Stratton, J. A., *Electromagnetic Theory*, McGraw-Hill, New York (1941).
Watson, J. H. P., "Magnetic Filtration," *J. App. Phys.*, **44**, 4209 (1973).
Zebel, G., "Deposition of Aerosol Flowing Past a Cylindrical Fiber in a Uniform Electrical Field," *J. Colloid Sci.*, **20**, 522 (1965).
Supplementary material has been deposited as Document No. 00000 with the National Auxiliary Publications Service (NAPS), c/o Microfiche Publications, 214-13 Jamaica Avenue, Queens Village, N.Y. 11428, and may be obtained for $4.00 for microfiche or $7.75 for photocopies.

Manuscript received May 28, 1981; revision received November 18, and accepted December 14, 1982.

One Variety of Quick Estimation Schemes for Parameters in Kinetic Models

Quick estimation of the parameters in a process model can play a significant role in kinetic studies. This work is aimed at developing a quick estimation strategy by means of process discretization and evaluating errors engendered by the strategy. A systematic error evaluation method and guidelines to select the best discretization relation used in the strategy are described.

K. WATANABE
and D. M. HIMMELBLAU
Department of Chemical Engineering
The University of Texas
Austin, TX 78712

SCOPE

We examine a number of aspects of estimating the coefficients in kinetic models that are not widely known. In particular we examine the class of methods in which the derivatives in the differential equations are approximated by a proper difference relation to yield a set of algebraic equations that can be solved for the coefficients. Such a procedure makes the estimation of the coefficients quite simple and is quicker than repeatedly integrating the differential equations themselves followed by least squares. But the discretization leads to an approximate model. To obtain reliable estimates of the coefficients one needs to know how the approximation incurred in the discretization influences the errors in the estimates. One also wants to control the size of the errors.

We describe how to develop a general and quick estimation strategy using commonly encountered discretization schemes and evaluate the errors in the coefficient estimates. We also provide guidelines as to how to select the best discretization scheme for a given estimation problem.

CONCLUSIONS AND SIGNIFICANCE

The analysis presented here leads to a straightforward interpretation of how errors in kinetic coefficients are engendered in quick estimation procedures in which differences are substituted for derivatives. The relative error in a discretization scheme was presented, and was utilized to check whether or not the coefficient estimates were reliable. Guidelines were cited as to how to select the best discretization scheme. From the results of this study, one can clearly determine the limitations of an estimation procedure and can select the best discretization scheme for a given estimation procedure.

INTRODUCTION

Estimation of unknown parameters in dynamic processes by means of quick and simple methods can make a significant contribution to adaptive process control (Landau, 1979) and process simulation as well as incipient process fault detection and diagnosis (Himmelblau, 1978). Such estimates can also be used to generate initial guesses of the unknown parameters for an optimal parameter seeking method (Glowinski and Stocki, 1981).

In general, parameter estimation techniques can be divided into three categories: (a) the extended Kalman filter (Jazwinski, 1970); (b) iterative methods in which iterative numerical integration of a set of ordinary differential equations is followed by an iterative optimization algorithm (Himmelblau et al., 1967; Froment, 1975; Pexidr, 1974; Seinfeld, 1970; Watanabe and Shimizu, 1975); and (c) noniterative methods in which a process model is discretized by a proper integration formula and then the set of resulting algebraic equations is solved simultaneously (Glowinski and Stocki, 1981).

The methods in categories (a) and (b) engender a heavy computational load, but if properly designed, can yield quite accurate

K. Watanabe is on leave from Hosei University at Koganei, Tokyo, Japan.

estimates of the model parameters. On the other hand, the methods in category (c) involve only a modest amount of computation so that they are attractive as methods of quick parameter estimation. We treat here only the methods in category (c). In particular, we will describe how the discretization scheme employed in the estimation strategy influences the error in the parameter estimates, and describe a method of checking whether or not the sampling interval of an integration formula violates the limitation imposed on the discretization scheme for a given process model. Further, we will provide guidelines as to how to select the best discretization relation for a given estimation problem.

ESTIMATION BY DISCRETIZATION

Consider a process model described by the following set of nonlinear differential equations:

$$\dot{\underline{x}}(t) = \underline{f}(\underline{x},t,\underline{p}), \quad \underline{x}(0) = \underline{x}_o \tag{1a}$$

$$\underline{y}(t) = \underline{x}(t) \tag{1b}$$

with an n-dimensional state vector $\underline{x}(t)$, an m-dimensional unknown parameter vector $\underline{p}$, an n-dimensional process function $\underline{f}(\ )$, and an n-dimensional full state noise free measurement vector $\underline{y}(t)$. To discretize Eq. 1a let us select a commonly used first-derivative multistep integration formula

$$\underline{x}(k+1) = \sum_{l=0}^{q} a_l \underline{x}(k-l) + h \sum_{l=-1}^{r} b_l \dot{\underline{x}}(k-l) \tag{2}$$

where h is the time step in integration formula 2 and corresponds to the sampling interval of the measurements $\underline{y}(t)$, and a_l and b_l are coefficients of the integration formula. (A more general integration formula than Eq. 2 would be a high-order derivative multistep integration method, a method that has more flexibility than Eq. 2, but involves a significantly greater computational load because of the calculation of the approximates for the high-order derivatives of $\underline{f}(\underline{x},t,\underline{p})$.)

Substitution of Eq. 1b into Eq. 1a and substitution of the resulting equation into Eq. 2 lead to a set of algebraic equations:

$$h \sum_{l=-1}^{r} b_l \underline{f}(\underline{y},k-l,\underline{p}) = \left[\underline{y}(k+1) - \sum_{l=0}^{q} a_l \underline{y}(k-l)\right] \tag{3a}$$

for $k = r, r+1, r+2, \ldots$.

