

Numerical Differentiation and Integration in Chemical Kinetics

KATE BAK

Chemical Laboratory IV, H. C. Ørsted Institute, The University of Copenhagen, Denmark

A procedure is described for determining the time when a chemical reaction has started from the measurements taken immediately after the start.

Formulae for smoothing kinetic data and obtaining time derivatives of the smoothed data are derived, using a least squares procedure. These formulae only apply to measurements taken at equidistant times.

Finally the evaluation of the integral $e^{-\lambda t} \int_0^t \varphi(t) e^{\lambda t} dt$, which occurs in kinetics, is given in terms of differences of φ .

1. The most common procedure for determining rate constants from kinetic data is to compare the measured values with the integrated rate expression. In certain cases, however, it is more convenient to use a differential method in which the rate expression is compared with a combination of concentrations and time derivatives of these, the latter being obtained by graphical or numerical differentiation. A discussion of such a procedure is given by Benson¹ who suggests that the most accurate way to obtain the derivative is by plotting the concentration as a function of time and drawing a smooth curve through the points. From this curve the slopes are obtained using a tangent meter. If the concentrations are measured at unequal time intervals, this is probably the simplest way to obtain the derivative, but if this is not the case, as for instance when automatic measuring devices are used, or when it is possible to plan the experiment such that measurements are taken at equal intervals, we believe that a purely numerical smoothing and differentiation procedure is preferable. We shall use here a smoothing based on the method of least squares,² but it is clear that under certain circumstances such as when very many measurements taken at short time intervals are available, a simpler procedure such as "smoothing by fives" might give equally good results.

Another numerical problem which is somewhat more complicated but also may occur in kinetic work is the problem of solving a differential equation in

which an experimentally determined function occurs. An example of this is given in the accompanying paper³ where it is shown that the solution of a certain second order differential equation can be expressed in terms of an integral

$$I(t) = e^{-\lambda t} \int_0^t \varphi(t) e^{\lambda t} dt \quad (1)$$

The function I is, as is seen, also the solution of the differential equation

$$\frac{dI}{dt} + \lambda I = \varphi(t) \quad (2)$$

with the initial condition $I(0) = 0$. Furthermore it is known about the measured function φ that $\varphi(0) = 0$, and since φ is a degree of reaction we must have $\varphi(\infty) = 1$, from which it follows that $I(\infty) = 1/\lambda$.

One could now attempt to calculate the function I by straight forward integration of (1) using Simpson's formula on smoothed values of φ . Since the integrand increases very rapidly with t , this is not a good procedure, and Hartree has suggested finding the value of the integral (1) by solving the differential eqn. (2) numerically. In the case discussed in Ref.³ this then effectively means that the solution of a second order differential equation (which contains two integrals of this type) is reduced to solving two very much simpler first order equations. If λ is very large, even the solution of the differential equation becomes inconvenient, and for that case we shall therefore develop a method of expanding the integral in terms of the differences of φ .

Whenever one wants to perform a numerical integration of a function measured at certain time intervals, it becomes of importance to know the correct time corresponding to the experimentally measured quantity. A constant additive error in the time, which is of minor importance for differentiation, can here influence the result greatly. The problem therefore becomes how to determine precisely when the experiment started.

We believe that the procedure one should adopt here is to define the start of the time scale in some arbitrary fashion (such as the time when a pipette is half empty) and then record data at equal time intervals. When the experiment is over, the time which the recorded data show to be the time of the start of the reaction is determined in the following way:

The measured values of $\varphi(t')$ where t' now stands for this preliminary time scale, is fitted to a polynomial expression

$$\varphi = \varphi_0 + \varphi_1 t' + \varphi_2 (t')^2 + \dots \quad (3)$$

for small values of t' , but neglecting the first, or the first two or three measured values in the case where the curve has a sigmoid shape. Special care should be taken to avoid points which seem to indicate incomplete mixing. The number of points left out and the number of points included, as well as the degree of the polynomial is obviously open to a certain degree of arbitrariness. It is our experience, however, that in the end it turns out that the choice does not change the results very much. The start of the reaction is then defined as the

numerically smallest value of t' for which the polynomial is zero. Denoting this value by t_0 , we have that the correct time is equal to $t = t' + t_0$. After the measured data have been smoothed using a least squares procedure one can then use ordinary interpolation technique⁶ to determine φ at equally spaced values of t , including $t = 0$.

