An Extension of the Spline Fit Technique and Applications to Thermodynamic Data

R. L. KLAUS and H. C. VAN NESS
Rensselaer Polytechnic Institute, Troy, New York

The spline fit technique is extended to permit the smoothing of experimental data. It is applied to thermodynamic data which are otherwise difficult to treat.

There is always a need for methods to smooth, interpolate, differentiate, or otherwise treat experimental data. Sometimes equations are available from theoretical considerations which contain a certain number of undetermined constants whose best values may be calculated from the experimental data. In other cases empirical equations have come into use where experience has shown them to correlate adequately a certain kind of experimental data. A familiar example is the use of the Antoine equation to correlate vapor pressure data as a function of temperature (*1*). In the absence of either a theoretical or an empirical correlating equation, polynomials have frequently been used. Experience has shown that they often correlate experimental data very well. Moreover, if the polynomial coefficients are viewed as the unknown quantities to be determined from experimental data, the fitting procedure is linear, a distinct computational advantage. There are a number of numerical procedures for this type of calculation. One of the most powerful is that due to Forsythe (*2*). This has been generalized and applied to various types of thermodynamic data (*3*).

To be acceptable, a representation should be reasonably low-ordered, correlate the data to within the experimental error, and introduce no unwarranted inflection points. The latter is particularly important if the data are to be differentiated. Unfortunately there are a number of instances where the search for an adequate polynomial representation of the complete set of data has been unsuccessful. Certain alcohol-hydrocarbon heat of mixing and vapor pressure data as a function of liquid mole fraction are of this type. The difficulty in representing these data seems to be related to the steepness of the required curves in one or more parts of the composition interval as compared with other parts.

THE SPLINE FIT TECHNIQUE

One way to circumvent these problems is to use the spline fit technique which is discussed by Landis and Nilsen (*4*). This method puts a different cubic between every two successive data points such that the curves pass exactly through each data point and that the first two derivatives of the curve on the right-hand side of a data point are equal, respectively, to the first two derivatives of the curve on the left-hand side of the data point, all derivatives evaluated at the data point. Since some of the same equations apply in the extended method to be presented later, the derivations of the equations of the original method are summarized below.

Since the curves in every interval* are cubics, the second derivative in a given interval is linear; that is

$$f''_{k-1}(x) = C_{k-1}\left(\frac{x_k - x}{L_k}\right) + C_k\left(\frac{x - x_{k-1}}{L_k}\right) \qquad (1)$$

The intervals are numbered according to the number of the left-most data point. It is clear that this equation is linear and that the proper values of the second derivatives are given at the interval boundaries.

Equation (1) may be integrated twice and the two integration constants evaluated by setting $f_{k-1}(x_{k-1}) = A_{k-1}$ and $f_{k-1}(x_k) = A_k$, where A_k is the value given by the cubic equation at x_k. This results in the following equation:

$$f_{k-1}(x) = \frac{C_{k-1}}{6L_k}(x_k - x)^3 + \frac{C_k}{6L_k}(x - x_{k-1})^3 + \left[\frac{A_k}{L_k} - \frac{L_k C_k}{6}\right](x - x_{k-1}) + \left[\frac{A_{k-1}}{L_k} - \frac{L_k C_{k-1}}{6}\right](x_k - x) \qquad (2)$$

which may also be written as

$$f_{k-1}(x) = \left(\frac{C_k - C_{k-1}}{6L_k}\right)(x - x_{k-1})^3 + \frac{C_{k-1}}{2}(x - x_{k-1})^2 + \left[\left(\frac{A_k - A_{k-1}}{L_k}\right) - \left(\frac{C_k + 2C_{k-1}}{6}\right)L_k\right](x - x_{k-1}) + A_{k-1} \qquad (3)$$

The first derivative is

$$f'_{k-1}(x) = \frac{-C_{k-1}}{2L_k}(x_k - x)^2 + \frac{C_k}{2L_k}(x - x_{k-1})^2 + \left(\frac{C_{k-1} - C_k}{6}\right)L_k + \left(\frac{A_k - A_{k-1}}{L_k}\right) \qquad (4)$$

