J. Phys. Chem. A **1999,** *103,* 337−343

ARTICLES

Deduction of Chemical Mechanisms from the Linear Response around Steady State

Rubén Díaz-Sierra,[†] José Bernardo Lozano,[‡] and Víctor Fairén*

Departamento de Física Fundamental, Universidad Nacional de Educación a Distancia,
Apartado 60141, 28080 Madrid, Spain

Received: July 16, 1998; In Final Form: September 25, 1998

The determination of abstract models of complex chemical mechanisms from the analysis of the experimental time series of the concentration of species is a topic of undoubted interest. Doing so from time series in the linear regime offers certain advantages because the equations of evolution can be easily retrieved. This contrasts with the much greater ambiguity associated to the corresponding process in the nonlinear regime. However, until presently, the procedure has not been formalized, and known methods of inference of the mechanism still rely too much on guess. The aim of the present article is to remedy to this problem by offering a systematic procedure for carrying out the deduction of the mechanism within the framework of the law of mass action. The number of independent steps and actual values of both the stoichiometric coefficients and of the rate constants are calculated from the linear equations governing the time response of chemical species to a small pulse perturbation of the steady state. Also, the deduction of these linear equations from that same response is outlined.

I. Introduction

When confronted to an unknown chemical mechanism, an interesting question may be raised: is it possible to devise single-handedly a prototypical reaction network model from the observed time evolution of reacting species? In other words, do experimental time series convey enough information for producing, in a more or less systematic way, a sensible mechanism? Although the question has been already affirmatively answered in the literature,[1] the topic is ripe for further elaboration and improvement. The purpose of the present article is to contribute to this line of action.

Surprisingly, chemically sensible modeling expectations may be satisfied if we define the problem within the context of time series in the linear regime. Here, there exists a one-to-one correspondence between an algebraic representation of the trajectory fitting the time series and the linear equations of motion around steady state

$$\frac{d\delta\mathbf{X}}{dt} = \mathbf{J}\cdot\delta\mathbf{X} \qquad (1)$$

where $\delta\mathbf{X}$ is a small perturbation and $\mathbf{J}$ is the Jacobian matrix, taken to be the linearization at steady state of the chemical rate equations. From the knowledge of its solution—a combination of exponentials—eq 1 can be formally reconstructed. This idea has been exploited by Chevalier et al.[1] They review different useful techniques for obtaining the Jacobian matrix elements (JMEs). Among them we emphasize that which makes use of the time series generated in response to a small pulse perturbation in some reactant.[2] Once the JMEs have been determined, a close examination of their sign pattern offers a hint on how the different species react: for example, a positive JME, say J_{ij}, may be interpreted as a direct participation of species X_j in the production of species X_i. The heuristics has been successfully applied[1] to the DOP model,[3] although little can be inferred on the specific values of both the stoichiometric coefficients and rate constants.

The previous procedure for the inference of a set of reactional steps from the signs of the JMEs is not devoid of great sources of ambiguity,[1,4] even in those models, as the DOP, to which it can be applied with relative success. In short, it yields acceptable skeletal model networks whenever each one of the off-diagonal JMEs is reasonably allocable to a single step or, at least, to a single dominating step. Even so, the method relies more on guess—and luck—than on a well-prescribed set of working rules.[1,4]

We present here a systematic method for providing a complete specification of a chemical mechanism from the knowledge of the JMEs. This is done by calculating the number of steps, stoichiometric coefficients, and rate constants. The method is systematic in the sense that it rephrases the problem in the context of the solution to a set of algebraic equations for the reaction rates. If the JMEs, computed from an experimental time series, are well defined, and we consider that all reactions proceed according to the law of mass action, then the compatibility conditions for that set of equations lead to the deduction of (1) the minimal set of independent steps compatible with the Jacobian matrix and (2) the stoichiometric coefficients of the reactants. Once these have been calculated, the equations are solved and the rate constants obtained.

---

* Corresponding author. E-mail:vfairen@apphys.uned.es. Permanent address: Departamento de Física Matemática y Fluidos, Universidad Nacional de Educación a Distancia, Apdo. 60141, 28080 Madrid, Spain. On sabbatical leave for the academic year 98−99 at: Service de Physique Statistique, CP231, Faculté de Sciences, ULB, Blvd. du Triomphe, 1050 Brussels, Belgium.

[†] E-mail: sierra@apphys.uned.es

[‡] E-mail: blozano@fisfun.uned.es.

10.1021/jp983041e CCC: $18.00

From a very general point of view, each one of the two necessary, sequential steps for getting the chemical mechanism from an experimental time series in the linear regime around steady state raises its own problems. The first step is the obtainment of the JMEs themselves, which must be accurate enough to capture the structure of the mechanism. This raises inevitably the question of to what extent can we trust those calculated JMEs, which is of course a question the answer of which involves many issues concerning both the sampling of data and their subsequent numerical treatment. The second step, which concerns the inference of the mechanism itself, depends strongly on the amount of information available. What is clear is that the more reliable the obtained Jacobian is, the better shall we be able to pinpoint systematically at a plausible, well-defined mechanism from a single Jacobian. However, within the bounds set by a limited accuracy in delivering the JMEs, the inference procedure faces an optimization problem, which is that of defining what amount of experimental measurements is necessary in order to compensate for possible fuzzy entries of a single, empirical Jacobian. Unfortunately, the whole issue is manifold and cannot be encapsulated in a single, uninterrupted treatment.

