Vapor-Liquid Equilibrium:

Part I. An Appraisal of Data Reduction Methods

Correlations of the excess Gibbs function and the activity coefficients of a binary, liquid-phase system at constant temperature depend on reduction of *P-x-y* data taken for the system in vapor-liquid equilibrium. It is shown that use of a full *P-x-y* data set is for all practical purposes based on just the *P-y* values. Alternative procedures require only *P-x* or *x-y* data. In any case, effective use of a full *P-x-y* data set can be made only if redundent experimental values of *x*,*y*, or *P* are compared for consistency with calculated values. This cannot be accomplished by means of the popular area test for thermodynamic consistency.

HENDRICK C. VAN NESS
STANLEY M. BYER and
RICHARD E. GIBBS

Division of Fluid, Chemical,
and Thermal Processes
Rensselaer Polytechnic Institute
Troy, New York 12181

SCOPE

The feasibility of vapor-liquid equilibrium (VLE) determination from measurements at constant temperature of just the solution vapor pressure as a function of liquid composition has already been shown (Ljunglin and Van Ness, 1962; Van Ness et al., 1967), and an apparatus has been developed (Gibbs and Van Ness, 1972) for the rapid accumulation of such data. The question arises as to whether this method, which does not require measurement of vapor compositions, is in some way deficient in comparison with methods based on measurement of all the variables *P*, *x*, and *y*. In fact, measurements of any pair of these variables provide sufficient data for the calculation of the third variable through exact relations provided by thermodynamics. When all three variables are measured the data set contains redundant information, and thermodynamics provides a means to test their internal consistency. Correlation of VLE data is invariably accomplished indirectly through thermodynamic functions for the liquid phase, for only in such a form do the data take on any theoretical significance. Thus the procedures used for reduction of VLE data are crucial, and it is important that credibility not be sacrificed. We have therefore critically examined all of the factors that play a part in the various reduction methods for constant-temperature VLE data. Illustrative results are based on a set of published VLE data.

CONCLUSIONS AND SIGNIFICANCE

There are three distinctly different procedures which may be used to reduce a set of *P-x-y* data for a binary system at constant temperature to a correlation of the liquid-phase activity coefficients. Each makes use of data for just two of the variables—*x-y*, *P-x*, or *P-y*. Although the usual method of data reduction is based on all the data, *P*, *x*, and *y*, for practical purposes it makes effective use of only the *P-y* data. Since experimental uncertainty is likely greatest for *y*, prudence suggests that the preferred procedure for data reduction is the one based on just the *P-x* data. Values of *y* calculated from the correlation can then be compared with measured values as a check on the thermodynamic consistency of the data. The consistency test is the only possible advantage to be gained from redundant measurements of *y*, for it is not possible to use them so as to improve the reliability of the correlation. Unless a consistency test is considered essential, experimental effort is better spent on improvement of the accuracy of *P-x* measurements than on measurement of redundant data. The attention given to consistency tests in the past has been largely misdirected, for the commonly applied area test is so limited as to have little value.

Any set of VLE data should include as an integral part of the set the vapor pressures of the pure components. The use of values for these vapor pressures that are discordant with the solution vapor pressures can cause severe distortion in the final correlation of results. It is likely that many correlations in the literature suffer from this fault and that undue complication of correlating equations is a consequence.

Vapor-liquid equilibrium data have most commonly been measured with a recirculating still. The data of Fried et al. (1967) for the system pyridine-tetrachloroethylene at 60°C are typical, and they are reproduced in Figure 1. We will examine this set of data in detail, employing various methods to find correlations for the excess Gibbs function and the activity coefficients of the liquid phase. The pertinent equations, which are restricted to binary systems at low pressure and constant temperature, are (Van Ness, 1964)

R. E. Gibbs is with the New York State Department of Environmental Conservation, Albany, N.Y.

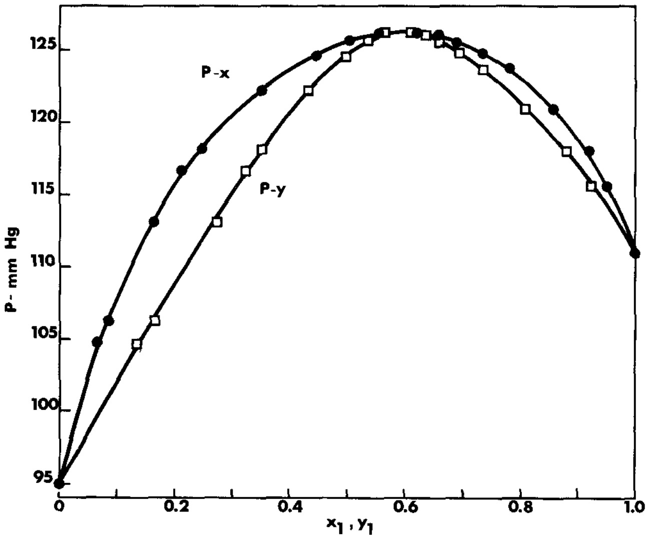

Fig. 1. P-x-y data for the system pyridine(1)—tetrachloroethylene(2) at 60°C (Fried et al., 1967).

$$\gamma_i = \frac{y_i P}{x_i P_i^{\text{sat}}} \tag{1}$$

$$\frac{g}{x_1 x_2} = \frac{\ln \gamma_1}{x_2} + \frac{\ln \gamma_2}{x_1} \tag{2}$$

$$\ln \gamma_1 = g + x_2 \frac{dg}{dx_1} \tag{3a}$$

$$\ln \gamma_2 = g - x_1 \frac{dg}{dx_1} \tag{3b}$$

$$\ln \frac{\gamma_1}{\gamma_2} = \ln \frac{y_1 x_2 P_2^{\text{sat}}}{y_2 x_1 P_1^{\text{sat}}} = \frac{dg}{dx_1} \tag{4}$$

where the identity $g \equiv G^E/RT$ has been employed. We have assumed that the vapor phase is an ideal gas and that liquid properties are independent of pressure. These assumptions are not essential but are appropriate here, simplifying the discussion and influencing the results to a negligible extent.