Through Equation (1) it is already ensured that the second derivatives will match at the interval boundaries. By choosing

$$A_k = y(x_k) \qquad (5)$$

where $y(x_k)$ is the value of the dependent variable at x_k, one ensures that the curves will go exactly through each data point. By setting

$$f'_{k-1}(x_k) = f'_k(x_k)$$

one ensures that the first derivatives match at each data point. This results in the following set of equations to be solved for $\{C_k\}(k = 1, 2, \ldots N$, where N is the number of data points):

---

*The k^th^ interval in these derivations is taken to be $x_k \le x \le x_{k+1}$, where x_k is the independent variable for the k^th^ data point.

$$\frac{L_k}{6} C_{k-1} + \left[\frac{L_k + L_{k+1}}{3}\right] C_k + \frac{L_{k+1}}{6} C_{k+1} = \left[\frac{A_{k+1} - A_k}{L_{k+1}}\right] - \left[\frac{A_k - A_{k-1}}{L_k}\right] \quad (6)$$

Equations such as this one can be written for $k = 2, 3, \ldots N$-1. Unique determination of the $\{C_k\}$ requires two more relationships. If the slopes of the curve at the end points are known, the two following equations may be written:

$$\begin{aligned} \frac{L_2}{3} C_1 + \frac{L_2}{6} C_2 &= \frac{A_2 - A_1}{L_2} - y'(x_1) \\ \frac{L_N}{6} C_{N-1} + \frac{L_N}{3} C_N &= y'(x_N) - \left(\frac{A_N - A_{N-1}}{L_N}\right) \end{aligned} \quad (7)$$

These slopes are rarely known, however. In the absence of other information one might assume that in the first and last intervals the curves are parabolic rather than cubic, in which case

$$C_1 = C_2 \text{ and } C_N = C_{N-1} \quad (8)$$

The N-2 equations of the form of Equation (6) may be used in conjunction with either Equation (7) or (8) to determine uniquely the set $\{C_k\}$. Since the equations are linear in $\{C_k\}$ and since the resulting array contains only three nonzero diagonals, the calculation is performed quite readily by a digital computer even for a fairly large number of data points.

THE EXTENDED SPLINE FIT METHOD

The method described above is particularly appropriate for data which are quite smooth. The $\{C_k\}$ may be calculated once and for all, and whenever an interpolated value of y or its derivative is needed, Equations (2), (3), and (4) provide the means to calculate them. Since the curves pass exactly through each data point, the spline fit accomplishes no smoothing of the data, and represents scatter and random error at least as faithfully as it does the large-scale trend of the data. In fact, where marked scatter exists, the spline fit representation may actually exaggerate it.

If, instead of defining the interval boundaries to pass through every data point, interval boundaries are determined arbitrarily such that each interval may contain a number of data points, the method can be extended to provide smoothing of the data. This is the basis of the extension to be discussed here.

Suppose constraints are placed on the cubics such that at every interval boundary the adjacent cubics give the same values for the dependent variable and its first two derivatives. Then Equations (1) through (4) again apply, with $\{x_k\}$ now being the values of the independent variables on each of the interval boundaries. The $\{A_k\}$ are now also unknowns. Suppose in addition that one requires that the sum of the squares of the deviations between the experimental points and the analytical representation be minimized; that is

$$\sigma^2 = \sum_{k=1}^{K-1} \sum_{i=1}^{n_k} \varepsilon_{ki} [y(x_{ki}) - f_k(x_{ki})]^2 = \text{minimum} \quad (9)$$

where

$\varepsilon_{ki} = 1$ if x_{ki} is on any interval boundary except the first or last; = 2 otherwise.

The introduction of ε_{ki} ensures that no data point will be weighted more than any other. Since points on interval boundaries other than the first and last are counted in two intervals, the squares of the deviations for these points are only weighted half as much as the others. Equations of the form of Equation (6) must also apply; so the problem resolves itself to that of the determination of $\{C_k\}$ and $\{A_k\}$ such that σ^2 is minimized under the constraints imposed by (K-2) equations of the form of Equation (6).

