$M$ = number of state variables in the objective function
$m$ = number of state variables
$N$ = last stage or the total number of stages
$n$ = $n^{\text{th}}$ stage
$\mathbf{p}$ = vector defined by Equation (8)
$q$ = flow rate of raffinate stream
$u$ = outlet flow rate of extract stream
$v$ = state variable
$w$ = control variable or feed rate of extract solvent
$x$ = concentration of component $C$
$y$ = concentration of component $A$
$\lambda$ = Lagrange multiplier

**Subscripts**

$E$ = extract stream
$k$ = $k^{\text{th}}$ iteration, assumed known
$k+1$ = $(k+1)$st iteration, assumed unknown
$R$ = raffinate stream

## LITERATURE CITED

1. Lee, E. S., *AIChE J.*, 14, No. 6, 908 (Nov., 1968).
2. ———, "Quasilinearization and Invariant Imbedding," Academic Press, New York (1968).
3. Bellman, R., "Introduction to Matrix Analysis," McGraw-Hill, New York (1960).
4. Lee, E. S., *Ind. Eng. Chem. Fundamentals*, 3, 373 (1964).
5. Bellman, R., and S. Dreyfus, "Applied Dynamic Programming," Princeton Univ. Press, N. J. (1962).

*Manuscript received November 18, 1968; revision received January 20, 1969; paper accepted January 22, 1969. Paper presented at AIChE New Orleans meeting.*

---

# Analysis of Kinetic Parameters from Batch and Integral Reaction Experiments

**JOHN H. SEINFELD and GEORGE R. GAVALAS**

**California Institute of Technology, Pasadena, California**

**The estimation of chemical reaction rate constants from experimental batch or integral reactor data is examined. Consideration is given to the effect on the accuracy of the parameter estimates of the number of concentrations measured, extending the range of measurements to higher conversions, and experimental errors. The study is based on estimates of two rate constants obtained by quasilinearization in the pyrolytic dehydrogenation of benzene. It is found that if fewer than the number of independent components are to be measured, certain components produce better parameter estimates than others. The level of experimental errors is by far the most significant factor in obtaining reliable parameter estimates.**

Several recent studies have been concerned with the nonlinear least-squares analysis of reaction rate data (*3, 7, 8, 9*). In each of these methods the sum of squares of the reaction rates $r$ are minimized. If differential reactor data are available, this represents an acceptable method of obtaining rate constants. Often, however, considerable problems arise in the experimental determination of differential reaction rates due to the small concentration changes involved. If the study is conducted in a batch or integral flow reactor, the experimental data are in the form of concentration vs. time (or reciprocal space velocity) measurements. The reaction rates $r$ must then be obtained by numerical differentiation of the experimental data, a highly inaccurate procedure.

The direct utilization of integral reactor data is a problem of estimation of parameters in ordinary differential equations. Several general techniques have been used for the solution of this problem (*2, 11*), and one of them, quasilinearization, has been shown to be particularly attractive because of its quadratic convergence characteristics (*2*). Bellman et al. (*1*) used quasilinearization to estimate the forward and reverse rate constants in the oxidation of nitric oxide from the classic data of Bodenstein and Lindner. Lee (*10*) estimated the Peclet number in a tubular reactor with axial mixing by means of quasilinearization. Donnelly and Quon (*4*) used quasilinearization to estimate rate constants for a series of monomolecular first-order reactions and the nonisothermal pyrolysis of propane.

An important part of the estimation problem is the assignment of confidence intervals or other measures of the reliability or accuracy of the estimated kinetic constants. An appropriate procedure for obtaining confidence intervals for parameter estimates in differential equations is presented by Rosenbrock and Storey (*11*), and a thorough example of the use of such an analysis is given by Heineken et al. (*5*).

The present note represents the results of a detailed numerical analysis with quasilinearization used on the estimation of rate constants for the pyrolytic dehydrogenation of benzene [two independent reactions in a tubular flow reactor for which experimental data are available in the literature (*6*)]. The following are studied:

1. The efficiency and range of convergence of the quasilinearization technique.
2. The number of components to be measured. If the number of reactions is $R$, there are $R$ independent components. However, it may not always be necessary to measure all $R$ concentrations, and, if not, it is important to determine which ones will yield the best estimates.
3. The effect of extending the range of measurements to higher conversions. To what extent should a particular reaction be carried so that estimates derived from the experimental observations are most meaningful.
4. The effect of the level of experimental errors on the accuracy of the estimates.

