

Using CSP to Understand Complex Chemical Kinetics*

S. H. Lam

Department of Mechanical and Aerospace Engineering
Princeton University

September 1992

Abstract

The conventional methods of simplified kinetics modeling through the use of partial-equilibrium and quasi-steady approximations are reviewed and critiqued. The method of *computational singular perturbation* (CSP) is then presented with special emphasis on the interpretation of CSP data to obtain physical insights on massively complex reaction systems. A simple example is used to demonstrate how CSP deals with complex chemical kinetics problems without the benefits of intuition and experience.

1 Introduction

An ideal scenario in the (near) future for the study of chemical kinetics would be that a comprehensive, reliable and up-to-date database of validated reaction rates is readily available to any researcher interested in any reasonable reaction system of interest. Using a suitable stiff solver, one can routinely compute for the numerical solution of a massively complex chemical kinetics system. However, the extraction of *physical insights* about the

*This work is supported by NASA Langley's Aerothermodynamics Branch, Space Systems Division.

reaction system from the massive printouts is a different matter, and remains a formidable task. Generally speaking, answers to the following questions are highly valued:

1. How can a *reduced reaction system*, involving a much smaller set of chemical species and elementary reactions, be derived which can approximate the original massive reaction system both qualitatively and quantitatively?
2. How can the rate-controlling elementary reactions be identified? What rate constants must be known accurately? What rate constants need not be known accurately provided that they are “sufficiently large”?
3. Which elementary reactions can be considered fast reactions in the time interval of interest so that appropriate approximations can be applied to obtain useful approximate algebraic relations between the species?
4. Which chemical species can be considered as “intermediaries” or “radicals?” in the time interval of interest so that their concentrations can be obtained from these approximate algebraic relations?

In the present paper, we shall call approximate algebraic relations between the species “*equations of state*.” In addition, we shall use the word “radical” to connote a special meaning in the CSP context instead of its normal chemical structure context. An operational definition for a CSP radical based on its mathematical role in the CSP theory will be given later. In most cases, a CSP radical will be found to be a chemical radical, and *vice versa*, but not always. To avoid confusion, we shall not use the term “intermediaries” altogether.

Item #1 is important pragmatically because computational cost is a strong function of the number of chemical species. Questions in item #2 needs no justification; one can claim insights on the reaction system only if these questions can be satisfactorily answered. The conventional methodology deals with items #3 and #4 first, mainly by *guessing* based on experience and intuition. Some talented investigators simply know which reactions are fast and which chemical species are radicals under what conditions for certain reaction systems (Peters, 1985, 1991; Peters and Williams, 1987; Chelliah and Williams, 1990; Bilger *et. al.* 1990, 1991). For the less gifted, data from

numerical solutions of the full kinetics equations can be used to provide hints in the identifications. Once the fast reactions and the radicals are correctly identified, standard techniques are available to obtain answers to questions in items #1 and #2.

The method of *computational singular perturbation* (CSP) (Lam, 1985; Lam *et. al.*, 1988a, 1988b, 1989, 1991a, 1991b; Goussis *et. al.*, 1990, 1992) provides a programmable mathematical algorithm to proceed routinely without the benefits of experience and intuition. It can be used to verify the validity of simplified models derived by *ad hoc* methods, and it can be used to deal with massively complex problems beyond the reach of such methods. Physical understanding of the reaction system under investigation can easily be extracted from the CSP data generated.

2 Statement of the Problem

Consider a reaction system of N unknowns¹ denoted by the column *state vector* $\mathbf{y} = [y^1, y^2, \dots, y^N]^T$. The governing system of ODE is:

$$\frac{d\mathbf{y}}{dt} = \mathbf{g}(\mathbf{y}) \quad (1)$$

where

$$\mathbf{g}(\mathbf{y}) \equiv \sum_{r=1}^R \mathbf{s}_r F^r(\mathbf{y}), \quad (2)$$

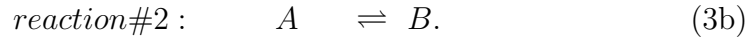
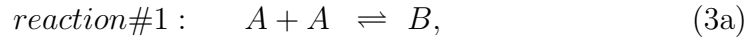
R is the number of elementary reactions being included in the reaction system, \mathbf{s}_r and $F^r(\mathbf{y})$ are the *stoichiometric vector* and the *reaction rate* of the r -th elementary reactions, respectively. The N -dimensional column vector \mathbf{g} is the overall reaction rate vector, and can be interpreted as the “velocity vector” of \mathbf{y} in the N -dimensional y -space. For a massively complex problem, N and R can be large numbers, and the accuracy or reliability of the available rate constants is usually less than ideal. Assuming that adequate computing power is available, the computation of the numerical solution of (1) is not an issue. The challenge is to obtain answers to questions in items #1 and #2 in §1.

¹The N -dimensional column vector may include temperature, total density, etc. in addition to chemical species

As previously indicated, the conventional method relies completely on the skill of the investigator to identify the fast reactions and the radicals. Once the identifications are made, the subsequent theoretical development proceeds relatively routinely. The interested reader is referred to Appendix B of Williams (1985) for a detailed exposition of the conventional methodology.

2.1 The Conventional Methodology by Example

We shall use a simple hypothetical reaction system to demonstrate the conventional methodology. 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 two elementary reactions:



The stoichiometric vectors and the reaction rates are:

$$\mathbf{s}_1 = [-2, \quad 1, \quad \Delta H_1]^T, \quad F^1 = k_1(A^2 - K_1 B), \quad (4a)$$

$$\mathbf{s}_2 = [-1, \quad 1, \quad \Delta H_2]^T, \quad F^2 = k_2(A - K_2 B). \quad (4b)$$

where the reaction rate coefficients k_1, k_2 and the equilibrium constants K_1, K_2 are known and for the sake of simplicity their dependence on C is assumed negligible. The heat of reaction (in the proper unit) for the two reactions are denoted by ΔH_1 and ΔH_2 , respectively.

The system of ODE is:

$$\frac{d\mathbf{y}}{dt} = \mathbf{s}_1 F^1 + \mathbf{s}_2 F^2 \quad (5)$$

which can be written out as follows:

$$\frac{dA}{dt} = -2F^1 - F^2, \quad (6a)$$

$$\frac{dB}{dt} = F^1 + F^2, \quad (6b)$$

$$\frac{dC}{dt} = \Delta H_1 F^1 + \Delta H_2 F^2. \quad (6c)$$

The initial conditions, $A(0)$, $B(0)$ and $C(0)$, are given. It can easily be shown that as $t \rightarrow \infty$, $A \rightarrow K_1/K_2$ and $B \rightarrow K_1/K_2^2$. Normalizing the unknowns intelligently, we have $\mathbf{y}_* = [A_*, B_*, C_*]^T$:

$$A_* = AK_2/K_1, \quad (7a)$$

$$B_* = BK_2^2/K_1, \quad (7b)$$

$$C_* = CK_2/(\Delta H_2 K_1). \quad (7c)$$

The system of ODE for A_* , B_* and C_* is:

$$\frac{dA_*}{dt} = -\frac{2}{K_2} K_1 k_1 (A_*^2 - B_*) - k_2 (A_* - B_*), \quad (8a)$$

$$\frac{dB_*}{dt} = K_1 k_1 (A_*^2 - B_*) + K_2 k_2 (A_* - B_*), \quad (8b)$$

$$\frac{dC_*}{dt} = \frac{\Delta H_1}{K_2 \Delta H_2} K_1 k_1 (A_*^2 - B_*) + k_2 (A_* - B_*). \quad (8c)$$

The initial conditions are:

$$A_*(0) = A(0)K_2/K_1, \quad (9a)$$

$$B_*(0) = B(0)K_2^2/K_1, \quad (9b)$$

$$C_*(0) = C(0)K_2/(\Delta H_2 K_1). \quad (9c)$$

As $t \rightarrow \infty$, we have $A_* \rightarrow 1$ and $B_* \rightarrow 1$.

Because this reaction system is a hypothetical one, the concept of conservation of atomic species cannot be applied—we do not know what atomic species are involved. However, it can easily be verified mathematically that:

$$(\Delta H_1 - \Delta H_2)A + (\Delta H_1 - 2\Delta H_2)B + C = \text{Constant} \quad (10)$$

is an exact “integral of motion” valid for all time, and can be physically interpreted as a statement of conservation of total energy.

For this simple example, if the “speeds” of the two reactions are competitive, there is no simplification available and little or no general statement can be made. However, when one reaction is much faster—in some sense—than the other, then certain mathematical simplifications are available and certain useful general statements can be made. Intuitively, we expect a rapid transient period in which the fast reaction would *dominate*, followed by a slower

evolutionary period in which the slow reaction would be *controlling*. In general, we expect to be able to neglect some of the slower reactions in the rapid transient period. However, it is well known that the handling of the nearly exhausted fast reactions in the slow evolutionary period requires considerable care—it suffices to say here that they cannot be summarily neglected.

2.2 The Case of reaction #1 being faster ($\frac{K_2 k_2}{K_1 k_1} \ll 1$)

In the rapid transient period, the fast reaction #1 dominates. Hence, the slow reaction #2 can be neglected to yield the simplified model:

$$\frac{dA_*}{dt} \approx -\frac{2}{K_2} K_1 k_1 (A_*^2 - B_*), \quad (11a)$$

$$\frac{dB_*}{dt} \approx K_1 k_1 (A_*^2 - B_*), \quad (11b)$$

$$\frac{dC_*}{dt} \approx \frac{\Delta H_1}{K_2 \Delta H_2} K_1 k_1 (A_*^2 - B_*). \quad (11c)$$

with (9a), (9b) and (9c) serving as the initial conditions. It is easy to show that in addition to (10), the following

$$K_2 A_* + 2B_* \approx \frac{K_2^2}{K_1} (A(0) + 2B(0)) \quad (12)$$

is approximately “conserved” (remains constant) during this period.