This problem may be solved by the well-known method of Lagrange's undetermined multipliers (5). If a function F of M variables, designated $v_1, v_2, \ldots v_M$ is to be minimized under J constraints of the form

$$\phi_j = 0 \quad (j = 1, 2, \ldots J) \quad (10)$$

where ϕ_j is a function of the M variables, the problem becomes one of solving the J equations of the form of Equation (10) along with M equations of the form

$$\frac{\partial F}{\partial v_m} + \sum_{j=1}^{J} \lambda_j \frac{\partial \phi_j}{\partial v_m} = 0 \quad (m = 1, 2, \ldots M) \quad (11)$$

where the $\{\lambda_j\}$ are the Lagrange multipliers. Since there are J Lagrange multipliers and M variables, and since there are J equations of the form of Equation (10) and M of the form of Equation (11), all the unknowns may be uniquely determined.

In this problem there are $2K$ unknowns (the $\{A_k\}$ and the $\{C_k\}$ for all k) and K-2 constraints of the form of Equation (6). There will thus be K-2 Lagrange multipliers. The function F is σ^2, and its derivatives are given by

$$\frac{\partial(\sigma^2)}{\partial A_j} = -2 \sum_{k=1}^{K-1} \sum_{i=1}^{n_k} \varepsilon_{ki} [y(x_{ki}) - f_k(x_{ki})] \frac{\partial f_k}{\partial A_j}(x_{ki})$$

$$\frac{\partial(\sigma^2)}{\partial C_j} = -2 \sum_{k=1}^{K-1} \sum_{i=1}^{n_k} \varepsilon_{ki} [y(x_{ki}) - f_k(x_{ki})] \frac{\partial f_k}{\partial C_j}(x_{ki})$$

Because of Equation (3), when $j = 1$ all partial derivatives on the right of the above two equations are zero except for $k = 1$, and for $j = K$ the partials vanish except for $k = K$-1. For all other values of j all partial derivatives vanish except for $k = j - 1$ and $k = j$. Thus these two equations become

$$\frac{\partial(\sigma^2)}{\partial A_j} = -2 \sum_{i=1}^{n(j-1)} \varepsilon_{(j-1)i} [y(x_{(j-1)i}) - f_{(j-1)}(x_{(j-1)i})] \times \left[\frac{x_{(j-1)i} - x_{j-1}}{L_j}\right] - 2 \sum_{i=1}^{n_j} \varepsilon_{ji} [y(x_{ji}) - f_j(x_{ji})] \times \left[1 - \left(\frac{x_{ji} - x_j}{L_{j+1}}\right)\right] \quad (12)$$

$$\frac{\partial(\sigma^2)}{\partial C_j} = -2 \sum_{i=1}^{n(j-1)} \varepsilon_{(j-1)i} [y(x_{(j-1)i}) - f_{(j-1)}(x_{(j-1)i})] \times \left[\frac{(x_{(j-1)i} - x_{j-1})^3}{6L_j} - \frac{L_j (x_{(j-1)i} - x_{(j-1)})}{6}\right] - 2 \sum_{i=1}^{n_j} \varepsilon_{ji} [y(x_{ji}) - f_j(x_{ji})] \cdot \left[\frac{-L_{j+1}(x_{ji} - x_j)}{3} + \frac{(x_{ji} - x_j)^2}{2} - \frac{(x_{ji} - x_j)^3}{6L_{j+1}}\right] \quad (13)$$

where it is understood that for $j = 1$ the first summations on the right of Equations (12) and (13) are not included and for $j = K$ the second summations on the right of those two equations are not included. From Equation (6) the derivatives of ϕ_j for use in Equation (11) are found to be

$$\frac{\partial \phi_{m-1}}{\partial A_m} = \frac{-1}{L_m} \quad (14a)$$

$$\frac{\partial \phi_m}{\partial A_m} = \frac{1}{L_{m+1}} + \frac{1}{L_m} \quad (14b)$$

$$\frac{\partial \phi_{m+1}}{\partial A_m} = \frac{-1}{L_{m+1}} \quad (14c)$$

TABLE 1. ETHANOL-n-HEPTANE. VAPOR PRESSURES (mm. Hg) AT 30°C.