As a general rule for choosing the number of terms in the polynomial and the number of points we would suggest using a second order polynomial, leaving out measurements during the first minutes or one tenth of the apparent half life, whichever is the smaller, and using six to ten points chosen from times less than one fourth half life in fitting the polynomial.

After this brief introduction, which certainly does not exhaust the possibilities of numerical methods in kinetics, we shall show how the formulae mentioned above are derived.

2. The purpose of smoothing is to eliminate random errors in the experimental values. A large number of methods to achieve this has been proposed, but the only method which seems to have some theoretical foundation is the fitting of a function to the data, using the method of least squares, and then replacing the measured values by the corresponding values of the function. If one knows in advance what the functional relationship between the measured values should be, this is a perfectly straight-forward operation, but unfortunately also in most cases quite uninteresting. Smoothing primarily becomes important when the proper functional relationship is not known. For these cases one must expect that the final interpretation of the data to some extent depends on the functional relationship chosen; therefore, in order to obtain a consistent procedure one should always use the same functions. The simplest functions to use are polynomials, and in order to obtain a smoothing, the degree of the polynomial should be considerably lower than the number of points.

Let the measured experimental points be $y_0, y_1, y_2, \dots, y_n$ and the corresponding values of a parameter be $x_0, x_1, x_2, \dots, x_n$. As usual we shall assume that the values of the parameter are known with a vanishingly small error, and consequently we find the smoothed out values by minimizing

$$S = \sum_{j=0}^n [y_j - \bar{y}(x_j)]^2 \quad (4)$$

in which $\bar{y}(x) = a_0 + a_1x + \dots + a_mx_m$ is an m 'th degree polynomial ($m < n$). Once the polynomial $\bar{y}(x)$ is known, it is a simple matter to calculate the improved values of \bar{y}_j ($j = 0, 1, \dots, n$) as well as the derivatives in these points.

Instead of determining the polynomial $\bar{y}(x)$ by setting the derivatives of S with respect to a_j ($j = 0, 1, \dots, m$) equal to zero we use the method of orthogonal polynomials, which entails far less work². The only restriction this method imposes on the values to use is that they have to be equally spaced in x . By putting $m = 3$ and $n = 4$, *i.e.* smoothing five points by a third degree polynomial, we obtain

$$\begin{aligned}
 I(t) &= e^{-\lambda t} \int_0^t \varphi e^{\lambda t} dt \approx e^{-\lambda t} \int_0^t (\varphi_1 t + \varphi_2 t^2 + \varphi_3 t^3) e^{\lambda t} dt \\
 &= \varphi_1 \left[\frac{t}{\lambda} - \frac{1}{\lambda^2} (1 - e^{-\lambda t}) \right] \\
 &+ \varphi_2 \left[\frac{t^2}{\lambda} - \frac{2t}{\lambda^2} + \frac{2}{\lambda^3} (1 - e^{-\lambda t}) \right] \\
 &+ \varphi_3 \left[\frac{t^3}{\lambda} - \frac{3t^2}{\lambda^2} + \frac{6t}{\lambda^3} - \frac{6}{\lambda^4} (1 - e^{-\lambda t}) \right]
 \end{aligned} \tag{5}$$

After the first few points of $I(t)$ have been obtained, the solution is continued by using, for instance, Milne's method of solving differential equations numerically⁵.

If λ is large, this method becomes less accurate than if λ is small, but on the other hand we see that for this case the value of $I(t)$ can only depend strongly on the behaviour of φ at times a little smaller than t .