At the end of this period, the fast reaction #1 will have become “exhausted,” and A_* and B_* will be approximately related by $B_* \approx A_*^2$. In the slow evolutionary period which follows, an approximate equation of state is obtained by setting F^1 to zero. This is called the *partial-equilibrium* approximation and a special follow-up procedure is recommended. First, one of the participants of reaction #1 is somehow declared a radical. If B_* is so declared, (8b) is used to eliminate the contribution from reaction #1 (*i.e.* $K_1 k_1 (A_*^2 - B_*)$) from the rest of the equations. Then, the approximate equation of state is differentiated with respect to time and used to eliminate dB_*/dt . The following simplified model is obtained:

$$B_* \approx A_*^2, \quad (13a)$$

$$\frac{dA_*}{dt} \approx \frac{K_2 k_2}{K_2 + 4A_*} (A_* - B_*), \quad (13b)$$

$$\frac{dC_*}{dt} \approx [1 - \frac{\Delta H_1}{\Delta H_2} (\frac{K_2 + 2A_*}{K_2 + 4A_*})] k_2 (A_* - B_*). \quad (13c)$$

It can easily be verified that the same answer is obtained (in contrast to results in §2.3.2 later) if A_* is chosen as a radical. It is important to note that (13b) and (13c) are *not* obtained by substituting (13a) into (8a) and (8c). In fact, (13a) must *never* be used in (8a), (8b) or (8c). The initial conditions can be found from $B_* \approx A_*^2$ and (12):

$$A_*(0+) \approx \frac{K_2}{4} [\sqrt{1 + \frac{8(A(0) + 2B(0))}{K_1}} - 1] \quad (14)$$

with $B_*(0+)$ and $C_*(0+)$ given by (13a) and (10). Note that k_1 does not appear at all in (13a), (13b) or (13c), hence its value must be unimportant provided that it is “sufficiently large.” In this time period, (13a), (13b) and (13c) together behave as a “one-step” reaction with an *effective stoichiometric vector* and an *effective reaction rate*, and generate approximate solutions which satisfy (10) exactly.

2.3 The Case of reaction #2 being faster ($\frac{K_1 k_1}{K_2 k_2} \ll 1$)

In the rapid transient period, reaction #1 is neglected to yield the simplified model:

$$\frac{dA_*}{dt} \approx -k_2 (A_* - B_*), \quad (15a)$$

$$\frac{dB_*}{dt} \approx K_2 k_2 (A_* - B_*), \quad (15b)$$

$$\frac{dC_*}{dt} \approx k_2 (A_* - B_*). \quad (15c)$$

again with (9a), (9b) and (9c) serving as the initial conditions. It is easy to show that

$$K_2 A_* + B_* \approx \frac{K_2^2}{K_1} (A(0) + B(0)) \quad (16)$$

is approximately conserved (remains constant) during this period.

At the end of this period, A_* and B_* will be approximately related by $B_* \approx A_*$. For this case, we shall present both the partial-equilibrium approximation and the quasi-steady approximation in the slow evolutionary period. We shall see that they produce different answers.

2.3.1 The Partial-Equilibrium Approximation

In the slow evolutionary period, we can proceed as before by assuming elementary reaction #2 to be in partial-equilibrium. Using the same recommended procedure, we obtain the following simplified model:

$$B_* \approx A_*, \quad (17a)$$

$$\frac{dA_*}{dt} \approx \frac{K_1 k_1}{1 + K_2} (B_* - A_*^2), \quad (17b)$$

$$\frac{dC_*}{dt} \approx \left[\frac{K_2 + 2}{K_2 + 1} - \frac{\Delta H_1}{\Delta H_2} \right] \frac{K_1 k_1}{K_2} (B_* - A_*^2). \quad (17c)$$

with initial conditions:

$$A_*(0+) \approx \frac{K_2^2 (A(0) + B(0))}{K_1 (1 + K_2)} \quad (18)$$

with $B_*(0+)$ and $C_*(0+)$ given by (17a) and (10). It is emphasized again that (17b) and (17c) are *not* obtained by substituting (17a) into (8b) and (8c). It can be shown that the partial-equilibrium approximation above can be formally justified when $K_1 k_1 \ll K_2 k_2$ and $A_*(0) = O(1)$, with no restriction on the magnitude of K_2 , and the resulting approximate solutions satisfy (10) exactly.

2.3.2 The Quasi-Steady Approximation

Alternatively, we can declare either the species B_* or A_* to be a radical in this period. The *quasi-steady approximation* neglects the time derivative term from the ODE of the radical to yield an approximate equation of state. If B_* is so declared, the approximate equation of state obtained from (8b) is used to simplify (8a) and (8c). The resulting set of equations (without (8b)) is then augmented by the approximate equation of state to yield the simplified model. We obtain:

$$B_* \approx A_*, \quad (19a)$$

$$\frac{dA_*}{dt} \approx \frac{K_1 k_1}{K_2} (B_* - A_*^2), \quad (19b)$$

$$\frac{dC_*}{dt} \approx \left[1 - \frac{\Delta H_1}{\Delta H_2} \right] \frac{K_1 k_1}{K_2} (B_* - A_*^2). \quad (19c)$$

with initial condition (18). Note that (19b) and (19c) agree with (17b) and (17c) and satisfy (10) *only* when $K_2 \gg 1$. If $K_2 \ll 1$, then A_* should have been declared a radical instead, and a different result valid only in this limit is obtained. If K_2 is of $O(1)$, then the quasi-steady approximation is simply incorrect—while the partial-equilibrium approximation is valid for all K_2 .

Again, the value of k_2 does not appear in the simplified model in the slow evolutionary period. The inference is that its value is unimportant provided that it is sufficiently large.

2.4 Comments on Conventional Asymptotics

The above example prompts the following questions and observations:

1. Exact algebraic “conservation laws” such as conservation of atomic species and conservation of energy frequently exist in chemical kinetics problems. In the above example, (10) is such a conservation law. Can such conservation laws, when they exist, be mathematically identified and derived?
2. In the rapid transient period, additional temporary approximate conservation laws may exist. In the above example, (12) and (16) are temporary approximate conservation laws valid for the respective cases. Can such temporary approximate conservation laws, when they exist, also be mathematically derived?
3. When a fast reaction exhausts itself, its net reaction rate is *not* negligible in general from the governing equations. For example, when reaction #1 is the faster reaction, the exhausted F^1 is not zero in the slow evolutionary period but is, to “leading order,” given by:

$$K_1 k_1 (A_*^2 - B_*) \rightarrow -\frac{(K_2 + 2A_*)}{(K_2 + 4A_*)} K_2 k_2 (A_* - B_*), \quad (20a)$$

or

$$F^1 \rightarrow -\frac{K_1 + 2A}{K_1 + 4A} F^2, \quad (20b)$$

which can be derived by comparing (8a) with (13b). In other words, when a fast elementary reaction is exhausted, its *net* reaction rate may be competitive with the slower reactions.

4. In any slow evolutionary period, the temporary approximate conservation laws valid in the previous periods are replaced by approximate equations of state obtained by applying the partial-equilibrium or the quasi-steady approximation. In our example, (13a), (17a) and (19a) are such equations of state for the respective cases. Can these equations of state be mathematically derived? How is (13a) related to (20b)?
5. In a slow evolutionary period, the values of the exhausted fast reaction rate constants are not important provided that they are fast enough. In our example, the fastest rate coefficient (either k_1 or k_2) does not appear in the simplified models when $M = 1$. From the computation point of view, their absence means the simplified model is no longer “stiff” (Aiken, 1985).
6. To get started on the conventional method, one must somehow decide that certain reactions are faster than others based on some vague, intuitive judgment. Unlike linear problems, the “speed” of a “reaction mode” in a non-linear problem can depend strongly on the state of the system. For example, even if $K_2k_2 \gg K_1k_1$, reaction #1 can still dominate reaction #2 initially provided that $A_*(0)$ is sufficiently large. Since we are clearly interested in non-linear problems, a mathematical criterion for the identification of fast reactions is very much needed.
7. Even after the fast reactions have somehow been identified, there is still the question on whether the partial-equilibrium or the quasi-steady approximation should be used to generate the approximate equations of state, and which species—radicals—can be eliminated by using these algebraic equations.
8. In Williams’ book (1985), a radical is described as a chemical species which is “neither initial reactants nor principal products,” a description too vague to be useful. In §2.3, we saw that whether A_* or B_* could justifiably be declared a radical depends on the magnitude of K_2 , and that when $K_2 = O(1)$ the quasi-steady approximation is simply wrong.
9. We shall adopt a pragmatic but precise definition for a radical in the CSP context: *a chemical species is a CSP-radical if its ODE can be replaced by an approximate equation of state.* For the moment, we

shall leave the questions of how to identify the radical(s) and how to solve for them from CSP-derived equations of state until the concepts of “radical pointer” and “radical correction” are introduced later.

10. We emphasize here again that the approximate equations of state obtained from the partial-equilibrium approximation must *never* be used directly in the original governing equations—the so-called recommended procedure must be followed. However, the approximate equations of state obtained from the quasi-steady approximation may be used more freely, but the validity of the results is more restricted. In fact, the quasi-steady approximation can be viewed as an additional *ad hoc* restriction imposed on the results of an appropriately applied partial-equilibrium approximation.
11. Conventional asymptotics assumes that the investigator is not only capable of non-dimensionalizing the variables intelligently, but also competent to take advantage of the presence of large or small parameter(s) in the resulting formulation. Can a theoretical structure be developed such that intelligent non-dimensionalization is not required?
12. In general, the conventional “magnitude” of a dimensional vector \mathbf{y} computed using the standard inner product may not make physical sense. In the above example, $\|\mathbf{y}\|_2 = (A^2 + B^2 + C^2)^{1/2}$ makes no physical sense at all. In the language of mathematics, the “norm” of the vector space of \mathbf{y} needs to be defined. Hence if intelligent non-dimensionalization is not assumed, the concept of a “good approximation” needs to be explicitly clarified.

3 The Goal and the Idea of CSP

The conventional method is only viable for relatively simple problems for which adequate amount of experience and intuition have been accumulated, and that the algebra involved is manageable. For massively complex problems with large values of N and R , a better method is clearly needed.

3.1 The Goal

The goal of CSP is to computationally derive time-resolved simplified models for massively complex chemical systems assuming that the appropriate reaction rate database is available. The *novel* feature of CSP is that the power of the modern computer is exploited not only to construct the numerical solutions, but also *to derive the approximate equations*.

The output of CSP is a set of numbers called the CSP data. An essential attribute of CSP data is that it must be easy to interpret. Answers to questions such as:

- how many approximate equations of state are available and how can they be derived?
- which species can be considered radicals—in the CSP context—so that they can be computed from these equations of state?
- which elementary reactions are controlling the reaction rate of the system? and
- what is the minimum reduced reaction system that will generate an approximate solution with a user-specified accuracy?

and others will be provided by the CSP data.

3.2 The Idea

The physical problem is completely specified by $\mathbf{g}(\mathbf{y})$, a non-linear function of \mathbf{y} obtained by summing all the physical processes which contribute to the time rate of change of \mathbf{y} . The question is: is there a better representation than the physical representation?

