A New Technique Using the Polynomials Approximation and Trapezoidal Average Method for Estimating Kinetic Rate Constants in Systems of Linear or Nonlinear Differential Equations

Katsutoshi OHKUBO, Kohji TSUCHIHASHI, and Hiroyuki SAKAMOTO

Department of Synthetic Chemistry, Faculty of Engineering, Kumamoto University, Kurokami, Kumamoto 860

(Received August 9, 1973)

A new technique using the polynomials approximation and the trapezoidal average method, has been proposed for estimating kinetic rate constants. It is implemented by minimizing the error of the concentration (of a product) obtained from the polynomial-approximated least square-error analysis of the observed concentrations during the initial reaction stage relative to that obtained theoretically from the trapezoidal average of the calculated concentration (of a product) on the basis of the law of mass action. The technique is useful for easily and exactly estimating the rate constants in systems of linear or nonlinear differential equations. A kinetic model, $A \longrightarrow B \rightleftarrows C$, was used to illustrate and develop the present technique.

In usual kinetic studies of chemical-reaction mechanisms, we often find it difficult to estimate exactly the kinetic rate constants of the elementary steps of the reaction. Some numerical integration methods, such as the Runge-Kutta integration,[1] Taylor's series expansion,[2] and Picard's method,[3] are usually used to estimate the rate constants. These methods, however, require a precision fit to dense data or much computation time.

The new technique presented in this paper requires only a few kinetic data (usually concentrations) about the species participating in the reaction. It is based on the polynomial approximation of the given kinetic data and the trapezoidal average of the minute change in the concentration of the species (typically, that of a product). The following kinetic model was used to illustrate and develop the present technique:

$$A \xrightarrow{k_1} B \underset{k_3}{\overset{k_2}{\rightleftarrows}} C \qquad (1)$$

where A=reactant; B=intermediate; C=product; k_i= rate constant i, and i=1, 2, and 3. The type of reaction depicted by Eq. (1) is found in such an organic reaction progress as chlorosulfonation:

$$RH \xrightarrow{ClSO_2OH} RSO_2OH + HCl \overset{ClSO_2OH}{\rightleftarrows} RSO_2Cl + H_2SO_4 \quad (R=\text{alkyl or aryl}).$$

Computation Method

The kinetic model illustrated by Eq. (1) can be described by the following differential equations drawn from the law of mass action:

$$d(A)/dt = -k_1(A) \qquad (A)_0 \neq 0 \qquad (2a)$$

$$d(B)/dt = k_1(A) - k_2(B) + k_3(C) \qquad (B)_0 = 0 \qquad (2b)$$

$$d(C)/dt = k_2(B) - k_3(C) \qquad (C)_0 = 0 \qquad (2c)$$

where the parentheses denote concentrations and where the subscript zeros stand for the initial concentrations of the species.

We will now explain the computation method of the present technique for estimating the rate constants, basing on the above differential equations. According to Eqs. (2a—c), the increments for the three components, A, B, and C, at the reaction time of t, $D(A)_t$, $D(B)_t$, and $D(C)_t$, are defined as:

$$D(A)_t = -k_1(A)_t dt, \quad (A)_{t+dt} = (A)_t + D(A)_t \qquad (3a)$$

$$D(B)_t = k_1(A)_t dt - k_2(B)_t dt + k_3(C)_t dt, \quad (B)_{t+dt} = (B)_t + D(B)_t \qquad (3b)$$

$$D(C)_t = k_2(B)_t dt - k_3(C)_t dt, \quad (C)_{t+dt} = (C)_t + D(C)_t \qquad (3c)$$

where the notations of $(\text{Component})_t$ and $(\text{Component})_{t+dt}$ indicate the concentrations of the components at t and $t+dt$ respectively.

The concentrations of the components at $t+mdt$ (m=1,2,3,$\cdots$), $(\text{Component})_{t+mdt}$ are, then, defined by:

$$(\text{Component})_{t+mdt} = (\text{Component})_t + \sum_{p=1}^{m} D(\text{Component})_{t+(p-1)dt} \qquad (4)$$

The $D(\text{Component})_{t+(p-1)dt}$ values in Eq. (4) are evaluated by the following trapezoidal average:

$$\begin{aligned} D(\text{Component})_{t+(p-1)dt} &= (D(\text{Component})_{t+(p-1)dt} + D(\text{Component})_{t+pdt})/2 \\ &= 1.5D(\text{Component})_{t+(p-1)dt} - 0.5D(\text{Component})_{t+(p-2)dt} \end{aligned} \qquad (5)$$

where $D(\text{Component})_{t+(p-1)dt}$ is defined by

$$D(\text{Component})_{t+(p-1)dt} = (D(\text{Component})_{t+pdt} + D(\text{Component})_{t+(p-2)dt})/2.$$

We will use the product of C as the Component and assume that p=2. Then, the above relationships can be rewritten as:

$$\begin{aligned} D(C)_{t+dt} &= (D(C)_{t+dt} + D(C)_{t+2dt})/2 \\ &= 1.5D(C)_{t+dt} - 0.5D(C)_t \end{aligned} \qquad (5')$$

where $D(C)_{t+dt}=(D(C)_{t+2dt}+D(C)_t)/2$.

