

01 Aug 1967

Numerical Differentiation Of Equally Spaced And Not Equally Spaced Experimental Data

Harry C. Hershey
Missouri University of Science and Technology

Jacques L. Zakin
Missouri University of Science and Technology

Robert Simha

Follow this and additional works at: https://scholarsmine.mst.edu/che_bioeng_facwork

 Part of the [Biochemical and Biomolecular Engineering Commons](#)

Recommended Citation

H. C. Hershey et al., "Numerical Differentiation Of Equally Spaced And Not Equally Spaced Experimental Data," *Industrial and Engineering Chemistry Fundamentals*, vol. 6, no. 3, pp. 413 - 421, American Chemical Society, Aug 1967.

The definitive version is available at <https://doi.org/10.1021/i160023a015>

This Article - Journal is brought to you for free and open access by Scholars' Mine. It has been accepted for inclusion in Chemical and Biochemical Engineering Faculty Research & Creative Works by an authorized administrator of Scholars' Mine. This work is protected by U. S. Copyright Law. Unauthorized use including reproduction for redistribution requires the permission of the copyright holder. For more information, please contact scholarsmine@mst.edu.

SUBSCRIPTS

C = continuous phase
D = dispersed phase
d = dispersed phase

Literature Cited

(1) Angelo, J. B., Lightfoot, E. N., Howard, D. W., *A.I.Ch.E. J.* **12**, 751 (1966).
 (2) Baird, M. H. I., Davidson, J. F., *Chem. Eng. Sci.* **17**, 87 (1962).
 (3) Beyaert, B. O., Lapidus, L., Elgin, J. C., *A.I.Ch.E. J.* **7**, 46 (1961).
 (4) Calderbank, P. H., *Trans. Inst. Chem. Engrs.* **36**, 433 (1958).
 (5) Calderbank, P. H., Korchinski, I. J. O., *Chem. Eng. Sci.* **6**, 65 (1956).
 (6) Christiansen, R. M., Hixon, A. N., *Ind. Eng. Chem.* **49**, 1017 (1957).
 (7) Conkie, W. R., Savic, P., Division of Mechanical Engineering National Research Council of Canada, Rept. **MT-23** (Oct. 22, 1952).
 (8) Evnohides, S., Thodos, G., *A.I.Ch.E. J.* **7**, 78 (1961).
 (9) Frössling, N., *Beitr. Geophysik.* **52**, 170 (1938).
 (10) Garner, F. H., Foord, A., Tayeban, M., *J. Appl. Chem.* **9**, 315 (1959).
 (11) Garner, F. H., Suckling, R. D., *A.I.Ch.E. J.* **4**, 114 (1958).
 (12) Griffith, R. M., *Chem. Eng. Sci.* **12**, 198 (1960).
 (13) Hammerton, D., Garner, F. H., *Trans. Inst. Chem. Engrs.* **32**, 518 (1954).
 (14) Hayworth, C. B., Treybal, R. E., *Ind. Eng. Chem.* **42**, 1174 (1950).
 (15) Heertjes, P. M., Holve, W. A., Talsma, H., *Chem. Eng. Sci.* **3**, 122 (1954).
 (16) Hsu, N. T., Sato, K., Sage, B. H., *Ind. Eng. Chem.* **46**, 870 (1954).

(17) Hu, S., Kintner, R. C., *A.I.Ch.E. J.* **1**, 42 (1955).
 (18) Johnson, A. I., Braida, L., *Can. J. Chem. Eng.* **35**, 165 (1957).
 (19) Keith, F. W., Hixson, A. N., *Ind. Eng. Chem.* **47**, 258 (1955).
 (20) Laddha, G. S., Smith, J. M., *Chem. Eng. Progr.* **46**, 195 (1950).
 (21) Langmuir, I., *Phys. Rev.* **12**, 368 (1918).
 (22) Li, P., West, F. B., Vance, W. H., Moulton, R. W., *A.I.Ch.E. J.* **11**, 581 (1956).
 (23) Null, H. R., Johnson, H. F., *Ibid.*, **4**, 273 (1958).
 (24) Pasternak, I. S., Gauvin, W. H., *Can. J. Chem. Eng.* **38**, 35 (April 1960).
 (25) Powell, R. W., *Trans. Inst. Chem. Engrs.* **18**, 36 (1940).
 (26) Ranz, W. E., Marshall, W. R., *Chem. Eng. Progr.* **48**, 173 (1952).
 (27) Rose, P. M., Kintner, R. C., *A.I.Ch.E. J.* **12**, 530 (1966).
 (28) Ruby, C. L., Elgin, J. C., *Chem. Eng. Progr. Symp. Ser.* **51**, No. 15, 17 (1955).
 (29) Schroeder, R. R., Ph.D. thesis, Illinois Institute of Technology, Chicago, Ill., 1964.
 (30) Skelland, A. H. P., Cornish, A. R. H., *A.I.Ch.E. J.* **9**, 73 (1963).
 (31) Steinberger, R. S., Treybal, R. E., *Ibid.*, **6**, 227 (1960).
 (32) Thorsen, G., Terjesen, S. G., *Chem. Eng. Sci.* **17**, 137 (1962).
 (33) Treybal, R. E., "Liquid Extraction," 2nd ed., p. 467, McGraw-Hill, New York, 1963.
 (34) Warshay, M., Bogusz, E., Johnson, M., Kintner, R. C., *Can. J. Chem. Eng.* **37**, 29 (1959).
 (35) Weaver, R. E. C., Ph.D. dissertation, Princeton University, Princeton, N. J., 1955.
 (36) Weaver, R. E. C., Lapidus, L., Elgin, J. C., *A.I.Ch.E. J.* **5**, 533 (1959).
 (37) Zenz, F. A., *Petrol Refiner* **36**, No. 8, 147 (1957).

RECEIVED for review October 19, 1966
 ACCEPTED April 25, 1967

NUMERICAL DIFFERENTIATION OF EQUALLY SPACED AND NOT EQUALLY SPACED EXPERIMENTAL DATA

HARRY C. HERSHEY,¹ JACQUES L. ZAKIN, AND ROBERT SIMHA²

Department of Chemical Engineering, University of Missouri at Rolla, Rolla, Mo.

Procedures are given for smoothing and differentiating experimental data with both equal and nonequal spacing in the independent variable. Selection of the number of points to be included in the "movable strip" technique and of the degree of the polynomial is discussed. Equations are given to estimate the error by calculating a confidence interval on each slope. A technique for handling certain types of non-random errors is presented.

THE scientist is often called upon to obtain derivatives of functions representing his experimental data. With the advent of the digital computer there is no longer any excuse for plotting the data by hand and then using "optical" methods (Simons, 1941) which are often unintentionally biased as well as tedious. On the other hand, the digital computer can generate such a quantity of numbers that the correct interpretation of the results is obscured or overlooked. This article emphasizes procedures for numerical differentiation of experimental data and estimation of the error in the slope, reviews the assumptions behind the methods, and gives some techniques for numerical differentiation of data with non-random errors.