For the vast majority of systems the empirical functional relation between g and x is found to be quite simple. For the system considered here, it can be faithfully represented by the three-constant equation

$$\frac{g}{x_1 x_2} = A x_2 + B x_1 - D x_1 x_2 \tag{5}$$

When this equation is substituted into Equations (3) and (4), they become

$$\ln \gamma_1 = x_2^2 [A + 2(B - A - D) x_1 + 3 D x_1^2] \tag{6a}$$

$$\ln \gamma_2 = x_1^2 [B + 2(A - B - D) x_2 + 3 D x_2^2] \tag{6b}$$

$$\ln \frac{\gamma_1}{\gamma_2} = A x_2^2 - B x_1^2 + 2 x_1 x_2 [B - A + D(x_1 - x_2)] \tag{7}$$

These equations are often called the four-suffix Margules equations.

PROCEDURES FOR DATA REDUCTION

The most direct procedure for correlation of the set of VLE data shown by Figure 1 is to calculate γ_1 and γ_2 by Equation (1) for each data point and then to apply Equation (2) to generate a set of values for $g/x_1 x_2$. All of the data—P, x, and y—are used in this procedure. The values of $g/x_1 x_2$ are then fit by a least-squares technique to a suitable empirical expression, in this case, Equation (5).

Fried et al. (1967) did not follow this procedure. Rather, they calculated values of $\ln(\gamma_1/\gamma_2)$ from their data and fit these by Equation (7). The interesting feature of this procedure is that it made no use whatever of their carefully measured values of P. It is clear from the first equality of Equation (4) that the pressure cancels from the ratio γ_1/γ_2, and that it might just as well never have been measured.

It is well known that measurement of P, x, and y at constant T represents an over-determination of the thermodynamic properties of a binary system in vapor-liquid equilibrium. It is, in fact, possible to base all calculations not only on x-y data alone, but equally well on P-x data alone or on P-y data alone. With a perfect set of P-x-y data, one should be able to generate the same correlation through use of all the data or from any pair of the variables. With a set of real data, however, there are likely to be discrepancies.

The calculational methods which require only P-x or P-y data are based on Equations (8) and (9).

$$\frac{dy_1}{dP} = \frac{y_1 y_2}{P(y_1 - x_1)} \tag{8a}$$

This is a form of the coexistence equation (Van Ness, 1964), and it may be solved for x_1:

$$x_1 = y_1 \left(1 - \frac{y_2}{P} \frac{dP}{dy_1}\right) \tag{8b}$$

An additional relation follows directly from Equation (1):

$$P = x_1 P_1^{\text{sat}} \gamma_1 + x_2 P_2^{\text{sat}} \gamma_2 \tag{9a}$$

When combined with Equation (3), this becomes

$$P = x_1 P_1^{\text{sat}} \exp\left(g + x_2 \frac{dg}{dx_1}\right) + x_2 P_2^{\text{sat}} \exp\left(g - x_1 \frac{dg}{dx_1}\right) \tag{9b}$$

Equations (8a) and (9b) provide two distinctly different calculational procedures that require just P-x data. Equation (8a) may be integrated by a marching procedure that starts at an appropriate point where limiting values are given by the data (Van Ness, 1970). This yields values for y, which in combination with the measured values of P and x, allow calculation of the activity coefficients and of $g/x_1 x_2$ or $\ln(\gamma_1/\gamma_2)$.

With Equation (9b) one uses some technique to search for the g versus x relation which allows this equation to reproduce the measured P versus x data as closely as possible. The most flexible technique is numerical, and the one proposed by Mixon et al. (1965) has proved to be entirely satisfactory. Either method requires an initial fit of the P versus x_1 data, and the least-squares spline fit of Klaus and Van Ness (1967) is especially useful. For the same input, the two methods produce identical results. For effective application both require use of a digital computer.

Finally, Equation (8b) allows the use of just P-y data. One first obtains a least-squares spline fit of the P versus y_1 values so as to produce reliable values for the derivative dP/dy. Values for x_1 then result from direct substitution into Equation (8b). These, together with the measured P-y data, allow calculation of the activity coefficients and of g/x_1x_2 or $\ln(\gamma_1/\gamma_2)$.