A representation is said to be a *physical representation* when the theoretician formulating the problem can explain and interpret each term physically. Equation (2) is such a representation because as written each term represents the contribution of each of the R reactions included in the full model. Must we always write \mathbf{g} in this particular form? Since \mathbf{g} is a N -dimensional vector, it can always be expressed in terms of a set of arbitrarily chosen N linearly

independent column *basis vectors*, \mathbf{a}_i , $i = 1, 2, \dots, N$. When some other set of basis vectors is used, \mathbf{g} has an *alternative representation*:

$$\mathbf{g} = \sum_{i=1}^N \mathbf{a}_i f^i \quad (21)$$

where f^i , called the *amplitude* of the i -th *mode*, is given by:

$$f^i(\mathbf{y}) \equiv \mathbf{b}^i \odot \mathbf{g} = \sum_{r=1}^R B_r^i F^r, \quad i = 1, 2, \dots, N, \quad (22)$$

where

$$B_r^i \equiv \mathbf{b}^i \odot \mathbf{s}_r, \quad i = 1, 2, \dots, N, \quad r = 1, 2, \dots, R, \quad (23)$$

and \odot is the dot product operator of the N -dimensional vector space. When properly normalized, f^i can also be considered as a “progress variable” for the i -th mode. The set of N row vectors \mathbf{b}^i are the inverses of \mathbf{a}_i ; together they satisfy the following orthonormal condition:

$$\mathbf{b}^i \odot \mathbf{a}_j = \delta_j^i, \quad i, j = 1, 2, \dots, N. \quad (24)$$

Note that once the set \mathbf{a}_i is chosen, the associated set \mathbf{b}^i is straightforwardly computed. Note also that B_r^i is not necessarily dimensionless unless deliberately made so.

In our example, we have $N = 3$. The physical representation chooses (by default) the following column basis vectors:

$$\mathbf{a}_1 = \mathbf{s}_1, \quad (25a)$$

$$\mathbf{a}_2 = \mathbf{s}_2, \quad (25b)$$

$$\mathbf{a}_3 = \mathbf{s}_3 = [0, \quad 0, \quad 1]^T. \quad (25c)$$

where we have added \mathbf{s}_3 to form a complete set. Any vector which is linearly independent of \mathbf{a}_1 and \mathbf{a}_2 may also be used. The associated row basis vectors are obtained by solving (24):

$$\mathbf{b}^1 = [-1, \quad -1, \quad 0], \quad (26a)$$

$$\mathbf{b}^2 = [1, \quad 2, \quad 0], \quad (26b)$$

$$\mathbf{b}^3 = [\Delta H_1 - \Delta H_2, \quad \Delta H_1 - 2\Delta H_2, \quad 1]. \quad (26c)$$

We can readily verify that:

$$f^1 = \mathbf{b}^1 \odot \mathbf{g} = F^1. \quad (27a)$$

$$f^2 = \mathbf{b}^2 \odot \mathbf{g} = F^2. \quad (27b)$$

$$f^3 = \mathbf{b}^3 \odot \mathbf{g} = F^3 = 0. \quad (27c)$$

The CSP idea is very simple: instead of using the physically meaningful stoichiometric vectors as the default basis vectors, let's exploit the theoreticians' prerogative of trying different alternatives—may be something else works better.

3.3 Same Problem, Different Basis Vectors

We shall rework the first case of our example using dimensional formulation.

Instead of the physically meaningful stoichiometric vectors, we shall use, *without comments at this point*, the following set of basis vectors:

$$\mathbf{a}_1 = [-2, \quad 1, \quad \Delta H_1]^T = \mathbf{s}_1, \quad (28a)$$

$$\mathbf{a}_2 = \frac{1}{K_1 + 4A} [K_1, 2A, (K_1 + 4A)\Delta H_2 - (K_1 + 2A)\Delta H_1]^T, \quad (28b)$$

$$\mathbf{a}_3 = [0, \quad 0, \quad 1]^T. \quad (28c)$$

The following row vectors are their inverses and together they satisfy (24):

$$\mathbf{b}^1 = \frac{1}{K_1 + 4A} [-2A, \quad K_1, \quad 0], \quad (29a)$$

$$\mathbf{b}^2 = [1, \quad 2, \quad 0], \quad (29b)$$

$$\mathbf{b}^3 = [\Delta H_1 - \Delta H_2, \quad \Delta H_1 - 2\Delta H_2, \quad 1]. \quad (29c)$$

The amplitudes of the modes are:

$$f^1 = \mathbf{b}^1 \odot \mathbf{g} = F^1 + \frac{K_1 + 2A}{K_1 + 4A} F^2, \quad (30a)$$

$$f^2 = \mathbf{b}^2 \odot \mathbf{g} = F^2, \quad (30b)$$

$$f^3 = \mathbf{b}^3 \odot \mathbf{g} = F^3 = 0. \quad (30c)$$

In terms of these basis vectors, the original reaction system becomes:

$$\frac{d\mathbf{y}}{dt} = \mathbf{a}_1 f^1 + \mathbf{a}_2 f^2 + \mathbf{a}_3 f^3 \quad (31)$$

We can rewrite (31) in long-hand notation as follows:²

$$\frac{dA}{dt} = -2f^1 + \frac{K_1}{K_1 + 4A}f^2, \quad (32a)$$

$$\frac{dB}{dt} = f^1 + \frac{2A}{K_1 + 4A}f^2, \quad (32b)$$

$$\frac{dC}{dt} = [\Delta H_1]f^1 + [\Delta H_2 - (\frac{K_1 + 2A}{K_1 + 4A})\Delta H_1]f^2. \quad (32c)$$

Taking the inner product of \mathbf{b}^i with (31), we obtain:

$$\frac{1}{K_1 + 4A} \frac{d}{dt} [-A^2 + K_1 B] = f^1, \quad (33a)$$

$$\frac{d}{dt} [A + 2B] = f^2, \quad (33b)$$

$$\frac{d}{dt} [(\Delta H_1 - \Delta H_2)A + (\Delta H_1 - 2\Delta H_2)B + C] = f^3. \quad (33c)$$

We shall assume for the moment that somehow the modes have been ordered in descending speed—mode #1 is faster than mode #2. This is physically equivalent to saying that F^1 is estimated to be faster than F^2 . Reaction rate f^3 of mode #3 is, for this problem, identically zero.

3.3.1 Exact Conservation Law

Whenever the amplitude of a slow mode is considered “negligible,” it can be summarily neglected. Since for this problem $f^3 = 0$, mode #3 is always negligible, (33c) recovers (10), the conservation law for energy obtained earlier. However, a dormant mode does not always yield a conservation law. See §7.2 later.

3.3.2 The Rapid Transient Period

In the rapid transient period, we assume that the dominant mode is f^1 , while f^2 and f^3 are negligible in comparison.³ The simplified model is:

$$\frac{d\mathbf{y}}{dt} \approx \mathbf{a}_1 f^1. \quad (34)$$

²Compare them with (6a), (6b) and (6c).

³A slower mode may not necessarily be negligible in this period; if it is not, it *must* be kept.

Neglecting the right-hand side of (33b) and (33c), we recover (12) and (10), the temporary and permanent conservation laws obtained earlier. The solution of (34) automatically satisfies these constraints. As time progresses, mode #1 decays rapidly and becomes exhausted as $f^1 \rightarrow 0$, and \mathbf{y} approaches a fixed point.

3.3.3 The Slow Evolutionary Period

In the slow evolutionary period, (12) ceases to be valid, but the exhausted mode #1 replaces it with a new approximate equation of state:

$$f^1(\mathbf{y}) = \sum_{r=1}^R B_r^1 F^r(\mathbf{y}) \approx 0, \quad (35a)$$

$$= F^1 + \frac{K_1 + 2A}{K_1 + 4A} F^2 \approx 0, \quad (35b)$$

which is the counterpart of (13a), and agrees with it in the $|F^1| \gg |F^2|$ limit. However, unlike (13a), which we emphasized must never be substituted directly into the original governing equations, (35a) needs no such admonition and can be so used freely. Using it in (31), we obtain the simplified model for this period:

$$\frac{d\mathbf{y}}{dt} \approx \mathbf{a}_2 f^2. \quad (36)$$

Solution of (36) automatically satisfies (35a) if it is satisfied initially. No decision on which species should be eliminated (*i.e.* considered as a radical) was needed so far. All results obtained can be shown to be consistent with the conventional method in the asymptotic limit of reaction #1 being much faster than reaction #2.

3.3.4 Response to a Slow Third Reaction

For our simple example with only two reactions and three unknowns, the reaction system reaches its steady state after both modes #1 and #2 are exhausted. What happens if there is a third reaction? Denote the stoichiometric vector and reaction rate of the third reaction by \mathbf{s}_3 and F^3 respectively. In addition, assume that the three stoichiometric vectors are linearly independent, and that $|F^1| \gg |F^2| \gg |F^3|$ so that the added reaction is the slowest.

Using the same basis vectors in §3.3, we can proceed as before. The main impact of this complication is that $f^i = \mathbf{b}^i \odot \mathbf{g}$ now contain contribution from F^3 . After mode #2 is exhausted, a even slower mode #3 will take over. One can easily verify that the simplified model for this period is:

$$\frac{d\mathbf{y}}{dt} \approx \mathbf{a}_3 f^3. \quad (37)$$

Note that in this time period, the response of A and B to the slow third reaction is negligible; only C responds significantly.

3.4 The Obvious Questions

It can be shown that the three simplified models (34), (36) and (37) displayed above are the leading-order approximations in the asymptotic limit of $|F^1| \gg |F^2| \gg |F^3|$ for the indicated time periods, as obtained earlier in §2.2. Hence, the alternative representation of \mathbf{g} using the special set of basis vectors ((28a), (28b) and (28c)) is clearly a good idea in this limit.

It can easily be verified that (28a), (28b) and (28c) are the right eigenvectors of the Jacobian of \mathbf{g} in the limit of $k_1 K_1 \gg k_2 K_2$. Why do these special basis vectors allows us to summarily neglect the exhausted fast modes, while the same casualness would not be tolerated previously? What is the relationship between the conventional methods and the eigen-vectors of the Jacobian? How does one know that the basis vectors which worked well for one set of initial conditions will work for a different set of initial conditions? What happens when $k_2 K_2 / k_1 K_1$ is only moderately small?

Additional questions are: How does one extract physically interesting information from these basis vectors? Which species can be considered a radical? How can the accuracy of these models be improved? How can the whole procedure be generalized to deal with a massively complex problem?

4 The Ideal Basis Vectors

Given any set of N linearly independent basis vectors, one can always decompose the N -dimensional vector \mathbf{g} into N additive components or modes. Most investigators use the physically meaningful stoichiometric vectors as basis vectors by default. When the set used is non-ideal, the speed ranking

of the modes is unclear, and the residual net reaction rate of an exhausted fast reaction mode can not be neglected in general, and must be treated with great care. We have seen in the above example that the use of an intelligently chosen alternative set of basis vectors can make a difference. A good set of basis vectors not only should order the modes according to their speed, but also guarantee that the residual net reaction rate of an exhausted mode be negligible.