Solution of the set of algebraic equations (Eq. 3a) for the unknown parameters $\underline{p}$ yields the desired estimates.

Suppose the process model (1) is linear in the parameters. Then the process function $\underline{f}(\underline{y},t,\underline{p})$ can be expressed as

$$\underline{f}(\underline{y},t,\underline{p}) = [F(\underline{y},t)]\underline{p} \tag{3b}$$

where $F(\underline{y},t)$ is an $n \times m$ dimensional matrix containing the elements of the nonlinear parameter free functions. Substitution of Eq. 3b into Eq. 3a leads to

$$\left[h \sum_{l=-1}^{r} b_l F(\underline{y},k-l)\right]\underline{p} = \left[\underline{y}(k+1) - \sum_{l=0}^{q} a_l \underline{y}(k-l)\right] \tag{3c}$$

for $k = r, r+1, r+2, \ldots$ which is a set of linear algebraic equations. It is known that if

$$\text{rank}\left[h \sum_{l=-1}^{r} b_l F^T(\underline{y},r-l) \quad h \sum_{l=-1}^{r} b_l F^T(\underline{y},r+1-l) \ldots\right]^T = m \tag{3d}$$

Equation 3c can have a unique solution, and the solution can be obtained by matrix inversion or by simple linear elementary operations.

Now, suppose that Eq. 1 is nonlinear in the parameters. Then we would have to employ iterative calculations such as the Newton-Raphson method to solve the set of nonlinear algebraic equations (Eq. 3a), and a unique solution is not guaranteed. Even a local solution may require extensive computation. Nevertheless, the procedures outlined above are very simple to execute even if an iterative computation to solve the algebraic equations is involved because the procedures avoid solving a set of ordinary differential equations iteratively, certainly a more complex technique. The elements in the estimate $\underline{p}$, the solution of Eq. 3a, are not the best estimates of the unknown parameters because of errors in the estimates due to:

(a) The numerical solution $\underline{p}$ of Eq. 3a possibly being nonoptimal if $\underline{f}(\underline{y},k-l,\underline{p})$ is nonlinear in the parameters

(b) The solution $\underline{p}$ being biased even if $\underline{f}(\underline{y},k-l,\underline{p})$ is linear in the parameters because the measurements $\underline{y}(k)$ may include noise with a large variance

(c) The discrete model (3a) obtained by discretizing Eq. 1 is only an approximate model so that the parameters estimated by using the approximate discrete model can yield erroneous values even if $\underline{f}(\underline{y},k-l,\underline{p})$ is linear, i.e., $\underline{f}(\underline{y},k-l,\underline{p}) = A(\underline{p})\underline{x}(k-l)$, and the measurements $\underline{y}(k)$ are noise free

Among the above three causes, (c) is most likely to cause difficulty in practice for the inexperienced analyst, particularly if the equations in process (1) are stiff. We will focus our attention on (c), and show how the approximate discretization degrades the accuracy of the parameter estimates as well as how to choose the best discretization scheme (integration formula) to discretize a given process model. For theoretical strictness and simplicity, the analysis will assume that the process model is linear. The theoretical results can be applied to nonlinear processes if the measured variables fall within the region in which linearization of the process is justified.

RELATION OF THE ERRORS IN THE PARAMETER ESTIMATES VS. APPROXIMATE DISCRETIZATION

Consider a process model described by the following set of linear differential equations:

$$\dot{\underline{x}}(t) = A(\underline{p})\underline{x}(t), \quad \underline{x}(0) = \underline{x}_o \tag{4a}$$

$$\underline{y}(t) = \underline{x}(t) \tag{4b}$$

where $A(\underline{p})$ is an $n \times n$ dimensional system coefficient matrix.

Our basic viewpoint in evaluating errors in the parameter estimates is that the unknown process parameters should be estimated so that the estimated characteristic roots of the process model will have values close to the true roots of the model. We select this criterion because in the estimation of dynamic processes the characteristic roots are the basic factors that determine the dynamic character of the process.

Since model (4) is linear, without any loss of generality, Eqs. 4 can be diagonalized and can be rewritten as a set of n-independent scalar equations:

$$\dot{z}_i(t) = \lambda_i(\underline{p}) z_i(t), \quad z_i(0) = z_{io} \quad \text{for } i = 1,2,\ldots,n \tag{5a}$$

$$\underline{y}(t) = T(\underline{p})[z_1(t) z_2(t) \ldots z_n(t)]^T \tag{5b}$$

where the $z_i(t), i = 1,2,\ldots,n$, are state variables of the diagonalized model, $\lambda_i(\underline{p}), i = 1,2,\ldots,n$, are the eigenvalues (characteristic roots) of $A(\underline{p})$, and $T(\underline{p})$ is an $n \times n$ dimensional diagonalization matrix for $A(\underline{p})$. Let $\lambda_i(\underline{p})$ and $z_i(t)$ be λ and $z(t)$, respectively. We will suppress the subscript i in what follows to simplify the notation. Then the theoretical discretization of any one of the n-independent scalar equations in Eq. 5a can be written as

$$z(k+1) = e^{\lambda h} z(k), \quad z(0) = z_o \tag{6a}$$

or in general

$$z(k-l) = e^{-(l+1)\lambda h} z(k+1) \tag{6b}$$

Substitution of Eq. 5a into the integration formula 2 and substitution of Eq. 6b into the resulting relation lead to

$$z(k+1) \cong \sum_{l=0}^{q} a_l e^{-(l+1)\lambda h} z(k+1) + \lambda h \sum_{l=-1}^{r} b_l e^{-(l+1)\lambda h} z(k+1) \tag{7a}$$

in which the approximately equal sign ($\simeq$) is used because of the approximation introduced by the discretization relation. Equation 7a can be solved implicitly for λh

$$\lambda h \cong \frac{1 - \sum_{l=0}^{q} a_l e^{-(l+1)\lambda h}}{\sum_{l=-1}^{r} b_l e^{-(l+1)\lambda h}} = L(\lambda h) = \bar{\lambda} h \qquad (7b)$$

where $\bar{\lambda}$ is the approximate of λ and is determined by the value of time step h, and by the coefficients $a_l, l = 0,1, \ldots, q$, and $b_l, l = -1,0, \ldots, r$, in the approximate discretization relation. We can interpret Eq. 7b as showing how the approximate discretization of a linear ordinary differential equation by the discretization relation deforms the value of a characteristic root λ into another value $\bar{\lambda}$ which is determined by the integration relation employed.