Our approach to 2 to 4 is analogous to that of (*5*).

## PYROLYSIS OF BENZENE

Hougen and Watson (*6*) present integral reactor data on the pyrolytic dehydrogenation of benzene to diphenyl and tripheryl:

$$2C_6H_6 \leftrightharpoons C_{12}H_{10} + H_2$$

$$C_6H_6 + C_{12}H_{10} \leftrightharpoons C_{18}H_{14} + H_2$$

Data are given for the concentrations of benzene and diphenyl as functions of reciprocal space velocity at $T$ = 1,265°F. and 1 atm.

The kinetic model proposed by Hougen and Watson (*6*) is

$$\frac{dx_1}{dt} = -r_1 - r_2 \quad ; x_1(0) = 1 \qquad (1)$$

$$\frac{dx_2}{dt} = \frac{r_1}{2} - r_2 \quad ; x_2(0) = 0 \qquad (2)$$

where

$$r_1 = k_1 \, [x_1^2 - x_2 \, (2 - 2x_1 - x_2)/3K_1] \qquad (3)$$

$$r_2 = k_2 \, [x_1x_2 - (1 - x_1 - 2x_2)(2 - 2x_1 - x_2)/9K_2] \qquad (4)$$

with $x_1$ and $x_2$ the pound-moles of benzene and diphenyl per pound-mole of pure benzene feed.

The experimental measurements $y_1$ and $y_2$ at values of the reciprocal space velocity $t$ as given by Hougen and Watson are:

| $t \times 10^4$ | 5.63 | 11.32 | 16.97 | 22.62 | 34.0 | 39.7 | 45.2 | 169.7 |
|---|---|---|---|---|---|---|---|---|
| $y_1$ | 0.828 | 0.704 | 0.622 | 0.565 | 0.499 | 0.482 | 0.470 | 0.443 |
| $y_2$ | 0.0737 | 0.113 | 0.1322 | 0.1400 | 0.1468 | 0.1477 | 0.1477 | 0.1476 |

The measurements are related to the states $x_1$ and $x_2$ by

$$y_1 = x_1 + \eta_1 \quad ; \quad y_2 = x_2 + \eta_2 \qquad (5)$$

where $\eta_1$ and $\eta_2$ indicate random experimental errors. The equilibrium constants $K_1$ and $K_2$ were determined from the run at lowest space velocity to be 0.242 and 0.428.

TABLE 1. EFFECT OF INITIAL GUESS ON CONVERGENCE

| Case | Initial guess | | Final values | | Iterations |
|---|---|---|---|---|---|
| | $k_1$ | $k_2$ | $k_1$ | $k_2$ | |
| 1 | 100 | 100 | Diverged | | — |
| 2 | 300 | 300 | 347.443 | 403.132 | 4 |
| 3 | 500 | 500 | 347.402 | 402.751 | 4 |
| 4 | 1,000 | 1,000 | Diverged | | — |

TABLE 2. EFFECT OF WEIGHTING ON $k_1$ AND $k_2$

| Case | Initial guess | | | | Final values | | Iterations |
|---|---|---|---|---|---|---|---|
| | $k_1$ | $k_2$ | $q_1$ | $q_2$ | $k_1$ | $k_2$ | |
| 5 | 300 | 300 | 1 | 5 | 347.386 | 404.196 | 5 |
| 6 | 300 | 300 | 1 | 0 | 351.124 | 366.366 | 4 |
| 7 | 300 | 300 | 0 | 1 | 346.826 | 401.076 | 7 |

TABLE 3. EFFECT OF FEWER DATA POINTS ON $k_1$ AND $k_2$

| Case | Initial guess | | Data points | Final values | | Iterations |
|---|---|---|---|---|---|---|
| | $k_1$ | $k_2$ | $M$ | $k_1$ | $k_2$ | |
| 8 | 300 | 300 | 6 | 346.421 | 402.129 | 4 |
| 9 | 300 | 300 | 4 | 346.890 | 398.742 | 4 |
| 10 | 300 | 300 | 3 | 346.853 | 406.032 | 4 |