APPLICATION TO EXPERIMENTAL DATA

All four calculational procedures have been applied to the data of Figure 1, and the results are summarized in Table 1, where the constants A, B, and D of Equations (5), (6), and (7) as determined by the four different treatments of the data are given. Figure 2 displays the four corresponding curves of g/x_1x_2 versus x_1. The different methods of treating the data clearly produce discordant results and indicate that the requirements of thermodynamic consistency are not satisfied. We may go one step further to show how the measured values of the properties P and y differ from their values as recalculated from the various correlations, as represented by Equation (5). The correlated values for γ_1 and γ_2 are then given by Equation (6), and once these are calculated for a particular value of x_1, they are substituted into Equation (9a) to give P. Equation (1) then gives y values. The values of P and y, so calculated at each experimental value of x_1 are compared with experimental values by means of the differences ΔP and Δy_1. The results appear in Figure 3. The root-mean square averages and the maximum absolute values of ΔP and Δy_1 are listed in Table 2.

Additional tests for thermodynamic consistency are now possible by comparison of calculated values of P or of y with experimental values that were not used in any way to generate the correlation. Where x-y data alone were used, the plot of ΔP versus x_1 provides a consistency check. Where P-x data alone were used the plot of Δy versus x_1 provides the test. In either case consistency requires that the Δ values scatter randomly about zero. Where there is bias from the $\Delta = 0$ axis, the data must be judged inconsistent. In both cases bias is evident. This could, of course, result also from use of a correlating equation of a form not capable of adequately representing the data. However, as discussed later, the evidence is that the equation used is suitable.

TABLE 1. CONSTANTS FOR EQUATIONS (5), (6), AND (7) AS DETERMINED BY DIFFERENT PROCEDURES

	From P-x-y Data[a]	From x-y Data[b]	From P-x Data[c]	From P-y Data[d]
A	0.93432	0.77882	0.82030	0.95665
B	0.84874	0.68925	0.77826	0.87698
D	0.48897	0.03721	0.09045	0.53546

[a] By Equations (1), (2), and (5).
[b] By Equations (4) and (7).
[c] By Equations (9b), (4), and (7).
[d] By Equations (8b), (4), and (7).

TABLE 2. DEVIATION BETWEEN COMPUTED AND MEASURED VALUES

Δ = (Value Predicted by Correlation) − (Experimental Value)

	Δy		ΔP, mm Hg	
Data used	RMS	Max	RMS	Max
P-x-y	0.0058	0.0119	0.43	0.99
x-y	0.0018	0.0036	1.35	1.87
P-x	0.0054	0.0092	0.25	0.50
P-y	0.0067	0.0136	0.47	1.10

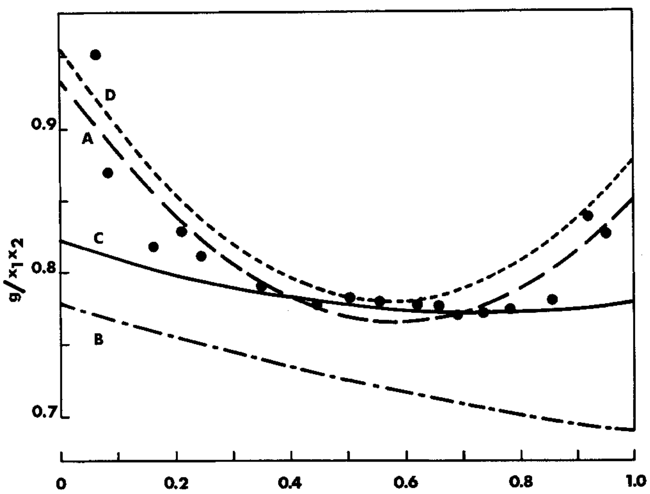

Fig. 2. g/x_1x_2 versus x_1 for the system pyridine(1)—tetrachloroethylene(2) at 60°C. Curve A is determined from P-x-y data; Curve B, from x-y data; Curve C, from P-x data; and Curve D, from P-y data. The filled circles are calculated from experimental data by Equation (2).

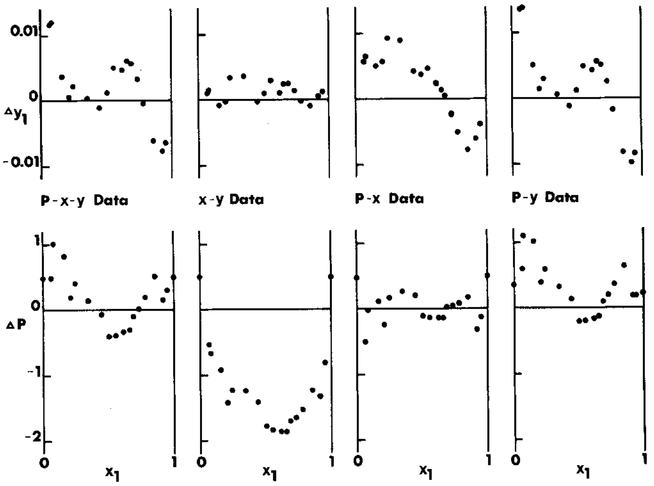

Fig. 3. Δy_1 and ΔP versus x_1 for the system pyridine(1)—tetrachloroethylene(2) at 60°C, where Δ denotes the difference between the value determined from the correlation of results and the experimental value.

PURE-COMPONENT VAPOR PRESSURES

It is important to note the influence of the values used for the vapor pressures of the pure constituents P_1^{sat} and P_2^{sat} in the calculation of activity coefficients. At constant T, these quantities are constants, and the values chosen influence all values of all thermodynamic functions calculated from the data. Moreover, since they are constants, use of inappropriate values can only introduce nonrandom bias into the calculated thermodynamic properties, and the effect can be severe.