x_1	Experimental vapor pressure	Fitted vapor pressure	Deviation, %	First derivative of vapor pressure	Second derivative of vapor pressure
0.0*	58.7	58.70	0.	4660	−381,000
0.0156	94.2	95.91	1.81	808	−113,000
0.0187	98.6	97.96	−0.65	542	−59,400
0.0216*	101.0	99.35	−1.63	442	−9,500
0.03884	105.4	105.66	0.25	296	−7,400
0.0518	108.4	108.93	0.48	211	−5,800
0.0953*	114.6	114.34	−0.23	76	−400
0.1464*	117.6	117.66	0.05	52	−550
0.3095*	121.24	121.24	0.	6	−20
0.5260*	122.10	122.10	0.	2	−11
0.6542*	121.90	121.92	0.01	−8	−160
0.7203	121.07	121.04	−0.03	−18	−119
0.7662*	120.10	120.11	−0.01	−23	−90
0.8072	118.94	119.00	0.04	−34	−512
0.8710*	115.40	115.28	−0.10	−88	−117
0.9265	107.41	107.62	0.20	−204	−3,010
0.9303	106.80	106.83	0.03	−216	−3,140
0.9630*	98.2	97.90	−0.31	−336	−4,220
0.9860	88.0	88.14	0.16	−551	−14,500
1.0000*	78.8	78.80	0.	−797	−20,700

*Indicates an interval boundary.
x_1 = mole fraction of ethanol.

$$\frac{\partial \phi_j}{A_m} = 0 \quad \text{for } j > m+1 \text{ and } j < m-1 \tag{14d}$$

$$\frac{\partial \phi_{m-1}}{\partial C_m} = \frac{L_m}{6} \tag{15a}$$

$$\frac{\partial \phi_m}{\partial C_m} = \frac{L_m + L_{m+1}}{3} \tag{15b}$$

$$\frac{\partial \phi_{m+1}}{\partial C_m} = \frac{L_{m+1}}{6} \tag{15c}$$

$$\frac{\partial \phi_j}{\partial C_m} = 0 \quad \text{for } j > m+1 \text{ and } j < m-1 \tag{15d}$$

This reduces to three the number of terms in the summation of Equation (11) for all j except $j = 1$ and $j = K$-1, in which cases the number of terms in the summation is two.

The following definitions are made:

$$s_{j+1}^{(k)} = \sum_{i=1}^{n_j} \varepsilon_{ji} (x_{ji} - x_j)^k \tag{16}$$

$$t_{j+1}^{(k)} = \sum_{i=1}^{n_j} \varepsilon_{ji}\, y(x_{ji})\, (x_{ji} - x_j)^k \tag{17}$$

If Equation (3) is combined with Equations (12) to (17), one eventually obtains

$$\frac{2A_{j-1}}{L_j}\left[s_j^{(1)} - \frac{s_j^{(2)}}{L_j}\right] + 2C_{j-1}\left[\frac{-s_j^{(2)}}{3} + \frac{s_j^{(3)}}{2L_j} - \frac{s_j^{(4)}}{6L_j^2}\right] - \frac{\lambda_{j-1}}{L_j} + 2A_j\left[\frac{s_j^{(2)}}{L_j^2} + s_{j+1}^{(0)} - \frac{2s_{j+1}^{(1)}}{L_{j+1}} + \frac{s_{j+1}^{(2)}}{L_{j+1}^2}\right] + \frac{C_j}{3}\left[\frac{s_j^{(4)}}{L_j^2} - s_j^{(2)} - 2L_{j+1}\, s_{j+1}^{(1)} + 5s_{j+1}^{(2)} - \frac{4s_{j+1}^{(3)}}{L_{j+1}} + \frac{s_{j+1}^{(4)}}{L_{j+1}^2}\right] + \lambda_j\left[\frac{1}{L_j} + \frac{1}{L_{j+1}}\right] + 2A_{j+1}\left[\frac{s_{j+1}^{(1)}}{L_{j+1}} - \frac{s_{j+1}^{(2)}}{L_{j+1}^2}\right] + \frac{C_{j+1}}{3}\left[-L_{j+1}\, s_{j+1}^{(1)} + s_{j+1}^{(2)} + \frac{s_{j+1}^{(3)}}{L_{j+1}} - \frac{s_{j+1}^{(4)}}{L_{j+1}^2}\right] - \frac{\lambda_{j+1}}{L_{j+1}} = 2\left[\frac{t_j^{(1)}}{L_j} + t_{j+1}^{(0)} - \frac{t_{j+1}^{(1)}}{L_{j+1}}\right] \tag{18}$$