In the evaluation of the $D(C)_{t+dt}$ values, the rate constants, k_i, are supplied by:

$$k_i = \min k_i + \sum_{n}^{n} 10^{(\gamma-\beta)} \min k_i, \quad \gamma = \text{gauss}\left[\frac{n-1}{9\alpha}\right] \qquad (6)$$

(n=1, 2, 3, $\cdots$, β=0, 1, 2, $\cdots$, $\alpha=10^{\beta}$)

where min k_i=lower limit of k_i ($1.0\times10^{\delta}$; δ=constant); γ=integral number; β=constant; α=parameter determining the number of significant figure for k_i (*viz.*, k_i=(β+1) significant figures at $\alpha=10^{\beta}$).

On the other hand, the time conversion of the observed concentration of C, $(C)_{t,obsd}$, is represented by the following series of polynomials, resulting from the least square-error analysis of the $(C)_{t,obsd}$ values:

$$(C)_{t,obsd} = a_0 + a_1 t + a_2 t^2 + a_3 t^3 + \cdots + a_n t^n \quad (n=0, 1, 2, \cdots) \tag{7}$$

where a_n (n=0,1,2,$\cdots$) denote coefficients. The average increment observed at $t+dt$, $D(C)_{t+dt,obsd}$, is, then, defined as:

$$D(C)_{t+dt,\,obsd} = (C)_{t+2dt,\,obsd} - (C)_{t+dt,\,obsd} \tag{8}$$

where the $(C)_{t+2dt,obsd}$ and $(C)_{t+dt,obsd}$ values are calculated by Eq. (7).

The best rate constants are estimated by minimizing the following relative error, ε:

$$\varepsilon = |D(C)_{t+dt} - D(C)_{t+dt,\,obsd}| / D(C)_{t+dt,\,obsd} < \delta\varepsilon \tag{9}$$

where $\delta\varepsilon$ is the error norm for converging the computation. In the evaluation of the ε value, the iteration procedure is necessary: that is, the rate constants, k_j (j=1, 2, or 3), are changed in turn until one gets the smallest ε value while the other constants, $k_{i(i\neq j)}$, are fited.

The $D(C)_{t+dt}$ values in Eq. (9) are computed by the following equation, derived from Eqs. (3c), (4), and (5′):

$$\begin{aligned} D(C)_{t+dt} = {} & [k_2((A)_0 - (A)_{t+dt}) - (k_2+k_3)(C)_{t+dt})]dt \\ & + 1.5[k_2(k_1+k_2+k_3)(A)_{t+dt} \\ & + (k_2+k_3)^2(C)_{t+dt} - k_2(k_2+k_3)(A)_0]dt^2 \end{aligned} \tag{10}$$

In Eq. (10), the following approximation is taken for simplicity of computation:

$$(A)_{t+dt} \approx (A)_t \tag{11}$$

In order to avoid any technical error caused by the above approximation, a very small dt value should be used for the computation. However, it is better to discard the above approximation when the $(A)_{t+dt}$ value can be evaluated from the time conversion of (A).

The conditions used throughout in this paper are dt=0.01—0.1, min k_i=1.0×10^{-5}, n=1—991, β=1, and α=10.

Computed Results and Discussion

Table 1 shows the kinetic data, $(Component)_{t,obsd}$, obtained by the simulation[4] of the kinetic model with specifications of $(A)_0$=10.0, $(B)_0=(C)_0$=0, k_1=0.1, k_2=0.6, k_3=0.2, and dt=0.001. We used the simulation data as the given kinetic data, because the preliminarily known k_i values facilitate the checking of the accuracy of the estimated rate constants, $k_{i,estimated}$. The coefficients (a_n) of the polynomials of Eq. (7) were, then, determined by the least square-error analysis of the $(C)_{t,obsd}$ values at t=0, 1, 2, $\cdots$, n in Table 1 (n corresponds to the subscript n of a_n in the highest-order term of Eq. (7)). Table 2 lists the a_n values,