In the present article, we restrict ourselves to the description of the basic structure of the algorithm that permits us to carry out the deduction of the mechanisms from the knowledge of the JMEs. Here, those JMEs shall be considered exactly known. Although it is an ideal situation, the bulk of the method will still be applicable to more realistic cases, albeit with appropriate adaptations. The issues concerning the calculation of the JMEs from noisy data, their reliability, and the possible adaptations in the method require separate analysis and publication.

In section II, we outline our procedure for calculating the JMEs from the response in time of the concentrations of reactants following a small pulse perturbation to a stable steady state. Section III is devoted to the development of the general method for reconstructing the chemical mechanisms from the knowledge of the JMEs. A simple example illustrates it. In section IV, we disclose certain features of the method that were not apparent in section III. Section V contains a discussion on directions for future research. We conclude in section VI.

II. Retrieving the Jacobian from Experimental Time Series

The first problem is that of obtaining the Jacobian, **J**, in (1), from experimental measurements. The issue has been reviewed by Chevalier et al.[1] The methods that seem to offer the best results use the measured time response of the reactants to a pulse perturbation around some reference steady state. One of them is the method of multilinear least-squares fit, which seems particularly attractive as far as it involves information of the complete time response. In order to avoid instabilities due to numerical differentiation, Chevalier et al.[1] suggest a fit of the time series to the general solution of the linear equation (1):

$$\delta \mathbf{X}(t) = \sum_{i=1}^{n} \rho^{(i)} e^{\lambda_i t} \tag{2}$$

The main drawback of this approach is that a straightforward fit to a sum of exponentials, with undetermined exponents, λ_i, is sometimes numerically problematic. To overcome this difficulty, Sorribas et al.[4] proposed to reconsider the target problem of the least-squares fit procedure. The idea is the following: if t_i, $i = 1, ..., N$, denote the sampling times, then the formal solution to (1) can be written as

$$\delta \mathbf{X}(t_{i+1}) = \exp[(t_{i+1} - t_i)\cdot\mathbf{J}]\cdot\delta\mathbf{X}(t_i) \tag{3}$$

If the sampling period is a constant, h, (3) yields

$$\delta \mathbf{X}(t_i + h) = \exp[h\cdot\mathbf{J}]\cdot\delta\mathbf{X}(t_i) = \Phi\cdot\delta\mathbf{X}(t_i) \tag{4}$$

The problem is then reduced to a simple multilinear regression with matrix Φ in (4) as the output. A nonnegligible advantage is that the correlation coefficient of the fit gives the possibility of checking whether the measured series is within the linear regime or not. Upon knowing Φ in (4), the eigenvalues of the Jacobian, λ_i, are easily calculated in terms of those of matrix Φ, Λ_i:

$$\lambda_i = \frac{\ln(\Lambda_i)}{h}$$

Once the λ_i have been computed and substituted in (2), the process of fitting a combination of exponential functions to the time series becomes an easy task. The JMEs are hereupon computed by a simple procedure, the technical details of which can be found in the literature.[1,4]

The advantages of the variant just presented are easy to enumerate: (1) no differentiation is needed; (2) eigenvalues and multiplicities are straightforward to calculate; (3) both regressions, on (4) and (2), are linear with unique solution, thus eliminating the sensitivity to an initial guess for problem parameters, as happens in nonlinear regression.

The method yields very satisfactory results in the case of a nondegenerate spectrum but fails at producing the right multiplicity of eigenvalues in the degenerate case. The computation of eigenvalues of matrix Φ is very sensitive to noise and round-off errors, making the approach not to be recommended unless the multiplicities of the eigenvalues are exactly known. In order to avoid the rather problematic computation of eigenvalues, we can resort to a slight variation of the previous method. (4) can formally be rewritten as

$$\mathbf{J} = \frac{1}{h}\ln\Phi = \frac{1}{h}\ln[\mathbf{I} + (\Phi - \mathbf{I})] \tag{5}$$

and expanded in its Taylor series

$$\mathbf{J} = \frac{1}{h}\left[(\Phi - \mathbf{I}) - \frac{(\Phi - \mathbf{I})(\Phi - \mathbf{I})}{2} + ...\right] \tag{6}$$

which holds for $||\Phi - \mathbf{I}|| \leq 1$. The norm is understood in the usual sense for operators.[5] By virtue of the properties of the norm, (6) is valid whenever

$$||\Phi - \mathbf{I}|| \leq e^{||\mathbf{J}||} - 1 \leq 1 \tag{7}$$

Since $||\mathbf{J}||$ is always greater than the modulus of its greatest eigenvalue, $\lambda_{\max}$, (6) is valid whenever

$$h \leq \frac{\ln 2}{|\lambda_{\max}|} \tag{8}$$

Condition 8 means that the sampling frequency must be such that the greatest time scale in the relaxation process toward steady state is caught. By applying (6), we can directly obtain the Jacobian from matrix Φ, without any intermediate search for eigenvalues. The shortcut is certainly beneficial, and the approach yields a faster convergence to the JMEs.