$$A_{j-1}\left[\frac{-s_j^{(4)}}{L_j^2} + \frac{s_j^{(3)}}{L_j} + s_j^{(2)} - s_j^{(1)} L_j\right] + C_{j-1}\left[\frac{-s_j^{(6)}}{6L_j^2} + \frac{s_j^{(5)}}{2L_j} - \frac{s_j^{(4)}}{6} - \frac{s_j^{(3)} L_j}{2} + \frac{s_j^{(2)} L_j^2}{3}\right] + \frac{\lambda_{j-1} L_j}{2} + A_j\left[\frac{s_j^{(4)}}{L_j^2} - s_j^{(2)} + \frac{s_{j+1}^{(4)}}{L_{j+1}^2} - \frac{4s_{j+1}^{(3)}}{L_{j+1}} + 5s_{j+1}^{(2)} - 2s_{j+1}^{(1)} L_{j+1}\right] + C_j\left[\frac{s_j^{(6)}}{6L_j^2} - \frac{s_j^{(4)}}{3} + \frac{s_j^{(2)} L_j^2}{6} + \frac{s_{j+1}^{(6)}}{6L_{j+1}^2} - \frac{s_{j+1}^{(5)}}{L_{j+1}} + \frac{13}{6} s_{j+1}^{(4)} - 2s_{j+1}^{(3)} L_{j+1} + \frac{2}{3} s_{j+1}^{(2)} L_{j+1}^2\right] + \lambda_j [L_j + L_{j+1}] + A_{j+1}\left[\frac{-s_{j+1}^{(4)}}{L_{j+1}^2} + \frac{3s_{j+4}^{(3)}}{L_{j+1}} - 2s_{j+1}^{(2)}\right] + C_{j+1}\left[\frac{-s_{j+1}^{(6)}}{6L_{j+1}^2} + \frac{s_{j+1}^{(5)}}{2L_{j+1}} - \frac{s_{j+1}^{(4)}}{6} - \frac{s_{j+1}^{(3)} L_{j+1}}{2} + \frac{s_{j+1}^{(2)} L_{j+1}^2}{3}\right] + \frac{\lambda_{j+1} L_{j+1}}{2}$$

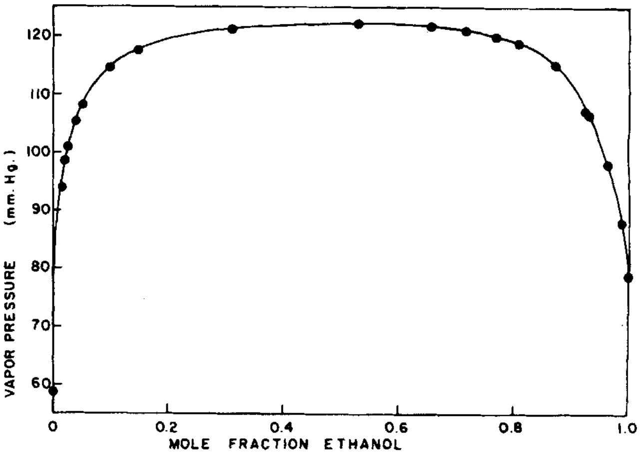

Fig. 1. Ethanol–n-heptane. Vapor pressure at 30°C. Data of Van Ness et al. (6).

