 The Theory of CSP

The basic idea of CSP was first published by Lam^[7] in 1985, the algorithm was tested and reported in a series of conference papers^{[13],[14],[15]}, a summary was given by Lam and Goussis^[10] in 1991, and its use as a tool to extract physical understanding was advocated by Lam^[12] in 1993. Because of its mathematical nature, the theory of CSP can appear to be rather abstract. The present section reviews the CSP theory and strategy—using the example problem to illustrate whenever possible—and presents a new procedure for the refinement of the CSP-basis vectors.

6.1 Basis Vectors

The vector \mathbf{g} contains all the physics of the problem, and is usually given by the investigator formulating the problem using the physical representation, (2.2).

In general, an N -dimensional vector may be expressed in terms of any set of N linearly independent basis vectors^[17]. CSP exploits the theoretician's prerogative to express \mathbf{g} in an *alternative representation*, and look for basis vectors with special properties.

Let $\mathbf{a}_i(t)$, $i = 1, 2, \dots, N$, be a set of N linearly independent column basis vectors. The set of inverse row basis vectors, $\mathbf{b}^i(t)$, $i = 1, 2, \dots, N$, can be computed from the orthonormal relations:

$$\mathbf{b}^i \odot \mathbf{a}_j = \delta_j^i, \quad i, j = 1, 2, \dots, N, \quad (6.1)$$

where \odot is the dot product operator of the N -dimensional vector space. The column vector \mathbf{g} can now be expressed as:

$$\mathbf{g} = \sum_{i=1}^N \mathbf{a}_i f^i, \quad (6.2)$$

where

$$f^i \equiv \mathbf{b}^i \odot \mathbf{g} = \sum_{r=1}^R B_r^i F^r, \quad i = 1, 2, \dots, N, \quad (6.3)$$

and

$$B_r^i \equiv \mathbf{b}^i \odot \mathbf{s}_r, \quad i = 1, 2, \dots, N. \quad (6.4)$$

Each of the additive terms in (6.2) represents a *reaction mode*, or simply *mode*. The *amplitude* and *direction* of the i -th mode are f^i and \mathbf{a}_i , respectively. Eventually, CSP provides an algorithm to compute an approximation to the “ideal” set of basis vectors for the derivation of the simplified models.

The physical representation of \mathbf{g} uses the physically meaningful (and time-independent) stoichiometric vectors as the default column basis vectors. For our example, the default set is:

$$\mathbf{a}_1 = \mathbf{s}_1 = [-2, \quad 1, \quad \Delta H_1]^T, \quad (6.5a)$$

$$\mathbf{a}_2 = \mathbf{s}_2 = [-1, \quad 1, \quad \Delta H_2]^T, \quad (6.5b)$$

$$\mathbf{a}_3 = \mathbf{s}_3 = [1, \quad -2, \quad \Delta H_3]^T. \quad (6.5c)$$

Using this set, the inverse row vectors can easily be computed:

$$\mathbf{b}^1 = [2\Delta H_2 + \Delta H_3, \quad \Delta H_2 + \Delta H_3, \quad 1]/H, \quad (6.6a)$$

$$\mathbf{b}^2 = [-2\Delta H_1 - \Delta H_3, \quad -\Delta H_1 - 2\Delta H_2, \quad -3]/H, \quad (6.6b)$$

$$\mathbf{b}^3 = [-\Delta H_1 + \Delta H_2, \quad -\Delta H_1 + 2\Delta H_2, \quad -1]/H, \quad (6.6c)$$

where

$$H \equiv \Delta H_1 - 3\Delta H_2 - \Delta H_3. \quad (6.6d)$$

Using the given input numerical data, we have:

$$H = 10^3 \text{cc}^\circ K/\text{mole}. \quad (6.7)$$

In spite of this quite respectable dimensional value, H is actually very nearly “zero,” because it is a small number in comparison to ΔH_1 , ΔH_2 and ΔH_3 . If H is identically zero, then the vectors \mathbf{s}_1 , \mathbf{s}_2 and \mathbf{s}_3 are not linearly independent. In that case an alternative \mathbf{a}_3 must be provided. We shall take advantage of $H \neq 0$ here and proceed with caution.

It is straightforward to verify that at $t = 0$, we have:

$$f^1 = \mathbf{b}^1 \odot \mathbf{g} = F^1 = 2.14 \times 10^{-4} \text{mole/cc-second}, \quad (6.8a)$$

$$f^2 = \mathbf{b}^2 \odot \mathbf{g} = F^2 = 1.39 \times 10^{-5} \text{mole/cc-second}, \quad (6.8b)$$

$$f^3 = \mathbf{b}^3 \odot \mathbf{g} = F^3 = -1.19 \times 10^{-8} \text{mole/cc-second}. \quad (6.8c)$$

6.2 The Speed Ranking of the Modes

Differentiating (6.3) with respect to time along a solution trajectory $\mathbf{y}(t)$, we obtain:

$$\frac{df^i}{dt} = \sum_{j=1}^N \Lambda_j^i f^j, \quad i = 1, 2, \dots, N, \quad (6.9)$$

where

$$\Lambda_j^i \equiv \left[\frac{d\mathbf{b}^i}{dt} + \mathbf{b}^i \odot \mathbf{J} \right] \odot \mathbf{a}_j, \quad i, j = 1, 2, \dots, N, \quad (6.10a)$$

$$\mathbf{J} \equiv \frac{\partial \mathbf{g}}{\partial \mathbf{y}} = N \times N \text{ Jacobian matrix}. \quad (6.10b)$$

The non-linear nature of the original problem is manifested by the fact that \mathbf{J} is in general not a constant matrix. At any moment in time, the eigen-values of \mathbf{J} can be computed. For problems arising from chemical kinetics, they are usually essentially real—when the problems are of the boundary-layer type. The reciprocal of an eigen-value, called the *time scale*, has the dimension of time, and shall be denoted by $\tau(i)$. Ordering them in increasing magnitudes, we have:

$$|\tau(1)| < \dots < |\tau(i)| < \dots < |\tau(N)|, \quad (6.11)$$

which provides an approximate speed ranking of the “eigen-modes.”

The question what is an ideal set of basis vectors now has an obvious answer: *ideal basis vectors* should diagonalize Λ_j^i , thereby uncoupling all the modes. This is the standard strategy for analyzing linear problems, but it obviously needs some modifications for non-linear problems because there is no finite algorithm to diagonalize Λ_j^i . A full discussion of this subtle point will be given in a future paper. When the set of basis vectors used is non-ideal, the modes are coupled, and each mode will *not* have a distinct characteristic time scale. As shown in our example in §5.4, the fastest mode F^1 evolves with its own characteristic time scale $|\tau(1)|$ only initially. As it becomes exhausted, it eventually follows F_∞^1 which evolves with the characteristic time scales of the slower modes. This *mode mixing* is the price we pay for not diagonalizing Λ_j^i , and is an intrinsic issue of non-linearity which must be dealt with: the fast modes behave as slow modes when they are near exhaustion. Viewed in this light, the task of deriving simplified models is reduced to finding basis vectors such that the fast modes mix with the slow modes as little as possible. From the pragmatic point of view, however, it is not really necessary to uncouple all the modes; it will be sufficient to uncouple the fast modes from the slow modes approximately—so that the residual coupling can be neglected in accordance with some user-specified accuracy threshold. Mode mixing among the fast modes or among the slow modes causes no difficulty and can be tolerated.

6.3 The Classification of Fast and Slow Modes

First of all, we need a more precise classification of fast and slow modes. Usually, an investigator has a definite idea on the desired time resolution Δt of the solutions. In other words, events whose time scales are shorter than Δt are *not of interest*. Hence, the group of M modes which satisfy:

$$|\tau(m)| < \Delta t, \quad m = 1, 2, \dots, M, \quad (6.12)$$

are considered *fast modes*, and all others are considered the *slow modes*. The fastest group of active slow modes are the *rate-controlling modes*. Slow modes with negligible amplitudes are called *dormant modes*.