Let $\epsilon_{\lambda h}$ be the permissible relative error of the value of the product (deformed value of a characteristic root) × (time step), i.e., $\bar{\lambda} h$, from the true value λh. Then the set of points λh on the λh-plane, which satisfy

$$\frac{|\lambda h - L(\lambda h)|}{|\lambda h|} \leqq \epsilon_{\lambda h} \qquad (7c)$$

determine the region R in which the relative error of $\bar{\lambda} h$ is less than $\epsilon_{\lambda h}$. From Eq. 7b, since the value λh is uniquely mapped to $\bar{\lambda} h$, we can determine the same region R on the $\bar{\lambda} h$-plane. Of course, each discretization scheme involves a different-size region, and the accuracy of the scheme is proportional to the size of the region R.

EXAMPLE

Suppose the measurements are noise-free and are given by

$$1, \ldots, e^{\lambda k h}, e^{\lambda(k+1)h}, e^{\lambda(k+2)h}, \ldots$$

and suppose the model is known to be

$$\dot{x}(t) = px(t), \quad x(0) = x_o$$

with one unknown parameter p. The true value of p is equal to λ. Discretization of the model by the trapezoidal rule

$$x(k+1) = x(k) + \frac{h}{2}[\dot{x}(k+1) + \dot{x}(k)]$$

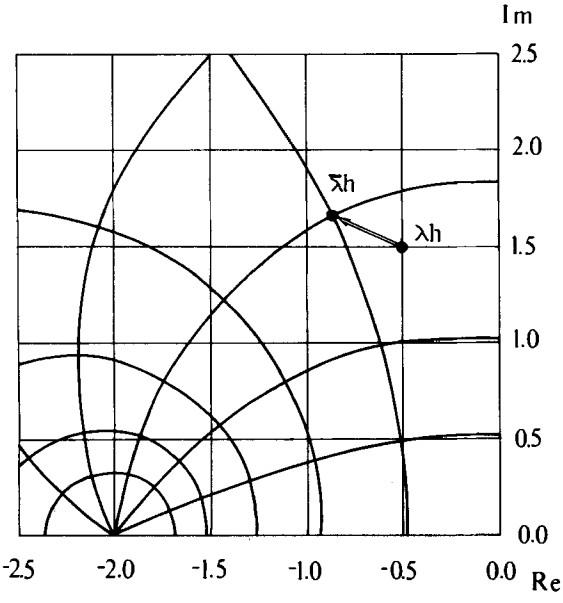


Figure 1. The grid represents locus of a characteristic root λh; the curves, the locus of the deformed characteristic root $\bar{\lambda} h$.

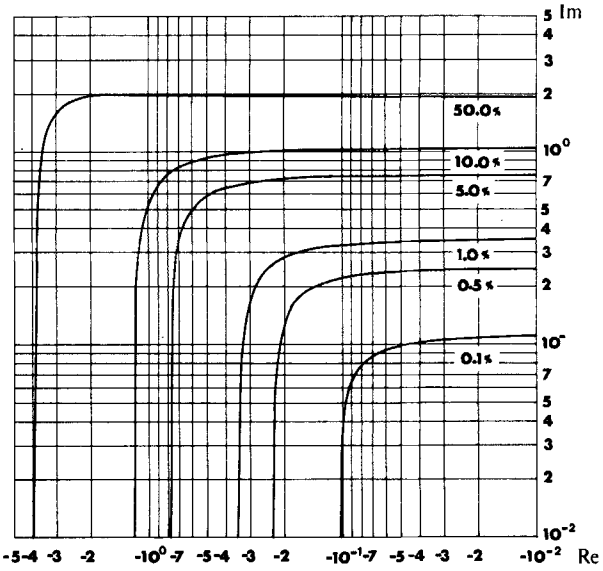


Figure 2(a). The curves form the boundaries of the regions R in the λh-plane for the trapezoidal rule for various permissible relative errors $\epsilon_{\lambda h}$ ranging from 0.1 to 50%.

leads to

$$x(k+1) = x(k) + \frac{h}{2}p[x(k+1) + x(k)]$$

Substitution of the measurements $e^{\lambda k h}$ and $e^{\lambda(k+1)h}$ for $x(k)$ and $x(k+1)$ leads to

$$e^{\lambda(k+1)h} - e^{\lambda k h} = \frac{h}{2}p[e^{\lambda(k+1)h} + e^{\lambda k h}]$$

which yields for the estimate of p

$$\hat{p} = \frac{2}{h} \cdot \frac{e^{\lambda h} - 1}{e^{\lambda h} + 1} = \bar{\lambda}$$

Note the same relation can be obtained directly from Eq. 7b by letting $a_o = 1$, $b_{-1} = 1/2$, $b_o = 1/2$ and $a_l = 0$, $b_l = 0$ for $l > 1$.