To find out how the amplitude of each mode evolves with time, we differentiate (1) with respect to time to obtain:

$$\frac{d\mathbf{g}}{dt} = \mathbf{J} \odot \mathbf{g} \quad (38)$$

where

$$\mathbf{J} = \frac{\partial \mathbf{g}}{\partial \mathbf{y}} = N \times N \text{ Jacobian matrix of } \mathbf{g} \text{ with respect to } \mathbf{y}. \quad (39)$$

Since the solution $\mathbf{y}(t)$ and therefore $\mathbf{J}(t)$ are both known functions of time, (38) can be considered a *linear* ODE (with known time-dependent coefficients) for \mathbf{g} . Taking the inner product of \mathbf{b}^i with (38) and using (21), we obtain:

$$\frac{df^i}{dt} = \sum_{j=1}^N \Lambda_j^i f^j, \quad i = 1, 2, \dots, N, \quad (40)$$

where

$$\Lambda_j^i \equiv \left[\frac{d\mathbf{b}^i}{dt} + \mathbf{b}^i \odot \mathbf{J} \right] \odot \mathbf{a}_j, \quad i, j = 1, 2, \dots, N. \quad (41)$$

A set of basis vectors $\mathbf{a}_i(t)$ is said to be *ideal* if (i) the inverse row vectors $\mathbf{b}^i(t)$ can be accurately computed from (24) for all time interval of interest, (ii) $\Lambda_j^i(t)$ is diagonal, and (iii) the diagonal elements of $\Lambda_j^i(t)$ are ordered in descending magnitudes. For linear problems where \mathbf{J} is a constant matrix, the ideal basis vectors would be the (constant) ordered eigen-vectors of \mathbf{J} . For non-linear problems, the eigen-vectors of \mathbf{J} are time-dependent, and they do not diagonalize Λ_j^i .

5 The CSP Refinement of Basis Vectors

The method of CSP does not attempt to find the ideal set of basis vectors—even when \mathbf{g} is linear. Instead, it assumes that, at any moment in time, a

trial set of ordered basis vectors is somehow available, that the first M fastest modes are exhausted as measured by some criterion (to be specified later), and generates from this trial set a new *refined* set of basis vectors, \mathbf{a}_i^o and \mathbf{b}_o^i , $i = 1, 2, \dots, N$ using a two-step *refinement* procedure (Lam and Goussis, 1991a). When recursively applied, the refinement procedure successively weakens the coupling between the fast modes and the slow modes.

The step #1 refinement is:

$$\mathbf{b}_o^m(M) \equiv \mathbf{b}^m + \sum_{J=M+1}^N p_J^m(M) \mathbf{b}^J, \quad m = 1, 2, \dots, M, \quad (42a)$$

$$\mathbf{a}_J^o(M) \equiv \mathbf{a}_J - \sum_{n=1}^M \mathbf{a}_n p_J^n(M), \quad J = M+1, \dots, N. \quad (42b)$$

The step #2 refinement is:

$$\mathbf{b}_o^I(M) \equiv \mathbf{b}^I - \sum_{n=1}^M q_n^I(M) \mathbf{b}_o^n(M), \quad I = M+1, \dots, N, \quad (43a)$$

$$\mathbf{a}_m^o(M) \equiv \mathbf{a}_m + \sum_{J=M+1}^N \mathbf{a}_J^o(M) q_m^J(M), \quad m = 1, 2, \dots, M. \quad (43b)$$

The matrices $p_J^m(M)$ and $q_m^J(M)$ above are defined by:

$$p_J^m(M) \equiv \sum_{n=1}^M \tau_n^m(M) \Lambda_J^n, \quad m = 1, 2, \dots, M, \quad J = M+1, \dots, N, \quad (44a)$$

$$q_m^J(M) \equiv \sum_{n=1}^M \Lambda_n^J \tau_m^n(M), \quad m = 1, 2, \dots, M, \quad J = M+1, \dots, N, \quad (44b)$$

and $\tau_m^n(M)$ is the inverse of $\Lambda_m^n(M)$:

$$\sum_{k=1}^M \tau_k^n(M) \Lambda_m^k(M) = \sum_{k=1}^M \Lambda_k^n(M) \tau_m^k(M) = \delta_m^n, \quad n, m = 1, 2, \dots, M. \quad (45)$$

The refinement procedure is readily programmable, and the refined basis vectors after each step satisfy the orthonormal relations, (24). It is essentially a generalization of the so-called “power method” for computing eigen-vectors, and produces a block diagonal Λ_j^i when converged.

The amplitude of the m -th fast mode satisfies the following differential equation (Lam and Goussis, 1991a):

$$\frac{df^m}{dt} = \sum_{n=1}^M \Lambda_n^m (f^m - f_\infty^m), \quad m = 1, 2, \dots, M, \quad (46a)$$

where

$$f_\infty^n \equiv - \sum_{J=M+1}^N p_J^n f^J, \quad m = 1, 2, \dots, M. \quad (46b)$$

$$f_o^m \equiv \mathbf{b}_o^m \odot \mathbf{g} = f^m - f_\infty^m, \quad m = 1, 2, \dots, M. \quad (46c)$$

Note that f_∞^m is some linear combination of the trial slow mode amplitudes. Solving for f_∞^m from (46c), we obtain:

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

which shows clearly that f_∞^m is proportional to the change of \mathbf{b}^m to \mathbf{b}_o^m as a consequence of the step #1 refinement procedure. In general, f_∞^m , which is computed using the trial set of basis vectors, can not be considered “small.”

Let $\tau(M)$ denote the magnitude of the time scale of the slowest of the fast modes. Let $\tau(M+1)$ denote the magnitude of the currently active time scale. Theoretically, they can be estimated by the M -th and $(M+1)$ -th eigen-values of \mathbf{J} . In the limit of small $\tau(M)/\tau(M+1)$ —large time scale separation—the formal asymptotic solution of (46a) is:

$$f^m = f_\infty^m + f_{o,\infty}^m + \dots, \quad m = 1, 2, \dots, M, \quad (47a)$$

$$f_{o,\infty}^m \equiv \sum_{n=1}^M \tau_n^m(M) \frac{df_\infty^n}{dt}, \quad m = 1, 2, \dots, M. \quad (47b)$$

The order of magnitude of $f_{o,\infty}^m$ can be estimated by:

$$f_{o,\infty}^m \approx O(f_\infty^m \tau(M)/\tau(M+1)), \quad m = 1, 2, \dots, M. \quad (48)$$

Assuming that $\tau(M+1) \gg \tau(M)$, the order of magnitude of $f_{o,\infty}^m$ is now small in comparison to f_∞^m —*i. e.* it is a “higher order” term. We have adopted the notation that superscript or subscript o indicates a variable evaluated with refined basis vectors. The above derivation can be found in Lam and Goussis (1991a).

Maas and Pope (1992) recognizes the importance of the eigen-vectors of \mathbf{J} , but recommends the use of its Schur vectors (Noble, 1988)—which transform \mathbf{J} into a lower-triangular form—as basis vectors. The advantage of Schur vectors for dealing with *dimensional* vectors is not clear.

5.1 The Trial Set of Basis Vectors

To get started, CSP needs to have a trial set of basis vectors. A number of options are available.

For sufficiently simple problems, one may wish to proceed analytically. N linearly independent stoichiometric vectors chosen from the reactions actually included in the full model (with supplementary basis vectors added as required) may be used as the default trial basis vectors. Educated guesses are then needed to establish the speed ranking of the modes.

A fool-proof procedure is to compute for the eigen-vectors of \mathbf{J} at $t = 0$ and use them as time-independent trial basis vectors for $t \geq 0$. The reciprocal of the eigen-values, denoted by $\tau(i)$, is an approximate measure of the characteristic time scales of the modes. We shall assume that the $\tau(i)$'s are essentially real and are ordered in ascending magnitudes.

When the refinement process is performed numerically on a computer, the refined basis vectors used in the previous time-step can be used as the trial basis vectors for the new time-step. Under this strategy, the initial choice of trial basis vectors at $t = 0$ is not critical.

5.2 The Number of Exhausted Modes

The refinement procedure requires the knowledge of M , the number of exhausted modes. In general, none of the fast modes are negligible at $t = 0$ (*i.e.* $M = 0$). As time progresses, the faster modes will eventually decay and become exhausted (*i.e.* $f_o^m \rightarrow f_{o,\infty}^m$). The number of exhausted fast modes, $M(t)$, is determined by requiring that their contributions to \mathbf{g} (see (54) and (55a) later) over the next time interval of $O(\tau(M+1))$ are negligible:

$$|\mathbf{a}_m^o f_o^m \tau(M)| < \mathbf{y}_{error}, \quad m = 1, 2, \dots, M, \quad (49a)$$

$$|\mathbf{a}_m^o f_{o,\infty}^m \tau(M+1)| < \mathbf{y}_{error}, \quad m = 1, 2, \dots, M, \quad (49b)$$

where \mathbf{y}_{error} is a dimensional column vector representing a user-specified threshold of absolute error allowed, and the vector inequality is individually enforced for *every* components of the vector. If one relaxes the requirement that the accuracy threshold for each component of \mathbf{y} be individually enforced, then the following are acceptable alternatives to (49a) and (49b):

$$|f_o^m| < O\left(\left|\frac{\mathbf{b}_o^m \odot \mathbf{y}_{error}}{\tau(M)}\right|\right), \quad m = 1, 2, \dots, M. \quad (49c)$$

$$|f_{o,\infty}^m| < O\left(\left|\frac{\mathbf{b}_o^m \odot \mathbf{y}_{error}}{\tau(M+1)}\right|\right), \quad m = 1, 2, \dots, M. \quad (49d)$$

Either of the above *exhaustion criterion* can readily be implemented computationally. The value of $\tau(M+1)$ can be estimated by the current integration step-size selected by any automatic (explicit) ODE integration package. Note that the value of M need not be a monotonic function of time.

Actually, the exhaustion criterion for the so-called CSP radicals defined in §6.4 can be made more lenient when the so-called radical correction (to be discussed in §6.5 later) is applied. A full discussion of this subtle point, however, is beyond the scope of this paper.

For most problems, the desired time resolution Δt is usually known. If

$$\Delta t > \tau(M), \quad (50)$$

then the M fastest modes must be nearly exhausted in the Δt time scale of interest. If we are not interested in the details of what happens in the next $\tau(M)$ seconds, we may use the simplified model with a non-zero value for M starting at $t \approx 0$, provided that we also adjust the initial conditions appropriately as demonstrated in (14) and (18). See §6.5 later.