If one is interested in all time scales, then $\Delta t = |\tau(M + 1)|$, where the integer M changes with time. Under this definition, the modes with time scales shorter than the current time scale are exhausted. We shall deal with

a precise definition of exhaustion later. In §5.4, we showed how to obtain the “asymptotic” large time solution for F^1 . We shall generalize the procedure for all M fast modes in §6.7 below.

6.4 The CSP Refinement Strategy

CSP assumes that at any moment in time a set of trial basis vectors is available, and its Λ_j^i is, in general, *not* diagonal. CSP does not attempt to diagonalize it; instead, it provides a programmable algorithm to refine the basis vectors so that the new Λ_j^i of the refined basis vectors is more block-diagonal than before.

For non-linear problems, the eigen-vectors of \mathbf{J} are time-dependent, and thus they do not diagonalize Λ_j^i . They are, however, excellent trial basis vectors. As demonstrated in the previous sections, the conventional method chooses from the set of default stoichiometric vectors, and requires a good guess of the speed ranking of the reactions.

The CSP theory uses the ratio

$$\epsilon_M(t) = \left| \frac{\tau(M)}{\tau(M+1)} \right| \quad (6.13)$$

as a small dimensionless parameter, and develops the refinement algorithm by an asymptotic analysis in the small ϵ_M limit. Physically, ϵ_M is a measure of the *time scale separation* of the fast and slow modes. Each formal application of the CSP refinement procedure depresses the magnitude of the off-diagonal blocks of Λ_j^i by $O(\epsilon_M)$.

6.5 The CSP Refinement Procedure

We shall assume that at any moment in time (on a solution trajectory), the current value of M and a set of trial basis vectors for the M fast modes are known:

$$\mathbf{a}_m, \mathbf{b}^m, \quad m = 1, 2, \dots, M. \quad (6.14)$$

With only the fast basis vectors available, we can compute for the $M \times M$ upper-left block of Λ_j^i , denoted by ω_n^m :

$$\omega_n^m \equiv \Lambda_n^m = \left(\frac{d\mathbf{b}^m}{dt} + \mathbf{b}^m \odot \mathbf{J} \right) \odot \mathbf{a}_n, \quad m, n = 1, 2, \dots, M. \quad (6.15)$$

The inverse of ω_n^m shall be denoted by τ_n^m :

$$\sum_{n'=1}^M \omega_{n'}^m \tau_n^{n'} = \sum_{n'=1}^M \tau_{n'}^m \omega_n^{n'} = \delta_n^m, \quad m, n = 1, 2, \dots, M. \quad (6.16)$$

We shall mark the refined basis vectors and all entities evaluated with refined basis vectors by either a superscript or a subscript \circ , or both.

The refinement procedure presented below is new and is an alternative to the refinement procedure presented by Lam and Goussis^[10] in 1991. The main advantage of the new method is that only the basis vectors for the fast subspace needs to be computed. Either method will work. A brief outline of the theory underlying the new method will be presented later in §6.7.

To refine the \mathbf{b}^m row vectors, we use the so-called *step #1* procedure:

$$\mathbf{b}_o^m = \sum_{n=1}^M \tau_n^m \left(\frac{d\mathbf{b}^n}{dt} + \mathbf{b}^n \odot \mathbf{J} \right), \quad m = 1, 2, \dots, M, \quad (6.17a)$$

$$\mathbf{a}_m = \text{unchanged}, \quad m = 1, 2, \dots, M. \quad (6.17b)$$

The *step #1* procedure depresses the upper-right $M \times (N - M)$ block of Λ_j^i —it makes the fast modes “purer” by weakening their couplings with the slow modes (See §6.7 later). To refine the \mathbf{a}_m column vectors, we use the so-called *step #2* procedure::

$$\mathbf{b}_o^m = \text{unchanged}, \quad m = 1, 2, \dots, M, \quad (6.18a)$$

$$\mathbf{a}_m^o = \sum_{n=1}^M \left(-\frac{d\mathbf{a}_n}{dt} + \mathbf{J} \odot \mathbf{a}_n \right) \tau_{m,o}^n, \quad m = 1, 2, \dots, M, \quad (6.18b)$$

where the definition of $\tau_{m,o}^n$ should be obvious. The *step #2* procedure depresses the lower-left $(N - M) \times M$ block of Λ_j^i —it makes the slow modes purer. Note that at each refinement step the orthonormality condition is always satisfied. These two steps may be performed singly, or in tandem in any order, or recursively any number of times—*provided the most current τ_n^m is used always*. This refinement procedure is mathematically equivalent to that presented in Lam and Goussis^[10], the difference being that in this formulation the slow basis vectors are not involved. If the time derivative terms in (6.17a) and (6.18b) were omitted, the procedure would be identical to the so-called *Mises Power Method* for finding eigen-vectors associated with the

largest eigen-values (See Carnahan, Luther and Wilkes^[18]). In essence, by allowing mode mixing among the fast modes, CSP extends the Mises Power Method to compute the next iterant for the fast basis vectors. Since the time scales of the fast modes are all faster than the current time scale, the time derivative terms are always small corrections.

In practice, the first guess of the trial basis vectors is usually time-independent. The subsequent refined sets will in general be time-dependent because they are, by construction, \mathbf{y} -dependent. In any case, their time derivatives can be evaluated accordingly. In a computer program, these time derivatives can be evaluated approximately using either stored or predicted data already available in the integration routine. Such programming issues, however, are beyond the scope of this paper.

6.6 The Fast Subspace Projection Matrix

We can form a $N \times N$ matrix $\mathbf{Q}(M)$ as follows:

$$\mathbf{Q}(M) \equiv \sum_{m=1}^M \mathbf{a}_m \mathbf{b}^m, \quad (6.19)$$

and call it the *fast subspace projection matrix*. When evaluated with refined basis vectors, it will be denoted by $\mathbf{Q}_o^o(M)$.

Any column vector \mathbf{c} or row vector \mathbf{r} can be decomposed into its fast and slow components using either $\mathbf{Q}(M)$ or $\mathbf{Q}_o^o(M)$. Using $\mathbf{Q}_o^o(M)$, we obtain:

$$\mathbf{c} = \mathbf{c}_o^{o,fast}(M) + \mathbf{c}_o^{o,slow}(M), \quad (6.20a)$$

$$\mathbf{r} = \mathbf{r}_o^{o,fast}(M) + \mathbf{r}_o^{o,slow}(M), \quad (6.20b)$$

where

$$\mathbf{c}_o^{o,fast}(M) \equiv \mathbf{Q}_o^o(M) \odot \mathbf{c}, \quad (6.21a)$$

$$\mathbf{c}_o^{o,slow}(M) \equiv (\mathbf{I} - \mathbf{Q}_o^o(M)) \odot \mathbf{c}, \quad (6.21b)$$

$$\mathbf{r}_o^{o,fast}(M) \equiv \mathbf{r} \odot \mathbf{Q}_o^o(M), \quad (6.21c)$$

$$\mathbf{r}_o^{o,slow}(M) \equiv \mathbf{r} \odot (\mathbf{I} - \mathbf{Q}_o^o(M)). \quad (6.21d)$$

We can decompose $\mathbf{Q}_o^o(M)$ into its M components:

$$\mathbf{Q}_o^o(M) = \sum_{m=1}^M \mathbf{Q}_{m,o}^o, \quad (6.22)$$

where

$$\mathbf{Q}_{m,o}^o \equiv \mathbf{a}_m^o \mathbf{b}_o^m, \quad m = 1, 2, \dots, M. \quad (6.23)$$

We shall call $\mathbf{Q}_{m,o}^o$ the *m-th fast mode projection matrix*.