The published set of VLE data considered here does not include measured values for P_1^{sat} and P_2^{sat} at 60°C. Rather, the authors refer to other work for these values and rely on the Antoine equation as an interpolation formula.* This has been fairly common practice, and it is a serious mistake. Vapor pressures are highly sensitive to

---

* Table 3 of the reference cited, which gives the Antoine constants, contains a typographical error. The column headings A and B should be interchanged.

experimental conditions and to the purity of the materials used. In any proper experimental determination of VLE it is essential that the vapor pressures of the pure constituents be measured with the same apparatus and for the same lots of materials as are used for the other measurements so that they are an integral part of the data set. Otherwise they may be, and almost always are, discordant in some degree with the solution vapor pressure measurements. This is not readily evident from a plot such as Figure 1 but shows up clearly when a spline fit is made of the P-x_1 data. Thus in Figure 3 the ΔP plot for the case where only P-x data are used reflects primarily the deviations of the experimental pressures from the spline-fit representation. The values at the ends stand out as not fitting in with the other measurements. The vapor pressures from which these two points were determined come from the Antoine equations referenced by the original authors, and they were treated as part of the P-x_1 data set. The spline fit yields vapor pressures P_1^{sat} and P_2^{sat} that are both about 0.5 mm Hg higher than those from the Antoine equations.

The effects on g/x_1x_2 of changes in the values of the vapor pressures used for its calculation from experimental data are given by the derivatives:

$$\frac{\partial(g/x_1x_2)}{\partial P_1^{sat}} = \frac{-1}{x_2P_1^{sat}} \quad \text{and} \quad \frac{\partial(g/x_1x_2)}{\partial P_2^{sat}} = \frac{-1}{x_1P_2^{sat}}$$

One of these derivatives approaches negative infinity at each end of the composition range, and when either x_1 or x_2 is small the value of g/x_1x_2 is very sensitive to the value used for one or the other of the vapor pressures. Since this sensitivity is much less in the midcomposition range, the result of use of inappropriate values for P_1^{sat} and P_2^{sat} is a marked distortion of the g/x_1x_2 versus x_1 curve. It is therefore best to use values of P_1^{sat} and P_2^{sat} as determined from a least-squares spline fit of pressure-composition data rather than directly measured or literature values. This is the choice we have made, and it ensures that the pure-component vapor pressures are at least in reasonable accord with the rest of the data.

The function $\ln(\gamma_1/\gamma_2)$, on the other hand, is affected only by the ratio of the vapor pressures:

$$\frac{\partial \ln(\gamma_1/\gamma_2)}{\partial(P_2^{sat}/P_1^{sat})} = \frac{P_1^{sat}}{P_2^{sat}}$$

Since this effect is independent of x, the experimental points on a plot of $\ln(\gamma_1/\gamma_2)$ are merely shifted up or down a constant amount by any change in the ratio of vapor pressures. However, there is no way to adjust the constants on the right-hand side of Equation (7) so as to produce a constant shift in the representation of $\ln(\gamma_1/\gamma_2)$, and changes in the vapor-pressure ratio therefore cause a distortion in the fitted function. Nevertheless, it has been our experience that less distortion results from fitting $\ln(\gamma_1/\gamma_2)$ than from fitting g/x_1x_2, and this has been our choice where the option is available.

RANDOM ERRORS OF MEASUREMENT

We must also consider the effect of random error in the measured variables on the error in the calculated thermodynamic functions. With temperature held constant, the variables subject to such measurement error are P, x, and y. Temperature is also subject to random fluctuation, but this shows up as error in the other variables. Although the vapor pressures P_1^{sat} and P_2^{sat} are functions of temperature, fixed values are used with the entire data set, and random error is not a consideration. Estimates of the standard deviations in the functions g/x_1x_2 and $\ln(\gamma_1/\gamma_2)$ when these quantities are calculated for a finite data set are given by equations derived in the manner of Ulrichson and Stevenson (1972).

For g/x_1x_2:

$$s^2 = \left(\frac{\ln\gamma_1}{x_2^2} - \frac{\ln\gamma_2}{x_1^2}\right)^2 s_x^2 + \left(\frac{x_1 - y_1}{x_1x_2y_1y_2}\right)^2 s_y^2 + \frac{s_P^2}{(Px_1x_2)^2} \tag{10}$$

For $\ln(\gamma_1/\gamma_2)$:

$$s^2 = \frac{s_x^2}{(x_1x_2)^2} + \frac{s_y^2}{(y_1y_2)^2} \tag{11}$$

In Equation (10) the coefficient of the s_x term is relatively very small, and it remains finite as x_1 and x_2 approach zero. Thus the error in g/x_1x_2 is insensitive to errors in x, and reflects primarily errors in y and P. The coefficients of s_y and s_P become infinite at the composition extremes.

Equation (11) indicates that the function $\ln(\gamma_1/\gamma_2)$ is equally sensitive to measurement errors in x and y, and in both cases the error contribution approaches infinity as x_1 or x_2 approaches zero. A term for error in pressure does not appear, because values of $\ln(\gamma_1/\gamma_2)$ do not depend on pressure measurements.