Suppose $h = 1$, and we select four successive values of λ, namely $\lambda = -0.1$, $\lambda = -1.0$, $\lambda = -10.0$ and $\lambda = -100.0$. The estimated parameters are

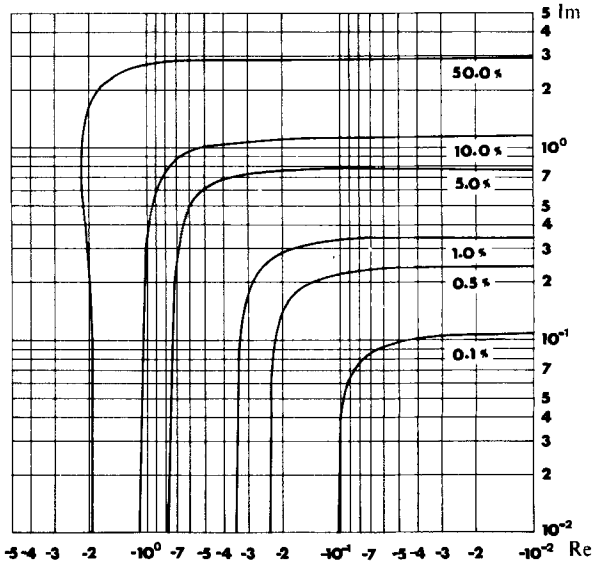


Figure 2(b). The curves form the boundaries of the regions R in the $\bar{\lambda} h$-plane for the trapezoidal rule for various permissible relative errors $\epsilon_{\lambda h}$ ranging from 0.1 to 50%.

$$\text{for } \lambda = -0.1 \quad \hat{p} = \bar{\lambda} = -0.0999168$$

$$\text{for } \lambda = -1.0 \quad \hat{p} = \bar{\lambda} = -0.9242343$$

$$\text{for } \lambda = -10.0 \quad \hat{p} = \bar{\lambda} = -1.9998184$$

$$\text{for } \lambda = -100.0 \quad \hat{p} = \bar{\lambda} = -2.0$$

Observe the asymptotic trend in $\bar{\lambda}$.

More generally, if a characteristic root λ of the given equations is a complex number such as represented in Figure 1 by the grid itself, the root obtained via the estimation procedure outlined above has the value $\bar{\lambda}$ shown in Figure 1 by the curved lines. Figures 2(a) and 2(b) show the region R on the λh-plane and the $\bar{\lambda} h$-plane for several values of $\epsilon_{\lambda h}$. We conclude that the use of an approximate discretization for derivatives and parameter estimation using the resulting algebraic equations leads to estimation error, and the error is proportional to the magnitude of the value of characteristic roots.

In the above example, we illustrated only a case involving use of the trapezoidal rule. One could analyze the error for any integration formula (discretization scheme) by following the same procedure outlined above.

SELECTION OF APPROPRIATE DISCRETIZATION

In the previous section, we showed how the discretization scheme used for parameter estimation is derived from the function $L(\lambda h)$ in Eq. 7b. Picking a good discretization scheme involves determining the appropriate width of sampling interval, the number of steps (q,r), and the values of the coefficients $a_l, l = 0,1, \ldots, q$ and $b_l, l = -1,0, \ldots, r$. We next summarize some remarks that are useful in deciding on the choice of the width of the sampling interval, the number of steps, and the values of the coefficients.

Width of Sampling Interval

Determination of the width of a sampling interval in process estimation itself is a difficult problem (Box and Jenkins, 1976). We will describe a method of selecting the sample interval to reduce the error in the parameter estimates. For a discretization scheme and process model, we want to find the maximum sampling interval which guarantees a specified accuracy for the discrete model. For a given sampling interval, we would like to determine whether or not the parameter estimates are satisfactory.

First, derive a discrete model equivalent to Eq. 3c for a selected sampling interval h, and estimate the unknown parameters in the linearized process model using the discrete model and sampled measurements. Next, obtain the product (characteristic root) $\times$ (sample interval) $= \bar{\lambda}_i h, i = 1,2, \ldots, n$, for the linearized process model. Then, based on the discussion in the previous section, we have the following guideline:

- For the discretization scheme employed, if all the $\bar{\lambda}_i h, i = 1,2, \ldots, n$, values fall within the region R under the specified $\epsilon_{\lambda h}$, the discretization error in the estimation procedure is permissible. Otherwise, the estimates are not reliable. For a permissible case, if h is not fixed, we can elongate the sampling interval h up to a value such that one of the $\bar{\lambda}_i h, i = 1,2, \ldots, n$, reaches the boundary of the region R.

The above guideline assists in discriminating between reliable and unreliable estimates, especially when the measurements involve nonequal sampling intervals.

Number of Discretization Steps

In general, the accuracy of the discretization scheme increases in proportion to the number of steps (q,r) for a proper choice of coefficients. But the computation load is also increased in proportion to (q,r). In solving Eq. 3a, an increase in r yields a heavier computational load than an increase in q because the frequency of generation of the Jacobian matrix $[\partial \underline{f}(\underline{y},t,\underline{p})/\partial \underline{p}^T]$ (or $F(\underline{y},k-l)$ in Eq. 3b if $\underline{f}(\underline{y},k-l,\underline{p})$ is linear in the parameters) and the summation of the matrices which require long computation times increases in proportion to r, whereas the frequency of products of vectors and scalars, and summations of the vectors which can be calculated quickly, increases in proportion to q. (See Eq. 3b.) Note if we choose the number of steps (q,r) such that $q \leq 1$ and $r \leq 1$, the discretization scheme can be applied to data with nonequal sampling intervals; otherwise, the discretization scheme should be applied to data with same sampling intervals. Further, if the process model is nonlinear, since most integration relations (Eq. 7b) are developed assuming that the process model is linear, the numbers of steps (q,r) should be determined such that the set of measured variables $\underline{x}(k-q)$, or $\underline{x}(k-r), \ldots, \underline{x}(k)$, are within the valid linearization region.