The *radical pointer* of the m-th mode, $Q_m(i)$, is given by^[12]:

$$Q_m(i) \equiv \text{the } i\text{-th diagonal element of } \mathbf{Q}_{m,o}^o, \quad i = 1, 2, \dots, N. \quad (6.24)$$

Note that $Q_m(i)$ is dimensionless, and its sum over all N components is 1.0. Geometrically, $Q_m(k)$ is a measure of the projection of the k-th unit vector in the m-th mode. Hence, *whenever $Q_m(k)$ is not a small number, species k is said to be a CSP radical*.

The *fast reaction pointer* of the m-th mode, $P_m(r)$, is given by:

$$P_m(r) \equiv (\mathbf{s}_r)^{-1} \odot \mathbf{Q}_{m,o}^o \odot \mathbf{s}_r, \quad r = 1, 2, \dots, R, \quad (6.25)$$

which is a dimensionless number nominally of order unity. Here, $(\mathbf{s}_r)^{-1}$ is a set of row vectors orthonormal to the \mathbf{s}_r set of (linearly independent) column vectors. Geometrically, $P_m(r)$ is a measure of the projection of the r-th stoichiometric vector in the m-th mode. *Whenever $P_m(r)$ is not a small number, the r-th reaction is said to be a fast reaction*.

It should be emphasized that the quality of the information provided by these pointers is dependent on the quality of the fast subspace basis vectors used. Hence, pointers computed from “guessed” trial basis vectors without any refinement at all have no theoretical standing. In other words, only pointers computed from refined basis vectors can provide reliable information.

6.7 The Amplitudes of the Fast Modes

Using the trial fast subspace basis vectors, the original system of ODE’s can be written as follows:

$$\frac{d\mathbf{y}}{dt} = \mathbf{g}^{fast}(M) + \mathbf{g}^{slow}(M), \quad (6.26)$$

where

$$\mathbf{g}^{fast}(M) \equiv \mathbf{Q}(M) \odot \mathbf{g} \quad (6.27a)$$

$$= \sum_{r=1}^R \mathbf{a}_m f^m, \quad (6.27b)$$

$$f^m = \mathbf{b}^m \odot \mathbf{g}, \quad (6.27c)$$

and

$$\mathbf{g}^{slow}(M) \equiv (\mathbf{I} - \mathbf{Q}(M)) \odot \mathbf{g}. \quad (6.28)$$

The equation governing the time evolution of the amplitude f^m of the trial fast mode can be straightforwardly obtained by differentiating (6.27c):

$$\frac{df^m}{dt} = \left(\frac{d\mathbf{b}^m}{dt} + \mathbf{b}^m \odot \mathbf{J} \right) \odot (\mathbf{g}^{fast}(M) + \mathbf{g}^{slow}(M)). \quad (6.29)$$

Using (6.15), (6.16), (6.17a), (6.27b) and (6.28), we obtain:

$$\frac{df^m}{dt} = \sum_{n=1}^M \omega_n^m (f^n - f_\infty^n), \quad m = 1, 2, \dots, M, \quad (6.30)$$

where

$$f_\infty^m \equiv (\mathbf{b}^m - \mathbf{b}_o^m) \odot \mathbf{g}, \quad m = 1, 2, \dots, M, \quad (6.31)$$

Equation (6.30) is the generalized version of (5.9a). It is clear that $f^m \rightarrow f_\infty^m$ for large time (assuming that the eigenvalues of ω_n^m are essentially negative), but there is *no* assurance that f_∞^m , which evolves with time with the current active time scale, is small and therefore negligible.

Using the refined fast basis vectors \mathbf{b}_o^m 's, the amplitude f_o^m of the refined fast modes are:

$$f_o^m \equiv \mathbf{b}_o^m \odot \mathbf{g} = f^m - f_\infty^m, \quad m = 1, 2, \dots, M. \quad (6.32)$$

The governing equation for f_o^m can be obtained from (6.30):

$$\frac{df_o^m}{dt} = \sum_{n=1}^M \omega_n^m (f_o^n - \sum_{n'=1}^M \tau_{n'}^n \frac{df_\infty^{n'}}{dt}), \quad m = 1, 2, \dots, M, \quad (6.33)$$

It is clear from (6.33) that for large time (again assuming that the eigenvalues of ω_n^m are essentially negative), we have:

$$f_o^m \rightarrow \sum_{n'=1}^M \tau_{n'}^m \frac{df_\infty^{n'}}{dt}. \quad (6.34)$$

Since the fast time scales of τ_n^m are all small in comparison to the currently active time scale, the magnitude of the f_o^m 's are now theoretically small. In

other words, by executing the step #1 refinement procedure, the amplitudes of the refined fast modes are now smaller than before.

The step #2 refinement procedure makes no contribution toward making the large time fast mode amplitudes smaller. Its role is solely to purify the amplitude of the slow modes, and is included in the refinement procedure for the sake of symmetry. Our experience showed no adverse effects when the step #2 refinement procedure was not implemented.

6.8 The Simplified Model

After executing both step #1 and step #2 refinements, we obtain the simplified model of the reaction system by neglecting $\mathbf{g}_o^{o,fast}$:

$$\frac{d\mathbf{y}}{dt} \approx \mathbf{g}_o^{o,slow} \quad (6.35)$$

where

$$\mathbf{g}_o^{o,slow} = \sum_{r=1}^R \mathbf{s}_{o,r}^{o,slow} F^r, \quad (6.36a)$$

$$\mathbf{s}_{o,r}^{o,slow} = (\mathbf{I} - \mathbf{Q}_o^o(M)) \odot \mathbf{s}_r. \quad (6.36b)$$

The “initial” conditions for (6.35) at $t = 0^+$ must be adjusted to satisfy the following approximate equations of state:

$$f_o^m = \mathbf{b}_o^m \odot \mathbf{g} \approx 0, \quad m = 1, 2, \dots, M. \quad (6.37)$$

The radical correction algorithm described below can be used to enforce this condition as shall be numerically demonstrated using our example later. Note that (6.35) provides N ODE’s for the N unknowns, and its solutions are guaranteed to satisfy (6.37) for all $t > 0^+$ automatically provided (6.37) is satisfied initially and all fast eigen-values of \mathbf{J} are essentially negative.

The amplitudes of the exhausted fast modes can be numerically monitored during a computation. When the amplitude of a previously declared exhausted fast mode is found to be no longer negligible according to the user-specified accuracy threshold, the mode can be straightforwardly declared active again.

6.9 The Radical Correction

Treating (6.30) as a linear equation for f^m , we can express its solution as the sum of a homogeneous solution and a particular solution:

$$f^m = f_{hom}^m + f_{part}^m. \quad (6.38)$$

We continue to assume that the eigen-values of ω_n^m are essentially negative (*i.e.* the fast modes are of the boundary-layer type). The homogeneous solution is responsible for satisfying the initial conditions, and for boundary-layer type problems f_{hom}^m rapidly becomes exponentially small in time measured in units of $\tau(M)$. The total adjustment of \mathbf{y} in this brief time period is given approximately by:

$$(\Delta \mathbf{y})_{rc} = - \sum_{m,n=1}^M [\mathbf{a}_m \tau_n^m(M) (f^n - f_\infty^n)]_{t=0}, \quad (6.39)$$

which is derived under the *ad hoc* assumption that $\mathbf{a}_m \tau_n^m(M)$ is approximately time-independent. Equation (6.39) is called the *radical correction* by Lam^[12].

7 The CSP Method on the Example

We shall step through the example numerically using CSP.