## ESTIMATION OF $k_1$ AND $k_2$

Our first task is the estimation of the rate constants $k_1$ and $k_2$ by minimizing the least-square expression

$$J = \sum_{i=1}^{M} [\mathbf{y}(t_i) - \mathbf{x}(t_i, \mathbf{k})]^T \mathbf{Q}_i [\mathbf{y}(t_i) - \mathbf{x}(t_i, \mathbf{k})] \qquad (6)$$

where $\mathbf{y} = [y_1, y_2]$, $\mathbf{x} = [x_1, x_2]$, and $\mathbf{k} = [k_1, k_2]$. $\mathbf{x}(t, \mathbf{k})$ is the solution of Equations (1) and (2). $\mathbf{Q}_i$ are a series of 2 × 2 symmetric, positive-definite weighting matrices, which for simplicity will be taken diagonal. $M$ is the number of measurements, in this case $M = 8$. The minimization of $J$ with respect to $\mathbf{k}$ can be carried out by the quasilinearization algorithm, for the details of which we refer the reader to (*1*). The convergence criterion used is

$$\frac{|k_1^{i+1} - k_1^i|}{k_1^i} + \frac{|k_2^{i+1} - k_2^i|}{k_2^i} \leq \epsilon \qquad (7)$$

where the superscript $i$ indicates the $i^{\text{th}}$ iteration, and $\epsilon$ is chosen as small as desired.

First, the range of initial guesses for which the quasilinearization algorithm converges was examined. In the following numerical results, with all eight data points, the elements $q_1$ and $q_2$ of $\mathbf{Q}_i$ were $q_1 = 1$ and $q_2 = 1$, with $\epsilon = 0.01$ (see Table 1).

Figure 1 has been prepared from Table 1 to illustrate the approximate region of convergence of initial guesses in the $k_1 - k_2$ plane. Obviously, many further trials with other sets of initial $k_1$, $k_2$ values would more closely define the two-dimensional region in the $k_1 - k_2$ plane within which all initial guesses will produce convergence of the algorithm. In any event, the convergence of the algorithm depends on the goodness of the initial guesses. Although Figure 1 is not to be construed as an exact region of convergence, it is evident that the initial guesses should be the same order of magnitude as the convergence values. Often, experimenters have no idea about the order of magnitude of the parameter values. Thus, in such a case it is

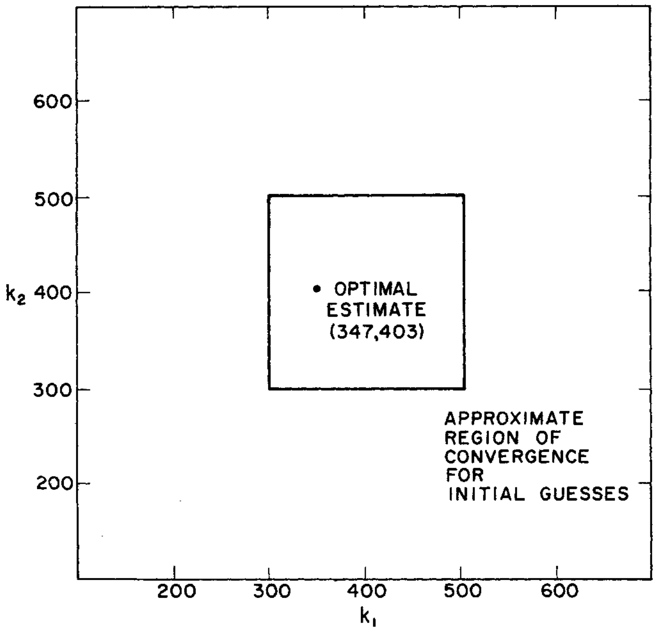

Fig. 1. Approximate region of convergence for initial guesses of $k_1$ and $k_2$ in the quasilinearization algorithm.

Table 4. Effect of Distribution of Data Points on $k_1$ and $k_2$

| Case | Initial guess | | Data points | Final values | | Iterations |
|---|---|---|---|---|---|---|
| | $k_1$ | $k_2$ | $M$ | $k_1$ | $k_2$ | |
| 11 | 300 | 300 | 4 | 347.722 | 403.329 | 3 |
| 12 | 300 | 300 | 4 | 345.318 | 403.260 | 4 |

necessary to affix to the quasilinearization algorithm a procedure for generating initial guesses from the raw experimental data. Such a procedure would be the following.