Experimental data are always subject to errors. These errors prevent a simple difference technique from being used to obtain derivatives, as is graphically shown by Ralston (1965). Such a scheme of approximating dY/dX by $\Delta Y/\Delta X$ at the point of interest results in the loss of significant figures in both numerator and denominator. Thus even the sign of the derivative is often incorrect, and the simple difference technique is completely unsatisfactory.

In numerical differentiation no attempt is made to represent all the data by a single function of high order. Although this technique is often satisfactory for integration, it usually leads to results as erroneous as those from the simple difference technique. The best estimate of the derivative at a point is to fit a function to several data points on both sides of the particular point and then differentiate the resulting function analytically at that point. The best procedure is to start at the top of the table of data and move stepwise down, evaluating

¹ Present address, Department of Chemical Engineering, Ohio State University, Columbus, Ohio.

² Present address, Department of Chemistry, University of Southern California, Los Angeles, Calif.

the derivative at each center point by a "movable strip" or "moving arc" technique (Lanczos, 1956; Sasuly, 1934). Thus the problem becomes one of deciding on the type of function to represent the data, its degree if it should be some type of polynomial, how many points to be included at each step, how to reduce the error present in the data by "smoothing" before differentiation, and how to estimate the error in the derivative.

Type of Function

Many types of functions have been introduced into the literature, but most of them lead to useful approximations only in particular applications. Common examples are Fourier functions (for periodic phenomena) and exponential functions. But for general application, polynomials are the most useful functions for approximations. In fact, polynomials will usually be satisfactory for numerical differentiation of both periodic and exponential data.

Polynomial approximations are of three general types: (1) interpolation approximations, of which Newton's interpolation formulas with divided differences and Lagrangian functions are two of many examples (Hildebrand, 1956; Lanczos, 1956; Ralston, 1965); (2) least squares approximations where the sum of the squares of the differences between $f(X)$ and its approximation is minimized over a discrete set of points (Lapidus, 1962); and (3) minimization of the maximum error approximations (Ralston, 1965). Of these, the least squares approximations are most useful for representing and differentiating experimental data (Ralston, 1965).

In least squares analysis the following model is assumed to represent the data:

$$f(X) = \sum_{j=0}^m \alpha_j P_j(X) + \epsilon \quad (1)$$

where m is the degree of the polynomial, α_j is the j th polynomial coefficient, $P_j(X)$ is some as yet unspecified function of the independent variable X and is of degree j , and ϵ is the error which is independently distributed at each point (random) with mean zero and variance σ^2 . In least squares analysis all the error is assumed to be in the dependent variable, f . This simplifies the error analysis.

A discussion of least squares polynomials is often divided into two sections: simple power functions and orthogonal polynomials. These are discussed separately here.

Least Squares Power Functions. For this case Equation 1 is simplified into a power series by equating $P_j(X)$ to X^j :

$$f(X) = \sum_{j=0}^m \beta_j X^j + \epsilon \quad (2)$$

where β_j are the power coefficients. It is easily shown that minimization of the sum of the squares of the residuals leads to the following set of simultaneous equations written in matrix form (Graybill, 1961):

$$(\mathbf{X}'\mathbf{X})(\hat{\mathbf{B}}) = (\mathbf{X}'\mathbf{Y}) \quad (3)$$

where the unknown coefficients $\hat{\beta}_j$ of the vector $\hat{\mathbf{B}}$ ($\hat{\beta}_j$ is thus the best linear unbiased estimate of β_j) are found by Gauss-Jordan elimination (Hildebrand, 1963) with a pivot picker (Orden, 1962) [or some other suitable method] and the definitions of matrices $(\mathbf{X}'\mathbf{X})$ and $(\mathbf{X}'\mathbf{Y})$ are:

$$(\mathbf{X}'\mathbf{X}) = \begin{pmatrix} N & \sum X_i & \sum X_i^2 & \dots & \sum X_i^m \\ \sum X_i & \sum X_i^2 & \sum X_i^3 & \dots & \sum X_i^{m+1} \\ \sum X_i^2 & \sum X_i^3 & \sum X_i^4 & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ \sum X_i^m & \sum X_i^{m+1} & \dots & \dots & \sum X_i^{2m} \end{pmatrix} \quad (4)$$

$$(\mathbf{X}'\mathbf{Y}) = \begin{pmatrix} \sum f_i \\ \sum X_i f_i \\ \sum X_i^2 f_i \\ \dots \\ \sum X_i^m f_i \end{pmatrix} \quad (5)$$

The computational problems in finding the least squares coefficients (vector $\hat{\mathbf{B}}$) and the reasons for them have been discussed at length (Bright and Dawkins, 1965; Lapidus, 1962; Orden, 1962; Ralston, 1965). The rule of thumb usually quoted (Bright and Dawkins, 1965; Lapidus, 1962) is that serious problems arise from roundoff error for $m \geq 6$. The authors' experience is that computational problems can start at $m \geq 3$, if the data have little error and all numbers are rounded off after eight digits at every stage of the calculation.

Table I shows the least square coefficients for a cubic fit of water density vs. temperature ("Handbook of Chemistry and Physics," 1955) as calculated in single precision arithmetic on an IBM 1620 Model II (8 digits) and in double precision after scaling f by subtracting the average of the f 's (to four digits). There were 21 points in this set of data between 20° and 40° C., with the independent variable ranging from 20.00 to 40.00 and the dependent variable from 0.99262 to 0.99823. A Gauss-Jordan subroutine with a pivot picker was used to solve the simultaneous equations. The results in Table I show that roundoff error can reduce single precision coefficients (word length of 8) to meaningless digits if the data are sufficiently precise. [As ϵ becomes smaller in Equation 2, the roundoff error becomes more serious, because the determinant of $(\mathbf{X}'\mathbf{X})$ approaches zero.]

To be conservative as to the number of digits needed in the computation to remove the possibility of roundoff error, a criterion of the largest number of significant digits in the independent variable times $2m - 1$ is proposed, since this is the number of digits in the largest element of $(\mathbf{X}'\mathbf{X})$ —i.e., $\sum X_i^{2m}$. However, for any given set of data a little experimentation with the highest degree polynomial to be selected, using a variety of word lengths in the calculation, will show exactly how many digits are needed.