III. Deriving the Reactional Mechanism from the Jacobian

A. General Procedure. We shall now detail the general procedure that permits the derivation of the unknown mechanism starting from the knowledge of the Jacobian. For that we assume the mechanism to be constituted of m generic reactional steps

$$\sum_{i=1}^{n}\alpha_{is}X_i \xrightarrow{k_s} \sum_{j=1}^{n}\beta_{js}X_j, \quad s = 1, ..., m \tag{9}$$

where the values of m, of the stoichiometric coefficients, and of the constants of reaction are to be determined. In order to define the problem in a proper mathematical context, let us write the equations of evolution for species concentrations, according to the law of mass action

$$\frac{dX_i}{dt} = \sum_{s=1}^{m}\gamma_{is}\nu_s + \phi_i, \quad i = 1, ..., n \tag{10}$$

In (10) ν_s stands for the rate of reaction of the sth step

$$\nu_s = k_s\prod_{l=1}^{n}X_l^{\alpha_{ls}}, \quad s = 1, ..., m \tag{11}$$

$\gamma_{is} = \beta_{is} - \alpha_{is}$, and ϕ_i means a controlled constant rate of feeding of the reaction with species X_i, either by chemical or mechanical means.

If system 10 has a steady-state solution, $X_{i,0}$, $i = 1, ..., n$, a straightforward calculation leads to the following relations

$$\sum_{s=1}^{m}\gamma_{is}\nu_{s,0} = -\phi_{i,0}, \quad i = 1, ..., n \tag{12a}$$

$$\sum_{s=1}^{m}\alpha_{js}\gamma_{is}\nu_{s,0} = J_{ij}X_{j,0}, \quad i, j = 1, ..., n \tag{12b}$$

where J_{ij} are the entries of the Jacobian matrix in (1)

$$J_{ij} = \frac{\partial}{\partial X_j}\left(\frac{dX_i}{dt}\right)\Big|_{X_0} \tag{13}$$

We can now define properly the problem. Our purpose is to solve (12) in order to uncover of the *unknown* mechanism 9. That is, we intend to obtain the following:

(a1) the number of steps, m;

(a2) the stoichiometric coefficients, α_{is}, γ_{is}, $i = 1, ..., n$, $s = 1, ..., m$;

(a3) the rate constants, k_s, $s = 1, ..., m$.

This is to be performed by assuming that we know the following:

(b1) the number of species, n;

(b2) the steady-state concentrations $X_{i,0}$, $i = 1, ..., n$;

(b3) the entries of the Jacobian matrix, J_{ij}, which have been supposedly obtained from experimental measurements;

(b4) the constant rates of feeding ϕ_i, $i = 1,.... ,n$.

The key to the solution of this problem is to consider (12) as a set of $n^2 + n$ linear equations for the variables $\nu_{s,0}$, $s = 1, ..., m$, with low integer coefficients, $\gamma_{is}\alpha_{js}$, and known right-hand side ($\phi_{i,0}$, $J_{ij}X_{j,0}$). The first objective is the determination of m (number of steps).

1. Determination of the Number of Steps, m. Let fix our attention on (12b) and handle it as a linear system where m is the number of problem variables, $\nu_{s,0}$, and therefore the number of equations needed to find an explicit solution. If n^2, the actual number of equations (12b), is higher than m, there must be $p \equiv n^2 - m$ redundant equations. As the stoichiometric coefficients, α and γ, are low integers, these p redundant equations must be linear combinations, with low integer coefficients, of the m independent ones. Upon looking at (12b), we can easily infer that all this means that only m entries $J_{ij}X_{j,0}$ are actually independent, and the $p = n^2 - m$ remaining entries must in fact be related to those m independent ones by linear combinations with low integer coefficients.

Thus, to determine the value of m, one has to identify the independent entries in the matrix of coefficients $J_{ij}X_{j,0}$. This can be done by a simple automatic procedure that spots at possible integer dependences among those entries. The result is a set of p relations of the form

$$\sum_{i,j}^{n} I_{ij}^{(r)}(J_{ij}X_{j,o}) = 0, \quad r = 1,..., p \tag{14}$$

where the $I_{ij}^{(r)}$ are in the form of low integers. Consequently, $m = n^2 - p$.