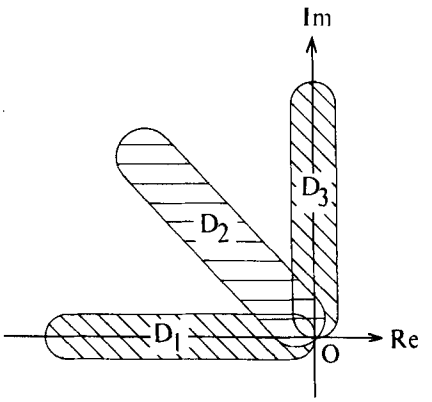
Determination of Values of Coefficients

The specific technique to calculate the values of the coefficients $a_l, l = 0,1, \ldots, q$, and $b_l, l = -1,0, \ldots, r$, of the discretization relation (Eq. 2 or 7b) depends on the characteristics of the trajectory of the measurements. If the series of measurements $\{\underline{y}(k), k = 0,1, \ldots\}$ has a nonoscillatory trajectory, the coefficients should be determined such that the deformed root $\bar{\lambda}$ of Eq. 7b is a good approximation of λ in domain D_1 of Figure 3. If the series of measurements has a damped oscillatory trajectory, the coefficients should be determined such that $\bar{\lambda}$ is good approximation of λ in domain D_2, and if the series of measurements has an oscillatory trajectory, the coefficients should be determined such that $\bar{\lambda}$ is good approximation of λ in domain D_3. The values of the coefficients can be determined either by: (a) minimizing a cost function via iterative nonlinear programming; or (b) an analytical function approximation method. Neither of these methods yields parameter estimates that have the optimal statistical properties of consistency and unbiasedness.

Determination of Values of Coefficients by Minimizing Cost Function. The usual cost functions in category (a) are either (i) the squared error function

$$J_o(a_o,a_1, \ldots, a_q, b_{-1}, b_o, \ldots, b_r) = \iint_{D_l} w(\lambda h) \left| \frac{1 - \sum_{l=0}^{q} a_l e^{-(l+1)\lambda h}}{\sum_{l=1}^{r} b_l e^{-(l+1)\lambda h}} \right|^2 d\xi d\eta \quad (8a)$$

where $\lambda h = \xi + i\eta, w(\lambda h)$ is a weighting function that can be se-



D_1: domain of roots corresponding to nonoscillatory behavior

D_2: domain of roots corresponding to damped oscillatory behavior

D_3: domain of roots corresponding to oscillatory behavior

Figure 3. Characteristic roots corresponding to various trajectory behavior.

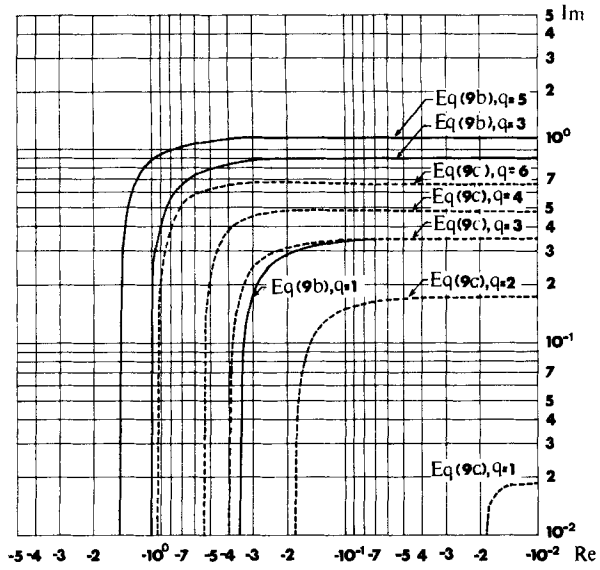


Figure 4(a). The curves form the boundaries of the regions R on λh-plane for $\epsilon_{\lambda h} = 1\%$ using Eqs. 9b and 9c.

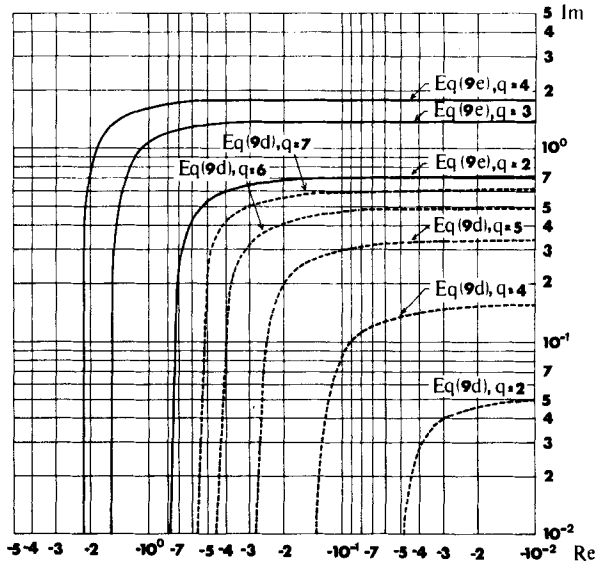


Figure 4(b). The curves form the boundaries of the regions R on λh-plane for $\epsilon_{\lambda h} = 1\%$ using Eqs. 9d and 9e.

lected by the analyst, and D_i is the domain for integration, or (ii) the maximum error function

$$J_1(a_o,a_1,\ldots,a_q,b_{-1},b_o,\ldots,b_r) = \max_{\lambda h}\left|\lambda h - \frac{1-\sum_{l=0}^{q} a_l e^{-(l+1)\lambda h}}{\sum_{l=-1}^{r} b_l e^{-(l+1)\lambda h}}\right| \quad (8b)$$

for the specified domain of λh. In the minimization of one of the two cost functions, we need at least the following constraints:

$$L(\lambda h)\Big|_{\lambda h=0} = 0 \text{ i.e., } \sum_{l=0}^{q} a_l = 1. \quad (8c)$$

$$\left.\frac{d^{(i)}L(\lambda h)}{d(\lambda h)^{(i)}}\right|_{\lambda h=0} = \begin{cases} 1 & \text{for } i = 1 \\ 0 & \text{for } i \geq 2 \end{cases} \quad (8d)$$

Only if the discretization scheme requires very high accuracy in the vicinity of $\lambda h = 0$ are the constraints (Eq. 8d) needed. These constraints guarantee that $\overline{\lambda h} \to \lambda h$ as $h \to 0$, which means that the choice of a smaller sampling interval for the measurements of $y(k)$ leads to more accurate discretization.