Orthogonal Least Squares Polynomials. The definition of a set of orthogonal polynomials $Q_j(X)$ is:

$$\sum_{i=1}^N Q_j(X_i) Q_k(X_i) = 0 \text{ if } j \neq k \quad (6)$$

Table I. Least Squares Cubic Coefficients for Density of Water vs. Temperature

Word Length	Scaling	$\hat{\beta}_0$	$\hat{\beta}_1 \times 10^6$	$\hat{\beta}_2 \times 10^6$	$\hat{\beta}_3 \times 10^7$	Standard Deviation, S
8	No	0.99282590	79.162407	-3.2574219	3.1187873	0.00005054
16	Yes	1.00025036	1.42258121	-0.61528822	0.19782108	0.00000239

The restriction of Equation 6 can be applied to the model of Equation 1 and the result expressed as:

$$f(X) = \sum_{j=0}^m \gamma_j Q_j(X) + \epsilon \quad (7)$$

where γ_j become the least squares orthogonal coefficients.

Perhaps the easiest set of orthogonal polynomials for use on a digital computer is that proposed by Forsythe (1957), which is defined as:

$$\begin{aligned} Q_{-1}(X) &= 0 \\ Q_0(X) &= 1 \\ Q_1(X) &= (X - a_1)Q_0(X) \\ Q_2(X) &= (X - a_2)Q_1(X) - b_1Q_0(X) \\ Q_3(X) &= (X - a_3)Q_2(X) - b_2Q_1(X) \\ &\dots \\ Q_j(X) &= (X - a_j)Q_{j-1}(X) - b_{j-1}Q_{j-2}(X) \end{aligned} \quad (8)$$

where a_j , b_j , and $\hat{\gamma}_j$ (the best linear unbiased estimate of γ_j) are calculated from:

$$a_j = \frac{\sum_{i=1}^N X_i [Q_{j-1}(X_i)]^2}{\sum_{i=1}^N [Q_{j-1}(X_i)]^2} \quad (9)$$

$$b_j = \frac{\sum_{i=1}^N [Q_j(X_i)]^2}{\sum_{i=1}^N [Q_{j-1}(X_i)]^2} \quad (10)$$

$$\hat{\gamma}_j = \frac{\sum_{i=1}^N f_i Q_j(X_i)}{\sum_{i=1}^N [Q_j(X_i)]^2} \quad (11)$$

It is obvious from Equations 8 through 11 that the result of the orthogonalization is to eliminate the need to solve a linear system of equations by reduction. Instead the solution has become trivial, in that successive substitution yields all the unknown coefficients. Thus the higher coefficients are calculated without changing the values of the lower coefficients, in contrast to the β_j of the power function least squares which are changed when the degree of the polynomial in the model is increased. Least squares orthogonal polynomials have been treated in depth elsewhere (Bright and Dawkins, 1965; Forsythe, 1957; Hildebrand, 1956; Lanczos, 1956; Lapidus, 1962; Ralston, 1965; Sasuly, 1934). But it should be emphasized that the least squares orthogonal polynomials, when fitted to a particular set of data, will yield the same coefficients $\hat{\beta}_j$ in Equation 2 as the solution to Equation 3, when all the constants in Equation 7 have been evaluated and factored into coefficients of simple powers of X . This is natural, since the sum of the squares of the differences between $f(X)$ and its approximation $\sum_{j=0}^m \hat{\gamma}_j Q_j(X)$ [hereafter called $Y(X)$] was minimized in the derivation, and only one equation of degree m will fit a given set of data in the least squares sense.

The principal advantages of orthogonal polynomials in numerical differentiation are for data equally spaced in the independent variable. [The case of data equally spaced in $f(X)$ can also be handled.] For unevenly spaced data, computation is as long for the low order orthogonal polynomials as for the power function least squares (Bright and Dawkins, 1965), and in computing centers where library routines are available for the Gauss-Jordan reduction the programming of orthogonal polynomials is somewhat more tedious.

If we agree to the restriction of an odd number, N' , of equally

spaced data points to be represented by least squares orthogonal polynomials, a change in variables from X to s where s is given by:

$$i = L + s \quad (12)$$

and

$$L = \frac{N' - 1}{2} \quad (13)$$

lets s go from $-L, -L + 1, \dots, 0, \dots, L - 1, L$ as i goes from 0 to N' . Since a dummy variable with integer values has replaced X , the least squares orthogonal polynomial becomes (Ralston, 1965):

$$f(s) = \sum_{j=0}^m \gamma_j Q_j(s, 2L) + \epsilon \quad (14)$$

where

$$\begin{aligned} Q_0 &= 1 \\ Q_1 &= s/L \\ Q_2 &= \frac{3s^2 - L - L^2}{L(2L - 1)} \\ &\dots \\ Q_j(s, 2L) &= \sum_{k=0}^j (-1)^{k+j} \frac{(j+k)^{(2k)}}{(k!)^2} \cdot \frac{(L+s)^{(k)}}{[2L]^{(k)}} \end{aligned} \quad (15)$$

with the factorial functions $(j+k)^{(2k)}$, $(L+s)^{(k)}$, and $[2L]^{(k)}$ defined according to the following example of $X^{(k)}$:

$$X^{(k)} = X(X-1)(X-2)(X-3) \dots (X-k+1) \quad (16)$$

and

$$X^{(0)} = 1 \quad (17)$$

etc., and:

$$\hat{\gamma}_j = \frac{\sum_{s=-L}^L f_s Q_j(s, 2L)}{\frac{(2L+j+1)!(2L-j)!}{(2j+1)[(2L)!]^2}} \quad (18)$$

To reconvert Equation 14 into terms of X , the substitution:

$$s = \frac{X - X_0}{h} - L \quad (19)$$

where X_0 is the value of X at $i = 0$ and h is $X_{i+1} - X_i$, yields the desired equation.

As can be seen from Equations 14 through 18, much of the computation involved is done only once after the number of points and the degree of fit have been chosen. For example, the third-degree seven-point formula for Y evaluated at $s = 0$ is

$$Y_0 = \frac{1}{21} (-2f_{-3} + 3f_{-2} + 6f_{-1} + 7f_0 + 6f_1 + 3f_2 - 2f_3) \quad (20)$$

The computing time to find Y_0 in Equation 20 is several orders of magnitude less than to find Y_0 by calculating $(\mathbf{X}'\mathbf{X})$ and $(\mathbf{X}'\mathbf{Y})$ from Equations 4 and 5, then \mathbf{B} in Equation 3, and lastly the generated function, \mathbf{Y} , from Equation 2. The final formulas for the derivative of Equation 14 are equally simple.

Smoothing

Practically all experimental data require smoothing before differentiation is performed. Smoothing (Hildebrand, 1956; Lanczos, 1956; Ralston, 1965) consists of passing a least

squares polynomial through the data points in order to determine α_j and the constants in $P_j(X)$ and then replacing f_i with the generated function $Y(X_i)$ as calculated from

$$\sum_{j=0}^m \alpha_j P_j(X_i)$$

As discussed for the general case in numerical differentiation, the best method is to select a small, odd number of points, evaluate Y at the center, and proceed using the "movable strip" technique. At the top and bottom of the table it will be necessary to calculate the points between the center and the extremity from off-center formulas (Hildebrand, 1956). Hildebrand (1956) points out that the amount of smoothing increases with the number of points in the movable strip and decreases with increasing m , the degree of the polynomial.