7.1 Choosing the Trial Basis Vectors

We assume that at the beginning we have no idea which reaction is fast. The eigen-values $\lambda(i)$ and eigen-vectors of \mathbf{J} at $t = 0$ can be computed numerically. We have:

$$\lambda(1) = -1.27 \times 10^2/\text{second}, \quad (7.1a)$$

$$\lambda(2) = -0.173/\text{second}, \quad (7.1b)$$

$$\lambda(3) = 0.00/\text{second}, \quad (7.1c)$$

indicating that there is a fast mode with time scale of the order of 10^{-2} seconds, followed by a slower mode with time scale of the order of about 10^1

seconds. The right (column) eigen-vectors α_i and left (row) eigen-vectors β^i , ranked in order of decreasing speed, are:

$$\alpha_1 = [-1.90, \quad 0.993, \quad 1.87 \times 10^4]^T, \quad (7.2a)$$

$$\alpha_2 = [-0.960, \quad -2.46 \times 10^{-2}, \quad -0.901 \times 10^5]^T, \quad (7.2b)$$

$$\alpha_3 = [0.000, \quad 0.000, \quad 1.00]^T, \quad (7.2c)$$

and

$$\beta^1 = [-2.46 \times 10^{-2}, \quad 0.960, \quad 0.00], \quad (7.3a)$$

$$\beta^2 = [-0.993, \quad -1.90, \quad 0.000], \quad (7.3b)$$

$$\beta^3 = [-0.890 \times 10^5, \quad -1.89 \times 10^5, \quad 1.00]. \quad (7.3c)$$

These may be used as our time-independent trial basis vectors for $t \geq 0$ but they diagonalize Λ_j^i only at $t = 0$. Since our time resolution of interest is in seconds, only the first mode can be considered fast. Hence, $M = 1$.

Alternatively, we can use the stoichiometric vectors of the three reactions as our trial basis vectors (*i.e.* the default set in §6.1). Since they are time-independent, Λ_j^i at $t = 0$ is easily computed:

$$\Lambda_j^i = \begin{pmatrix} -1.16 \times 10^2 & -1.13 \times 10^2 & 2.23 \times 10^2 \\ -1.12 \times 10^1 & -1.11 \times 10^1 & 2.21 \times 10^1 \\ 2.16 \times 10^{-3} & 2.08 \times 10^{-3} & -4.08 \times 10^{-3} \end{pmatrix}. \quad (7.4)$$

This matrix has significant off-diagonal terms. Nevertheless, its diagonal elements can be used to estimate the time scales of the modes, and they indicate that mode #1 (*i.e.* reaction #1) is fastest and its time scale is possibly below the time resolution of interest—suggesting that $M = 1$.

The first eigen mode (which has the largest eigenvalue magnitude) is obviously the fastest. Hence, α_1 and β^1 are excellent trial fast basis vectors for $M = 1$. In what follows, however, we shall choose (6.5a) and (6.6a) instead as our trial fast basis vectors for $t \geq 0$, *but we shall verify this choice using the reaction pointer later*. We have:

$$\mathbf{a}_1 = [-2.00, \quad 1.00, \quad 1.10 \times 10^4]^T, \quad (7.5a)$$

$$\mathbf{b}^1 = [-9.00 \times 10^1, \quad -1.90 \times 10^2, \quad 1.00 \times 10^{-3}]. \quad (7.5b)$$

As mentioned previously, radical and reaction pointers computed using trial basis vectors are theoretically unreliable, and therefore their results should

be ignored. Nevertheless, we shall compute them here for later comparisons. The N elements of the unrefined radical pointer of mode #1 are (taking advantage of the fact that $H \neq 0$):

$$Q_1(1) = 1.80 \times 10^2, \quad Q_1(2) = -1.90 \times 10^2, \quad Q_1(3) = 1.10 \times 10^1. \quad (7.6)$$

The R elements of the unrefined fast reaction pointer of mode #1 are, as expected:

$$P_1(1) = 1.00, \quad P_1(2) = 0.00, \quad P_1(3) = 0.00. \quad (7.7)$$

It is emphasized again that these results from unrefined trial basis vectors are without merit.

7.2 Skipping the Rapid Transient Period

Since we are not interested in the rapid transient period which lasts tens of milli-seconds, the main issue now is to find the adjusted initial conditions for the simplified model which governs the slow evolutionary period.

In the rapid transient period, \mathbf{y} adjusts rapidly in such a way that the amplitude of the fastest mode approaches zero. This adjustment should, in general, be computed by integrating the full model ODE until $\mathbf{g}_o^{o,fast}$ is sufficiently small to be neglected. It can also be computed approximately by the radical correction given by (6.39), provided that there is adequate reason to believe that the trial fast basis vectors are good choices for the whole rapid transient period. For the example here, the amplitude of the fastest mode at $t = 0$ is $f^1 = F^1 = 2.14 \times 10^{-4}$. Making the radical correction using the trial fast basis vectors, we obtain the following adjusted initial condition at $t = 0^+$:

$$\mathbf{y}(0^+) = \mathbf{y}(0) + \Delta\mathbf{y}_{rc} = [1.46 \times 10^{-4}, \quad 1.95 \times 10^{-6}, \quad 300.02]^T, \quad (7.8)$$

which yields a much smaller amplitude, $f^1 = F^1 = -1.36 \times 10^{-7}$. Note that the relative magnitude of the correction to B is much larger than those to A and C .

7.3 The CSP Refinement

Using the trial fast basis vectors (7.5a) and (7.5b), we can decompose \mathbf{g} into its fast and slow components:

$$\mathbf{g} = \mathbf{g}^{fast} + \mathbf{g}^{slow}, \quad (7.9a)$$

where

$$\mathbf{g}^{fast} = \mathbf{Q}(1) \odot \mathbf{g} = [2.68 \times 10^{-6}, -1.34 \times 10^{-7}, -1.47 \times 10^{-2}]^T, \quad (7.9b)$$

$$\mathbf{g}^{slow} = (\mathbf{I} - \mathbf{Q}(1)) \odot \mathbf{g} = [6.87 \times 10^{-6}, -6.90 \times 10^{-6}, -0.667]^T. \quad (7.9c)$$

This \mathbf{g}^{fast} is of the “same order” as \mathbf{g}^{slow} (component by component), and *its neglect will cause an order unity error* since it was computed using the trial fast basis vectors which have never been refined.

Using the same trial basis vectors (7.5a) and (7.5b), the refined fast basis vectors at $t = 0^+$ are:

$$\mathbf{a}_1^o = [-1.92, 1.00, 1.90 \times 10^4]^T, \quad (7.10a)$$

$$\mathbf{b}_o^1 = [-2.52 \times 10^{-2}, 0.950, 0.000], \quad (7.10b)$$

with $\tau_o^o(1) = (\mathbf{b}_o^1 \odot \mathbf{J} \odot \mathbf{a}_1^o)^{-1} = -0.788 \times 10^{-2}$ seconds. Note that \mathbf{a}_1^o and \mathbf{b}_o^1 are quite close to $\boldsymbol{\alpha}_1$ and $\boldsymbol{\beta}^1$, the right and left eigen-vectors of the fastest mode. Note also that \mathbf{b}^1 and \mathbf{b}_o^1 bear no resemblance to each other. The interested readers may readily verify that if a pair of random vectors were used instead of (7.5a) and (7.5b) as the starting trial basis vectors, essentially the same results would be obtained after two or three refinements.

If we use the refined basis vectors (*i.e.* $(\mathbf{b}^1)_o^{o,fast}$, $(\mathbf{b}^2)_o^{o,slow}$, $(\mathbf{b}^3)_o^{o,slow}$) and their corresponding inverse column vectors) at $t = 0$ as constant basis vectors for $t > 0^+$, the new Λ_j^i at $t = 0$ is:

$$\Lambda_{j,o}^{i,o} = \begin{pmatrix} -1.27 \times 10^2 & -1.80 \times 10^{-1} & 5.40 \times 10^{-1} \\ -1.49 \times 10^{-2} & -1.67 \times 10^{-1} & 5.01 \times 10^{-1} \\ 7.59 \times 10^{-5} & 8.51 \times 10^{-4} & -2.55 \times 10^{-3} \end{pmatrix}. \quad (7.11)$$

Note that the off-diagonal terms are much smaller than before.