Analytical Function Approximation. As the following relationship exists

$$\lambda h = ln(e^{\lambda h}) \quad (9a)$$

one can show that the rational polynomial equation of $e^{\lambda h}$ in Eq. 7b is an approximation of the function $ln(e^{\lambda h})$. Therefore, we can obtain an analytic function approximation such as Eq. 7b by sets of approximate logarithmic functions.

From the Taylor expansion around $z = 0$ for the functions $ln(1 + z)$ and $ln\{(1 + z)/(1 - z)\}$, we have

$$ln(e^{\lambda h}) \cong (e^{\lambda h} - 1)\left\{\sum_{l=0}^{q-1} \frac{[-(e^{\lambda h}-1)]^l}{l+1}\right\}_3 \quad |e^{\lambda h} - 1| < 1 \quad (9b)$$

$$ln(e^{\lambda h}) \cong 2\left\{\sum_{l=0}^{(q-1)/2} \frac{1}{2l+1}\left(\frac{e^{\lambda h}-1}{e^{\lambda h}+1}\right)^{2l+1}\right\} \quad (q-l)/2 \text{ integer}, \quad 0 < e^{\lambda h} \quad (9c)$$

and from the Padé approximation for the function $z^{-1}ln(1 + z)$ and $zln(1 + 1/z)$, we have

$$ln(e^{\lambda h}) = (e^{\lambda h} - 1) + \frac{a_1}{a_o}(e^{\lambda h}-1)^2 + \frac{a_2}{a_o}(e^{\lambda h}-1)^3 + \ldots + \frac{a_{q-1}}{a_o}(e^{\lambda h}-1)q \quad 1 < e^{\lambda h} < 3 \quad (9d)$$

$$ln(e^{\lambda h}) \cong (e^{\lambda h}-1)\frac{\frac{1}{r_{q-1}}\left\{a_o + a_1\left(\frac{1}{e^{\lambda h}-1}\right) + a_2\left(\frac{1}{e^{\lambda h}-1}\right)^2 + \ldots + a_n\left(\frac{1}{e^{\lambda h}-1}\right)^{q-1}\right\}}{\frac{1}{b_o}\left\{b_o + b_1\left(\frac{1}{e^{\lambda h}-1}\right) + b_2\left(\frac{1}{e^{\lambda h}-1}\right)^2 + \ldots + b_n\left(\frac{1}{e^{\lambda h}-1}\right)^{q-1}\right\}} \quad |arg(e^{\lambda h})| < \pi \quad (9e)$$

The coefficients can be found in the literature (Luke, 1969).

To evaluate the performance of each of the two discretization relations obtained by the function expansion method, we calculated the apparent values $\overline{\lambda h}$ of the characteristic root for a complex root λh and obtained the region R on the λh-plane in which the relative error of the apparent root, $\epsilon_{\lambda h}$ $(= |\lambda h - \overline{\lambda h}|/|\lambda h|)$, was less than 1%. Figure 4 shows the regions R for each of the relations 9b, 9c, 9d and 9e. From these figures it appears that relation 9e has the best properties for $\lambda h \leq 0$. Among the relations with $r = 0$, i.e., Eqs. 9b and 9d, Eq. 9b is slightly better than Eq. 9d.

NUMERICAL EXPERIMENTS

Consider the kinetic model used by Seinfeld and Gavalas (1970). These equations are not stiff.

$$\dot{x}_1 = -r_1 - r_2; \quad x_1(0) = 1.0 \quad (10a)$$

$$\dot{x}_2 = (r_1/2) - r_2; \quad x_2(0) = 0. \quad (10b)$$

where

$$r_1 = k_1[x_1^2 - x_2(2 - 2x_1 - x_2)/0.726] \quad (10c)$$

$$r_2 = k_2[x_1x_2 - (1 - x_1 - 2x_2)(2 - 2x_1 - x_2)/3.852] \quad (10d)$$

The known values of the kinetic parameters are $k_1 = 347$ and $k_2 = 403$. The measurements include noise $e_i(t_j)$ such that

$$y_i(t_j) = x_i(t_j) + e_i(t_j), \quad i = 1,2; \quad j = 1,2,\ldots,8 \quad (10e)$$

The sampling times and data are listed in the left side of Table 1; we assume the data at $t_5 = 34.0$ are missing. Model (10) is dicretized by using the trapezoidal rule and the kinetic parameters are estimated at each sampling interval using the sets of data $\{y_1(t_j), y_2(t_j), y_1(t_{j+1})$, and $y_2(t_{j+1})\}$, $j = 0,1,\ldots,7$ (excluding $j = 5$). The values of the estimated parameters listed in the fourth and fifth columns

TABLE 1. ERROR IN THE ESTIMATES

$t\cdot10^4$	Test Data y_1	Test Data y_2	Estimation Results $\hat{k}_1$	Estimation Results $\hat{k}_2$	$\epsilon_{\lambda h}$
0.0	1.0	0.0	350.924	491.247	1.10%
5.73	0.828	0.0737	354.581	408.804	1.02%
11.32	0.704	0.1130	343.924	377.416	0.99%
16.97	0.622	0.1322	338.455	424.042	1.01%
22.62	0.565	0.1400	307.765	379.192	10.40%
34.0					
39.7	0.482	0.1477	334.749	431.423	1.03%
45.2	0.470	0.1477	82.091	107.535	75.40%
169.7	0.433	0.1476			

of Table 1 are the same as those obtained by Glowinski and Stocki (1981) except for the time interval $22.64 \leq t \leq 39.7$ because of the assumed missing data.