5.3 The Refinement Process

The CSP refinement procedure consists of the following two steps.

1. Refinement of the fast row vectors using (42a) and the slow column vectors using (42b). This step depresses the magnitude of the upper-right block of Λ_K^m , ($m = 1, \dots, M$, $K = M + 1, \dots, N$), by the factor $\tau(M)/\tau(M+1)$, and thus weakens the coupling of the fast mode amplitudes from the slow.
2. Refinement of the slow row vectors using (43a) and the fast column vectors using (43b). This step depresses the magnitude of the lower-left block of Λ_m^K , ($K = M + 1, \dots, N$, $m = 1, \dots, M$), by the factor $\tau(M)/\tau(M+1)$, and thus weakens the coupling of the slow mode amplitudes from the fast.

The full cycle of the two-step refinement process renders the new Λ_j^i matrix calculated from the refined basis vectors more nearly block diagonal

than before. Step #1 improves the accuracy of the simplified model, while step #2 ensures that the simplified model is not stiff. The speed ranking of the modes can be estimated by the diagonal elements of the Λ_j^i matrix.

In our example in §3.3, the set of alternative basis vectors (28a), (28b) and (28c) used was obtained by adopting (25a), (25b) and (25c) as the ordered trial set, and then refining them analytically through step #1 using $M = 1$. We obtain:

$$\mathbf{b}_o^1 = \frac{1}{K_1 + 4A} [-2A, \quad K_1, \quad 0]^T, \quad (51a)$$

$$\mathbf{a}_2^o = \frac{1}{K_1 + 4A} [K_1, \quad 2A, \quad (K_1 + 4A)\Delta H_2 - (K_1 + 2A)\Delta H_1]^T, \quad (51b)$$

$$\mathbf{a}_3^o = [0, \quad 0, \quad 1]^T. \quad (51c)$$

These refined vectors are significantly modified from their trial counterparts. Performing step #2, we obtain:

$$\mathbf{b}_o^2 = [1 + \frac{2A\epsilon}{K_1 + 4A}, \quad 2 - \frac{K_1\epsilon}{K_1 + 4A}, \quad 0], \quad (52a)$$

$$\mathbf{b}_o^3 = [\Delta H_1 - \Delta H_2, \quad \Delta H_1 - 2\Delta H_2, \quad 1], \quad (52b)$$

$$\begin{aligned} \mathbf{a}_o^1 = [-2 + \frac{K_1\epsilon}{K_1 + 4A}, \quad 1 + \frac{2A\epsilon}{K_1 + 4A}, \\ \Delta H_1 - (\frac{K_1 + 2A}{K_1 + 4A}\Delta H_1 - \Delta H_2)\epsilon], \end{aligned} \quad (52c)$$

where ϵ is a dimensionless parameter defined by

$$\epsilon \equiv \frac{k_2(K_2 + 2)}{k_1(K_1 + 4A)}. \quad (53)$$

These refined vectors are only slightly modified from their trial counterparts for $\epsilon \ll 1$. Generally speaking, whenever the set of trial \mathbf{a}_i basis vectors are intelligently chosen, step #2 will provide only small corrections. If the trial set was randomly chosen, more than one full cycle of the two-step refinement process may be necessary to generate the correct leading order result in the limit of $\epsilon \rightarrow 0$.

The trial set is usually chosen initially to be time-independent for the sake of convenience. For non-linear problems, the refined set will in general

be time-dependent, and the associated time derivatives in Λ_j^i must be properly evaluated. Analytically, this is a straightforward step which could be quite tedious algebraically. From the programming point of view, however, good approximation to $d\mathbf{b}^i/dt$ can be obtained in a number of ways, such as utilizing the availability of stored values in the previous time steps and/or the predicted values in the next time step used in most integration routines.

6 Using the Refined Basis Vectors

Using the refined basis vectors, the governing system of ODE's become:

$$\frac{d\mathbf{y}}{dt} = \mathbf{g}_o^{o,fast} + \mathbf{g}_o^{o,slow} \quad (54)$$

where

$$\mathbf{g}_o^{o,fast} \equiv \sum_{m=1}^M \mathbf{a}_m^o f_o^m = \left(\sum_{m=1}^M \mathbf{a}_m^o \mathbf{b}_o^m \right) \odot \mathbf{g}, \quad (55a)$$

$$f_o^m = \mathbf{b}_o^m \odot \mathbf{g}, \quad m = 1, 2, \dots, M, \quad (55b)$$

$$\mathbf{g}_o^{o,slow} = \mathbf{g} - \mathbf{g}_o^{o,fast} = \sum_{r=1}^R \mathbf{s}_{o,r}^{o,slow} F^r, \quad (55c)$$

$$\mathbf{s}_{o,r}^{o,slow} \equiv \left(\mathbf{I} - \sum_{m=1}^M \mathbf{a}_m^o \mathbf{b}_o^m \right) \odot \mathbf{s}_r, \quad r = 1, 2, \dots, R. \quad (55d)$$

The column vector $\mathbf{s}_{o,r}^{o,slow}$ is mathematically the projection of \mathbf{s}_r in the so-called *slow subspace* or *manifold*, and can be interpreted physically as the *effective stoichiometric vector* of the r-th elementary reaction. Unlike the original chemical \mathbf{s}_r , the components of $\mathbf{s}_{o,r}^{o,slow}$ are not necessarily integers or rational numbers, and may involve species which do not appear in $F^r(\mathbf{y})$.

When the M fast modes satisfy the exhaustion criterion described in the next section, the simplified model is simply:

$$\frac{d\mathbf{y}}{dt} \approx \mathbf{g}_o^{o,slow} \quad (56a)$$

$$= \sum_{K=M+1}^N \mathbf{a}_K^o f_o^K. \quad (56b)$$

6.1 Classification of Modes

In general, there are three kinds of modes: exhausted modes, currently active modes, and dormant modes. Let $\tau(i)$ represent the characteristic time scale of the i -th mode. It is assumed that:

$$\tau(m) \ll \tau(K), \quad m = 1, 2, \dots, M, \quad K = M + 1, \dots, N. \quad (57)$$

The ordering among $\tau(m)$ and $\tau(K)$ is not important.

Exhausted Modes: By applying the exhaustion criterion (49b), the number M of exhausted fast modes is determined, yielding a set of M approximate equations of state:

$$f_o^m = \mathbf{b}_o^m \odot \mathbf{g} = \sum_{r=1}^R B_{o,r}^m F^r \approx 0, \quad m = 1, 2, \dots, M, \quad (58)$$

where

$$B_{o,r}^i \equiv \mathbf{b}_o^i \odot \mathbf{s}_r, \quad i = 1, 2, \dots, N, \quad r = 1, 2, \dots, R. \quad (59)$$

Currently Active Modes: The remaining $N - M$ modes are kept in (56b). Using (55c), we rewrite the simplified model as follows:

$$\frac{d\mathbf{y}}{dt} \approx \sum_{r=1}^R \mathbf{s}_{o,r}^{o,slow} F^r \quad (60)$$

Equation (60) can be numerically integrated without the need of a stiff solver. The integration time-step used should be a fraction of $\tau(M + 1)$, the characteristic time scale of the $(M + 1)$ th mode. The initial condition of (60) must satisfy (58) in the sense of (49c). If the initial trial fast column basis vectors were randomly chosen, the accuracy of (60) is $O((\tau(M)/\tau(M + 1))^c)$ where c is the number of full two-step refinement cycles. If they were initially “intelligently” chosen, then c is the number of step #1 refinements.

Dormant Modes: The amplitude of some of the slow modes kept in (56b) may be extremely small or even identically zero. When a slow mode is deemed negligible—when its contribution to $\mathbf{g}_o^{o,slow}$ in the current

time period is estimated to be less than \mathbf{y}_{error} component by component, it is said to be a *dormant mode*. In the example, mode #3 is a permanent dormant mode. If mode #N is a dormant mode, then:

$$f_o^N = \mathbf{b}_o^N \odot \mathbf{g} \approx 0, \quad (61)$$

or,

$$f_o^N = \sum_{r=1}^R B_{o,r}^N F^r \approx 0. \quad (62)$$

However, unlike (58), (62) is *not* a useful equation—it *can not be used as an approximate equation of state*. Instead, (61) is the more useful equation, indicating that the vector \mathbf{g} has negligible projection in a specific “direction.” If \mathbf{b}_o^N can be shown to be the gradient of a scalar—a condition known as *holonomic constraint* in classical mechanics—then that scalar is a conserved “integral of motion.” Conservation laws of atomic species manifests themselves as permanent dormant modes with *constant* row vectors (which are always holonomic). Equation (33c) is a conservation law since \mathbf{b}^3 is holonomic and f^3 is identically zero.

The solution of (60) automatically satisfy (58), (61) and (62). The distinction between (58) and (61) will be further discussed later.

6.2 Equation of State and Participation Index

The left-hand side of (58) usually consists of large positive and negative terms which nearly cancel each other. Which elementary reactions participate most strongly in this balancing act? The *participation index*, denoted by $P_{r,o}^i$, is designed to provide this information and is defined as follows:

$$P_{r,o}^i \equiv \frac{B_{o,r}^i F^r}{\sum_{r=1}^R |B_{o,r}^i F^r| + \left| \frac{\mathbf{b}_o^i \odot \mathbf{y}_{error}}{\tau(M+1)} \right|}, \quad i = 1, 2, \dots, N, \quad r = 1, 2, \dots, R. \quad (63)$$

where it is assumed that forward and reverse reactions are counted as distinct so that no cancellation occurs within any F^r . $P_{r,o}^i$ is a measure of the participation of the r-th elementary reaction to the balancing act of the i-th

mode. To get an idea of which elementary reactions are the major participants of each fast mode, we just need to sort the elementary reactions by their participation indices. A exhausted fast mode usually has several significant participation indices, a slow dormant mode may not have any, while a permanently dormant mode representing a conservation law has none. In our example worked out in §3.3, the participation index data for mode #1 will show that reaction #1 is the main participant of f^1 as defined by (30a), reaction #2 is the main participant of f^2 as defined by (30b), while $f^3 = 0$ has no participant at all and can be interpreted as a conservation law.