2. Equations for the Stoichiometric Coefficients and Their Determination. Once we have obtained the value of m, all we have to do is to insert (14) into (12b) and, by considering all $\nu_{s,0}$ linearly independent, obtain

$$\sum_{i,j}^{n} I_{ij}^{(r)}(\alpha_{js}\gamma_{is}) = 0, \quad r = 1, ..., p, \quad s = 1, ..., m \tag{15}$$

What can be inferred from (15) is not yet the actual numerical values of the coefficients α's and γ's. In fact, for each of the m reactional steps, (15) only provides p nonlinear relationships among $2n$ variables. This implies that the α's and γ's can actually be produced from (15) as parameter-dependent functions

$$\alpha_{is}(\mu_1, ..., \mu_R), \quad \gamma_{is}(\mu_1, ..., \mu_R) \tag{16}$$

In practice, many of those functions (16) are found to be identically null because most reactional steps do not involve more than a fraction of all n species. On the other hand, those that are nonnull are in the form of very simple linear combinations of the parameters μ_R.

The specific values of the μ_R's are found in the following way:

(1) Substitute (16) into the subset of m independent equations in (12b). The result is in the form of m functions

$$\nu_{s,0}(\mu_1, ..., \mu_R) \tag{17}$$

(2) Put (17) into (12a) and solve for the μ_R's in terms of the known feeding rates, ϕ_i. This yields

$$\mu_R(\phi_i, ..., \phi_n) \tag{18}$$

which solves, together with (16), the problem of the stoichiometric coefficients.

3. Determination of the Rate Constants, k_s. The only remaining unknowns are the rates $\nu_{s,0}$, that can now be straightforwardly recovered from (17). These values, together

TABLE 1: Input Data in Example 1

Constant feeding rates: $\phi_1 = 4; \phi_2 = 1.11; \phi_3 = 0.3$
Steady state: $X_{1,0} = 0.804674; X_{2,0} = 0.715431; X_{3,0} = 3.36161$
Jacobian: $\mathbf{J} = \begin{pmatrix} -6.58031 & 0 & -0.804674 \\ 3.2187 & -3.36161 & -0.715431 \\ -3.36161 & 3.36161 & -0.0892429 \end{pmatrix}$
Matrix $\boldsymbol{\Psi}$: $\boldsymbol{\Psi} = \begin{pmatrix} -5.295 & 0 & -2.705 \\ 2.59 & -2.405 & -2.405 \\ -2.705 & 2.405 & -0.3 \end{pmatrix}$

with (11), lead to the computation of the rate constants k_s. The mechanism is then fully retrieved and the procedure terminated.

All this process can be illustrated by the following example.

B. Example 1. We may understand the previous procedure in all its details by checking what it gives in the simple system[6]

$$2X_1 \xrightarrow{k_1} 2X_2 \tag{19a}$$

$$X_2 + X_3 \xrightarrow{k_2} 2X_3 \tag{19b}$$

$$X_1 + X_3 \xrightarrow{k_3} P \tag{19c}$$

Let us assume that we are supplied with the values $X_{i,0}$ and ϕ_i, $i = 1, 2, 3$, and with the Jacobian matrix, as shown in Table 1. We wish to recover from those data the complete mechanism (19). We start by constructing matrix Ψ, of entries $\Psi_{ij} = J_{ij}X_{j,0}$, as displayed in Table 1. With this information at hand we proceed by applying the previous procedure step by step.

It is quite easy to verify that just three of the entries of matrix Ψ are independent: $J_{11}X_{1,0}$, $J_{22}X_{2,0}$, and $J_{33}X_{3,0}$. Thus the number of steps, m, is equal to three as expected.

Additionally, we get $p = 6$ linear relationships (see (14)).

$$\begin{gathered}
J_{12}X_{2,0} = 0 \\
J_{13}X_{3,0} = J_{22}X_{2,0} + J_{33}X_{3,0} \\
J_{21}X_{1,0} = -J_{11}X_{1,0} + J_{22}X_{2,0} + J_{33}X_{3,0} \qquad (20) \\
J_{23}X_{3,0} = J_{22}X_{2,0} \\
J_{31}X_{1,0} = J_{22}X_{2,0} + J_{33}X_{3,0} \\
J_{32}X_{2,0} = -J_{22}X_{2,0}
\end{gathered}$$

By inserting (20) into (12b), we find the particular realization of (15) for the present case:

$$\begin{gathered}
\alpha_{2s}\gamma_{1s} = 0 \\
\alpha_{3s}\gamma_{1s} = \alpha_{2s}\gamma_{2s} + \alpha_{3s}\gamma_{3s} \\
\alpha_{1s}\gamma_{2s} = -\alpha_{1s}\gamma_{1s} + \alpha_{2s}\gamma_{2s} + \alpha_{3s}\gamma_{3s} \\
\alpha_{3s}\gamma_{2s} = \alpha_{2s}\gamma_{2s} \qquad (21) \\
\alpha_{1s}\gamma_{3s} = \alpha_{2s}\gamma_{2s} + \alpha_{3s}\gamma_{3s} \\
\alpha_{2s}\gamma_{3s} = -\alpha_{2s}\gamma_{2s}
\end{gathered}$$

where

$$s = 1, 2, 3$$

We get three parameter-dependent solutions to (21)

$$(\alpha_{1s}, \alpha_{2s}, \alpha_{3s}, \gamma_{1s}, \gamma_{2s}, \gamma_{3s}) = \begin{cases} (a, & 0, & 0, & -d, & d, & 0) \\ (0, & b, & b, & 0, & -e, & e) \\ (c, & 0, & c, & f, & 0, & f) \end{cases} \tag{22}$$

where

$$s = 1, 2, 3$$

with a, b, c being natural numbers, and d, e, f, integers.