Each set of data must be judged as to the proper N' and m , and so generalization is difficult. However, selection of N' and m is not so critical in smoothing as in the actual differentiation, because the number of times that smoothing is applied is also controlled. The authors have had good results with $m = 3$ and $N' = 5$ or $N' = 7$ formulas. Often no significant improvement is found after smoothing has been applied more than four times. However, the smoothing has to be terminated at the stage where the random error is eliminated and yet the basic shape and character of the data have been unaltered.

Selection of Degree of Fit in Numerical Differentiation

It is advantageous to perform numerical differentiation with m as low as possible, because as m approaches $N - 1$, the least squares polynomial approximates the data points more closely at the expense of increasing the variations in the derivative. When $m = N - 1$, the polynomial fits the data exactly (passes through each point) and may have as many as $N - 2$ points at which the slope changes sign. Hence its derivative often fluctuates widely and in general is a poor representation of the slope of the points.

A second-degree least squares polynomial will give excellent results on almost all sets of data, unless the points are unusually far apart or sparse. However, a statistical test can be made if there is reason to suspect that a higher degree would produce better results. If ϵ in Equation 1 can be assumed to be distributed normally as well as independently with mean zero and variance σ^2 , an analysis of variance (AOV) can be performed to test the hypothesis $\alpha_{m+1} = 0$, given two sets of least squares polynomials of degree m and $m + 1$ (Graybill, 1961), and thereby determine if the polynomial of degree m adequately represents the data.

Since the AOV requires that both the $m + 1$ degree series of least squares coefficients and the m degree series of coefficients be determined in order to test the hypothesis $\alpha_{m+1} = 0$, the m degree coefficients will be called $\Phi_0, \Phi_1, \dots, \Phi_m$ and the $m + 1$ coefficients $\psi_0, \psi_1, \dots, \psi_{m+1}$. The first step is to find the least squares coefficients in the following equations:

$$Y = \sum_{j=0}^m \Phi_j X^j \quad (21)$$

$$Y = \sum_{j=0}^{m+1} \psi_j X^j \quad (22)$$

The equations for the AOV are:

$$F_{m+1} = \frac{A_{m+1}}{E_{m+1}} \quad (23)$$

where F_{m+1} is the ratio of the mean sum of squares for ψ_{m+1} set of coefficients, adjusted for the variation that can be removed by the Φ_j coefficients in Equation 21, A_{m+1} :

$$A_{m+1} = [\psi_0 \sum f_i + \psi_1 \sum X_i f_i + \psi_2 \sum X_i^2 f_i + \dots + \psi_{m+1} \sum X_i^{m+1} f_i] - [\Phi_0 \sum f_i + \Phi_1 \sum X_i f_i + \Phi_2 \sum X_i^2 f_i + \dots + \Phi_m \sum X_i^m f_i] \quad (24)$$

to the error in the fit of Equation 22, E_{m+1} :

$$E_{m+1} = [\sum f_i^2 - \psi_0 \sum f_i - \psi_1 \sum X_i f_i - \dots - \psi_{m+1} \sum X_i^{m+1} f_i] / [N - m - 2] \quad (25)$$

If F_{m+1} is less than the value of the central F statistic with 1 degree of freedom in the numerator and $(N - m - 2)$ degrees in the denominator at the desired confidence level ($p = 0.05$ is convenient), the data are adequately represented by Equation 21, since the α_{m+1} —i.e., ψ_{m+1} —coefficient makes no significant improvement in the fit.

The AOV to test the hypothesis $\alpha_{m+1} = 0$ frequently must be interpreted carefully, because sometimes ϵ is not normally distributed. For example, data taken on electrical instruments throughout several ranges, each range differing as a result of different size resistors in series at the time of the measurements, will not always have a common normal distribution. Often best results will be obtained by assuming $\alpha_{m+1} = 0$ if F_{m+1} is reasonably close to the F statistic, even when the F statistic is determined at $p = 0.05$ or $p = 0.01$. Some statisticians recommend that if α_{m+1} is found not to be significant, α_{m+2} be checked also, just in case the true function of the data were odd or even. However, this is usually unnecessary for experimental data.

If the AOV is used to determine the lowest order polynomial to be used in sloping the data, the movable strip technique will have to be applied in order to see that all portions are well represented by the degree finally used.

The AOV is even more sensitive to roundoff error in the computation than the solution of the least squares matrix. The subtractions in both Equations 24 and 25 are almost always between nearly equal numbers. After the subtraction, Equation 23 requires a division. Thus the AOV will require a careful choice of word length (discussed in conjunction with the solution of Equation 3).

Selection of Number of Points in Each Strip

There is no known test to assist in the selection of N' , the number of data points to be included in each strip. The larger N' , the less the error in any given data point affects the slope of the curve at that and adjacent points. A three-point formula is unsatisfactory for experimental data, because a second-degree polynomial fits all three points exactly. A seven-point polynomial often gives good results (Zakin *et al.*, 1966). A five-point formula works well when the data are not too "bumpy," whereas a nine-point or higher polynomial may be required for data with large errors, even after smoothing.

Numerical Differentiation of Equally Spaced Data

Smoothing Formulas. Orthogonal polynomials are the most convenient for evenly spaced data. The most useful smoothing formulas are the third-degree five-point and the third-degree seven-point. Table II supplies the coefficients for smoothing formulas to be used in Equation 26:

$$Y_s = \frac{1}{D} \sum_{i=-s}^{i=s} (C_i f_i) \quad (26)$$

Table II. Orthogonal Least Squares Coefficients for Third-Degree Smoothing Formulas

N'	S	D	C_{-3}	C_{-2}	C_{-1}	C_0	C_1	C_2	C_3
5	-2	70	...	69	4	-6	4	-1	...
	-1	35	...	2	27	12	-8	2	...
	0	35	...	-3	12	17	12	-3	...
	1	35	...	2	-8	12	27	2	...
7	-3	42	39	8	-4	-4	1	4	-2
	-2	42	8	19	16	6	-4	-7	4
	-1	42	-4	16	19	12	2	-4	1
	0	21	-2	3	6	7	6	3	-2

where the formulas for Y_1 , Y_2 , and Y_3 are obtained from symmetry by multiplying the subscript of every f_i by (-1) ; as shown in Table II for the five-point Y , Y_0 is the smoothed value at the center of the strip ($s = 0$), etc. The off-center equations are used only at the beginning and end of the data table. Equation 20 serves as an example of how to use Table II to construct the desired equations.