The elements of the new radical pointer using the fast refined basis vectors of mode #1 are:

$$Q_1(1) = 0.482 \times 10^{-1}, \quad Q_1(2) = 0.952, \quad Q_1(3) = 0.00. \quad (7.12)$$

This (refined) radical pointer is theoretically reliable, and it indicates that B alone is qualified to be a CSP radical, and that C must never be so identified. The elements of the new fast reaction pointer are:

$$P_1(1) = 0.906, \quad P_1(2) = 0.941 \times 10^{-1}, \quad P_1(3) = 3.58 \times 10^{-5}. \quad (7.13)$$

It is noted that $P_1(1)$ alone is not a small number, formally verifying that reaction #1 is the fast reaction of this fast mode.

We can fine tune the initial conditions using the refined basis vectors and apply the radical correction once more. We obtain:

$$\mathbf{y}(0^+) = [1.46 \times 10^{-4} \text{mole/cc}, 1.89 \times 10^{-6} \text{mole/cc}, 300.02^\circ \text{K}]^T. \quad (7.14)$$

This correction is mainly on B , the CSP radical. It is applied again here because (7.8) was computed using the unrefined trial basis vectors which were assumed to be good choices. In principle, the new initial conditions should be obtained by numerically integrating the original system of ODE's including $\mathbf{g}_o^{o,fast}$, and let the rapid transient do the adjustments naturally.

We can now compute the refined fast and slow components of \mathbf{g} at $t = 0^+$:

$$\begin{aligned} \mathbf{g}_o^{o,fast} &= \mathbf{Q}_o^o(1) \odot \mathbf{g} \\ &= [1.55 \times 10^{-5}, -8.09 \times 10^{-6}, -0.153]^T, \end{aligned} \quad (7.15a)$$

$$\begin{aligned} \mathbf{g}_o^{o,slow} &= (\mathbf{I} - \mathbf{Q}_o^o(1)) \odot \mathbf{g} \\ &= [-5.90 \times 10^{-6}, -1.57 \times 10^{-7}, -0.555]^T. \end{aligned} \quad (7.15b)$$

Equation (7.15a) must be interpreted with care—it should *not* be used to assess whether $\mathbf{g}_o^{o,fast}$ can be neglected in comparison with $\mathbf{g}_o^{o,slow}$. First of all, much cancellations occurred in the evaluation for $\mathbf{g}_o^{o,fast}$, thus *its accuracy is suspect*. Furthermore, the value of f_o^m used is the sum of $f_{o,homo}^m$ and $f_{o,part}^m$, but the $f_{o,homo}^m$ contribution is known to decay rapidly with time. Hence whenever the accuracy of directly evaluated $\mathbf{g}_o^{o,fast}$ is suspect, its large time order of magnitude estimate should be made with $f_o^m \rightarrow f_o^m \approx f_{o,part}^m$, even when numerically $|f_{o,homo}^m| \gg |f_{o,part}^m|$, as mentioned earlier. The large time asymptotic approximation to the (refined) particular solution is:

$$f_{o,part}^m \approx f_{o,\infty}^m + \sum_{n=1}^M \tau_{n,o}^{m,o}(M) \frac{df_{o,\infty}^n}{dt} + \dots, \quad m = 1, 2, \dots, M, \quad (7.16)$$

where

$$f_{o,\infty}^m \equiv \sum_{n=1}^M \tau_n^m(M) \frac{df_{\infty}^n}{dt}, \quad n = 1, 2, \dots, M, \quad (7.17)$$

and f_{∞}^n was previously defined in (6.31). In a computer code, $f_{o,\infty}^m$ can readily be evaluated—the time derivative in (7.16) and (7.17) can be computed

either exactly or approximately (*e.g.* using backward finite difference); no severe cancellation is expected. An estimate of the order of magnitude of the exhausted amplitude f_o^m is:

$$f_o^m \rightarrow f_{o,\infty}^m \approx O\left(\frac{\tau_o^o(M)}{\tau_o^o(M+1)}\right)f_\infty^m, \quad m = 1, 2, \dots, M, \quad (7.18)$$

where $\tau_o^o(M)$ and $\tau_o^o(M+1)$ are the time scales of the slowest fast mode and the fastest slow mode, respectively. $\tau_o^o(M)$ can be estimated by the largest reciprocal diagonal element of a Λ_n^m refined at least once. $\tau_o^o(M+1)$ can be similarly estimated, or it can be taken to be the integration step size selected by an integration routine (*e.g.* RKQC^[19]). Hence, we have:

$$\mathbf{g}_o^{o,fast} = \sum_{m=1}^M \mathbf{a}_m^o f_o^m \approx O\left[\sum_{m=1}^M \mathbf{a}_m^o \left(\frac{\tau_o^o(M)}{\tau_o^o(M+1)}\right) f_\infty^m\right]. \quad (7.19)$$

For the example, we have $\tau_o^o(1) \approx -0.8 \times 10^{-2}$, $\tau_o^o(2) \approx -0.6 \times 10^1$, and $f_\infty^1 \approx 6.63 \times 10^{-6}$. Hence, this rough estimate yields:

$$\mathbf{g}_o^{o,fast} \approx O([-1.6 \times 10^{-8}, \quad 0.8 \times 10^{-8}, \quad 1.6 \times 10^{-4}]^T). \quad (7.20)$$

If all of its components are considered small enough (see next section), we can declare the refined fast mode exhausted—in spite of (7.15a) which does not appear negligible at all—and neglect $\mathbf{g}_o^{o,fast}$ in comparison to $\mathbf{g}_o^{o,slow}$ to yield the desired simplified model.

For our example, the amplitude of the exhausted mode #1 is given by:

$$f_o^1 = \mathbf{b}_o^1 \odot \mathbf{g} = \sum_{r=1}^3 B_{o,r}^1 F^r. \quad (7.21)$$

where $B_{r,o}^1$ at $t = 0^+$ are, numerically:

$$B_{o,1}^1 = \mathbf{b}_o^1 \odot \mathbf{s}_1 = 1.00, \quad (7.22a)$$

$$B_{o,2}^1 = \mathbf{b}_o^1 \odot \mathbf{s}_2 = 0.975, \quad (7.22b)$$

$$B_{o,3}^1 = \mathbf{b}_o^1 \odot \mathbf{s}_3 = -1.92. \quad (7.22c)$$

When f_o^1 can be neglected from (7.21) in comparison to the magnitude of the largest positive or negative terms on the right hand side, we have:

$$\sum_{r=1}^3 B_{o,r}^1 F^r \approx 0, \quad (7.23a)$$

or

$$\sum_{r=1}^3 B_{o,r}^1 (F_+^r - F_-^r) \approx 0. \quad (7.23b)$$

The above two equations, while theoretically identical, provide two separate pieces of useful information. Equation (7.23a) relates the three values of F^r 's, while (7.23b) shows where the major cancellations occur. The major cancellations can easily be identified by the use of the *participation index*^[12]—which will show for this problem that $F_+^1 \approx F_-^1$.

Equation (7.23b) can be used to solve for B , which was identified by the radical pointer to be a CSP radical, thus eliminating the need for its ODE. Equation (7.23a) can be used to solve for F^1 , which was identified by the fast reaction pointer to be a fast reaction, in terms of F^2 and F^3 (see (7.26a) later). It should be obvious to the perceptive readers that if (7.23a) were used instead to solve for F^3 in terms of F^1 and F^2 , the results obtained would be both incorrect and misleading.

We can now compute for $\mathbf{g}_o^{o,slow}$ as follows:

$$\mathbf{g}_o^{o,slow} = \sum_{r=1}^R \mathbf{s}_{o,r}^{o,slow} F^r = \mathbf{s}_{o,1}^{o,slow} F^1 + \mathbf{s}_{o,2}^{o,slow} F^2 + \mathbf{s}_{o,2}^{o,slow} F^3. \quad (7.24)$$

At $t = 0^+$, we have:

$$\mathbf{s}_{o,1}^{o,slow} = [-0.916, -0.250 \times 10^{-1}, -0.863 \times 10^5]^T \times 10^{-1}, \quad (7.25a)$$

$$\mathbf{s}_{o,2}^{o,slow} = [0.859, 0.234 \times 10^{-1}, 0.809 \times 10^5]^T, \quad (7.25b)$$

$$\mathbf{s}_{o,3}^{o,slow} = [-2.67, -0.728 \times 10^{-1}, -2.52 \times 10^5]^T, \quad (7.25c)$$

and

$$F^1 = -0.975F^2 + 1.92F^3, \quad (7.26a)$$

$$F^2 = -6.76 \times 10^{-6} \text{ mole/cc-second}, \quad (7.26b)$$

$$F^3 = 2.61 \times 10^{-8} \text{ mole/cc-second}. \quad (7.26c)$$

Equation (7.26a) should be compared with (5.1), (5.5) and (5.10), the analytical results valid in the limit $k_1 \rightarrow \infty$ obtained earlier.