Each one of the row vectors in (22) stands for one of the $m = 3$ independent steps of the mechanism. Consequently, from what we know up to now, we can already write a generic prototype mechanism

$$\begin{gathered}
aX_1 \rightarrow (a-d)X_1 + dX_2 \\
bX_2 + bX_3 \rightarrow (b-e)X_2 + (b+e)X_3 \\
cX_1 + cX_3 \rightarrow (c+f)X_1 + (c+f)X_3 \qquad (23)
\end{gathered}$$

$$d = 1, 2, ..., a; \quad e = -b, ... -1, 1, ..., b; \quad f = -1, -2, ..., -c$$

The range of values for the parameters d to f have been set such as to satisfy the requirements $\beta_{ij} \geq 0$ in (9).

Taking into account (12b), (22), and the values of Ψ in Table 1, we solve for the $\nu_{s,0}$ and get

$$\nu_{1,0} = \frac{2.59}{ad} \tag{24}$$

$$\nu_{2,0} = \frac{2.405}{be}$$

$$\nu_{3,0} = \frac{-2.705}{cf}$$

We substitute (22) and (24) into (12a)^{i}, only a, b, and c appear explicitly in the resulting system, which, once solved, yields

$$a = \frac{5.18}{\phi_1 - \phi_2 - \phi_3} \tag{25}$$

$$b = \frac{4.81}{\phi_1 + \phi_2 - \phi_3}$$

$$c = \frac{5.41}{\phi_1 + \phi_2 + \phi_3}$$

In Table 1 we find the values for the ϕ's. Accordingly, $(a, b, c) = (2, 1, 1)$ and, from (23), $(d, e, f) = (\{1, 2\}, \{-1, 1\}, -1)$. Parameters d and e are double-valued. The requirement of positive rates constants, k_s, will partially solve this indeterminacy, as we shall see now.

Once the stoichiometric coefficients have been found, system 11, together with (24) and the obtained values of the parameters a to f, permits the computation of the rate constants, k_s. The value $e = -1$ is thus eliminated if $k_2 > 0$. However, both values for d are compatible with $k_1 > 0$ and thus perfectly admissible. We therefore recover the original system 19, with the same values ($k_1 = k_2 = k_3 = 1$) initially given to the rate constants in order to obtain numerically our starting Jacobian. Actually, the derived mechanism is the same as (19), except for the first reaction in (23), which can take any one of the following forms:

$$2X_1 \xrightarrow{k_1} 2X_2 \quad (\textit{for } d = 2)\text{, or} \tag{26}$$

$$2X_1 \xrightarrow{2k_1} X_1 + X_2 \quad (\textit{for } d = 1)$$

In fact, both reactions in (26) are dynamically indistinguishable, as they lead to identical rate laws. We thus cannot discriminate between the two alternatives exclusively on information obtained from the dynamics.

IV. The Case of Missing Feeding Rates, ϕ_i

Let us assume that for a number l ($l < n$) of the reacting species, the feeding rates are null. From (12a)

$$\sum_{s=1}^{m} \gamma_{is}\nu_{s,0} = 0, \quad i = 1, ..., l \tag{27}$$

In fact, we can actually be more general and assume that there exist $l < n$ relationships of the form

$$\sum_{j=1}^{n} Q_{ij}\phi_j = 0, \quad i = 1, ..., l \tag{28}$$

with integer coefficients Q_{ij}, and such that the rank($\mathbf{Q}$) is maximal. From (28), we have

$$\sum_{s=1}^{m} (Q\cdot\gamma)_{is}\nu_{s,0} = 0, \quad i = 1, ..., l \tag{29}$$

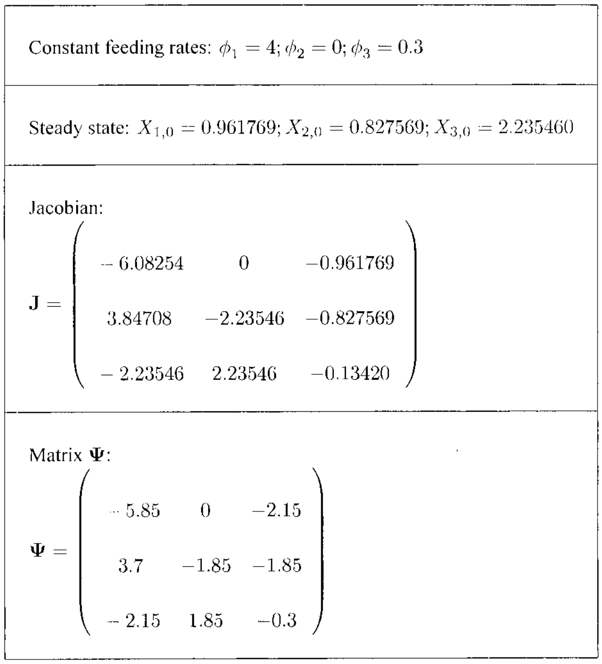
which contains (27) as a particular case.