6.3 Simplified Model and Importance Index

The right-hand side of (60) consists of terms which control the reaction rate of the system. For each component of the state vector \mathbf{y} , we can order the terms on the right-hand side in descending magnitudes. Which elementary reactions are most important in controlling the reaction rate of a particular species of interest? The *importance index*, denoted by $I_r^{i,o}$ and designed to provide this information, is defined as follows:

$$I_{o,r}^{i,o} \equiv \frac{s_{o,r}^{i,o,slow} F^r}{\sum_{r=1}^R \left| s_{o,r}^{i,o,slow} F^r \right| + \left| \frac{y_{error}^i}{\tau(M+1)} \right|}, \quad i = 1, 2, \dots, N, \quad r = 1, 2, \dots, R. \quad (64)$$

$I_{o,r}^{i,o}$ is a measure of the relative importance of the contribution of the r -th elementary reaction to the current reaction rate of the i -th species. Note that if I_r^i 's were computed using the default \mathbf{s}_r 's, we would obtain the misleading information that the fast reactions are rate-controlling even after they are exhausted. In general, the values of the importance indices change discontinuously at the moment when a new fast mode is declared exhausted—the $I_{o,r}^{i,o}$'s of the major participants in the new exhausted mode will drop, while those in the emerging rate-controlling modes will rise. In the example worked out in §3.3, the importance index data on (36) computed using $\mathbf{s}_{o,r}^{o,slow}$ (see §7.5) will show that reactions #1 and #2 are important in controlling the overall reaction rate but each in their own time period only.

6.4 The Radical Pointer

When M modes are exhausted, M algebraic equations are obtained from (58) which can be used to replace M of the ODE's in the subsequent epochs. The question is: *which* ODE's can be replaced? In other words, which species are the radicals?

CSP associates with each exhausted mode one or more species by the *radical pointer*. The radical pointer of the m -th mode, $Q_m(i)$, is defined by the N diagonal elements of the $N \times N$ matrix $\mathbf{a}_m \mathbf{b}^m$, refined basis vectors preferred. Geometrically, the magnitude of $Q_m(i)$, which is dimensionless and its sum over i is unity, is a measure of how “perpendicular” the i -th species axis is to the surface defined by the m -th equation of state in \mathbf{y} space. A species k is identified as a *CSP radical* associated the m -th exhausted mode whenever $Q_m(k)$ is not a small number.

For the example treated in §3.3, we have for mode #1:

$$\mathbf{a}_1 \mathbf{b}^1 = \frac{1}{K_1 + 4A} \begin{pmatrix} 4A & -2K_1 & 0 \\ -2A & K_1 & 0 \\ -2A\Delta H_1 & K_1\Delta H_1 & 0 \end{pmatrix}. \quad (65)$$

The set of its diagonal elements is the radical pointer:

$$Q_1(i) = \frac{1}{K_1 + 4A} [4A, \quad K_1, \quad 0]. \quad (66)$$

This radical pointer informs us that either A or B, but definitely not C, may be used as a radical. In other words, only the ODE of either A or B can be replaced by (35b).

Maas and Pope (1992) did not realize that the choice of what species to solve for from the equations of state is restricted, and suggested that they “...can be chosen quite arbitrarily” with special caution on uniqueness of solutions. This is not our experience (Lam and Goussis, 1991b), and (66) is developed to provide a quantitative criterion for radicals. We have tested our radical pointer criterion by purposely solving for the “wrong” species—according to our radical pointer—from (58) in a number of test problems, and obtained wrong answers as expected.

6.5 The Radical Correction

What if the values of f_o^m in (58) are small but not as small as we would like at some moment in time? Theoretically, all M modes should rapidly decay to some much smaller values in the next several $\tau(M)$ seconds, while the main reaction activity proceeds with the current characteristic time scale of $\tau(M+1)$ seconds. If we are not interested in the details of the decay process but are only interested in finding an approximate “initial” condition for the next time period, we can use essentially Newton’s method to find the change to the value of \mathbf{y} , $\Delta\mathbf{y}_{rc}$, which would zero out the values of f_o^m , $m = 1, 2, \dots, M$:

$$\Delta\mathbf{y}_{rc} \approx - \sum_{m,n=1}^M \mathbf{a}_m^o \tau_n^m f_o^n. \quad (67)$$

In other words, $\mathbf{y} \rightarrow \mathbf{y} + \Delta\mathbf{y}_{rc}$ as the amplitudes of the fast modes decay toward zero in the next several $\tau(M)$ seconds. We shall call $\Delta\mathbf{y}_{rc}$ the *radical correction*. In Lam and Goussis (1991a), the radical correction was referred to as “inclusion of the homogeneous solution” in evaluating the change of \mathbf{y} , but no details were given there.

When a simplified model is used to compute for an approximation solution by neglecting the M exhausted modes, the values of f_o^m ’s are small at the beginning of the time period and are theoretically expected to remain small in the slow evolutionary period. The radical correction can be used to ensure that the numerical solutions adequately satisfy the approximate equations of state as required.

Suppose the initial values in the example worked out in §3.3 does not satisfy the partial-equilibrium of mode #1:

$$f_o^1(0) = F^1(0) + \frac{K_1 + 2A(0)}{K_1 + 4A(0)} F^2(0) \neq 0, \quad (68)$$

and the solution in the rapid transient period is of no interest to the investigator. The radical correction can be used to obtain the effective initial values, the counter part of (14), for the next time period. With $M = 1$, $\tau_1^1 = -(k_1(K_1 + 4A(0)))^{-1}$, (67) gives:

$$A(0+) \approx A(0) - \frac{2f_o^1(0)}{k_1(K_1 + 4A(0))}, \quad (69a)$$

$$B(0+) \approx B(0) + \frac{f_o^1(0)}{k_1(K_1 + 4A(0))}, \quad (69b)$$

$$C(0+) \approx C(0) + \frac{\Delta H_1 f_o^1(0)}{k_1(K_1 + 4A(0))}. \quad (69c)$$

It can easily be verified that (10) is satisfied and $A + 2B$ is unchanged by this radical correction, while the new value of $f_o^1(0+)$ becomes much smaller than before. By recursive applications $A(0+)$ will converge to (14), and $f_o^1(0+)$ will converge to the nearest stable zero. The computation for the slow evolutionary time period can now commence with (36) with the new initial values.

6.6 Explosive Modes

For chemical kinetics, the eigen-values of \mathbf{J} are usually real, and mostly negative, signifying decaying modes. Occasionally, some modes may have positive eigen-values, signifying “*explosive modes*” which are often of interest in the study of “ignition” mechanisms (Trevino and Solorio, 1991; Trevino, 1991). In our example, whenever $A < K_1/(2K_2)$ an explosive mode exists which is clearly the manifestation of the *chain branching* of the B radical in the reaction system.

Because the matrix Λ_j^i is not exactly block diagonal, the non-zero coupling between the fast exhausted modes and the slow explosive modes can cause some exhausted modes to come alive again. From the point of view of computation and programming, this complication is easily handled.

6.7 The Minimum System

Using the participation and importance indices, it is a simple matter to identify the minimum set of species and unknowns (ignore the unimportant reactants but includes any species of special interest to the investigator) and the minimum set of elementary reactions (ignore the unimportant reactions) to form a minimum reduced reaction system which can represent \mathbf{g} to any reasonable desired user-specified threshold of accuracy for any time period(s) of interest.

7 Comments on CSP

In terms of elegance, there is no substitute for a conventional asymptotic analysis of an appropriately non-dimensionalized problem with a well defined small parameter ϵ , with analytical results expressed in terms of simple functions.

In dealing with practical problems, there exists few guidelines for order of magnitude estimates and the intelligent non-dimensionalization of the variables. The identification of the small parameter ϵ is also not straightforward; for a real problem it may be insufficiently small even when successfully identified. For moderately complex reaction systems, the formalism of asymptotics must be replaced by *ad hoc* quasi-steady and/or partial-equilibrium hypothesis, and the role of experience and intuition then becomes central. With sufficient algebraic skill, such methodology is indeed capable of generating analytical results. However, for massively complex reaction systems, this option is simply not viable.

The theory of CSP provides a new formalism to do asymptotics. Most importantly, it does not require non-dimensionalization of the variables and the identification of a small parameter. The time-dependent value of $\tau(M)/\tau(M+1)$ is the small parameter. The CSP algorithm can be performed analytically for sufficiently simple problems, as was demonstrated in §3.3, and it can be programmed to handle massively complex problems—provided the fast modes are of the boundary layer type and decay with time eventually.

CSP can be used to test an intuitive guess, provides a fool-proof way to obtain the correct leading approximation in the absence of good ideas, and allows the theoretician to concentrate on the task of extracting physical insights—because all the massively complicated algebra are left to be performed numerically on the computer.

In most cases, it is relatively easy to interpret a CSP-generated equation of state as either a partial-equilibrium or a quasi-steady approximation. However, there are also situations when such interpretation is not immediately obvious, and the CSP data becomes physically sensible only after further reflections.

From the point of view of asymptotics, CSP is distinctly different from the conventional analytical method of “matched asymptotic expansions,” for no asymptotic matching is required. A detailed discussion of the differences, however, is beyond the scope of this paper.

7.1 Local-Eigen Vectors vs. Refined Basis vectors

As mentioned previously, the set of local eigen-vectors of \mathbf{J} is a fool-proof choice for use as trial basis vectors. In fact, the special set of basis vectors ((28a), (28b) and (28c)) used in §3.3 is the leading order approximation of the right eigen-vectors of \mathbf{J} , with k_2K_2/k_1K_1 playing the role of a small parameter. If only so-called leading order approximation (for sufficiently small k_2K_2/k_1K_1) is desired, this set is totally adequate. But what if the value of k_2K_2/k_1K_1 of practical interest is only 0.15, and it is desired to develop a simplified model with accuracy around 3%? The CSP refinement procedure can be recursively applied to improve the accuracy.

The CSP theory formally guarantees that each full cycle of refinement increases the accuracy of the simplified model by the factor $\tau(M)/\tau(M+1)$. In the first full cycle, we can generate the leading approximation by using constant trial vectors so that no time derivatives are included. In most cases, the leading approximation to a complex problem is adequate to provide most physical insights. The second full cycle needs the evaluation of first time derivatives, and the third full cycle needs the evaluation of second time derivatives, *etc.* Mathematically, these time derivatives are all available. Hence CSP is capable of significantly higher accuracy than using just the local eigen-vectors. It is, of course, not at all necessary to use the local eigen-vectors as trial basis vectors; any informed, intelligent guess can be used. In a computer code, it is possible to implement two full refinement cycles using relatively little computational resource.

To continue beyond §3.3, we may choose (28a), (28b) and (28c) as our time-dependent trial set, respecting the presence of $A(t)$ in \mathbf{a}_2 and \mathbf{b}^1 . This (second) refinement proceeds routinely except that the time derivative term in (41) must be included in evaluating Λ_j^i . The new CSP-derived simplified models so generated will be accurate to $O(k_2K_2/k_1K_1)^2$, compared to $O(k_2K_2/k_1K_1)$ for the results obtained earlier. One more cycle of refinement will generate models of accuracy $O(k_2K_2/k_1K_1)^3$, *etc.*

7.2 Equation of State and Conservation Law

In the combustion literature, the concept of conserved scalars is linked with the assumption of equal mass diffusivity coefficients and unity Lewis numbers. Since no diffusion effects are considered here, these requirements be-

come irrelevant. In this section, we shall show how (approximate) conserved scalars for such systems can be found as a by-product of CSP.