Sloping Formulas. After the data have been smoothed, N' and m are selected and the suitable formulas applied. Table III supplies the coefficients for the more commonly used sloping formulas to be used in Equation 27:

$$\frac{dY_s}{dX} = \frac{1}{Dh} \sum_{i=-s}^{i=s} (C_i f_i) \quad (27)$$

where h is the spacing in X —i.e., $X_{i+1} - X_i$.

The formulas for dY_1/dX , dY_2/dX , etc., are obtained by multiplying the subscript of every f_i by (-1) and then changing the sign of the coefficient of f_i , as shown in Table III for the dY_1/dX second-degree, five-point formula.

Additional smoothing and differentiation formulas may be derived if needed from Equations 14 through 19.

Numerical Differentiation of Data Not Equally Spaced

For data not equally spaced the power function least squares are recommended for both smoothing and differentiation. Although the computing time is roughly the same for both the power function and the orthogonal cases, the power functions are easier to program. Furthermore, only a trivial change in the program is required to change N' and a few extra statements are required to change m in the power function case, whereas the orthogonal sets require more programming in addition to calculating the new polynomials from Equations 8 through 11.

Table III. Orthogonal Least Squares Coefficients for Sloping Formulas

N'	m	S	D	C_{-4}	C_{-3}	C_{-2}	C_{-1}	C_0	C_1	C_2	C_3	C_4	
5	2	-2	70	-54	13	40	27	-26	
		-1	70	-34	3	20	17	-6	
		0	10	-2	-1	0	1	2	
		1	70	6	-17	-20	-3	34	
7	2	-3	84	-39	-6	15	24	6	-21	...	
		-2	84	-29	-6	9	16	15	6	-11	...
		-1	84	-19	-6	3	8	9	6	-1	...
		0	28	-3	-2	-1	0	1	2	3	...
9	2	-4	4620	-1428	-511	166	603	800	757	474	-49	-812	
		-3	4620	-1148	-441	86	433	600	587	394	21	-532	
		-2	4620	-868	-371	6	263	400	417	314	91	-252	
		-1	4620	-588	-301	-74	93	200	247	234	161	28	
0	60	-4	-3	-2	-1	0	1	2	3	4	...		
7	3	-3	252	-446	311	374	72	-266	-311	266	...
		-2	252	-206	101	146	48	-74	-101	86	...
		-1	252	-50	-25	2	24	34	25	-10	...
		0	252	22	-67	-58	0	58	67	-22	...

The second-degree normal equations from Equations 3 to 5 are:

$$\begin{pmatrix} N & \sum X_i & \sum X_i^2 \\ \sum X_i & \sum X_i^2 & \sum X_i^3 \\ \sum X_i^2 & \sum X_i^3 & \sum X_i^4 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix} = \begin{pmatrix} \sum f_i \\ \sum X_i f_i \\ \sum X_i^2 f_i \end{pmatrix} \quad (28)$$

and so on.

For smoothing, coefficients $\hat{\beta}_j$ are found as in the least squares power functions, and $Y(X_i)$ are calculated from $\sum_{j=0}^m \beta_j X_i^j$ for the center or off-center points, as the case may be. For differentiation, Equation 28 is solved as above and the estimated derivative $dY(X_i)/dX$ is calculated from the derivative of $Y(X_i)$:

$$\frac{dY(X_i)}{dX} = \sum_{j=1}^m (j) \beta_j X_i^{j-1} \quad (29)$$

Estimate of Error in Slope

A convenient estimate of the error in the slope as calculated by the above methods is provided by a confidence set on the slope.

Two-sided confidence limits are defined as $\pm t(p/2, N - m - 1)S$, where $t(p/2, N - m - 1)$ has the Student's distribution with $N - m - 1$ degrees of freedom and S is defined in Equation 32. If ϵ is distributed normally and independently with mean zero and variance σ^2 , the probability that the true value of the slope, $dY_T(X_i)/dX$, lies in the interval:

$$\frac{dY(X_i)}{dX} - t(p/2, N - m - 1)S \leq \frac{dY_T(X_i)}{dX} \leq \frac{dY(X_i)}{dX} + t(p/2, N - m - 1)S \quad (30)$$

is $1 - p$. This statement is properly interpreted as follows: If points X_i are measured repeatedly in proper statistical replications, slopes $dY(X_i)/dX$ are expected to be in the confidence interval expressed by Equation 30 in the fraction of experiments equal to $1 - p$.

The estimation of the confidence interval of a slope $dY(X_i)/dX$ is really a special case for finding confidence intervals of any linear combination of β_j and is treated in the general case elsewhere (Graybill, 1961). First the following vector is defined:

$$g(\mathbf{X}) = \begin{pmatrix} 0 \\ 1 \\ 2X \\ 3X^2 \\ \vdots \\ mX^{m-1} \end{pmatrix} \quad (31)$$

where $\mathbf{g}(\mathbf{X})$ is the vector representing dY/dX . Then the confidence interval on dY/dX for any X is:

$$\frac{dY}{dX} - t(p/2, N - m - 1) \sqrt{\mathbf{g}'(X)(\mathbf{X}'\mathbf{X})^{-1} \mathbf{g}(X) \hat{\sigma}^2} \leq \frac{dY_T}{dX} \leq \frac{dY}{dX} + t(p/2, N - m - 1) \sqrt{\mathbf{g}'(X)(\mathbf{X}'\mathbf{X})^{-1} \mathbf{g}(X) \hat{\sigma}^2} \quad (32)$$

where $\mathbf{g}'(\mathbf{X})$ is the transpose of $\mathbf{g}(\mathbf{X})$, $(\mathbf{X}'\mathbf{X})^{-1}$ the inverse of the matrix $(\mathbf{X}'\mathbf{X})$ in Equation 4 and $\hat{\sigma}^2$ is given by:

$$\hat{\sigma}^2 = \frac{(\mathbf{Y}'\mathbf{Y}) - (\hat{\mathbf{B}})(\mathbf{X}'\mathbf{Y})}{N - m - 1} \quad (33)$$

with

$$(\mathbf{Y}'\mathbf{Y}) = \sum f_i^2 \quad (34)$$

and $(\hat{\mathbf{B}})$ and $(\mathbf{X}'\mathbf{Y})$ defined in Equation 3.