We took the estimates $\hat{k}_1$ and $\hat{k}_2$ and linearized the process model at each sampling point and obtained the product (characteristic root of the estimated model) $\times (t_{j+1} - t_j) = \overline{\lambda h}$. From Figure 2(b), or from Eqs. 7c and 7b, the relative errors of the maximum value of $\overline{\lambda h}$ at each sampling point were evaluated as shown in the right side of Table 1. The estimates for the sampling intervals $22.62 \leq t \leq 39.7$ and $45.2 \leq t \leq 169.7$ have larger relative errors than the other estimates. The average values of the estimates at each sampling interval excluding the above two estimates are $\hat{k}_1 = 365.53$ and $\hat{k}_2 = 426.59$ which are quite satisfactory.

Next, consider another example using the same kinetic model as above but with the different known values of the kinetic parameters, namely, let $k_1 = 1{,}000$ and $k_2 = 100$. We assume the measurements are sampled data containing observation noise whose standard deviation is 1.0×10^{-5} and the measurements are taken at equal sampling intervals $h = 10$. This model is relatively stiff and represents typical problems faced in data analysis. Since the sampling interval is relatively large for the characteristic roots of the linearized model, a one-step discretization scheme cannot discretize the process model correctly. For example, with $h = 10$, the parameters estimated using the trapezoidal rule were $\hat{k}_1 = 758.18$ and $\hat{k}_2 = 161.45$. Instead, we select the following multistep discretization schemes which use a set of four sampled data and are derived from Eqs. 9:

from Eq. 9b with $q = 3$

$$\underline{x}(k) = (1/2)\underline{x}(k-1) - \underline{x}(k-2) + (11/2)\underline{x}(k-3) + 3\underline{f}(\underline{x},k-3) \quad (11a)$$

from Eq. 9c with $q = 3$

$$\underline{x}(k) = \underline{x}(k-3) + (3/8)\underline{f}(\underline{x},k) + (9/8)\underline{f}(\underline{x},k-1) + (9/8)\underline{f}(x,k-2) + (3/8)\underline{f}(\underline{x},k-3) \quad (11b)$$

from Eq. 9d with $q = 3$

$$\underline{x}(k) = 6.3\underline{x}(k-1) - 17.1\underline{x}(k-2) + 11.8\underline{x}(k-3) + 7.5\underline{f}(\underline{x},k-3) \quad (11c)$$

from Eq. 9e with $q = 3$

$$\underline{x}(k) = -18\underline{x}(k-1) + 9\underline{x}(k-2) + 10\underline{x}(k-3) + 9\underline{f}(\underline{x},k-1) + 18\underline{f}(\underline{x},k-2) + 3\underline{f}(\underline{x},k-3) \quad (11d)$$

The estimation results using the above discretization formulas are listed in Table 2. In the incipient stages (from step 0 to step 2) of estimation, the accuracy of the estimates with respect to their true values is low because the state variables change drastically from (1., 0.) to (0.5, 0.221), and the variables are not within the region in which linearization of the process is justified. Among the four estimates at the incipient stages, the estimate using the discretization relation (Eq. 11d) gives the best results; the estimates obtained by using the other three relations (Eqs. 11a, 11b and 11c) include very large errors whereas the result by Eq. 11d does not. After the large transients have settled down (after five steps), all the estimates yield satisfactory results. The estimates by Eq. 11a, 11b and 11d have almost same accuracy, and the estimate by Eq. 11c has lower accuracy than the other three.

DISCUSSION

In the first example using the model presented by Seinfeld and Gavalas (1970), the estimates of k_1 and k_2 for the last interval differ considerably from the earlier estimates. Glowinski and Stocki (1981) explained that the phenomena occurred because the kinetic model equations near chemical equilibrium are different from the assumed model. Our analysis ascribes the phenomena to a different cause.

Since (a) the observation noise level is very low and since (b) the model in the vicinity of the last time interval can be linearized satisfactorily [the characteristic roots at $t = 45.2$ (−653.8, −605.1) and those at $t = 169.7$ (−661.2, −605.9) are respectively very close in value], the error evaluation approach described here can be applied almost exactly to the last interval in the data of Seinfeld and Gavalas. The predicted relative error at the last time interval is 75.4%. This large deviation means that the estimated values of the characteristic roots of the linearized process are very different from the true values. We believe that this large error in the apparent roots was induced by the choice of the large sampling interval. The characteristic roots of the linearized model using the erroneous estimates of $k_1 = 82.091$ and $k_2 = 107.535$ are calculated to be [−169.55, −169.78]. For the case in which the true roots are [−661.2, −605.9], the calculated λh (the deformed characteristic roots) are [−160.55, −160.47] using discretization relation (Eq. 7b) and the coefficients of the trapezoidal rule. Thus, the roots using the erroneous estimates have values very close to the calculated roots of the linearized model; but, of course, both sets of values are wrong.