Taking the dot product of \mathbf{b}_o^i with (54), we have:

$$\mathbf{b}_o^i \odot \frac{d\mathbf{y}}{dt} = f_o^i, \quad i = 1, 2, \dots, N. \quad (70)$$

If a row vector \mathbf{b}_o^n can be expressed as the gradient (with respect to \mathbf{y}) of a scalar function $\Theta_{n,o}(\mathbf{y})$ times a scalar function $\theta_{n,o}(\mathbf{y})$, then the mode shall be called an *holonomic* mode—a concept well known in classical mechanics. The corresponding equation is (70) is called an *holonomic constraint*, and can be rewritten as:

$$\theta_{n,o} \frac{d\Theta_{n,o}}{dt} = f_o^n, \quad \text{selected } n\text{'s}. \quad (71)$$

We shall call $\Theta_{n,o}$, when it exists, an *eligible conserved scalar*.

When the m -th fast mode becomes exhausted, an equation of state is obtained by setting $f_o^m \approx 0$. If this mode is also holonomic, the associated eligible conserved scalar $\Theta_{m,o}$ then becomes a *exhausted conserved scalar*—its value remains constant from the moment of exhaustion. It is likely that all exhausted refined fast modes are holonomic, and that $\Theta_{m,o}$ and f_o^m are linearly dependent—thus no new information is provided by the exhausted conserved scalars.

When the amplitude of the K -th slow mode is found “negligible,” setting this dormant mode $f_o^K \approx 0$ yields no information at all. We can only conclude from (70) that the component of the vector $d\mathbf{y}/dt$ “in the the direction of \mathbf{a}_K^o ” is negligible. However, if this dormant mode is holonomic (*e. g.* \mathbf{b}_o^K are constants, yielding $\theta_{K,o} = 1$, $\Theta_{K,o} = \mathbf{b}_o^K \odot \mathbf{y}$), then the associated $\Theta_{K,o}$ is a *dormant conserved scalar*—its value is approximately conserved whenever the K -th mode is dormant. Conserved scalars in the combustion literature are dormant ones, and they almost always have simple physical meanings, such as conservation of atomic species or total energy. While $\Theta_{K,o} \approx \text{constant}$ is also an algebraic relation between the species, we do *not* call it an equation of state because the *constant* is not universal.

In our example, the basis vectors used in §3.3 are the leading order results obtained by refining the trial basis vectors given in §3.2. Equations (33a), (33b) and (33c) correspond to (70) and they are all holonomic. The eligible

conserved scalars are ($M = 1$):

$$\Theta_{1,o} = -A^2 + K_1 B, \quad (72a)$$

$$\Theta_{2,o} = A + 2B, \quad (72b)$$

$$\Theta_{3,o} = (\Delta H_1 - \Delta H_2)A + (\Delta H_1 - 2\Delta H_2)B + C, \quad (72c)$$

with

$$\theta_{1,o} = \frac{1}{K_1 + 4A}, \quad \theta_{2,o} = 1, \quad \theta_{3,o} = 1. \quad (73)$$

All three refined modes in this example are exactly holonomic. In general, there is no guarantee that any CSP refined mode is holonomic⁴—it is not possible to determine from numerical CSP data whether a mode is holonomic or to deduce the form of the possible eligible conserved scalar except for the obvious case of time-independent \mathbf{b}_o^n .

7.3 The User-Specified \mathbf{y}_{error}

The role of a theory is to construct a model which should be as simple as possible and yet can generate valid predictions in some domain of the parameter space of interest. A good model need not predict *all* things correctly. For example, a certain simplified model may be very good indeed for the heat release and temperature history, but very bad for the predictions of certain pollutant concentrations. For the same reaction system, different simplified models may be appropriate depending on the desired scope and accuracy of the solutions.

The CSP theory addresses this issue by choosing not to directly define a vector norm for the N -dimensional state space of \mathbf{y} . Instead, it asks the investigator to supply the vector \mathbf{y}_{error} , which specifies the amount of absolute error considered tolerable for each component of \mathbf{y} . This step is intuitively sensible, and is conceptually equivalent to choosing a norm. The CSP theory relies on this \mathbf{y}_{error} to decide whether a decaying fast mode should be neglected. As a consequence, a CSP-derived simplified model which focuses

⁴A sufficient but *not* necessary condition for the n -th mode to be holonomic is that $\mathbf{b}^n = \text{constant}$. Hence constant trial row basis vectors always generate holonomic modes, but the resulting eligible conserved scalars may never become conserved scalars. For our example, if the default trial \mathbf{b}^1 in §3.2 were used, we would obtain $\Theta_1 = -A - B$ which, in contrast to $\Theta_{1,o}$, is not conserved after fast mode #1 is exhausted.

on temperature only may be quite different from one which claims additional predictive capability such as pollutant concentrations.

If too stringent a \mathbf{y}_{error} is chosen, CSP would simply report $M = 0$ and would fail to recommend any simplified model. As \mathbf{y}_{error} becomes more forgiving, CSP would recommend simpler models consistent with the specified tolerance. This posture of producing simplified models on demand is refreshing—when compared to the posture of conventional asymptotics.

7.4 Insights from CSP Numbers

In a theoretical analysis, no specific numerical values of the parameters in the problem need be specified; only their order of magnitude is assumed. While the accuracy of the analysis may be limited—because the small parameter may not be that small in practice—the parametric dependence of the results are available for inspection and interpretation. An experienced theoretician can speak knowledgeably about the reaction system and interpret the analytical results in physically meaningful and intuitive terms such as chain branching, radical formation, ignition temperature, alternative reaction paths, *etc.*

When results are obtained computationally, each set of numerical data is valid only for one specific set of parameters. For massively complex problems, the computer printouts are overwhelming. The interpretation of these numbers is a much more difficult matter. The CSP data generated, which is all numerical, is designed to be easy to use to extract qualitative and physically meaningful insights.

For non-linear problems, the CSP data is itself massive. However, it need not be computed at every time step; it is needed only when the theoretician wants to know: what is going on here? While the time dependence of the reaction system may appear to be very complex, the CSP-derived refined basis vectors may depend only weakly on time, and the simplified model derived at selected points in time can provide insights about the “physical mechanism” of the system for a finite time interval. In fact, significant change of the behaviors of basis vectors is a signal that the physical mechanism has changed, and a new simplified model is needed. For linear problems, a single set of CSP data—the eigen structure of the constant matrix J —will explain everything for all time. For our simple example (which is non-linear), a single set of time-independent trial basis vector is adequate to generate the correct

refined set for all time.

From the CSP viewpoint, the problem of simplified kinetics modeling reduces to finding a set of basis vectors which make Λ_j^i block-diagonal. The conventional methods guess at them, while CSP provides a rational and programmable way of refining a trial set to make the resulting Λ_j^i more diagonal than before. For massively complex problems, the computer is assigned the task of computing an adequate set, and the investigator uses common sense to extract physically interesting information from the computer generated numerical basis vectors. The participation index and the importance index are provided to quantify the roles played by each elementary reaction included in the model. Together they provide answers to most of the interesting questions about a reaction system. Any simplified model derived based on intuition and experience can now be compared with the CSP-derived results to verify that the various judgments and guesses used are correct.

Depending on the user-specified accuracy threshold, \mathbf{y}_{error} , certain chemical species and elementary reactions can be removed from the reaction system, producing the so-called reduced reaction system. The more lenient \mathbf{y}_{error} is, the simpler the CSP-derived reduced system—while the essential character of the full system is retained. In Goussis and Lam (1992), the methanol oxidation problem was studied with 30 species plus temperature and 173 elementary reactions. The CSP data generated can be roughly divided into four epochs, each characterized by a distinctly different set of CSP-derived basis vectors, and together they indicated (through the CSP indices) that a reduced reaction system with 15 species (including methanol) would have adequate accuracy for all 15 unknowns. This CSP-derived insight was computationally confirmed. This reduced reaction system, of course, makes no predictions on the discarded species.

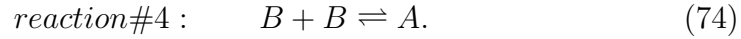
7.5 CSP vs. Sensitivity Analysis

The standard method of sensitivity analysis (Yetter, Dryer and Rabitz, 1991) is to linearize the problem and then to compute for the linearized response to various perturbations. For example, if one wishes to know what happens to y^{95} when k_{127} is changed by a small amount, one numerically computes $d(\log y^{95})/d(\log k_{127})$.

CSP provides an alternative to this brute-force method. First, we look at the participation indices $P_{127,o}^m$, $m = 1, 2, \dots, M$. If these indices indicate

that reaction #127 is a significant participant of the m -th exhausted fast mode, then k_{127} will directly affect the values of the radicals pointed to by the radical pointer $Q_m(i)$. Otherwise k_{127} has no effect on the radicals. Next, we look at the importance index $I_{o,127}^{95,o}$; its numerical value tells us how strongly k_{127} affects the reaction rate of y^{95} .

Frequently, it is desired to know what would happen if additional elementary reactions were added to a reaction system after a computation is already done. CSP can be used to provide useful answers. To illustrate, let reaction #4 be added to our example reaction system:



Its stoichiometric vector and reaction rate are:

$$\mathbf{s}_4 = [1, \quad -2, \quad \Delta H_4]^T, \quad F^4 = k_4(B^2 - K_4 A). \quad (75)$$

Since \mathbf{s}_4 is not linearly independent, we have:

$$\mathbf{s}_4 = \mathbf{s}_1 - 3\mathbf{s}_2 - (\Delta H_1 - 3\Delta H_2 - \Delta H_4)\mathbf{s}_3. \quad (76)$$

Hence:

$$B_4^i = B_1^i - 3B_2^i - (\Delta H_1 - 3\Delta H_2 - \Delta H_4)B_3^i, \quad (77)$$

$$\mathbf{s}_{o,4}^{o,slow} = \mathbf{s}_{o,1}^{o,slow} - 3\mathbf{s}_{o,2}^{o,slow} - (\Delta H_1 - 3\Delta H_2 - \Delta H_4)\mathbf{s}_{o,3}^{o,slow}. \quad (78)$$

Using the basis vectors in §3.3, we can computed the following CSP data:

$$B_1^1 = 1, \quad B_2^1 = \frac{K_1+2A}{K_1+4A}, \quad B_3^1 = 0, \quad (79a)$$

$$B_1^2 = 0, \quad B_2^2 = 1, \quad B_3^2 = 0, \quad (79b)$$

$$B_1^3 = 0, \quad B_2^3 = 0, \quad B_3^3 = 1, \quad (79c)$$

and (for the time period when $M = 1$ only),

$$\mathbf{s}_{o,1}^{o,slow} = [0, \quad 0, \quad 0]^T, \quad (80a)$$

$$\mathbf{s}_{o,2}^{o,slow} = \frac{1}{K_1 + 4A} [K_1, \quad 2A, \quad (K_1 + 4A)\Delta H_2 - (K_1 + 2A)\Delta H_1]^T, \quad (80b)$$

$$\mathbf{s}_{o,3}^{o,slow} = [0, \quad 0, \quad 1]^T. \quad (80c)$$

The effective stoichiometric vectors for the other two time periods are also available. It is now a simple matter to compute the participation and importance indices for reaction #4 from this CSP data and decide whether its previous omission can be justified in any time period of interest. If the indices $P_{4,o}^i$ and $I_4^{i,o}$ show that reaction #4 is a major player, then the old basis vectors should be refined to reflect the new character of the new reaction system.