With respect to the data set under consideration, and in view of the results in Table 2, we assign the values

$$s_x = s_y = 0.005 \text{ and } s_P = 0.5 \text{ mm. Hg.}$$

Calculations based on Equation (10) then provide estimates of the standard deviation in g/x_1x_2, and these results are reflected by the dashed lines of Figure 4, which show $g/x_1x_2 \pm s$ and provide the 68% confidence interval around the curve generated through use of just the P-x_1 data. This is curve C of Figure 2. The points represent the direct experimental results and are the same points shown in Figure 2. This curve was chosen in view of Table 2 and Figure 3, where it is seen that this fit produces the best compromise for goodness-of-fit with respect to both y and P. In addition, this curve represents an excellent correlation of the 10 data points lying within the narrowest region of the 68% confidence interval. The remaining 7 points lies close to the upper bound of the

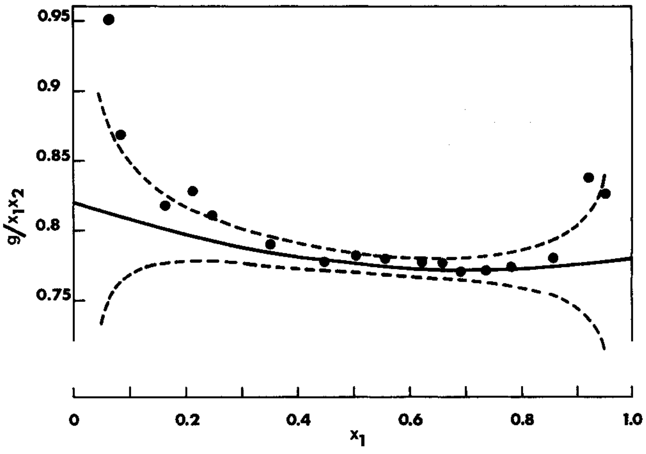

Fig. 4. g/x_1x_2 versus x_1 for the system pyridine(1)—tetrachloroethylene(2) at 60°C. The filled circles and Curve A of Figure 2 are reproduced here; in addition, the 68% confidence interval is indicated by the dashed lines.

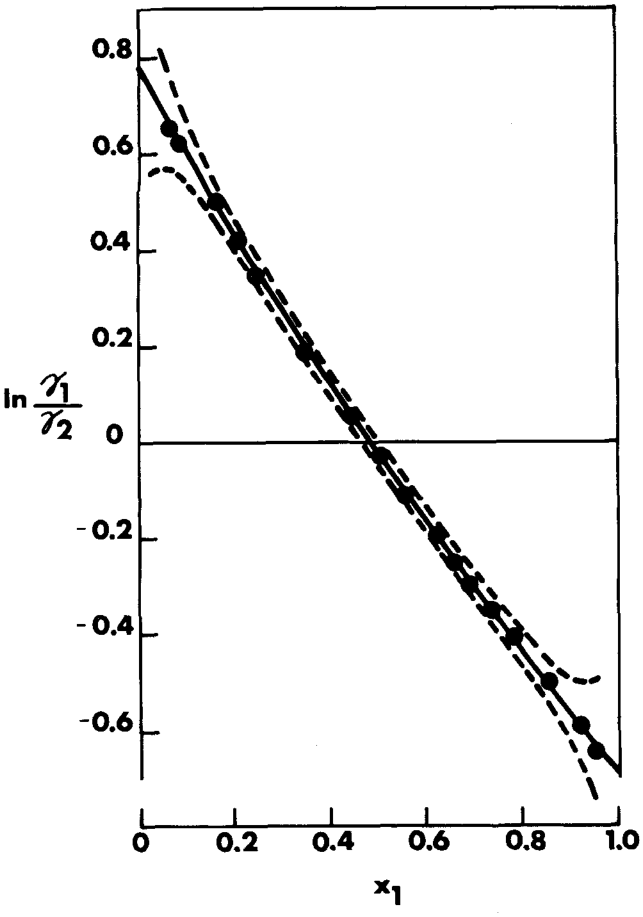

Fig. 5. $\ln(\gamma_1/\gamma_2)$ versus x_1 for the system pyridine(1)—tetrachloroethylene(2) at 60°C. The filled circles represent data points calculated from the experimental data. The solid curve is the least-squares fit of the data to Equation (7). The dashed lines indicate the 68% confidence interval.

confidence interval, and their deviations appear more systematic than random.

Figure 5 shows a plot of $\ln(\gamma_1/\gamma_2)$ versus x_1. The points are calculated from the experimental data, and the solid line represents the correlation based on just the x-y data, corresponding to curve B of Figure 2. The dashed lines indicate the 68% confidence interval as determined by Equation (11). Note that there is no tendency for the points here to deviate widely from the correlating curve, even where the confidence interval spreads.

ANALYSIS OF DATA-REDUCTION METHODS

With respect to Figure 3, it is evident that the plots based on the full set of P-x-y data bear a remarkable resemblance to those based on just the P-y data. This is not accidental and is a consequence of the fact that the quantity g/x_1x_2 upon which use of the P-x-y data depends is very insensitive to errors in x. This means that it is not possible to make effective use of the measurements of x because errors in x have little influence on the resulting correlation. The reliability of the correlation rests on the measurements of P and y. The measurement of x is superfluous and may as well be omitted. The procedure based on P-y data alone makes no use of the x values, and the similarity in results obtained by the two procedures as shown by Figures 2 and 3 naturally follows.

The correlation based on just P-y data does not admit to use of ΔP and Δy for a consistency check. However, we do have a set of calculated x_1 values from Equation (8b) which can be compared with the unused experimental values of x_1. The resulting plot of Δx_1 versus x_1 is shown by Figure 6. There is a clear bias of these results from the $\Delta x = 0$ axis, and again the data must be judged inconsistent in some degree. The root-mean square value of Δx is 0.0057.