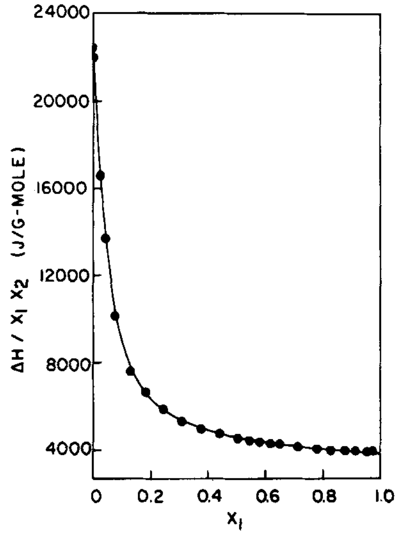

Fig. 2. Isopropanol–*n*-heptane. $\Delta H/x_1x_2$ at 45°C. Data of Van Ness et al. (*10*). x_1 = mole fraction of isopropanol.

$$= \frac{t_j^{(3)}}{L_j} - t_j^{(1)} L_j - \frac{t_{j+1}^{(3)}}{L_{j+1}} + 3t_{j+1}^{(2)} - 2t_{j+1}^{(1)} L_{j+1} \qquad (19)$$

Such equations may be written for $j = 1, 2, \ldots K$ with the understanding that for $j = 1$ the first three terms on the left of these two equations do not appear and that $L_1 = 0$; for $j = K$, the last three terms on the left do not appear and $L_{K+1} = 0$. One may write K equations each of the form of Equations (18) and (19) and K-2 constraining equations of the form of Equation (6). Since there are $3K$-2 unknowns (K values of $\{A_k\}$; K values of $\{C_k\}$; and K-2 values of the Lagrange multipliers $\{\lambda_k\}$), unique values for all the variables may be determined.

If all the data points lie on interval boundaries, Equations (18) and (19) reduce to

$$\frac{-\lambda_{j-1}}{L_j} + 4A_j + \lambda_j\left[\frac{1}{L_j} + \frac{1}{L_{j+1}}\right] - \frac{\lambda_{j+1}}{L_{j+1}} = 4y(x_j)$$

$$\frac{\lambda_{j-1}L_j}{2} + \lambda_j[L_j + L_{j+1}] + \frac{\lambda_{j+1}L_{j+1}}{2} = 0$$

except for $j = 1$ and $j = K$, in which cases the first and last terms, respectively, on the left of these equations do not appear. The solution for this special case is $\lambda_j = 0$ and $A_j = y(x_j)$ for all j. This is exactly equivalent to the basic spline fit method.

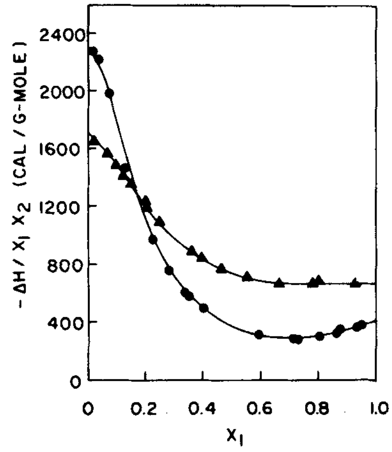

Fig. 3. Alcohol-water. $\Delta H/x_1x_2$ at 25°C. Data of Lama and Lu (*11*). x_1 = mole fraction of alcohol. Circle = ethanol-water. Triangle = methanol-water.

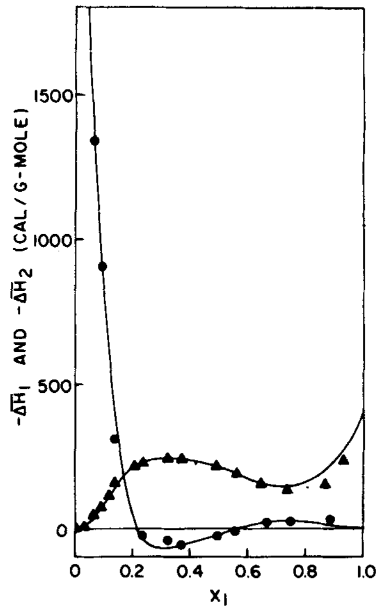

Fig. 4. Ethanol-water. $\overline{\Delta H}_i$ at 25°C. Data of Bertrand et al. (*13*). x_1 = mole fraction of ethanol. Circle = $\overline{\Delta H}_1$. Triangle = $\overline{\Delta H}_2$.