One could easily fit a polynomial to the first several data points and compute the value of the derivative of the polynomial at some point in this range, perhaps at $t = 0$. This derivative is then an estimate of the rate $r(\mathbf{x}, \mathbf{k})$. The values of the $k_i$ can then be determined by using the kinetic model. Although the estimates obtained in this manner are crude at best, they should be within the order of magnitude of the true values. If this procedure cannot be trusted alone, then a simple method of steep descent can be used to improve these initial estimates. The steep descent procedure, described by Rosenbrock and Storey (*11*), can be stopped when the percent change in the $k_i$ is less than some predetermined criterion. These values are then used as initial guesses for the quasilinearization algorithm.

We might point out that a somewhat different scheme for enhancing convergence has been suggested by Donnelly and Quon (*4*). In the event that the quasilinearization algorithm does not converge, these authors suggest perturbing the actual data $\mathbf{y}(t_i)$ to form new data $\mathbf{y}^{(1)}(t_i)$ that fall between the observed data $\mathbf{y}(t_i)$ and the data predicted from the initial guess, $\mathbf{x}(t_i,\mathbf{k}^*)$. A sequence of perturbed problems is solved until the desired solution is reached.

Either of these two schemes can be used to enhance the convergence of the algorithm. It is not readily apparent which of the two, or any modifications that can be proposed, will be computationally more advantageous. Indeed, the development of a general purpose estimation routine based on quasilinearization with a separate scheme for generating initial guesses would be of significant value in the evaluation of experimental kinetic data.

In the previous runs, each observation $y_1$ and $y_2$ is given equal weighting ($q_1 = 1$, $q_2 = 1$). To examine the effect of the weighting of the observations, $q_1 = 1$ and $q_2 = 5$ were used with $\epsilon = 0.01$ (case 5). See Table 2.

By comparison of cases 2 and 5 it is seen that the final values of $k_1$ and $k_2$ are not affected appreciably by different weighting of the states. The limit of different weighting factors is represented by cases 6 and 7, wherein only one output is observed. It appears that better estimates are obtained by observing $y_2$ than $y_1$. This point will be discussed in more detail later.

In all preceding cases the estimating procedure utilized each of the eight data points. It is interesting to examine the effect on the estimates of $k_1$ and $k_2$ of using experimental measurements at fewer than eight values of $t$. Insight should be provided into the problem of choosing a priori the number of values of the independent variable at which to make measurements. Values of $M = 6$, 4, and 3 corresponding to $t = 5.63$, 16.97, 22.62, 34.0, 42.2, and 169.7; $t = 5.63$, 16.97, 34.0, and 169.7; and $t = 5.63$, 34.0, and 169.7 ($\times 10^{-4}$), respectively, were used in cases 8, 9, and 10 with $q_1 = 1$, $q_2 = 1$, and $\epsilon = 0.01$ (Table 3).

There is some variation in the estimates of $k_2$ as the number of data points is reduced, but the estimates of $k_1$ remain generally constant. The results from cases 8 to 10 seem to indicate that for this example reasonable estimates are obtained with observations at as few as three values of $t$.

Table 5. Effect of Additional Errors on $k_1$ and $k_2$

| Case | $M$ | $\alpha$ | $w$ | Final values $k_1$ | Final values $k_2$ |
|---|---|---|---|---|---|
| 13 | 8 | 0.2 | 1 | 389.702 | 499.571 |
| 14 | 16 | 0.2 | 1 | 383.528 | 523.989 |
| 15 | 8 | 0.05 | 1 | 357.030 | 424.497 |
| 16 | 16 | 0.05 | 1 | 355.737 | 431.479 |
| 17 | 30 | 0.05 | 1 | 347.655 | 403.485 |
| 18 | 30 | 0.05 | 1,000 | 347.421 | 403.932 |
| 19 | 8 | 0.02 | 1 | 351.126 | 411.403 |

The next question one might raise is the following. If we know that a number of data points, for example four, are enough to yield acceptable (again, we will define acceptable later) estimates of $k_1$ and $k_2$, how should the four points be distributed over the range of the observed conversions? In particular, should the points be evenly distributed, or is a clustering near the low or high conversion more desirable? Cases 8 to 10 represent an even distribution of the points. Consider now cases 11 and 12, wherein four points are used at values of $t = 5.63$, 11.32, 16.97, and 22.62; and $t = 34.0$, 39.7, 45.2, and 169.7, respectively (Table 4).

At this point it is difficult to distinguish between the accuracy of the estimates in cases 11 and 12, although slightly better estimates are obtained in case 11 where the data are concentrated in the region of most rapid change of $x_1$ and $x_2$.