It is perhaps surprising that the points on the ΔP plot (Figure 3) in this case do not scatter randomly about zero, since the initial step is a least-squares spline fit of the experimental pressures. The reason is that the spline fit of P is with respect to the measured values of y_1, while the fit of $\ln(\gamma_1/\gamma_2)$ to provide the final correlation is necessarily with respect to the calculated values of x_1, because this function is a liquid-phase property. Thus bias appears in both ΔP and Δy. The same effect is produced when calculations are based on the full set of P-x-y data, because only P-y data are effectively used, and the final fit is to x_1. This split in the bias of results between ΔP and Δy reflects a slight mismatch between the form of the correlating function used and the form required for a perfect correlation of the measured P-y data. It is not possible to determine whether the deficiency lies with the form of the correlation or with the data. At best one can point out that curve A of Figure 2 represents the points based on experimental data as well as can be expected in view of the confidence interval associated with it. Similarly, Figure 5 suggests no serious deficiency in the form of the correlating function. It should be noted that the correlating function is used here merely as a convenience, not as an essential part of our consideration. All final correlations could be strictly graphical or numerical.

We conclude from this analysis of data-reduction methods that in reality only three truly different methods

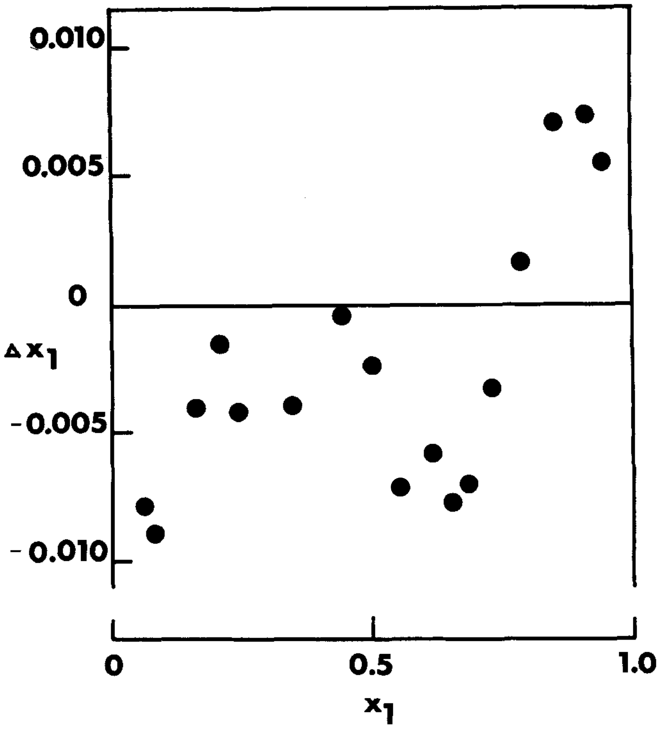

Fig. 6. Δx_1 versus x_1 for the system pyridine(1)—tetrachloroethylene(2) at 60°C as determined from P-y data, where Δ denotes the difference between the value calculated from the correlation of results and the experimental value.

exist for the correlation of a full set of P-x-y data, and each makes use of just two of the measured variables. Thus we have the options to base the correlation on x-y data, on P-x data, or on P-y data. We do not really have the option to make effective use of the full P-x-y data set, except through use of redundant measurements for testing of thermodynamic consistency.

CONSISTENCY TESTS

Thermodynamic consistency tests for VLE measurements are all based on the Gibbs-Duhem equation, and much has been written in the literature about various applications of this equation through the activity coefficients. We have in passing made reference to six different sensitive and comprehensive consistency tests that are obvious when one looks at the treatment of P-x-y data from the point of view adopted here but which have not previously been proposed. Comparison of the g/x_1x_2 versus x_1 curves of Figure 2 resulting from any two of the three basic methods of data treatment (curves A and D being considered equivalent) provide three such tests. Similarly, the comparison of calculated and experimental (but unused) data as shown by Figures 3 and 6—ΔP for x-y data, Δy for P-x data, and Δx for P-y data—represent three additional tests. Unfortunately, consistency tests, either alone or in concert, cannot indicate how or where the data are in error. In fact, a restricted consistency test may well fail to indicate error even when the data are inconsistent.

The widely used area test, which requires the area under a plot of $\ln(\gamma_1/\gamma_2)$ versus x_1 to be zero is a consistency test only in the most restricted sense, because the only data needed to construct the plot are x and y values and the ratio $P_2^{\text{sat}}/P_1^{\text{sat}}$. The pressures, which represent the redundant data, are not involved, except for the limiting values P_1^{sat} and P_2^{sat}. Thus the only testing done by the area test as applied to isothermal data is of whether the vapor pressure ratio $P_2^{\text{sat}}/P_1^{\text{sat}}$ is appropriate to the set of x-y data. With respect to the present data set, the area test is very nearly satisfied, as can be seen from Figure 5, where the solid line meets the requirement exactly. However, the comparison of experimental and calculated pressures based on the same x-y data and shown in Figure 3 indicate a clear departure of the data from consistency. A plot of $\ln(\gamma_1/\gamma_2)$ versus x_1 will be sensitive to scatter in the x-y data, but it will show nothing about the internal consistency of such data.