When following the procedure of section III, relations (29) do slightly affect the calculation of the number of steps, m. To see this, let us assume that we compute, as in section III, the number of independent entries $J_{ij}X_{j,0}$. We shall presently denote the result by m'. The reason by which m' is different from m—the actual number of steps—becomes apparent when we observe that, by virtue of (29), we have only $(m - l)$ independent rates $\nu_{s,0}$. This number has to coincide with the number m' of independent equations in (12b). That is, $m' = m - l$. We must then have a number of steps that is given by

$$m = m' + l \tag{30}$$

The following example explains the difference introduced with respect to what we had in section III.

TABLE 2: Input Data in Example 2



Constant feeding rates: $\phi_1 = 4; \phi_2 = 0; \phi_3 = 0.3$

Steady state: $X_{1,0} = 0.961769; X_{2,0} = 0.827569; X_{3,0} = 2.235460$

Jacobian:

$$\mathbf{J} = \begin{pmatrix} -6.08254 & 0 & -0.961769 \\ 3.84708 & -2.23546 & -0.827569 \\ -2.23546 & 2.23546 & -0.13420 \end{pmatrix}$$

Matrix $\mathbf{\Psi}$:

$$\mathbf{\Psi} = \begin{pmatrix} -5.85 & 0 & -2.15 \\ 3.7 & -1.85 & -1.85 \\ -2.15 & 1.85 & -0.3 \end{pmatrix}$$

A. Example 2. Let us deal with the same model (19) of section III. In order to illustrate what has been said at the beginning of this section, we make $\phi_2 = 0$. Obviously, in comparison to example 1, that choice changes the steady state values and the entries of the Jacobian matrix. In Table 2 we display the new data of interest for this case.

By inspecting matrix Ψ in Table 2 we find only two independent entries: $J_{22}X_{2,0}$ and $J_{33}X_{3,0}$, while the others comply to

$$\begin{aligned} J_{11}X_{1,0} &= 3J_{22}X_{2,0} + J_{33}X_{3,0} \\ J_{12}X_{2,0} &= 0 \\ J_{13}X_{3,0} &= J_{22}X_{2,0} + J_{33}X_{3,0} \\ J_{21}X_{1,0} &= -2J_{22}X_{2,0} \\ J_{23}X_{3,0} &= J_{22}X_{2,0} \\ J_{31}X_{1,0} &= J_{22}X_{2,0} + J_{33}X_{3,0} \\ J_{32}X_{2,0} &= -J_{22}X_{2,0} \end{aligned} \tag{31}$$

If we were to proceed for (31) as we did for (20), we would infer that the model has only two steps ($m = 2$). However, it is easy to check that this conclusion does lead to a system (15) that has no solution. As said before, this inconsistency is avoided if, at this point, we take into account the requirement

$$\sum_{s=1}^{m} \gamma_{2s}\nu_{s,0} = -\phi_{2,0} = 0 \tag{32}$$

Then, following the notation at the beginning of the section, $m' = 2$ and $l = 1$, and, by (30), the number of steps is actually $m = 3$. According to (32) one of the rates of reactions, say $\nu_{3,0}$, is linearly dependent on the others

$$\nu_{3,0} = -\frac{\sum_{s=1}^{2} \gamma_{2s}\nu_{s,0}}{\gamma_{23}} \tag{33}$$

There are still two independent entries $J_{ij}X_{j,0}$ in (12b), but three rates, $\nu_{s,0}$, related through (33).

Introducing (31) and (33) into (12b), and assuming $\nu_{1,0}$ and $\nu_{2,0}$ are independent, we obtain