Table 1. Given kinetic data

t[a]	(A)	(B)	(C)
0	10.00000	0.00000	0.00000
1	9.04838	0.72595	0.22567
2	8.18731	1.11407	0.69863
3	7.40818	1.34449	1.24733
4	6.70320	1.49872	1.79807
5	6.06531	1.61390	2.32079
6	5.48812	1.70717	2.80472
7	4.96585	1.78663	3.24752
8	4.49329	1.85632	3.65039
9	4.06570	1.91839	4.01592
10	3.67880	1.97410	4.34711
11	3.32871	2.02431	4.64698
12	3.01195	2.06965	4.91841
13	2.72532	2.11063	5.16405
14	2.46597	2.14770	5.38632
15	2.23131	2.18123	5.58746
16	2.01897	2.21157	5.76946
17	1.82684	2.23902	5.93414
18	1.65299	2.26386	6.08315
19	1.49569	2.28633	6.21798
20	1.35335	2.30666	6.33998
21	1.22457	2.32506	6.45037
22	1.10803	2.34171	6.55026
23	1.00259	2.35677	6.64064
24	0.90718	2.37040	6.72242
25	0.82085	2.38273	6.79642
26	0.74274	2.39389	6.86337
27	0.67206	2.40399	6.92395
28	0.60810	2.41313	6.97877
29	0.55023	2.42139	7.02837
30	0.49787	2.42887	7.07326

a) Arbitrary time unit.

Table 2. Coefficinets (a_n) for the polynomial approximation

a_n	Order of polynomials					
	n=2	n=3	n=4	n=5	n=6	n=10
a_0	1.934×10^{-16}	2.680×10^{-16}	6.413×10^{-15}	2.709×10^{-13}	-1.946×10^{-14}	8.648×10^{-10}
$10^2\,a_1$	10.202	4.4842	2.0379	0.95372	0.45605	0.039630
$10\,a_2$	1.2365	2.0942	2.5427	2.7686	2.8822	2.9879
$10^2\,a_3$		−2.8592	−5.3054	−6.8865	−7.8197	−8.8519
$10^3\,a_4$			4.0771	8.5946	12.120	17.256
$10^4\,a_5$				−4.5175	−10.738	−25.089
$10^5\,a_6$					4.1472	27.338
$10^5\,a_7$						−2.1602
$10^6\,a_8$						1.1613
$10^8\,a_9$						−3.7742
$10^{10}a_{10}$						5.5664

TABLE 3. $(C)_{t+dt}$ (AND/OR $(C)_{t+2dt}$) USED FOR THE PRESENT COMPUTATIONS

t	Order of polynomials						Simulation[a]
	$n=2$	$n=3$	$n=4$	$n=5$	$n=6$	$n=10$	
1	0.22567	0.22567	0.22567	0.22567	0.22567	0.22567	0.22567
1.01	0.22917	0.22946	0.22954	0.22957	0.22958	0.22958	0.22959
1.02	0.23270	0.23327	0.23344	0.23349	0.23351	0.23352	0.23353
1.04	0.23983	0.24097	0.24129	0.24140	0.24144	0.24146	0.24147
1.06	0.24706	0.24977	0.24925	0.24940	0.24945	0.24948	0.24950
1.08	0.25439	0.25666	0.25729	0.25749	0.25756	0.25760	0.25760
1.10	0.26219	0.26465	0.26542	0.26567	0.26534	0.26580	0.26581
1.12	0.26934	0.27272	0.27363	0.27392	0.27402	0.27408	0.27409
1.16	0.28469	0.28915	0.29032	0.29069	0.29081	0.29089	0.29090
1.20	0.30044	0.30592	0.30733	0.30777	0.30792	0.30801	0.30802

a) See Ref. 5 ($(A)_0=10.0$; $(B)_0=(C)_0=0$; $k_1=0.1$; $k_2=0.6$; $k_3=0.2$; $dt=0.001$).

and Table 3 shows the $(C)_{t+dt,obsd}$ and $(C)_{t+2dt,obsd}$ values evaluated by the polynomial approximation.

Now, let us examine the effect of the dt values on the estimation of the rate constants. The rate constants for the kinetic model were estimated using the polynomials ($n=10$) with the following specifications: $(A)_0=10.0$; $(B)_0=(C)_0=0$; $dt=0.01$—0.1; $(C)_t=(C)_{1.0}$; $(C)_{t+dt}=(C)_{1.0+dt}$; $(C)_{t+2dt}=(C)_{1.0+2dt}$. Table 4 shows the change in the $k_{i,estimated}$ values due to the dt values, together with the relative errors of ε and ε_i ($i=1$, 2, and 3), defined by Eq. (9) and (12) respectively:

$$\varepsilon_i = |k_{i,\text{estimated}} - k_{i,\text{given}}|/k_{i,\text{given}} \tag{12}$$

where $k_{i,given}$ stands for the exact k_i values (*viz.*, $k_1=0.1$, $k_2=0.6$, and $k_3=0.2$). It may be seen from Table 4 that the smaller the dt value becomes, the more the ε_i values are lowered. Since the best k_i values should be given by the $k_{i,estimated}$ values, which are not appreciably influenced by the dt value, it is desirable to establish the following relationship:

$$|(A)_{t+dt}-(A)_t|/(A)_t \approx 0 \tag{13}$$

where the $(A)_{t+dt}$ and $(A)_t$ values are, respectively, obtained from Eq. (3a) and Table 1.