Hougen and Watson (*6*) used numerical differentiation at each of the data points and then averaged the resulting values of $k_1$ and $k_2$, obtaining $k_1 = 348$ and $k_2 = 404$ [(lb.)(moles)/359 (cu. ft.) (hr.) (atm.$^2$)]. It is purely coincidental that these values obtained in such a crude fashion are so close to the best estimates.

To examine the effect of larger experimental error on the estimates of $k_1$ and $k_2$, additional artificial error was introduced in the original experimental data. The values of $k_1$ and $k_2$ from case 2 were selected as the true values and were used to generate true values of $x_1$ and $x_2$. Then, experimental values of $y_1$ and $y_2$ were generated according to

$$y_1^*(t_i) = x_1(t_i)(1 + \alpha \sin wt_i) \tag{8}$$

$$y_2^*(t_i) = x_2(t_i)(1 + \alpha \cos wt_i) \tag{9}$$

where $\alpha$ and $w$ are constants governing the amplitude and frequency of the artificial error. With given values of $\alpha$ and $w$, the cases in Table 5 were studied with $k_1$ and $k_2$ equal to 300 initially.

## ACCURACY ANALYSIS OF THE ESTIMATES

The next step is to obtain estimates of the accuracy or reliability for the values of $k_1$ and $k_2$ obtained in the preceding numerical experiments. We follow the procedure of Rosenbrock and Storey (*11*) and refer the reader to that book and to Heineken et al. (*5*) for details of the analysis.

The procedure to obtain confidence intervals assumes that the estimates are close to the true values of the parameters and that the experimental errors $\eta_1$ and $\eta_2$ are normally distributed with zero mean and covariance matrix $\mathbf{M}$, which normally is known no more exactly than as $\sigma^2\mathbf{I}$.

If $\mathbf{P}$ is the covariance matrix for the parameter deviations, then

TABLE 6. 95% CONFIDENCE INTERVALS FOR $k_1$ AND $k_2$

| Case | $P_{11}$ | $P_{12}$ | $P_{22}$ | $1.96\sigma_1$ | $1.96\sigma_2$ |
|---|---|---|---|---|---|
| 2 | 0.428 | −0.488 | 13.3 | 1.28 | 7.13 |
| 6 | 5.55 | −49.2 | 507.0 | 4.62 | 44.2 |
| 7 | 6.85 | 14.4 | 39.3 | 5.14 | 12.3 |
| 9 | 0.614 | −0.781 | 23.6 | 1.53 | 9.52 |
| 10 | 0.659 | −0.650 | 21.6 | 1.59 | 9.10 |
| 15 | 533.0 | −391.0 | 9626.0 | 45.3 | 192.0 |
| 16 | 90.4 | −99.9 | 2645.0 | 18.6 | 101.0 |

$$\mathbf{P} = \left[ \frac{1}{\sigma^2} \sum_{i=1}^{M} \mathbf{D}_i^T \mathbf{D}_i \right]^{-1} \tag{10}$$

where

$$\mathbf{D}_i = \left[ \frac{\partial \mathbf{x}(t_i, \mathbf{k})}{\partial \mathbf{k}} \right]_{\mathbf{k}=\mathbf{k}^*} \tag{11}$$

and $\mathbf{k}^* = [k_1^*, k_2^*]$ are the best or optimal values of the parameters. Having computed $\mathbf{P}$, we obtain confidence intervals by

$$\sigma_b{}^2 = \mathbf{b}^T\mathbf{P}\mathbf{b} \tag{12}$$

where $\mathbf{b}$ is an appropriate unit vector for which the associated 95% confidence interval is $1.96\sigma_b$. For example, if $\mathbf{b}^T = [1, 0]$

$$\sigma_1{}^2 = E[(k_1 - k_1{}^*)^2] = P_{11} \tag{13}$$

is the variance associated with the estimate $k_1$ (as distinguished from $\sigma^2$ the variance associated with the experimental error).

The results of the analysis for several of the previous cases are presented in Table 6. The results of case 2 were taken to be the optimal estimates $k_1^*$, $k_2^*$.

By comparing cases 2 and 6 and 2 and 7 it is seen that better estimates are achieved by observing $x_2$ alone than $x_1$ alone. It is difficult to attribute this result to the nature of $x_2$ (diphenyl) as an intermediate product or simply to the manner in which $x_2$ enters $r_1$ and $r_2$. When fewer data points are taken, cases 9 and 10, the confidence intervals are very close to those of case 2. This indicates that it is better to observe more state variables at fewer values of $t$ than fewer states at more values of $t$.

