The Continuous-Lumping Method for Vapor-Liquid Equilibrium Calculations

The continuous-lumping method is a refinement of continuous thermodynamics which does not depend upon the continuous-mixture hypothesis. The method utilizes a generalized Fourier series expansion of a functional formulation of the equilibrium relations obtained by embedding the classical discrete equilibrium equations into generalized function space. The resulting method provides an effective means of reducing the number of parameters needed to describe a many-component, complex mixture; furthermore, the method provides a straightforward means of converting the equilibrium equations to the new parameterized form. The error associated with the reduction in number of parameters can be easily estimated using the ratio convergence test for mathematical series. The theory of the method is first presented and the method is then used in an example calculation to illustrate the advantages of the method over pseudocomponent lumping and continuous thermodynamics. The method is useful particularly for mixtures described by chromatographic data.

Patrick K. Moore
Rayford G. Anthony
Kinetics, Catalysis and Reaction Engineering Laboratory
Department of Chemical Engineering
Texas A&M University
College Station, TX 77843

Introduction

Many engineering problems in industry require vapor-liquid equilibrium (VLE) calculations for complex mixtures, such as petroleum, coal-liquids, and polymer mixtures. The compositions of such mixtures are frequently described by approximate analytical techniques such as chromatography. For example, Figures 1 and 2 depict a size-exclusion chromatogram (SEC) of a coal liquid and a gas chromatogram of a diesel fuel sample, respectively. Before VLE calculations can be performed, the data from such analyses must be parameterized into a numerical form suitable for use in the VLE equations.

The traditional method for performing this parameterization process is pseudocomponent lumping. In pseudocomponent lumping, the chromatogram is divided into sections with all of the material in each section being lumped into a single component class called a pseudocomponent. In the subsequent VLE calculations, the pseudocomponents are treated as if they were real, physical components. The trouble with the pseudocomponent method is that there is no obvious optimum way to define the pseudocomponent lumps. Defining just a few broad pseudocomponents provides a simple description of the mixture at the risk of introducing errors into the VLE calculations. On the other hand, defining a large number of narrow pseudocomponents results in an accurate mixture description but also increases computation time. This increase in computation time may become excessive when the VLE calculations are used in computer simulations of petroleum reservoirs and petroleum refinery distillation columns.

The problems with the pseudocomponent method have led researchers to develop the method of continuous thermodynamics (Cotterman et al., 1985). In the continuous thermodynamic theory, the discrete mixture, described in the traditional manner using mole fractions $\{x_i\}_{i=1,c}$, is approximated by a hypothetical continuum of components, described using a component distribution function $x(u)$. If the component index u is identified with the retention time, the function $x(u)$ corresponds directly to the chromatogram of the mixture (provided response factor and normalization effects are taken into account).

In analogy to the discrete equilibrium relationship between the mole fractions of the vapor and liquid phases,

$$y_i = K_i x_i, \qquad (i = 1, 2, \ldots, c), \tag{1}$$

continuous thermodynamics assumes the following functional relationship:

$$y(u) = k(u)x(u). \tag{2}$$

Correspondence concerning this paper should be addressed to R. G. Anthony.
Current address of P. K. Moore: Dow Chemical, USA, Freeport, TX 77541.

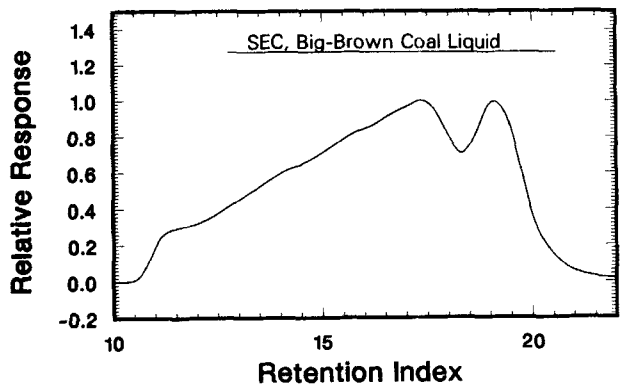

Figure 1. Size-exclusion chromatogram of big-brown coal liquid.

Similar functional equations can be written for the component material balances and normalization conditions involved in a phase-equilibrium calculation. The model equations are then solved, usually using quadrature methods. A detailed description of the continuous thermodynamic theory and corresponding calculational techniques is provided by Cotterman et al. (1985).

The continuous thermodynamic method is readily applicable to smooth chromatographic distributions such as the size-exclusion chromatogram shown in Figure 1. The distribution of material is a gradual, continuous function of the component index (retention time) and no individual components are identifiable. However, the applicability of the continuous thermodynamic method to the diesel-fuel chromatogram of Figure 2 is questionable. First, the chromatogram is detailed enough to show the discrete nature of the mixture; the continuous-mixture hypothesis may not be satisfied. Second, the chromatogram is highly discontinuous and is not suitable for the currently available calculational techniques which rely on numerical quadrature formulas. However, the diesel-fuel chromatogram still shows a great deal of order, as indicated by the gradual change in the height of the major chromatogram peaks. One would hope an approximate calculational technique could be developed to perform VLE calculations for mixtures described by such discontinuous chromatograms. In the following sections we describe the extension of continuous-thermodynamic theory to semicontinuous mixtures of this type.

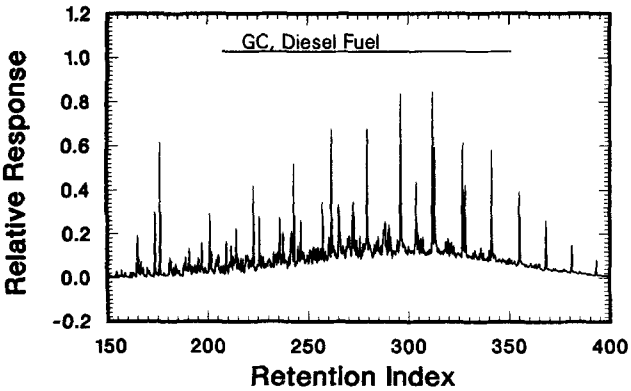

Figure 2. Capillary-column gas chromatogram of diesel fuel.