RECOMMENDATIONS FOR DATA REDUCTION

If a full set of P-x-y data is determined experimentally, one must choose one of the three basic methods for data reduction. The measurements of y are usually the least reliable, and it is therefore reasonable to base all calculations on just the P-x data. The calculated values of y are then available for comparison with the measured values for a comprehensive and directly meaningful consistency check. Where VLE data are represented by just the P-x trace, as measured in a static equilibrium cell, there is no alternative to use of the method based on just the P-x data, and no consistency test is possible because there is no redundancy in the data.

In either case one needs a standard procedure for data reduction. The steps recommended are as follows:

1. Generate a least-squares spline fit (Klaus and Van Ness, 1967) of the P versus x_1 data that exhibits smooth first derivatives and reasonably smooth second derivatives of P with respect to x_1. This may require several trials that depend on the number of data intervals chosen and on the location of the interval boundaries. The vapor pressures P_1^{sat} and P_2^{sat} to be used throughout the data reduction process are those values given at $x_1 = 0$ and $x_1 = 1$ by the spline fit.
2. Process the smoothed data by some method which is based on Equation (9b). The procedure described by Prausnitz et al. (1967) is certainly satisfactory and produces directly the constants in the correlating equation chosen. The disadvantage is that the form of the equation must be assumed from the start. We have found the more flexible method of Mixon, Gumowski, and Carpenter (1965) to be very efficient. It produces numerical values of the various functions, γ_1, γ_2, g/x_1x_2, and $\ln(\gamma_1/\gamma_2)$, and these may then be fit by the most appropriate correlating equation. The coexistence equation, Equation (8a) can certainly be integrated, but it is often a much more troublesome process than those based on Equation (9b).
3. If numerical values of the thermodynamic functions have been generated, we recommend that $\ln(\gamma_1/\gamma_2)$ be chosen for fitting, as our experience is that this procedure minimizes the distortion caused by nonrandom errors in the data set.

MISCELLANEOUS ITEMS

Two special topics deserve brief mention. There is a restriction on the terminal slopes which may be assumed by a P-x_1 relationship. This is illustrated for the state at $x_1 = y_1 = 0$, where the following relationships hold (Van Ness, 1964):

$$\lim_{x_1\to 0}\frac{g}{x_1x_2} = \lim_{x_1\to 0}\ln\frac{\gamma_1}{\gamma_2} = \lim_{x_1\to 0}\ln\gamma_1 = \ln\frac{P_2^{\text{sat}}}{P_1^{\text{sat}}}\left(\frac{dy_1}{dx_1}\right)_0$$
$$= \ln\left[\frac{P_2^{\text{sat}} + (dP/dx_1)_0}{P_1^{\text{sat}}}\right] \qquad (12)$$

Since the quantity in brackets in the last member of this set cannot be negative, its minimum values is zero. The restriction on $(dP/dx_1)_0$ is therefore

$$P_2^{\text{sat}} + (dP/dx_1)_0 \geqq 0$$

or

$$(dP/dx_1)_0 \geqq -P_2^{\text{sat}}$$

Similarly at $x_1 = 1$, we have the restriction

$$(dP/dx_1)_1 \leqq P_1^{\text{sat}}$$

Thus the terminal slopes of the P-x curve are restricted in such a way that this curve must lie above the dashed lines of Figure 7.

The difficulty is that when the data approach one of these restrictive conditions the values of g/x_1x_2 and of $\ln(\gamma_1/\gamma_2)$ in the dilute region become highly sensitive to the values of the slopes dP/dx determined by a fit of the data. If the spline fit of P versus x_1 is not exactly right in this dilute region, then the computed values of g/x_1x_2 and of $\ln(\gamma_1/\gamma_2)$ take a wild turn up or down. Such behavior must be completely discounted as an artifact. One way to circumvent the problem is to make further trials with the spline-fit in an effort to find a representation of the data that does not cause an erratic effect. Slight changes have an enormous influence on the results.

This problem is not unique to the treatment of P-x data; it appears no matter what data are measured or used. Comparison of the last two members of Equation (12)

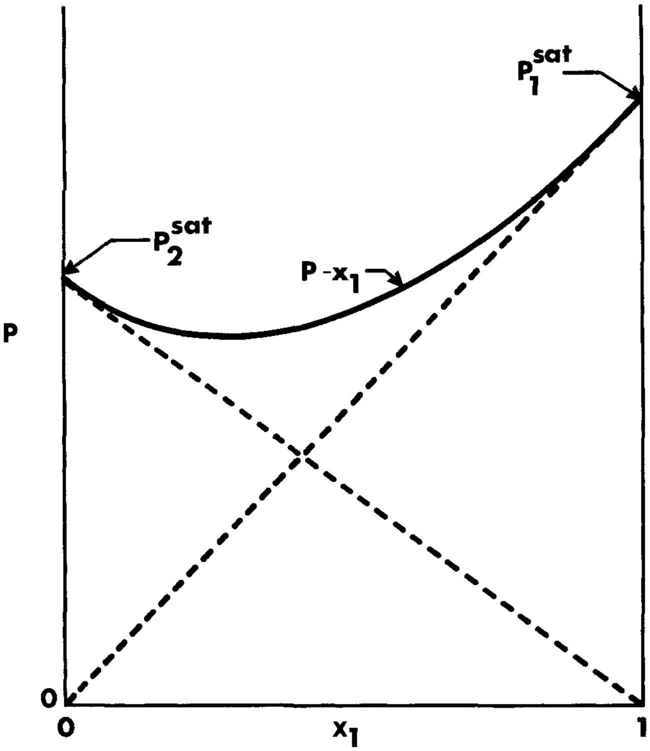

Fig. 7. Schematic representation of P versus x_1 for *VLE*, showing by the dashed lines the limiting values of the terminal slopes that are in theory possible.

shows that when the term in brackets of the last member is close to zero, then $(dy_1/dx_1)_0$ must also be very small. This means that near $x_1 = y_1 = 0$ there can be little change in y_1. Thus y_1 remains small over a considerable range of x_1. Reference to Equations (10) and (11) shows that the coefficients of the s_y terms are large whenever y_1 (or y_2) remains small, and the confidence intervals do not narrow until one is well away from the dilute region.