The confidence interval calculation in the case of the slope at center point of an orthogonal polynomial with equal X spacing is considerably shortened because $\mathbf{g}(\mathbf{X})$ simplifies to:

$$\mathbf{g}(\mathbf{X}) = \begin{matrix} 0 \\ 1 \\ 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{matrix} \quad (35)$$

Example. Table IV presents a set of experimentally determined length-temperature measurements (coded to save space) and typical results of smoothing and differentiating according to the techniques described previously. For the smoothing operations a third-degree seven-point orthogonal polynomial from Table II and Equation 26 was chosen:

$$Y \text{ smoothed } (X = 0.1) = \frac{1}{42} [(39)(0.551) + (8)(0.610) - (4)(0.663 + 0.717 - 0.875) + (1)(0.776) - (2)(0.880)] = 0.556$$

$$Y \text{ smoothed } (X = 0.5) = \frac{1}{21} [(-2)(0.610 + 0.935) + (3)(0.663 + 0.880) + (6)(0.717 + 0.875) + (7)(0.776)] = 0.787$$

The smoothed values were then resmoothed, and so on, with the results shown in Table IV.

A second-degree seven-point orthogonal polynomial (Table III and Equation 28) was used to differentiate the data. An example calculation is shown for the case of no smoothing and $X = 0.5$:

$$\frac{dY}{dX} (X = 0.5) = \frac{1}{(28)(0.1)} [(3)(0.935 - 0.610) + (2)(0.880 - 0.663) + (0.875 - 0.717)] = 0.560$$

The 16 points in Table IV were chosen from the middle of a data set of 125 points in order to illustrate the advantages of smoothing data before differentiation. The 16 points are in a region of relatively constant slope, yet at point $X = 0.6$ there is a glaring inconsistency in the data for which the scientists who took the data had no positive explanation. The inconsistent point affects $2N' - 1$ (or in this case 13) slopes in the moving strip technique. The resultant discontinuity in slopes for the no smoothing case ($NS = 0$) is clearly shown in Figure 1, whereas a much better estimate of the slope is obtained after four smoothings ($NS = 4$).

Table IV also includes values of Y after 20 smoothings. The important thing to note is that little change occurred between four and 20 smoothings. After four smoothings the "noise" in the data was reduced to a very reasonable level and little benefit was gained from further smoothings.

Method for Sloping Data with Nonrandom Errors and for Improving the Estimate of Transitions

In considering nonrandom errors in experimental data, no general statement can be made because each case must be scrutinized individually. The effects on sloping of some nonrandom errors can be at least minimized by a technique called "breaking."

Sometimes data taken on electric instruments are subject to unavoidable jumps or shifts. An example is length-temperature measurement on polymer samples (Zakin *et al.*, 1966). Differentiation of length-temperature data yields the location and magnitude of glassy transitions. One bad point affects $2N - 1$ slopes to a varying extent in the movable strip technique, although smoothing can handle the case of only one bad point, as shown in Figure 1. But if the instrument zero shifts (or if some other occurrence produces a similar discontinuity), smoothing will not significantly improve the estimate.

Table IV. Example Results^a

X	$Y(NS = 0)$	$Y(NS = 1)$	$Y(NS = 4)$	$Y(NS = 20)$	$\frac{dY}{dX}(NS = 0)$	$\frac{dY}{dX}(NS = 4)$
0.1	0.551	0.556	0.555	0.553	0.604	0.561
0.2	0.610	0.600	0.603	0.606	0.596	0.565
0.3	0.663	0.659	0.660	0.661	0.589	0.570
0.4	0.717	0.726	0.722	0.718	0.582	0.574
0.5	0.776	0.787	0.783	0.775	0.560	0.571
0.6	0.875	0.844	0.840	0.832	0.540	0.554
0.7	0.880	0.894	0.892	0.888	0.527	0.537
0.8	0.935	0.940	0.942	0.943	0.520	0.530
0.9	0.987	0.985	0.994	0.997	0.519	0.534
1.0	1.048	1.050	1.048	1.051	0.565	0.541
1.1	1.110	1.107	1.105	1.105	0.561	0.548
1.2	1.168	1.163	1.160	1.159	0.557	0.551
1.3	1.211	1.217	1.215	1.214	0.548	0.554
1.4	1.269	1.268	1.269	1.269	0.547	0.556
1.5	1.326	1.323	1.324	1.325	0.546	0.557
1.6	1.382	1.383	1.383	1.383	0.545	0.559

^a NS refers to number of smoothings.

Figure 2 shows some length-temperature data for a mixture of 52% polypropylene and 48% polyethylene measured with a linear voltage differential transformer (LVDT) (Zakin *et al.*, 1966). At about -137°C . there is an unexplained shift in the LVDT output voltage. Figure 3 compares expansion coefficients $(1/L_0) dL/dT$ calculated from these data by smoothing four times, using Equation 26 and Table II with $N' = 7$ and differentiating using Equation 28 and Table III with $N' = 7$ and $m = 2$, with those obtained by "breaking" the data at -137°C ., smoothing four times, and differentiating (using the same equations). The "break" consists of terminating the moving strip on one side of the discontinuity and starting the strip again on the other side. There is considerable improvement when the break is used as can be seen in Figure 3: The

discontinuity is not breached by any smoothing or sloping polynomial, and the smoothing formulas tend to straighten out the points near the discontinuity, so that the slopes form a smooth curve, despite the discontinuity.

A break in the smoothing and sloping process can also assist in the location of a minor transition and estimation of the magnitude of a major transition. Figure 4 shows length-temperature measurements for a polyurethane sample. Figure 5 shows the results of smoothing four times and sloping with breaks at -137° , -60° , and -39°C . Using a graphical method in Figure 4, the transition at the lowest temperature appears to occur at -132°C . However, the numerical differentiation techniques with a break show that -137°C . is the better value. Breaking at the higher temperatures

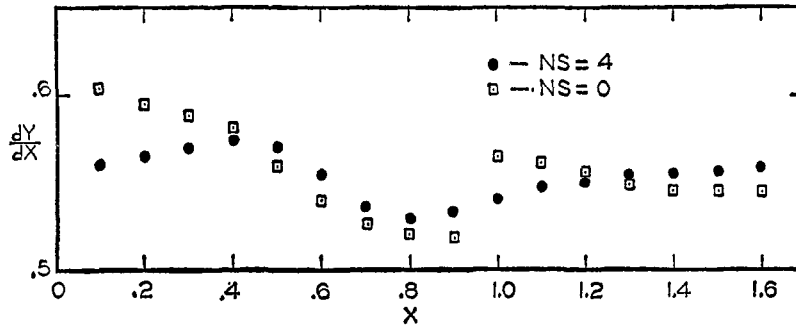


Figure 1. Effect of smoothing on slopes of data in Table IV

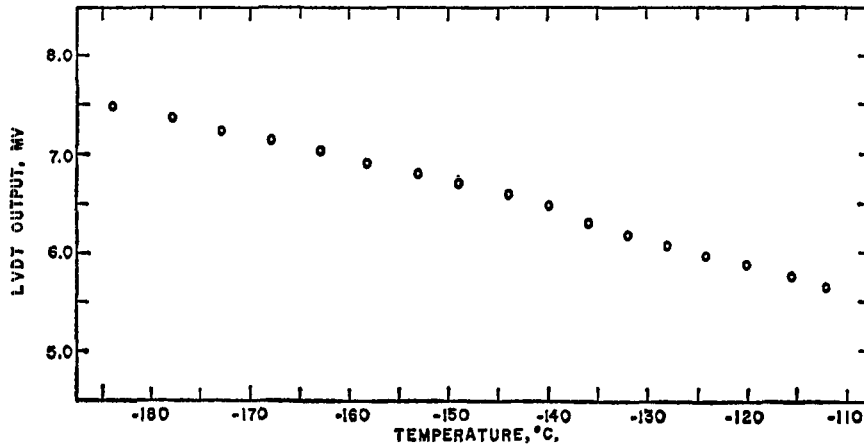


Figure 2. LVDT output vs. temperature data for mixture of 52% polypropylene and 48% polyethylene