It is sometimes desirable to anchor the end points of the curve at predetermined values. This may be accomplished by adding $A_1 = y(x_1)$ and $A_K = y(x_K)$ to the constraints and allowing for the two additional Lagrange multipliers. A computer program has been developed to perform all these calculations.*

APPLICATIONS TO THERMODYNAMIC DATA

The extended spline fit method is now applied to certain thermodynamic data for solutions which in most cases are not well represented over the whole composition range by a single polynomial. It is clear from the derivations that interval boundaries need not be defined to pass through data points, but in the calculations presented here it was found that this is the best choice.

The first application is to the vapor pressure data for ethanol-*n*-heptane at 30°C. as reported by Van Ness et al. (6) and illustrated in Figure 1. These data are required for the calculation of the composition of the vapor phase. Several methods are available to perform this calculation (7 to 9) and each requires many interpolated values. It is clear that the results of this calculation depend on the accuracy of these values.

The points chosen as interval boundaries are indicated in Table 1. This table contains the results of the calculations, namely, smoothed vapor pressures and values of the first two derivatives. The curve was anchored at the ends at values equal to the vapor pressures of the pure components. It is evident that the representation is smooth and passes very close to the data points. The curve in Figure 1 is drawn from the smoothed values.

The inclusion in Table 1 of values for the second derivative is not meant to imply that the spline fit technique will yield smooth and accurate results for this quantity. A criterion of suitability of fit is that the second derivative exhibit reasonable continuity. If this is the case, then the analytical representation and its first derivative will be smooth.

The extended spline fit method also lends itself to the

---

*For further information on this FORTRAN IV computer program write to Professor H. C. Van Ness, Chemical Engineering Department, Rensselaer Polytechnic Institute, Troy, New York 12181.

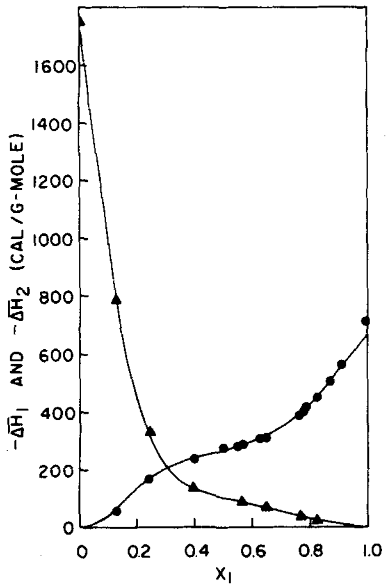

Fig. 5. Methanol-water. $\overline{\Delta H}_i$ at 25°C. Data of Bertrand et al. (13). x_1 = mole fraction of methanol. Triangle = $\overline{\Delta H}_1$. Circle = $\overline{\Delta H}_2$.

representation of thermodynamic excess functions for binary systems. Although considerable success has been achieved in representing these functions by sets of orthogonal polynomials (*3*), there are some sets of data for which such methods are either completely unsuccessful or at best marginal. Heat of mixing data for certain alcohol-hydrocarbon systems serve as good examples.

ΔH data for the system isopropanol-*n*-heptane at 45°C. have been reported by Van Ness et al. (*10*) and are illustrated in Figure 2. The spline fit method of Landis and Nilsen applied to the $\Delta H/x_1x_2$ values gave a number of unwarranted inflection points. Several points were then chosen as interval boundaries and the calculations were made with the extended spline fit method. Table 2* contains the resulting smoothed values of $\Delta H/x_1x_2$ and its first two derivatives. The curve in Figure 2 is drawn from the smoothed values, and is seen to correlate the data points very well.

Two final examples involve the heat of mixing data of Lama and Lu (*11*) for ethanol-water and methanol-water at 25°C. Results are given in Tables 3 and 4 and in Figure 3. For the interval boundaries chosen, excellent representations of the data are obtained. Because of some scatter in the data for ethanol-water near the mole fraction extremes, a smoothing curve was extrapolated to reasonable intercepts, and these values were used to anchor the end points of the spline fit representation. Partial molar heats of mixing were calculated from the spline fit representations of $\Delta H/x_1x_2$ (*12*) and are listed in Table 5. The curves of Figures 4 and 5 are drawn from these values. Direct measurements of these partial molar properties were reported by Bertrand et al. (*13*) and their data are also shown in Figures 4 and 5. Considering the sensitivity of the calculations, the unusual shape of the partial property curves, and the possible errors of measurement, one sees that the agreement is remarkable. Although these same data have also been successfully represented by single polynomials (*3*), this in no way detracts from the value of these illustrations in demonstrating the power of the extended spline fit method in the representation of data.