We have not included in this study any corrections for nonideal behavior of the vapor phase. Such corrections can readily be incorporated if the necessary data, such as virial coefficients, are available. However, if such corrections are applied, they should be complete. The common practice of making partial corrections that amount to the assumption of an ideal solution in the vapor phase is generally a mistake, for it often produces an overcorrection that is less realistic than no correction at all (Van Ness, 1964).

Once a correlation is formulated and published, its proper use becomes a consideration. The correlation is a compact representation of much valuable information in that it can be used to regenerate VLE data. In regenerating these data the user should remember that where the thermodynamic functions are most sensitive to errors in experimental data, the regenerated values are least sensitive to deficiencies in the correlation, that is, near the composition extremes. The user is also free to pick the vapor pressures P_1^{sat} and P_2^{sat} to be used with the correlation in the regeneration of data. He may for good reason wish to use values other than those on which the correlation was based, and he should be aware of the effect of this choice on his results. The following equations give the effect on the recalculated values of P and y of use of vapor pressures differing by δP_1^{sat} and δP_2^{sat} from those originally used:

$$\delta P = \gamma_1 x_1 \delta P_1^{sat} + \gamma_2 x_2 \delta P_2^{sat}$$

or

$$\frac{\delta P}{P} = y_1 \frac{\delta P_1^{sat}}{P_1^{sat}} + y_2 \frac{\delta P_2^{sat}}{P_2^{sat}}$$

and

$$\delta y_1 = y_1 y_2 \left[\frac{\delta P_1^{sat}}{P_1^{sat}} - \frac{\delta P_2^{sat}}{P_2^{sat}} \right]$$

ACKNOWLEDGMENT

Acknowledgment is made to the donors of the Petroleum Research Fund, administered by the American Chemical Society, for support of this research.

NOTATION

A, B, D = empirical constants in Equations (5), (6), and (7)
G^E = excess Gibbs function, liquid phase
g = G^E/RT
P = vapor pressure in general
P^{sat} = vapor pressure of a pure component
R = universal gas constant
s = estimate of standard deviation
s_P, s_x, s_y = estimate of standard deviations for the errors of measurement of P, x, and y
T = absolute temperature
x = mole fraction, liquid phase
y = mole fraction, vapor phase
γ = activity coefficient, liquid phase
Δ = signifies the difference between a calculated and an experimental value
1, 2 = as subscripts, identify the two components of a binary system

LITERATURE CITED

Fried, V., P. Gallant, and G. B. Schneier, "Vapor-Liquid Equilibrium in the System Pyridine-Tetrachloroethylene," *J. Chem. Eng. Data*, **12**, 504 (1967).

Gibbs, R. E., and H. C. Van Ness, "Vapor-Liquid Equilibria from Total-Pressure Measurements. A New Apparatus," *Ind. Eng. Chem. Fundamentals*, **11**, 410 (1972).

Klaus, R. L., and H. C. Van Ness, "An Extension of the Spline Fit Technique and Applications to Thermodynamic Data," *AIChE J.*, **13**, 1132 (1967).

Ljunglin, J. J., and H. C. Van Ness, "Calculation of Vapor-Liquid Equilibria from Vapor Pressure Data," *Chem. Eng. Sci.*, **17**, 531 (1962).

Mixon, F. O., B. Gumowski, and B. H. Carpenter, "Computation of Vapor-Liquid Equilibrium Data from Solution Vapor Pressure Measurements," *Ind. Eng. Chem. Fundamentals*, **4**, 455 (1965).

Prausnitz, J. M., C. A. Eckert, R. V. Orye, and J. P. O'Connell, *Computer Calculations for Multicomponent Vapor-Liquid Equilibria*, p. 76, Prentice-Hall, Englewood Cliffs, N. J. (1967).

Ulrichson, D. L., and F. D. Stevenson, "Effects of Experimental Errors on Thermodynamic Consistency and on Representation of Vapor-Liquid Equilibrium Data," *Ind. Eng. Chem. Fundamentals*, **11**, 287 (1972).

Van Ness, H. C., *Classical Thermodynamics of Non-Electrolyte Solutions*, pp. 122-147, Pergamon, Oxford (1964).

———., "On Integration of the Coexistence Equation for Binary Vapor-Liquid Equilibrium," *AIChE J.*, **16**, 18 (1970).

———., C. A. Soczek, and N. K. Kochar, "Thermodynamic Excess Properties for Ethanol-*n*-Heptane," *J. Chem. Eng. Data*, **12**, 346 (1967).

Manuscript received August 11, 1972; revision received September 22, 1972; paper accepted September 22, 1972.