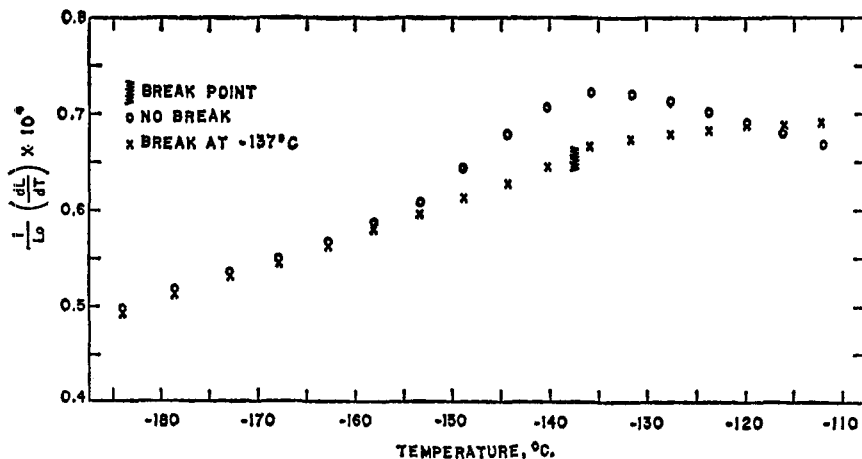


Figure 3. Effect of breaking on differentiation with smoothing on linear expansion coefficient as a function of temperature for data of Figure 2

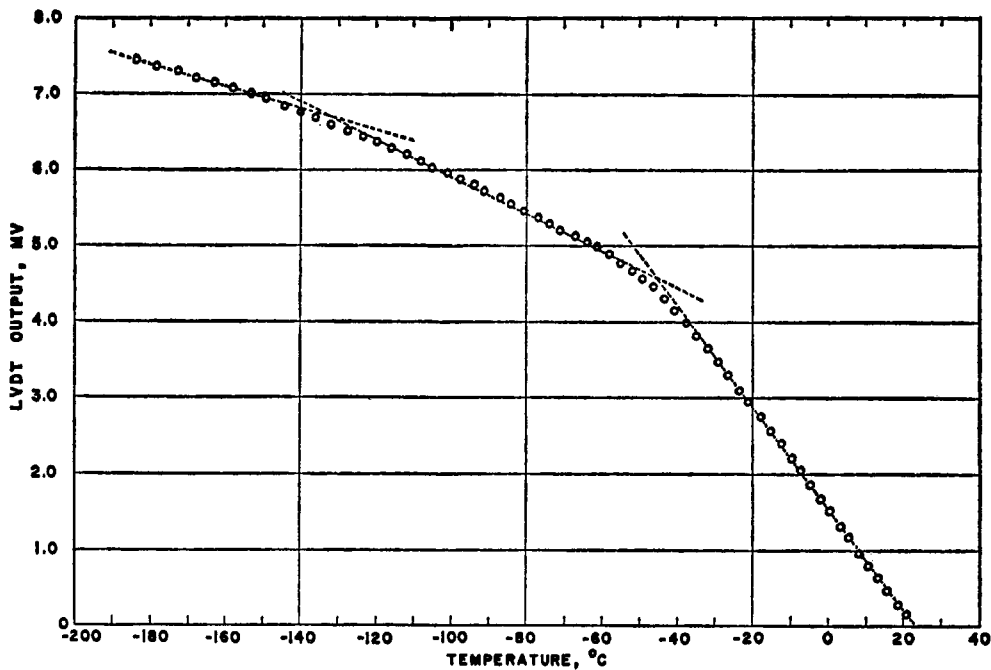


Figure 4. Graphical method for determining transition temperatures on LVDT output vs. temperature data for a polyurethane

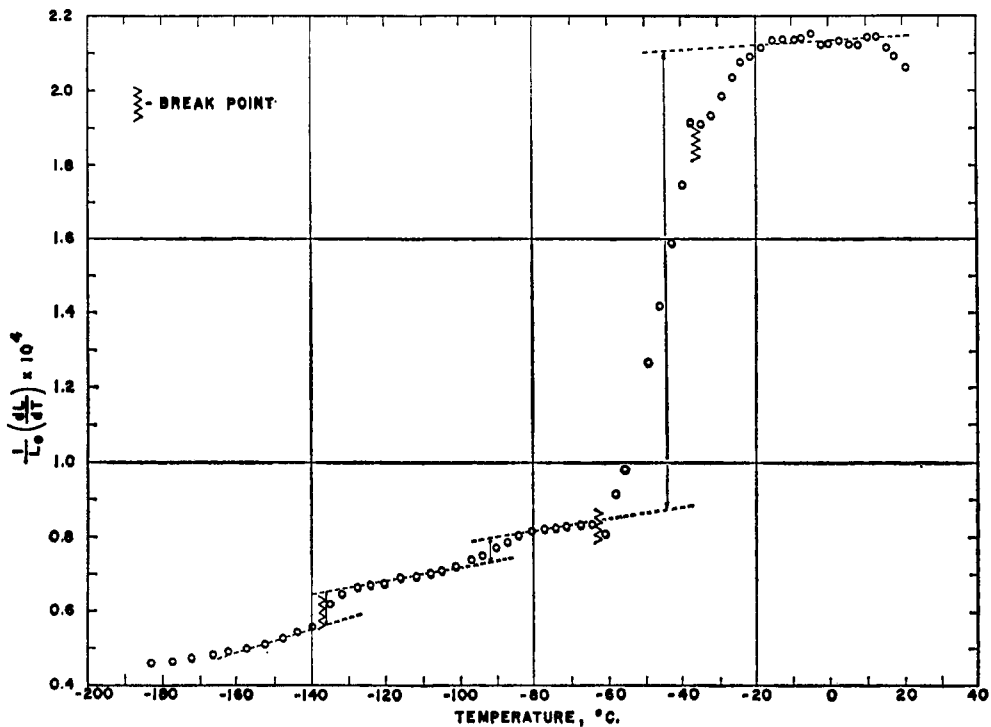


Figure 5. Differentiation with smoothing and breaks for determining transition temperatures from data of Figure 4

prevented the smoothing from "smearing out" the major transition and permitted a better estimate of the change in the magnitude of the expansion coefficient.