Finally, the effect of experimental error is considered. The confidence intervals are much larger, as can be predicted directly from Equations (10) and (12) wherein $\sigma_b$ is prportional to $\sigma$. It has been shown by Heineken et al. (5) that at constant level of error the confidence limits for parameter estimates asymptotically decrease as $M^{-1}$. This is verified in this example by comparing the confidence limits of cases 15 and 16, where $M$ was increased from 8 to 16.

For this example, several important conclusions can be drawn:

1. One should try to measure as many of the $R$ independent concentrations as possible. If a choice must be made, intermediates appear to give better parameter estimates than primary constituents.

2. One is confronted with a trade off between the number of data points and the level of experimental error. If the measurements are accurate, data do not have to be taken at a large number of values of $t$ to obtain good estimates. It appears that the most important consideration is to reduce the level of experimental error as much as possible. One is also confronted with a trade off if one tries to measure as many of the independent concentrations as possible, but some contain more error than others.

The preceding conclusions also apply to batch reactions. It would appear fruitful to follow the course of a batch reaction with measurements that can be made as accurately as possible. Total pressure represents such a measurement.

The quasilinearization algorithm has proved to be rapid and computationally efficient (*2, 4, 10*) within the stipulations on convergence noted earlier. As such, the parameter estimates obtained through quasilinearization can be used to distinguish among rival models. Once the parameter estimates have been obtained, classical statistical techniques can be used to select the most appropriate model.

Finally, we must stress that the conclusions above are based on results obtained from the particular example chosen. Although the trends observed are felt to be of a general nature applicable to the analysis of batch and integral reaction experiments, each new experimental situation should be approached freshly with the above points in mind.

## ACKNOWLEDGMENT

Acknowledgment is given to the donors of the Petroleum Research Fund, administered by the American Chemical Society, for partial support of J. H. Seinfeld during the course of this study.

## NOTATION

$\mathbf{D}$ = matrix defined in Equation (11)
$J$ = least-square criterion
$\mathbf{k}$ = parameter vector
$K_1, K_2$ = equilibrium constants
$M$ = number of values of $t$ at which data are taken
$\mathbf{M}$ = covariance matrix of the experimental error
$\mathbf{P}$ = covariance matrix of the parameter estimates
$\mathbf{Q}$ = weighting matrix
$q_1, q_2$ = diagonal elements of $Q$
$r_1, r_2$ = reaction rates
$t$ = time variable
$w$ = constant
= xstate vector
$\mathbf{y}$ = observation vector

### Greek Letters

$\alpha$ = constant
$\epsilon$ = convergence criterion
$\eta_1, \eta_2$ = experimental errors
$\sigma^2$ = variance of the experimental error
$\sigma_1{}^2, \sigma_2{}^2, \sigma_b{}^2$ = variances of the parameter estimates

## LITERATURE CITED

1. Bellman, R., et al., *Math. Biosciences*, **1**, 71 (1967).
2. Bellman, R., and R. Kalaba, "Quasilinearization and Nonlinear Boundary-Value Problems," American Elsevier, New York (1965).
3. Box, G. E. P., and H. L. Lucas, *Biometrika*, **46**, 77 (1959).
4. Donnelly, J. K., and D. Quon, paper presented at Am. Inst. Chem. Engrs. Tampa meeting (1968).
5. Heineken, F. G., H. M. Tsuchiya, and Rutherford Aris, *Math. Biosciences*, **1**, 115 (1967).
6. Hougen, O., and K. M. Watson, "Chemical Process Principles," Vol. 3, Wiley, New York (1948).
7. Kittrell, J. R., W. G. Hunter, and C. C. Watson, *AIChE J.*, **11**, No. 6, 1051 (1965).
8. *Ibid.*, **12**, No. 1, 5 (1966).
9. Lapidus, Leon, and T. I. Peterson, *ibid.*, **11**, 5, 891 (1965).
10. Lee, E. S., *Ind. Eng. Chem. Fundamentals*, **7**, 1, 152 (1968).
11. Rosenbrock, H. H., and C. Storey, "Computational Techniques for Chemical Engineers," Pergamon, New York (1966).

*Manuscript received July 26, 1968; revision received January 20, 1969; paper accepted January 22, 1969. Paper presented at AIChE Los Angeles meeting.*