7.6 Effects on Diffusion

If spatial diffusion is included, the governing equations will contain the Laplacian operator, and become a system of PDE's:

$$\frac{d\mathbf{y}}{dt} = \mathbf{g} + \mathbf{D} \odot \nabla^2 \mathbf{y} \quad (81)$$

where \mathbf{D} is—in most cases—a $N \times N$ diagonal matrix of diffusion coefficients which are considered constants here for the sake of simplicity.

Assuming that the spatial diffusion mechanism is not the fastest process, we can obtain the corresponding approximate equations of state, (58), when the fastest M modes are exhausted as before. In the slow evolutionary time period and spatial domain, the fast modes can be neglected from the right-hand side of (81) to obtain, instead of (60), the following simplified model:

$$\frac{d\mathbf{y}}{dt} \approx \sum_{r=1}^R \mathbf{s}_{o,r}^{o,slow} F^r + \mathbf{D}_o^o \odot \nabla^2 \mathbf{y} \quad (82)$$

where \mathbf{D}_o^o is the *effective diffusion matrix*, the projection of \mathbf{D} in the slow subspace, given by:

$$\mathbf{D}_o^o \equiv (\mathbf{I} - \sum_{m=1}^M \mathbf{a}_m^o \mathbf{b}_o^m) \odot \mathbf{D}. \quad (83)$$

Let the diffusion matrix \mathbf{D} in our example be diagonal, with D_A , D_B , and D_C as the diagonal elements:

$$\mathbf{D} = \begin{pmatrix} D_A & 0 & 0 \\ 0 & D_B & 0 \\ 0 & 0 & D_C \end{pmatrix}. \quad (84)$$

When $M = 1$, the effective diffusion matrix \mathbf{D}_o^o can be computed from (83) and (65):

$$\mathbf{D}_o^o = \frac{1}{K_1 + 4A} \begin{pmatrix} K_1 D_A & 2K_1 D_B & 0 \\ 2AD_A & 4AD_B & 0 \\ 2A\Delta H_1 D_A & -K_1\Delta H_1 D_B & (K_1 + 4A)D_C \end{pmatrix}. \quad (85)$$

This derivation of \mathbf{D}_o^o using CSP basis vectors is straightforward and programmable, and can be used to deal with massively complex PDE systems. Adjacent to solid boundaries, spatial boundary layers will exist, and *effective boundary values* for (82) are needed and must be carefully derived.

8 Discussion

Experience and intuition have always been central to the conventional derivation of simplified models. It was generally known that both the partial-equilibrium and the quasi-steady approximations were related to conventional analytical singular perturbation procedures—provided that the relevant small parameter(s) could be identified.

The goal of developing a general theory of singular perturbation which can handle any system of non-linear first order ODE's in a programmable manner without the identification of small parameter(s) seems very ambitious indeed. The CSP theory developed by Lam *et. al.* has achieved this goal, but only for *boundary-layer type* problems where all fast modes eventually decay exponentially. Fortunately, most problems in chemical kinetics are of this type.

The basic strategy of CSP is to uncouple the fast, exhausted modes from the slower, currently active modes through an intelligent choice of basis vectors. For a strictly linear problem, the ideal uncoupling basis vectors are the time-independent eigen-vectors of \mathbf{J} . For non-linear problems, the desired uncoupling basis vectors are time-dependent because Λ_j^i depends on $\mathbf{y}(t)$. The main concession made by CSP is to abandon the goal of diagonalizing Λ_j^i and be satisfied with a nearly block-diagonal one. A programmable recursive refinement procedure is provided to successively weaken the fast-slow couplings.

Pragmatically, the extraction of physical insights from the CSP data is advantageous only if the CSP data, such as the participation index and the

importance index, changes less rapidly with time than the solution $\mathbf{y}(t)$. In problems involving chemical reactions, the basic chemical process is usually described by a reaction rate F'' and a *constant* stoichiometric vector \mathbf{s}_r . Whatever simplifications which are available can usually be attributed to some reactions being much faster than others and had now spent themselves. As a consequence, good decoupling basis vectors in chemical kinetics are most likely to be weakly time-dependent, because the relevant chemically meaningful stoichiometric vectors of the fast reactions are time-independent. In the example studied in §3.3, the information extracted from the CSP data hardly change with time at all, because the “mechanism” of the problem did not change.

The CSP algorithm depend on the fast modes to become exhausted, which translates into assuming that the fast eigen-values of \mathbf{J} are “essentially real” and negative. Theoretically, if a highly oscillatory fast mode does not damp out (*i.e.* a Wentzel, Kramers and Brillouin (WKB) type problem), CSP will only be able to identify this mode, but will not be able to provide an algorithm to handle the evolution of \mathbf{y} in the next larger time scale. The extension of CSP to WKB type problems remains to be explored.

CSP can be used, in principle, on problems which have a large number of modes but only moderate amount of time scale separation (*i.e.* the ratio, $(\tau(M+1)/\tau(M))$, is only moderately large), provided that only moderate accuracy is desired. It thrives on large time scale separations, and performs best when the separations are “asymptotically large.”

9 Concluding Remarks

A computer code, called CSP91 and programmed by Dr. Dimitris Goussis, has implemented the CSP algorithm, and is available on request. The code can be used either as a diagnostics subroutine, or as an stiff ODE solver. It has been used to demonstrate in a number of test problems (Lam *et. al.*, 1989; Goussis *et. al.* 1990; Goussis and Lam, 1992) that numerical solutions generated by the CSP-derived simplified models are in quantitative agreement with that of the corresponding full models in accordance to the accuracy thresholds specified. Different simplified models are generated for different user-specified \mathbf{y}_{error} . The CSP-derived insights, such as which fast reactions are in partial-equilibrium with each other (through the participation indices),

which reactions are rate-controlling (through the importance indices), which species can be considered as radicals in the CSP context (through the radical pointers) and how they can be solved for from the equations of state (through the radical correction), have been thoroughly tested and verified. In most cases, the CSP-derived insights are consistent with the expectations of competent chemical kineticists knowledgeable in the reaction system under study. For the few cases when the CSP-derived insights appeared surprising at first sight, they all eventually became obvious after further reflections.

References

- Aiken, R. C. (1985). *Stiff Computation*, Oxford University Press.
- Bilger, R. W., M. B. Esler, and Starner, S. H. (1991). In *Reduced Kinetic Mechanisms and Asymptotic Approximations for Methane-Air Flames*, Chapter 5: "On Reduced Mechanisms for Methane-Air Combustion," Lecture Notes in Physics 384, M. Smooke, Ed., Springer-Verlag.
- Bilger, R. W., Starner, S. H. and Kee, R. J. (1990). "On the Reduced Mechanisms for Methane-Air Combustion in Non-Premixed Flames, *Combust. Flame* **80**, p. 135.
- Chelliah, H. K. and Williams, F. A. (1990). "Aspects of the Structure and Extinction of Diffusion Flames in Methane-Oxygen-Nitrogen Systems," *Combust. Flame*, **80**, pp. 17-48.
- Goussis, D. A., S. H. Lam and Gnoffo, P.A. (1990). "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.
- Goussis, D. A. and Lam, S. H. (1992). "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.
- Lam, S. H. (1985). "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.
- Lam, S. H. and Goussis, D. A. (1988a). "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.

Lam, S. H. and Goussis, D. A. (1988b). "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.

Lam, S. H., Goussis, D. A., and Konopka, D. (1989). "Time-Resolved Simplified Chemical Kinetics Modeling Using Computational Singular Perturbation", AIAA 27th Aerospace Sciences Meeting, (AIAA 89-0575), Reno, Nevada, January 9-12.

Lam, S. H. and Goussis, D. A. (1991a). 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.

Lam, S. H. and Goussis, D. A. (1991b). "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.

Maas, U. and Pope, S. B. (1992). "Simplifying Chemical Kinetics: Intrinsic Low-Dimensional Manifolds in Composition Space," *Combust. Flame*, **88**, pp. 239-264.

Noble, B. and Daniel, J. W. (1988). *Applied Linear Algebra*, Third Edition, Prentice Hall, Inc.

Peters, N. (1985). "Numerical Simulation of Combustion Phenomena", R. Glowinski, B. Larrourou, and R. Temam, Eds., Lecture Notes in Physics 241, pp. 90-109, Springer-Verlag, New York.

Peters, N. (1991). In *Reduced Kinetic Mechanisms and Asymptotic Approximations for Methane-Air Flames*, Chapter 3: "Reduced Mechanisms," Lecture Notes in Physics 384, M. Smooke, Ed., Springer-Verlag.

Peters, N. and Williams F. A. (1987). "The Asymptotic Structure of Stoichiometric Methane-Air Flames", *Combust. Flame*, **68**, pp. 185-207.

Trevino, C. and Solorio, F. (1991). "Asymptotic Analysis of the High Temperature Ignition of $CO/H_2/O_2$ Mixtures," *Combust. Flame*, **86**, pp. 285-295.

Trevino, C. (1991). "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.

Williams, F.A. (1985). *Combustion Theory, The Fundamental Theory of Chemically Reacting Systems*, 2nd edition, Benjamin/Cummings Publ. Co., Menlo Park, Ca.

Yetter, R. A., Dryer, F. L. and Rabitz, H. (1991). "A Comprehensive Reaction Mechanism of Carbon Monoxide/Hydrogen/Oxygen Kinetics," *Combust. Sci. Tech.*, **79**, pp. 97-128.