Care must be taken to ensure a break at the proper location. The authors' experience has been that if the wrong point were chosen initially it was obvious from the resulting plot.

Acknowledgment

The authors are indebted to the Computer Science Center, University of Missouri at Rolla, and the Computer Center,

Ohio State University, for making computer time available. Part of the work was done under NASA Grant NSG-343 to the University of Southern California. The helpful comments of Charles E. Antle are acknowledged.

Nomenclature

- a_j = defined by Equation 9
- A_{m+1} = defined by Equation 24
- AOV = analysis of variance
- (\hat{B}) = $N \times 1$ matrix containing $\hat{\beta}_j$
- b_j = defined by Equation 10

C = constant in Equation 26 or 27
 D = constant in Equation 26 or 27
 E_{m+1} = variance, defined by Equation 25
 F = central F statistic
 F_{m+1} = defined by Equation 23
 $f, f(X), f_s, f_i$ = dependent variable
 $\mathbf{g}(\mathbf{X})$ = $N \times 1$ matrix defined by Equation 31
 $\mathbf{g}'(\mathbf{X})$ = $1 \times N$ matrix, transpose of $\mathbf{g}(\mathbf{X})$
 h = spacing in equal-spaced independent variable, $X_{i+1} - X_i$
 i, j, k = summation indices
 (k) = factorial notation, defined by Equations 16 and 17
 L = defined by Equation 13
 m = degree of polynomial
 N = number of data points included in polynomial
 N' = odd N
 NS = number of smoothings
 $P_j(X)$ = least squares polynomial of degree j , general case, in Equation 1
 p = level of significance
 Q_j = orthogonal least squares polynomial of degree j , defined for the general case by Equation 8, for the equal-spaced independent variable case by Equation 15
 s = translated independent variable, of Equation 12
 S = $\sqrt{\mathbf{g}'(\mathbf{X})(\mathbf{X}'\mathbf{X})^{-1}\mathbf{g}(\mathbf{X})\hat{\sigma}^2}$
 $t(p/2, n - m - 1)$ = Student's t statistic, at $p/2$ level of significance with $n - m - 1$ degrees of freedom
 X, X_i, X_o = independent variable
 $(\mathbf{X}'\mathbf{X})$ = $N \times N$ matrix defined by Equation 4
 $(\mathbf{X}'\mathbf{X})^{-1}$ = inverse of $(\mathbf{X}'\mathbf{X})$
 $(\mathbf{X}'\mathbf{Y})$ = $N \times 1$ matrix, defined by Equation 5
 Y = approximation of f , calculated from $\sum_{j=0}^m \hat{\gamma}_j Q_j(X)$ in the case of equally spaced data and from $\sum_{j=0}^m \hat{\beta}_j X^j$ for nonequally spaced data
 $(\mathbf{Y}'\mathbf{Y})$ = defined by Equation 34
 α_j = j th coefficient of generalized least squares polynomial model in Equation 1

β_j = j th least squares power function coefficient of model in Equation 2
 $\hat{\beta}_j$ = best linear unbiased estimate (least squares) of β_j , calculated from Equation 3
 γ_j = j th least squares orthogonal coefficient of model in Equation 7
 $\hat{\gamma}_j$ = best linear unbiased estimate of γ_j , calculated from Equation 11 or 18
 Δ = difference, $i + 1$ point minus i point
 ϵ = random error
 σ^2 = variance
 $\hat{\sigma}^2$ = estimated variance, defined by Equation 33
 Φ_j = m th degree least squares coefficients in Equation 22
 ψ_j = $m + 1$ th degree least squares coefficients in Equation 21

Literature Cited

- Bright, J. W., Dawkins, G. S., *IND. ENG. CHEM. FUNDAMENTALS* **4**, 93 (1965).
 Forsythe, G. E., *J. Soc. Ind. Appl. Math.* **5**, 74 (1957).
 Graybill, F. A., "Introduction to Linear Statistical Models," Vol. I, McGraw-Hill, New York, 1961.
 "Handbook of Chemistry and Physics," 37th ed., p. 1972, Chemical Rubber Publishing Co., Cleveland, Ohio, 1955.
 Hildebrand, F. B., "Introduction to Numerical Analysis," McGraw-Hill, New York, 1956.
 Hildebrand, F. B., "Methods of Applied Mathematics," Prentice-Hall, Englewood Cliffs, N. J., 1963.
 Lanczos, C., "Applied Analysis," Prentice-Hall, Englewood Cliffs, N. J., 1956.
 Lapidus, L., "Digital Computation for Chemical Engineers," McGraw-Hill, New York, 1962.
 Orden, A., "Mathematical Methods for Digital Computers," Ralston Wilf, ed., Wiley, New York, 1962.
 Ralston, A., "First Course in Numerical Analysis," McGraw-Hill, New York, 1965.
 Sasuly, Max, "Trend Analysis in Statistics," Brookings Institute, Washington, D. C., 1934.
 Simons, H. P., *Ind. Eng. Chem., Anal. Ed.* **13**, 536 (1941).
 Zakin, J. L., Simha, R., Hershey, H. C., *J. Appl. Polymer Sci.* **10**, 1455 (1966).

RECEIVED for review August 22, 1966

ACCEPTED April 14, 1967

DESIGN AND EFFECTIVENESS OF FEEDFORWARD CONTROL SYSTEMS FOR MULTICOMPONENT DISTILLATION COLUMNS

T. W. CADMAN

Department of Chemical Engineering, University of Maryland, College Park, Md. 20740

R. R. ROTHFUS AND R. I. KERMODE¹

Department of Chemical Engineering, Carnegie Institute of Technology, Pittsburgh, Pa. 15213

The design and simulation of feedforward and feedback control systems for a five-tray, single-feed distillation column have been carried out for the multicomponent system benzene-toluene-xylene. For changes in feed composition, linear feedforward compensation may range from nearly perfect to detrimental, depending on column operating conditions.

WITHIN the past few decades, the increasing interest in higher product quality, safety of operation, and optimum economic performance has been the motivation for a revolution in the theory of process dynamics and control. Unfortunately,

the complexity of chemical and petroleum processing has severely hampered the parallel development of potential applications. Such is particularly the case for one of the most common of the unit operations, multicomponent distillation. In this paper, emphasis is placed on the evaluation of the potentials and application of the linear theory of feedforward and feedback control design to multicomponent distillation.

¹ Present address, University of Kentucky, Lexington, Ky. 40506