$$\begin{gathered}
(\alpha_{1s}\gamma_{1s} - 3\alpha_{2s}\gamma_{2s} - \alpha_{3s}\gamma_{3s}) + \left(-\alpha_{13}\frac{\gamma_{13}}{\gamma_{23}}\gamma_{2s} + 3\alpha_{23}\gamma_{2s} + \alpha_{33}\frac{\gamma_{33}}{\gamma_{23}}\gamma_{2s}\right) = 0 \\
(\alpha_{2s}\gamma_{1s}) + \left(-\alpha_{23}\frac{\gamma_{13}}{\gamma_{23}}\gamma_{2s}\right) = 0 \\
(\alpha_{3s}\gamma_{1s} - \alpha_{2s}\gamma_{2s} - \alpha_{3s}\gamma_{3s}) + \left(-\alpha_{33}\frac{\gamma_{13}}{\gamma_{23}}\gamma_{2s} + \alpha_{23}\gamma_{2s} + \alpha_{33}\frac{\gamma_{33}}{\gamma_{23}}\gamma_{2s}\right) = 0 \\
(\alpha_{1s}\gamma_{2s} + 2\alpha_{2s}\gamma_{2s}) + (-\alpha_{13}\gamma_{2s} - 2\alpha_{23}\gamma_{2s}) = 0 \qquad (34) \\
(\alpha_{3s}\gamma_{2s} - \alpha_{2s}\gamma_{2s}) + (-\alpha_{33}\gamma_{2s} + \alpha_{23}\gamma_{2s}) = 0 \\
(\alpha_{1s}\gamma_{3s} - \alpha_{2s}\gamma_{2s} - \alpha_{3s}\gamma_{3s}) + \left(-\alpha_{13}\frac{\gamma_{33}}{\gamma_{23}}\gamma_{2s} + \alpha_{23}\gamma_{2s} + \alpha_{33}\frac{\gamma_{33}}{\gamma_{23}}\gamma_{2s}\right) = 0 \\
(\alpha_{2s}\gamma_{3s} + \alpha_{2s}\gamma_{2s}) + \left(-\alpha_{23}\frac{\gamma_{33}}{\gamma_{23}}\gamma_{2s} - \alpha_{23}\gamma_{2s}\right) = 0 \\
s = 1, 2
\end{gathered}$$

The three simplest, chemically meaningful, parameter-dependent solutions to (34) are

$$(\alpha_{1s}, \alpha_{2s}, \alpha_{3s}, \gamma_{1s}, \gamma_{2s}, \gamma_{3s}) = \begin{cases} (2b, & 0, & 0, & -d, & d, & 0) \\ (0, & b, & b, & 0, & -e, & e) \\ (c, & 0, & c, & f, & 0, & f) \end{cases} \tag{35}$$

where

$$s = 1, 2, 3$$

with b, c being natural numbers, and d, e, f, integers.

In comparison to (22), (35) depends on five parameters—instead of six. Requirement $\phi_2 = 0$ is responsible for this decrease in the number of free parameters. The corresponding three-step prototype mechanism follows from (35):

$$\begin{gathered}
2bX_1 \rightarrow (2b - d)X_1 + dX_2 \\
bX_2 + bX_3 \rightarrow (b - e)X_2 + (b + e)X_3 \qquad (36) \\
cX_1 + cX_3 \rightarrow (c + f)X_1 + (c + f)X_3
\end{gathered}$$

$$d = 1, 2, ..., 2b; \quad e = -b, ... -1, 1, ..., b; \quad f = -1, -2, ..., -c$$

Notice that we obtain the same mechanism as in (23). The only difference is on the value of parameter a in (22), which is now restricted to $a = 2b$ in (36).

We can now reproduce the procedure of section III and determine parameters in (36) from the values of ϕ_1 and ϕ_3. We arrive at the same conclusions: that is, except for the indeterminacy (26) in the first reaction, we recover model (19), with rate constants $k_1 = k_2 = k_3 = 1$.

V. Directions for Future Research

We have developed in previous sections a procedure for the establishment of a detailed mechanism in chemical networks, starting from *well-defined* Jacobian matrix elements. However, as pointed out in the Introduction, a model-maker chemist faced with the task of obtaining a Jacobian from finite-precision or noisy experimental measurements will not always be able to satisfy this requirement. As we have already mentioned, a detailed solution to this problem demands a lengthy treatment that is somewhat out of the scope of this article. Nevertheless, this does not prevent us from discussing some of the issues at stake and outlining possible alternatives to reach the objective of a prototypical mechanism. Those issues are currently under study.

The methods of section II have been found to work well.[4] However, a performing data analysis will never compensate for an uninformative time series. This is the reason why, even within experimental limitations, the pulse perturbation and subsequent sampling must be conducted with a carefully designed strategy. Chevalier et al.[1] have already stressed the importance in discriminating *essential* from *nonessential* species, if one wishes to choose an appropriate species for the initial pulse perturbation. The response to a perturbation in an essential species will provide a balanced information on all time scales—eigenvalues—while that of a nonessential species will not. Even so, the perturbation strategy may well be further improved. In many cases, a “cocktail” of different species proves to yield a better response than a perturbation of a single species. A procedure may be devised for calculating the optimal direction of perturbation in phase space. For this purpose, a “learning” algorithm using eigenvectors of computed Jacobians is currently under development.

The sampling rate also deserves some attention, especially in cases in which a very fast time scale dominates the relaxation. A careless sampling may not provide a sufficient number of points for catching appropriately that fast time scale.