CONCLUSIONS

We have examined some of the factors that influence the success or failure of quick estimation methods for the coefficients in sets

TABLE 2. ESTIMATION OF PARAMETERS USING THE DISCRETIZATION RELATIONS EQS. 11a, 11b, 11c and 11d

$t\cdot10^4$	y_1	y_2	Eq. 11a $\hat{k}_1$	Eq. 11a $\hat{k}_2$	Eq. 11b $\hat{k}_1$	Eq. 11b $\hat{k}_2$	Eq. 11c $\hat{k}_1$	Eq. 11c $\hat{k}_2$	Eq. 11d $\hat{k}_1$	Eq. 11d $\hat{k}_2$
0.0	1.0000	0.0000	724.24	-7.6×10^5	906.13	107.09	636.45	-8.3×10^5	971.85	103.57
10.0	0.5632	0.2049	711.04	101.09	886.41	100.12	621.12	100.22	962.77	100.55
20.0	0.5002	0.2207	628.71	100.11	-1.3×10^4	99.98	512.67	99.05	2,069.90	99.97
30.0	0.4873	0.2133	1,560.59	99.91	1,020.74	100.03	1,693.40	98.84	1,009.61	99.96
40.0	0.4845	0.2043	1,042.17	99.68	997.72	100.07	1,030.54	98.72	1,002.10	100.14
50.0	0.4769	0.1961	1,013.70	100.37	983.21	100.12	991.89	99.23	980.89	100.10
60.0	0.4728	0.1891	967.04	99.91	976.18	100.03	952.83	98.84	974.86	100.03
70.0	0.4692	0.1831	976.50	100.00	981.09	100.02	957.24	98.90	982.93	100.04
80.0	0.4659	0.1780								
90.0	0.4630	0.1736								
100.0	0.4604	0.1699								

of ordinary differential equations. We indicated how the particular relation selected for discretization of the derivatives in the process model influences the accuracy of the estimates of the coefficients. It is quite possible for one discretization method to yield highly biased estimates if the equations are stiff whereas selection of another method and/or time interval for taking measurements would yield quite satisfactory results. Based on the examples used here, one can apply the guidelines of error evaluation to various other kinetic models that may or may not be linear in the coefficients.

NOTATION

a,b = coefficients in integration (discretization) formula
A = process coefficient matrix
D = complex domain for integration in a cost function
e = observation noise
$\underline{f}$ = process function vector
F = parameter free matrix of process
h = integration (discretization) time step and sampling interval
J_o,J_1 = cost function to obtain an optimal discretization relation
k = discrete time variable
k_1,k_2 = kinetic parameters
L = discretization relation
m = number of unknown parameters
n = number of state variables
$\underline{p}$ = vector of unknown process parameters
p = unknown process parameter
q,r = step number of integration (discretization)
r_1,r_2 = rate of first and second reactions
R = region in which relative error of apparent characteristic roots is less than a specified value
t = time variable
T = process diagonalization matrix
$\underline{x}$ = state vector
$\underline{y}$ = measurement vector
z = state variable of the diagonalized process
$\epsilon_{\lambda h}$ = relative error of the apparent characteristic root
λ = characteristic root of process
$\overline{\lambda}$ = apparent characteristic root
$\hat{}$ = estimate

LITERATURE CITED

Box, G. E. P., and G. M. Jenkins, *Time Series Analysis: Forecasting and Control*, Holden-Day, San Francisco (1976).
Froment, G. F., "Model Discrimination and Parameter Estimation in Heterogeneous Catalysis," *AIChE J.*, **21**, 1041 (1975).
Glowinski, J., and J. Stocki, "Estimation of Kinetic Parameters—Initial Guess Generation Method," *AIChE J.*, **27**, 1041 (1981).
Himmelblau, D. M., C. R. Jones, and K. B. Bischoff, "Determination of Rate Constants for Complex Kinetics Model," *Ind. Eng. Chem. Fund.*, **6**, 539 (1967).
Himmelblau, D. M., *Fault Detection and Diagnosis in Chemical and Petrochemical Processes*, Elsevier Scientific Publishing Co., New York (1978).
Jazwinski, A. H., *Stochastic Processes and Filtering Theory*, Academic Press, New York (1970).
Landau, Y. D., *Adaptive Control, The Model Reference Approach*, Marcel Dekker Inc., New York (1979).
Luke, Y. L., *The Special Functions and Their Approximation*, Academic Press, New York (1969).
Pexidr, V., "Estimation of Nonlinear-Regression Kinetic Data III. Evaluation of Integral Kinetic Data," *Chem. Prum.*, **24**, 254 (1974).
Seinfeld, J. H., "Nonlinear Estimation Theory," *Ind. Eng. Chem.*, **62**, 32 (1970).
Seinfeld, J. H., and G. H. Gavalas, "Analysis of Kinetic Parameters from Batch and Integral Reaction Experiments," *AIChE J.*, **16**, 644 (1970).
Watanabe, K., and H. Shimizu, "Parameter Optimization Software Using Padé's Computation of Transient Response," *Electrical Eng. in Japan*, **95**, 119 (1975).

Manuscript received March 1, 1982; revision received October 5, and accepted October 20, 1982.

Moving Boundary Problems in Simple Shapes Solved by Isotherm Migration

The Modified Isotherm Migration Method (MIMM) is developed for solving one-dimensional Stefan-type problems of multifront propagation in slabs, cylinders and spheres with a boundary condition of the third kind (radiation-type boundary condition). The method is illustrated for all three geometries with applications to the thawing of ice and to the freezing of a 50–50 dispersion of tetradecane in water, a system in which two freezing fronts propagate simultaneously.

Y. TALMON
Department of Chemical Engineering
Technion—Israel Institute of Technology
Haifa, Israel 32000
and
H. T. DAVIS and
L. E. SCRIVEN
Department of Chemical Engineering
and Materials Science
University of Minnesota
Minneapolis, MN 55455

SCOPE

Classical analytical solutions have been found to just a few simple cases of moving boundary problems. Although modern analysis by computer-aided finite element methods is now showing great promise (Comini et al., 1974; Bonnerot and Jamet, 1977), most numerical solutions have been based on the finite difference approximation. One such scheme suitable especially