The above numerical CSP-derived $\mathbf{s}_{o,r}^{o,slow}$'s for the example problem should be compared with (5.3a), (5.3b) and (5.3c), the partial-equilibrium analytical

results derived earlier:

$$[\mathbf{s}_{o,1}^{o,slow}]_{k_1 \rightarrow \infty} = [0, 0, 0]^T, \quad (7.27a)$$

$$[\mathbf{s}_{o,2}^{o,slow}]_{k_1 \rightarrow \infty} = \left[\frac{K_1}{K_1 + 4A}, \frac{2A}{K_1 + 4A}, \Delta H_2 - \Delta H_1 \frac{K_1 + 2A}{K_1 + 4A} \right]^T, \quad (7.27b)$$

$$[\mathbf{s}_{o,3}^{o,slow}]_{k_1 \rightarrow \infty} = \left[\frac{-3K_1}{K_1 + 4A}, \frac{-6A}{K_1 + 4A}, \Delta H_3 - \Delta H_1 \frac{2(K_1 + A)}{K_1 + 4A} \right]^T. \quad (7.27c)$$

It is seen that for the given finite value of k_1 of interest, the error of the above analytical results is about 10%.

The relative importance of the two slow reactions (F^2 or F^3 , or both) in the evaluation of $\mathbf{g}_o^{o,slow}$ can be determined from (7.24) by the *importance index* introduced in Lam^[12].

7.4 The User-specified Accuracy Threshold

A perceptive reader would have noticed in (7.20) that the error introduced by the neglect of $\mathbf{g}_o^{o,fast}$ to B , the CSP radical, appears to be the largest. The accuracy of the CSP radicals can be improved by using the radical correction^[12]. It can easily be shown that when $\mathbf{g}_o^{o,fast}$ is neglected and the initial conditions properly adjusted, the solution generated automatically satisfies $f^m \approx 0$, $m = 1, 2, \dots, M$. These M equations of state can thus be used to eliminate the ODE's of the M CSP radicals. The theory of CSP uses the radical pointer to identify one or more special unknowns as CSP radicals for each of the M exhausted fast mode. Theoretically, the values of the CSP radicals may be solved for from the equations of state using the radical correction. Hence, when the radical correction is applied, the accuracy of the CSP radicals is controlled by the accuracy of the equations of state, and *not* by the accuracy of the ODE's. A detailed discussion of this subtle point is beyond the scope of this paper. It suffices to state here that the error committed in neglecting $\mathbf{g}_o^{o,fast}$ can be estimated in the conventional way for all other unknowns, but for the CSP radicals the correct estimate is a factor $\tau(M)/\tau(M+1)$ smaller than indicated when the radical correction is used.

With this caveat, the error of neglecting $\mathbf{g}_o^{o,fast}$ can be computationally assessed, and fast modes are declared exhausted only when their neglect introduces an error estimated to be below the *user-specified accuracy threshold*,

\mathbf{y}_{error} . The *exhaustion criterion*^[12] for the m -th fast mode is:

$$\left| \mathbf{a}_m^o f_{o,\infty}^m \tau(M) \right| \ll \mathbf{y}_{error}, \quad m = 1, 2, \dots, M, \quad (7.28)$$

which must be satisfied component by component.

7.5 Exact or Approximate Conservation Laws

In a time period of interest, some reaction modes may be so slow that they are dormant—*i. e.* they can be neglected.

Let the $(N - M)$ -dimensional slow subspace be spanned by \mathbf{a}_J and \mathbf{b}^J ($J = M + 1, \dots, N$), the slow trial slow set which complements the fast trial set of basis vectors. The fast subspace refinement process described earlier will also affect these slow basis vectors as the coupling between the fast and slow modes is weakened. However, the modes within the fast and slow refined subspaces remain mixed.

One may specifically look for dormant modes among the $N - M$ slow modes using any reasonable algorithm. If the amplitude f_o^K of the K -th (slow) mode satisfies the user-specified accuracy threshold component by component:

$$\left| \mathbf{a}_K^o f_o^K \tau(M + 1) \right| \ll \mathbf{y}_{error}, \quad (\text{selected } K\text{'s}), \quad (7.29)$$

then it is considered a *dormant mode*.

Exact dormant modes occur frequently in chemical kinetics problems—they are called conservation laws of atomic species. Generations of students have been taught to look for such conservation laws and to use them to advantage in analysis. From the CSP point of view, all exact conservation laws are simply special cases of dormant modes. However, not all dormant modes represent conservation laws.

Dormant modes play no significant role in CSP—inclusion of dormant modes does not cause stiffness. Identification of dormant modes is of interest primarily because they may suggest physically interesting concepts. CSP provides no special technique to find them; it is assumed that they can be somehow found and be identified using (7.29).

Unlike exhausted fast modes, the smallness of the dormant mode amplitude,

$$f_o^K = \mathbf{b}_o^K \odot \mathbf{g} \approx 0, \quad (\text{selected } K\text{'s}), \quad (7.30)$$

does *not* produce an equation of state. In the case of conservation law of atomic species, for example, $\mathbf{b}_o^K \odot \mathbf{g} = 0$ is an identity: zero equals to zero.

However, if \mathbf{b}_o^K for any mode, exhausted, active, or dormant, can be expressed in the form:

$$\mathbf{b}_o^K = \theta_{K,o} \frac{\partial \Theta_{K,o}}{\partial \mathbf{y}}, \quad (\text{selected } K\text{'s}), \quad (7.31)$$

where $\theta_{K,o}$ and $\Theta_{K,o}$ are scalar functions of \mathbf{y} , then mode $\#K$ indeed may represent a conservation law. When $\Theta_{K,o}$'s exist, Lam^[12] called them *eligible conserved scalars*, and such modes *holonomic modes*—the latter name is borrowed from classical mechanics^[20]. At the present time, there is no known numerical method to determine from a computed \mathbf{b}_o^K whether the mode is holonomic—except when it is a time-independent constant.

All exhausted fast modes are approximately holonomic, with $\theta_{m,o} = 1$ and $\Theta_{m,o} \approx f_o^m$. When the K -th mode is found to be dormant *and* holonomic, $\Theta_{K,o}$, an algebraic function of the unknowns, is a *conserved scalar*, and is commonly known as a Shvab-Zel'dovich variable or *coupling function*^[1] if it depends linearly on the unknowns. Unlike equations of state obtained from exhausted fast modes, conserved scalars always contain a free parameter to be determined by initial or boundary conditions. For addition discussions, see Lam^[12].

In our example, the slow mode $\#3$ is a dormant mode. Analytically, it can be shown that its basis vectors are:

$$\mathbf{a}_3^o = [0, 0, 1]^T, \quad (7.32a)$$

$$\mathbf{b}_o^3 \approx [\Delta H_1 - \Delta H_2, \Delta H_1 - 2\Delta H_2, 1](1 + O(k_3 K_3/k_1 K_1)), \quad (7.32b)$$

In other words, this is a dormant holonomic mode, with $\theta_{3,o} = 1$ and its conserved scalar given by:

$$\Theta_{3,o} = \mathbf{b}_o^3 \odot \mathbf{y} \approx (\Delta H_1 - \Delta H_2)A + (\Delta H_1 - 2\Delta H_2)B + C, \quad (7.33)$$

which can be interpreted as the total energy of the system. Actually, $\Theta_{3,o}$ is not an absolute constant but in fact evolves with time with time scale of $O(H/(k_3 K_3 \Delta H_3))$ which is measured in hundreds of seconds; it only *appears* to be a conserved scalar quantity in the time period of $O(1/(k_2 K_2))$ which is measured in seconds.