The selection of interval boundaries in any particular application of this method is necessarily arbitrary, and varying results are obtained depending upon the choice. One wants the best possible representation of the data that does not introduce unwarranted inflection points. Beyond mere trial and error, experience appears to be the best guide.

ACKNOWLEDGMENT

This work was supported by National Science Foundation Grant GK 1038.

NOTATION

A_k = value of the analytical representation at the k^{th} interval boundary
C_k = value of the second derivative at the k^{th} interval boundary
F = any analytical function
$f_k(x)$ = analytical representation in the k^{th} interval
ΔH = enthalpy change of mixing
i, k = summation indices
J = number of constraints
K = number of interval boundaries
$L_k = x_k - x_{k-1}$
M = number of variables
N = number of data points; number of interval boundaries
n_k = number of data points in the k^{th} interval
$s_{j+1}^{(k)}$ = quantity as defined by Equation (16)
$t_{j+1}^{(k)}$ = quantity as defined by Equation (17)
v_i = i^{th} variable
x = independent variable
x_1, x_2 = mole fractions of components 1 and 2, respectively
x_k = k^{th} data point; k^{th} interval boundary
x_{ki} = i^{th} data point in the k^{th} interval
$y(x_k)$ = dependent variable for the k^{th} data point
= differentiation with respect to the displayed variable
{ } = entire set of constants

Greek Letters

ε_{ki} = factor as defined by Equation (9)
λ_j = j^{th} Lagrange multiplier
ϕ_j = j^{th} constraint
σ^2 = sum of the squares of deviations between experimental points and the analytical representation

LITERATURE CITED

1. Willingham, C. B., et al., *J. Res. Natl. Bur. Stds.*, 35, 219 (1945).
2. Forsythe, G. E., *Soc. Ind. Appl. Math. J.*, 5, 209 (1959).
3. Klaus, R. L., and H. C. Van Ness, *Chem. Eng. Progr. Symp. Ser.*, in press.
4. Landis, F., and E. N. Nilson, "Progress in International Research on Thermodynamic and Transport Properties," p. 218, Academic Press, New York (1962).
5. Sokolnikoff, I. S., and R. M. Redheffer, "Mathematics of Physics and Modern Engineering," p. 249ff., McGraw-Hill, New York (1958).
6. Van Ness, H. C., C. A. Soczek, and N. K. Kochar, *J. Chem. Eng. Data*, 12, 346 (1967).
7. Van Ness, H. C., "Classical Thermodynamics of NonElectrolyte Solutions," p. 136ff., Pergamon, New York (1964).
8. Tao, L. C., *Ind. Eng. Chem.*, 53, 307 (1961).
9. Mixon, F. O., B. Gumowski, and B. H. Carpenter, *Ind. Eng. Chem. Fundamentals*, 4, 455 (1965).
10. Van Ness, H. C., C. A. Soczek, G. L. Peloquin, and R. L. Machado, *J. Chem. Eng. Data*, 12, 217 (1967).
11. Lama, R. F., and B. C. Y. Lu, *ibid.*, 10, 216 (1965).
12. Van Ness, H. C., "Classical Thermodynamics of Non-Electrolyte Solutions," p. 101, Pergamon, New York (1964).
13. Bertrand, G. L., F. J. Millero, et al., *J. Phys. Chem.*, 70, 699 (1966).

Manuscript received January 24, 1967; revision received March 20, 1967; paper accepted March 23, 1967.

---

*Tables 2, 3, 4, and 5 are deposited as Document No. 9684 with the American Documentation Institute, Photo-duplication Service, Library of Congress, Washington, 25, D. C., and may be obtained for $1.25 for photoprints or 35-mm. microfilm.