On the other hand, we infer from sections III and IV the necessity of extracting two sequential pieces of information from the JMEs. One concerns the existing relationships among the entries $J_{ij}X_{j,0}$, which lead to the establishment of the minimal number of steps of the mechanism and the corresponding relationships among the stoichiometric coefficients, (21) and (34) and, consequently, to the prototype mechanisms (23) and (36). Some of those relationships among the $J_{ij}X_{j,0}$ can be spotted even in the case of uncertain JMEs. We have made some preliminary calculations in the case of time series with limited precision. Even from these inevitably inexact output values for the JMEs, and within a given tolerance, certain relationship patterns start to appear. However, a criterion is needed for distinguishing which are the real, structural equalities, as in (31), and which are only apparent and due to uncertainties in the computed JMEs. One possibility, which we are currently investigating, is to change the feeding constants in order to get Jacobians in some alternative steady states. It is expected that the comparison of several Jacobian matrices will provide the necessary discriminating information for uncovering the actual, structural relationships between entries and discard the misconceived ones. A strong argument in favor of this belief is that the latter will not be invariant under a change of Jacobian.

The second piece of information is more stringent on the exactness of the calculated values of the Jacobian and also demands a precise control of the fluxes, ϕ_i. We recall that it permitted to give specific numerical values to the stoichiometric coefficients and rate constants. We can assume that the fluxes can be eventually controlled and known with precision. However, if the values of the JMEs are know with limited precision, nothing can be done and we shall have to make do with it in order to roughly calculate the stoichiometric coefficients and rate constants. These latter values can serve as first estimate to guide the experimentalist in designing further tests on the system. To the theoretician, they may be useful as initial guesses in training algorithms designed for matching a parameter-dependent set of nonlinear equations to a time series, once we have a first approximate to the prototypical mechanism; it may not be as concise as (23) and (36), but it can provide workable parameter-dependent nonlinear equations of evolution. Any of the methods[7−10] developed for fitting nonlinear equations can then be applied more successfully to the present case.

VI. Conclusions

We have presented a method that delivers a complete reaction mechanism starting from the knowledge of a Jacobian and the constant feeding rates of the involved species. We thus improve previous approaches,[1] which relied essentially on guessing possible reactional steps from the signs of the JMEs.

The idea is the same that has motivated previous works on the issue:[1,2] to avoid the ambiguities on defining a model when trying to do so from nonlinear data. The correspondence between the time series and the equations of evolution is unambiguous in the linear regime. Thus the obtainment of the Jacobian is more a matter of technical detail than a really conceptual one. We have not addressed in depth the specific problem of the computation of the JMEs, though we have outlined some of the methods that we are currently developing. However, a later use of nonlinear time series should not be disregarded because, once a prototypical mechanism has been determined from the Jacobian, some of the model parameters may have been left undetermined. This may happen when the required information—Jacobian and/or feeding rates—is incomplete.

In order to be fully completed, our procedure in its present state requires unequivocal values in data. It can nevertheless arrive at a prototypical, basic mechanism from the single knowledge of relations among the JMEs, but not necessarily that of their exact values. Although, in that case, specific stoichiometric coefficients and rate constants cannot be computed, that prototypical mechanism may prove of great help and can be completed by some additional information—nonlinear, for example. A deeper treatment of these issues will be the matter of a future publication.

Acknowledgment. This work has been supported by the DGICYT (Spain), under Grant No. PB94-0390, and by the European Contract ESPRIT WG24490-CATHODE2. R.D.S. acknowledges a predoctoral fellowship from Comunidad Autónoma de Madrid (Spain).

References and Notes

(1) Chevalier, T.; Schreiber, I.; Ross, J. *J. Phys. Chem.* **1993**, *97*, 6776.
(2) Bar-Eli, K.; Geiseler, W. *J. Phys. Chem.* **1983**, *87*, 1352.
(3) Olsen, L. F.; Degn, H. *Biochem. Biophys. A.* **1978**, *523*, 321. Degn, H.; Olsen, L. F.; Perram J. W. *Ann. N. Y. Acad. Sci.* **1979**, *316*, 623.
(4) Sorribas, A.; Lozano, J. B.; Fairén, V. *Recent Res. Dev. Phys. Chem.* **1998**, *2*, 553−573.
(5) Hirsh, M.; Smale, S. *Differential equations, dynamical systems and linear algebra*; Academic Press: Orlando, FL, 1974.
(6) Hanusse P. Ph.D. Thesis. Université de Bordeaux, France, 1976.
(7) Breeden, J. L.; Hübler, A. *Phys. Rev. A* **1990**, *42*, 5817.
(8) Eisenhamer, T.; Hübler, A.; Packard, N.; Kelso, J. A. S. *Biol. Cybern.* **1991**, *65*, 107.
(9) Gouesbet, G. *Phys. Rev. A* **1991**, *43*, 5321.
(10) López, V.; Huerta, R.; Dorronsoro, J. R. *Neur. Comput.* **1993**, *5*, 795.