8 A Stiffness-Related Programming Issue

The *Associative Law of Addition* is one of the fundamental laws of algebra. It states that:

$$(x + y) + z = x + (y + z). \quad (8.1)$$

It is not generally appreciated that this law is not valid on a finite precision computer. One can easily confirm this observation by trying $x = 10^{15}$, $y = -10^{15}$, and $z = 1.23$ on any electronic calculator.

The numerical violation of the Associative Law of Addition is caused by loss of significant figures in subtracting large numbers or adding large and small numbers. In a chemical kinetics problem involving fast and slow reactions, the evaluation of the vector \mathbf{g} indeed involves subtraction and addition of large and small numbers. Hence one must pay special attention to the evaluation of \mathbf{g} inside a computer program.

Mathematically, \mathbf{g} can in principle be evaluated by any of the following three formally identical expressions:

$$\mathbf{g} := \sum_{r=1}^R \mathbf{s}_r F^r, \quad (8.2a)$$

$$\mathbf{g} := \sum_{r=1}^R \mathbf{s}_r F_+^r - \sum_{r=1}^R \mathbf{s}_r F_-^r, \quad (8.2b)$$

$$\mathbf{g} := \sum_{r=1}^M \mathbf{a}_m^o f_o^m + \mathbf{g}_o^{o,slow}, \quad (8.2c)$$

where f_o^m is given by (6.32):

$$f_o^m := \sum_{r=1}^R B_{o,r}^m F^r, \quad m = 1, 2, \dots, M, \quad (8.2d)$$

$$B_{o,r}^m \equiv \mathbf{b}_o^m \odot \mathbf{s}_r, \quad (8.2e)$$

and

$$\mathbf{g}_o^{o,slow}(M) := \sum_{r=1}^R \mathbf{s}_{o,r}^{o,slow} F^r, \quad (8.2f)$$

$$F^r := F_+^r - F_-^r, \quad r = 1, 2, \dots, R. \quad (8.2g)$$

For sufficiently stiff problems, all three evaluations (8.2a,b,c) are numerically

unreliable when the fast modes are near exhaustion. However, (8.2c) can be made reliable by applying CSP concepts as shall be shown presently.

The amount of cancellations in the evaluation of f_o^m and F^r can be measured by:

$$\delta f_o^m \equiv \frac{|f_o^m|}{\sum_{r=1}^R (|B_{o,r}^m F_+^r| + |B_{o,r}^m F_-^r|)}, \quad m = 1, 2, \dots, M, \quad (8.3a)$$

$$\delta F^r \equiv \frac{|F^r|}{|F_+^r| + |F_-^r|}, \quad r = 1, 2, \dots, R, \quad (8.3b)$$

which are both dimensionless numbers. Clearly, when either δf_o^m or δF^r is very small, the accuracy of f_o^m or F^r so evaluated is suspect. In our example, δf_o^1 and δF^1 are $O(k_2 K_2 / (k_1 K_1))$ after the rapid transient period, and—to dramatize the situation—if $k_2 K_2 / (k_1 K_1) = O(10^{-18})$ then even double-precision arithmetic would be inadequate if (8.2d) is used for all time.

CSP uses (8.2d), but only when no major cancellations occur. As the fast modes become exhausted, the f_o^m 's as evaluated by (8.2d) decrease in amplitude, and the magnitude of the δf_o^m 's can become small as a consequence of cancellations. When this occurs, the f_o^m 's so evaluated are suspect, and (6.34) for f_o^m is available to provide the leading order asymptotic approximation for f_o^m for “large” time. Since $f_{o,\infty}^m$ can be computed at any moment in time, the needed time derivative can be approximately evaluated numerically. CSP understands that the contribution of $\mathbf{g}_o^{o,fast}$ to \mathbf{g} in (8.2c) is at best a minor correction whenever substantial cancellations occur in the direct evaluation of f_o^m . Once (8.2d) is abandoned in favor of (6.34), the resulting system of ODE is no longer stiff.

Similarly, in the evaluation of $\mathbf{g}_o^{o,slow}$, all F^r 's are involved. The values of the exhausted fast δF^r 's evaluated by (8.3b) will be small, and their values are suspect. When this occurs, the theory of CSP provides M fast reaction pointers which can be used to identify M fast reactions. The net reaction rates of these M fast reactions can be solved for from (8.2d) in terms of the rest of the net reaction rates. Hence, the use of potentially inaccurate directly evaluated fast F^r 's are avoided. In our example, the value of F^1 after the rapid transient periods should *not* be computed from $F^1 = k_1(A^2 - K_1 B)$, the theoretically exact expression, but should be evaluated instead from the CSP-derived approximation

$$F^1 \approx -(B_{o,2}^1 F^2 + B_{o,3}^1 F^3) / B_{o,1}^1,$$

as was done in (7.26a).

The computation of the effective stoichiometric vector of the r -th reaction, $\mathbf{s}_{o,r}^{o,slow}$, involves $[\mathbf{I} - \mathbf{Q}_o^o(M)]$, and its dimensionless diagonal elements are obtained by a subtraction process which may lose significant figures. The following artifice has been found successful in eliminating specious errors in the direct evaluation of $\mathbf{s}_{o,r}^{o,slow}$: Whenever any diagonal element of this matrix falls below an appropriate threshold (*e.g.* 10^{-4} for four significant figures calculations), it is replaced by a zero.

9 Discussion

The traditional role of a theoretician has always been to simplify seemingly complex problems into their bare elements. Physical insights and intuition are at the root of this process, helped considerably by the methodology of asymptotics when a small parameter is available and can be identified. As a practical matter, asymptotic solutions usually consists of very few terms—the so-called leading order solution is usually all that could be expected from a massively complex reaction system because of the massive amount of algebra involved. In §5, the analytical results obtained are formally correct only in the limit of $k_2K_2/(k_1K_1) \rightarrow 0$, $k_3/(k_1K_2^2) \rightarrow 0$. No assurance of accuracy is provided when applied to problems with $k_2K_2/(k_1K_1) \approx 0.1$.

The theory of CSP welcomes physical insights and intuition, but is not dependent on them. It uses the eigen-values of \mathbf{J} to order the trial modes, and provides a refinement procedure to improve the decoupling of the trial fast and slow subspaces. From the viewpoint of CSP, users of conventional asymptotics simply make educated guesses at the fast basis vectors. The special procedures to apply the partial-equilibrium and quasi-steady approximations are just variants of the CSP refinement procedure. In the absence of insights and intuition, the conventional methodology cannot get started at all, but CSP can proceed routinely—the eigen-vectors of \mathbf{J} can always be used as the trial set^[21]. Most importantly, the refinement procedure can be recursively applied, allowing “higher-order” corrections to be included. The CSP user has the responsibility to specify the threshold of tolerable error \mathbf{y}_{error} , and CSP makes sure that the error introduced by the neglect of $\mathbf{g}_o^{o,fast}$ is below threshold. Different simplified models will be generated for different user-specified \mathbf{y}_{error} ; a lenient threshold will yield a simpler model than a

more stringent one.

The method of CSP is a significant advance over the method of *matched asymptotic expansions*^[22] for boundary-layer type problems. First of all, it can be applied without the need of non-dimensionalizations and identification of small parameters. In essence, CSP exploits the disparity of time scales between the exhausted fast modes and the currently active slower modes, and the small dimensionless parameter being exploited is ϵ_M as defined in (6.13). The simplified model is constructed by an iterative procedure; each iteration improves the accuracy of the model by $O(\epsilon_M)$. It is important to note that the time derivative term in (6.17a) and (6.18b) *must* be included in the refinement procedures if accuracy beyond “leading order” is desired. The superiority of the CSP approach to the conventional methodology, when applied analytically to simple problems, is clearly demonstrated in §7.3. For massively complex problems, the CSP approach has no peers.