It is of interest here to compare the $k_{i,estimated}$ values presented in Table 4 with those calculated by means of the following equation:

$$\begin{aligned} D(C)_{t+dt} &= (C)_{t+2dt} - (C)_{t+dt} \\ &= [k_2((A)_0-(A)_{t+dt})-(k_2+k_3)(C)_{t+dt}]dt \\ &\quad + [k_2(k_1+k_2+k_3)(A)_{t+dt}+(k_2+k_3)^2(C)_{t+dt} \\ &\quad -k_2(k_2+k_3)(A)_0]dt^2 \end{aligned} \tag{14}$$

where $(A)_{t+dt}$ is approximated by $(A)_t$. The $k_{i,estimated}$ values evaluated by Eq. (14) were found to be inaccurate in some measure as compared with those listed in Table 4; for example, $k_1=0.14$, $k_2=0.60$, $k_3=0.20$, and $\varepsilon=0.221\%$ at $dt=0.1$, and $k_1=0.13$, $k_2=0.60$, $k_3=0.20$, and $\varepsilon=0.028\%$ at $dt=0.01$ under the specifications mentioned before. It may be deduced, therefore, that the trapezoidal average method depicted by Eq. (5′) is superior to the usual technique illustrated by Eq. (14) for estimating the rate constants.

Next, let us notice the effect of the order (n) of the polynomial approximation on the evaluation of the $k_{i,estimated}$ values. Table 5 lists the $k_{i,estimated}$ values

TABLE 4. EFFECT OF THE dt VALUE ON THE ESTIMATION OF RATE CONSTANTS

dt	k_i			$\varepsilon_i(\%)$			$\varepsilon(\%)$
	k_1	k_2	k_3	ε_1	ε_2	ε_3	
0.1	0.097	0.60	0.20	3.0	0.0	0.0	0.063
0.08	0.097	0.60	0.20	3.0	0.0	0.0	0.003
0.06	0.098	0.60	0.20	2.0	0.0	0.0	0.048
0.04	0.098	0.60	0.20	2.0	0.0	0.0	0.026
0.02	0.098	0.60	0.20	2.0	0.0	0.0	0.004
0.01	0.10	0.60	0.20	0.0	0.0	0.0	0.029

Computations were carried out with the polynomials of $n=10$. (Iteration=2; Computation time with FACOM 230-25=1.7 min in every case.)

obtained by the trapezoidal method using the polynomials of $n=2$, 3, 4, 5, 6, and 10. As we expected, the increase in the n value of the polynomials decreases the ε_i ($i=1$—3) values as defined by Eq. (12), and the polynomials of $n=10$ offer the exact rate constants. The present technique requires a polynomial approximation involving a relatively high order of terms, because it takes the rough approximation depicted by Eq. (11).

TABLE 5. ESTIMATED RATE CONSTANTS AND RELATIVE ERRORS

n	k_i[a]			$\varepsilon_i(\%)$			$\varepsilon(\%)$
	k_1	k_2	k_3	ε_1	ε_2	ε_3	
2	0.01	0.55	0.19	90.0	8.33	5.0	0.0494
3	0.010	0.59	0.19	90.0	1.67	5.0	0.187
4	0.052	0.60	0.20	48.0	0.0	0.0	0.0068
5	0.077	0.60	0.20	23.0	0.0	0.0	0.0023
6	0.089	0.60	0.20	11.0	0.0	0.0	0.0034
10	0.10	0.60	0.20	0.0	0.0	0.0	0.0298

a) k_i=2 significant figures. Computations were carried out with $dt=0.01$. In every case, iteration=2 and computation time=1.7 min by FACOM 230-25.

Finally, it may be worth emphasis that the present technique will be useful for estimating the kinetic rate constants in the systems of linear or nonlinear differential equations.

The calculations were carried out on a FACOM

320–25 computer at the Engineering Research Center of Kumamoto University.

References

1) For instance, see H. Margenau and G. H. Murphy, "The Mathematics of Physics and Chemistry," D. van Nostrand, New York (1943), p. 450.
2) I. G. Darvey, S. J. Prokhovnik, and J. F. Williams, *J. Theor. Biol.*, **11**, 459 (1966).
3) R. Tanner, *Ind. Eng. Chem. Fundam.*, **11**, 1 (1972).
4) K. Ohkubo and T. Hashimoto, This Bulletin, **45**, 3350 (1972).