This is particularly clear when we write the integral as

$$I(t) = \int_0^t \varphi(\tau) e^{-\lambda(t-\tau)} d\tau = \int_0^t \varphi(t-x) e^{-\lambda x} dx \tag{6}$$

We see that I is the cumulative effect of φ except for an exponential damping factor. When this damping factor becomes important for large λ , only the region near $x = 0$ contributes. It should therefore be possible to use an expansion of the integral in differences of φ around t .

By repeated partial integration we obtain

$$\begin{aligned}
 \lambda I(t) &= \lambda e^{-\lambda t} \int_0^t \varphi e^{\lambda t} dt \\
 &= \sum_{n=0}^{\infty} (-\lambda)^{-n} D^n \varphi(t) - e^{-\lambda t} \sum_{n=0}^{\infty} (-\lambda)^{-n} D^n \varphi(0)
 \end{aligned} \tag{7}$$

where D is the operator d/dt . When λ is a large number, λt is large compared with unity even for small values of t ; then the last sum can be neglected, and we are left with the first sum which symbolically can be written

$$\sum_{n=0}^{\infty} (-\lambda)^{-n} D^n \varphi(t) = \frac{1}{1 + \lambda^{-1} D} \varphi(t) \tag{8}$$

The relations between D , the difference operator Δ defined by $\Delta \varphi(t) = \varphi(t+h) - \varphi(t)$, and the shift operator E defined by $E \varphi(t) = \varphi(t+h)$ are $E-1 = \Delta$ and $E = \exp(hD)$. Therefore, proceeding purely formally we obtain

$$\begin{aligned}
 \lambda I &= \frac{1}{1 + (\lambda h)^{-1} \ln(1 + \Delta)} \varphi \\
 &= [1 - (\lambda h)^{-1} \Delta + \frac{1}{2}(2 + \lambda h)(\lambda h)^{-2} \Delta^2 \\
 &\quad - \frac{1}{6}(3 + 3\lambda h + \lambda^2 h^2)(\lambda h)^{-3} \Delta^3 + \dots] \varphi
 \end{aligned} \tag{9}$$

where the series has been obtained by a Taylor expansion of the expression on the left hand side. To verify this expression in general would be quite complicated, but, as done for similar cases by Hartree,⁶ it can be verified when φ is equal to a polynomial or an exponential or combinations of these functions, and thereby for any function which can be approximated by such functions.

Using this expression for I , which is only valid when the second term in eqn. (7) vanishes, I can be directly found from a table of differences of the measured φ -values.

4. The main justification for the present note is that the methods developed have been put to practical use in the accompanying paper. Since the use of these methods is not limited to the specific example discussed there, it was thought desirable to record them separately. The major conclusion of this work is that numerical differentiation on experimental data can be performed with reasonable accuracy, provided a least squares method is used rather than the customary five point formulae recommended in many textbooks⁷. With the use of numerical differentiation, and more specifically solutions of kinetic integrals in terms of differences, the treatment of kinetic data becomes considerably more flexible, and this, for example, allows the calculation of rate constants by the comparison of experimental results with the rate expression rather than by comparison with the chromonal (the integrated rate expression).

Acknowledgements. The author's thanks are owing to Dr. Thor A. Bak for many helpful discussions. Financial support from *Statens almindelige Videnskabsfond* is gratefully acknowledged.

REFERENCES

1. Benson, S. W. *The Foundations of Chemical Kinetics*, McGraw Hill 1960, p. 82.
2. Milne, W. E. *Numerical Calculus*, Princeton University Press 1949, p. 275.
3. Bak, K. and Andersen, B. *Acta Chem. Scand.* **17** (1963) 992.
4. Hartree, D. R. *Numerical Analysis*, Clarendon Press 1958, p. 111.
5. Milne, W. E. *Numerical Calculus*, Princeton University Press 1949, p. 134.
6. Hartree, D. R. *Numerical Analysis*, Clarendon Press 1958, p. 53.
7. See for instance. Margenau, H. and Murphy, G. M. *The Mathematics of Physics and Chemistry*, van Nostrand 1944, p. 456.

Received December 21, 1962.