The CSP-derived simplified model remains a system of N ODE’s which is accompanied by a set of M equations of state. Theoretically, the solutions of the simplified model automatically satisfy the M equations of state for $t \geq 0^+$ if the initial conditions satisfy them at $t = 0^+$ —assuming that exhausted modes did not become active again. CSP uses all N ODE’s to march forward in time for all N components of \mathbf{y} , and uses the M equations of state only to apply the radical corrections to prevent the “drifting” of the exhausted fast mode amplitudes.

For each fast mode, CSP provides a radical pointer which identifies the CSP radicals—species which can be accurately solved for from the equation of state. It is extremely important to note that one may *not* arbitrarily select any M species to be solved for from the M algebraic equations. In the conventional approach, the algebraic difficulty of solving for the radicals frequently forces additional *ad hoc* approximations. Again, CSP deals with this obstacle using iteration; the programmable radical correction procedure can be recursively applied to solve for the CSP radicals. In fact, the radical correction should be applied after every integration step to counter the usually larger estimated error of neglecting $\mathbf{g}_o^{o,fast}$ for the CSP radicals. The refined basis vectors from the previous time step can always be used as trial basis vectors for the current time step. In addition, CSP provides a fast reaction pointer for each fast mode, and an algorithm to accurately evaluate the amplitudes of exhausted fast reactions.

The CSP method has no difficulty in identifying dormant modes, but does

not provide a method for finding the so-called eligible conserved scalars—except for the simplest case when the row basis vector of the mode in question is found to be time-independent. At the present time, CSP does not take advantage of any conservation laws which may be available.

Trevino *et. al.*^{[23],[24]} successfully studied ignition phenomena with the assistance of CSP data. CSP data can also be used to assess sensitivity of solutions with respect to the input rate coefficients^[25], and offers an attractive alternative to the conventional method of sensitivity analysis^[26].

References

- [1] Williams, F.A., *Combustion Theory, The fundamental theory of chemically reacting systems*, 2nd edition, Benjamin/Cummings Publ. Co., Menlo Park, Ca., 1985.
- [2] Benson, S. W., “The induction period in chain reactions,” *J. Chem. Phys.* **20** p. 1605, 1952.
- [3] Bowen, J. R., Acrivos, A. and Oppenheim, A. K., “Singular perturbation refinement to quasi-steady state Approximation in chemical kinetics,” *Chem. Eng. Sci.* **18**, p. 177, 1963.
- [4] Peters, N., “Numerical simulation of combustion phenomena,” Glowinski, R., Larrourou, B and Temem, R., Eds., *Lecture Notes in Physics 241*, p. 90-109, Springer-Verlag, New York, 1985.
- [5] Peters, N. and Williams F. A., “The asymptotic structure of stoichiometric methane-air Flames”, *Combust. Flame*, **68**, pp. 185-207, 1987.
- [6] Aiken, R. C., *Stiff Computation*, Oxford University Press, 1985.
- [7] Lam, S. H., “Singular perturbation for stiff equations using numerical methods”, in *Recent Advances in the Aerospace Sciences*, Corrado Casci, Ed., Plenum Press, New York and London, pp.3-20, 1985.
- [8] Lam, S. H. and Goussis, D. A., “Basic theory and demonstrations of computational singular perturbation for stiff equations”, *12th IMACS World Congress on Scientific Computation*, Paris, France, July 18-22, 1988; *IMACS Transactions on Scientific Computing '88, Numerical and*

Applied Mathematics, C. Brezinski, Ed., J. C. Baltzer A. G. Scientific Publishing Co., pp. 487-492, November, 1988.

- [9] Lam, S. H. and Goussis, D. A., "Understanding complex chemical kinetics with computational singular perturbation", *Proceedings of the 22nd International Symposium on Combustion*, pp. 931-941, University of Washington, Seattle, Washington, August 14-19, 1988.
- [10] Lam, S. H. and Goussis, D. A., in *Reduced Kinetic Mechanisms and Asymptotic Approximations for Methane-Air Flames*, Chapter 10: "Conventional asymptotics and computational singular perturbation for simplified kinetics modeling," *Lecture Notes in Physics 384*, M. Smooke, Ed., Springer-Verlag, 1991.
- [11] Goussis, D. A. and Lam, S. H., "A study of homogeneous methanol oxidation kinetics using CSP," *Proceedings of the 24th (International) Symposium on Combustion*, University of Sydney, Sydney, Australia, July 5-10, pp. 113-120, 1992.
- [12] Lam, S. H., "Using CSP to understand complex chemical kinetics," *Combust. Sci. and Tech.*, **89**, pp. 375-404, 1993.
- [13] Lam, S. H., Goussis, D. A., and Konopka, D., "Time-Resolved Simplified Chemical Kinetics Modeling Using Computational Singular Perturbation", AIAA 27th Aerospace Sciences Meeting, (AIAA 89-0575), Reno, Nevada, January 9-12, 1989.
- [14] Goussis, D. A., S. H. Lam and Gnoffo, P.A., "Reduced and Simplified Chemical Kinetics for Air Dissociation Using Computational Singular Perturbation", AIAA 28th Aerospace Sciences Meeting, (AIAA 90-0644), Reno, Nevada, January 9-12, 1990.
- [15] Lam, S. H. and Goussis, D. A., "Identification of Intermediaries and Fast Reactions in Complex Chemical Mechanisms Using CSP," presented at the 4th International Conf. Numerical Combustion, St. Petersburg Beach, Florida, December 2-4, 1991.
- [16] Trevino, C. and Williams, F.A., "An asymptotic analysis of the structure and extinction of methane-air diffusion flame," *11th International*

Colloquium on Dynamics of Explosions and Reactive Systems, Warsaw, Poland, 1987.

- [17] Noble, B. and Daniel, J. W., *Applied Linear Algebra*, Third Edition, Prentice Hall, Inc. 1988.
- [18] Carnahan, B, Luther, H. A. and Wilkes, J. O., *Applied Numerical Methods*, John Wiley and Sons, New York, 1969.
- [19] Press, W. H., Flannery, B. P., Teukolsky, S.A. and Vetterling, W. T., *Numerical Recipes, The Art of Scientific Computing*, Cambridge University Press, Cambridge, 1986.
- [20] Sommerfeld, A., *Mechanics*, Academic Press, N. Y., 1952.
- [21] Maas, U. and Pope, S. B., "Simplifying Chemical Kinetics: Intrinsic Low-Dimensional Manifolds in Composition Space," *Combust. Flame*, **88**, pp. 239-264, 1992.
- [22] Van Dyke, M., *Perturbation Methods in Fluid Mechanics*, Parabolic Press, Stanford, Ca., 1975.
- [23] Trevino, C. and Solorio, F., "Asymptotic Analysis of the High Temperature Ignition of $CO/H_2/O_2$ Mixtures," *Combust. Flame*, **86**, pp. 285-295, 1991.
- [24] Trevino, C., "Ignition Phenomena in $H_2 - O_2$ Mixtures," in *Dynamics of Deflagration and Reactive Systems and Flames*, A. L. Kuhl *et. al.*, Ed., *Progress in Aeronautics and Astronautics*, AIAA, 1991.
- [25] Lam, S. H. and Goussis, D. A., "Sensitivity analysis of complex simulations using basis vectors," Proceedings of the 13th IMACS World Congress on Computation and Applied Mathematics, R. Vichnevetsky, J. .J. H. Miller, Eds., **3**, p. 1292, Criterion Press, Dublin, Ireland, 1991.
- [26] Yetter, R. A., Dryer, F. L. and Rabitz, H., "A Comprehensive Reaction Mechanism of Carbon Monoxide/Hydrogen/Oxygen Kinetics," *Combust. Sci. Tech.*, **79**, pp. 97-128, 1991.