The Continuous-Lumping Method

The continuous thermodynamic method is based on the assumption that the original discrete mixture can be approximated as a continuum of components, and the theory is subsequently developed for such "continuous mixtures." In the development of the continuous-lumping method, we have required that the final method be applicable to the original discrete mixture; the continuous-mixture hypothesis is not made. The continuous-thermodynamic equations can be shown to be consistent with the original mixture, provided the molar distribution function is defined using the generalized function $\delta(u)$ known as the Dirac delta function:

$$x(u) = \sum_{j=1}^{c} x_j \delta(u - a_j), \tag{3}$$

where x_j is the mole fraction of the jth component and a_j is the jth component index. Substituting equations of the form given in Eq. 3 for $x(u)$ and $y(u)$ into Eq. 2 and integrating over a sufficiently small domain about the component index a_i result in the discrete mixture equation given by Eq. 1. [Note that $K_i = k(a_i)$.] Thus, the continuous thermodynamic formulation of the equilibrium relation (Eq. 2) is equivalent to the discrete equilibrium relation provided that the mole-fraction distribution functions $x(u)$ and $y(u)$ are defined as in Eq. 3.

The second step in the development of the continuous-lumping method is to note that Eq. 2 can be written as follows,

$$y(w) = \int_{\Omega} K(w, u)x(u)du, \tag{4}$$

where

$$K(w, u) = k(u)\delta(w - u). \tag{5}$$

Equation 4 can be shown to be equivalent to Eq. 2 by substituting in Eq. 5 and noting the nature of the dirac delta function. The form for the equilibrium relation given by Eq. 4 was originally obtained by formally embedding the discrete equilibrium equations (Eq. 1) into generalized function space (Moore, 1988).

The chief advantage of the form given in Eq. 4 is that it allows the separate approximation of the functions $K(w, u)$ and $x(u)$ using generalized Fourier series:

$$x(u) = \sum_{i=1}^{\infty} \bar{x}_i \phi^i(u), \tag{6}$$

where

$$\bar{x}_i = \int_{\Omega} x(u)\phi_i(u)du; \tag{7}$$

and for $K(w, u)$,

$$K(w, u) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \bar{K}_i^j \phi^i(w)\phi_j(u), \tag{8}$$

where

$$\bar{K}_i^j = \int_{\Omega} \int_{\Omega} \phi_i(w)K(w, u)\phi^j(u)\, dw\, du$$
$$= \int_{\Omega} \phi_i(u)k(u)\phi^j(u)\, du. \tag{9}$$

The generalized Fourier expansions in Eqs. 6 through 9 have utilized a linearly independent set of basis functions $\{\phi^i(u)\}_{i-1,\infty}$ in addition to the dual set $\{\phi_i(u)\}_{i-1,\infty}$ defined as follows:

$$\int_\Omega \phi_i(u)\phi^j(u)du = \delta_{ij}. \tag{10}$$

The dual-function concept is an extension of the orthogonal function technique. Expressing the Fourier expansions in terms of dual functions allows the representation of nonorthogonal expansions such as piecewise-linear approximations to the functions. Substituting Eqs. 8 and 6, and a similar expression for $y(u)$ into Eq. 4 and simplifying using standard techniques results in the following relationship between the vapor-phase coefficients $\{\bar{y}_i\}_{i-1,\infty}$ and the liquid-phase coefficients $\{\bar{x}_i\}_{i-1,\infty}$:

$$\bar{y}_i = \sum_{j-1}^{\infty} \bar{K}_i^j \bar{x}_j, \ (i = 1, 2, \ldots, \infty). \tag{11}$$

The final step in the development of the continuous-lumping method is the truncation of the infinite series in Eqs. 6 and 8 to n terms. Thus, the continuous-lumping method consists of a two-part approximation: (1) the approximation of the exact discrete distribution $x(u)$ given by Eq. 3 with a truncated series expansion; and (2) the approximation of the exact series form of the equilibrium relationship (Eq. 11) by neglecting terms higher than n. The effects of each of these approximations are illustrated in the following sections.

Approximation of the Component Distribution Function

In this section, the effect of a generalized Fourier serie expansion upon a discrete component distribution of the forn given by Eq. 3 is illustrated by expanding a hypothetical 25 component n-alkane distribution function in terms of Legendr polynomials (Figures 3a–3d). The Legendre expansion approxi mations $x_n(u)$ to the original discrete distribution $x(u)$ ar defined as follows:

$$x_n(u) = \sum_{i-1}^{n} \bar{x}_i \phi^i(u). \tag{12}$$

The basis functions $\{\phi^i(u)\}_{i-1,n}$ are taken to be the orthonormal form of the Legendre polynomials. An orthonormal function equals its dual function: $\phi^i(u) = \phi_i(u)$. Hence, the Legendre coefficients are given by Eq. 7:

$$\bar{x}_i = \int_\Omega x(u)\phi_i(u)\,du = \int_\Omega x(u)\phi^i(u)\,du. \tag{13}$$

Substituting in Eq. 3 for $x(u)$ and integrating results in

$$\bar{x}_i = \sum_{j-1}^{c} x_j \phi^i(a_j). \tag{14}$$

Each Legendre coefficient $\bar{x}_i$ can be calculated in a straightfor-ward manner provided the mole fractions x_i and component

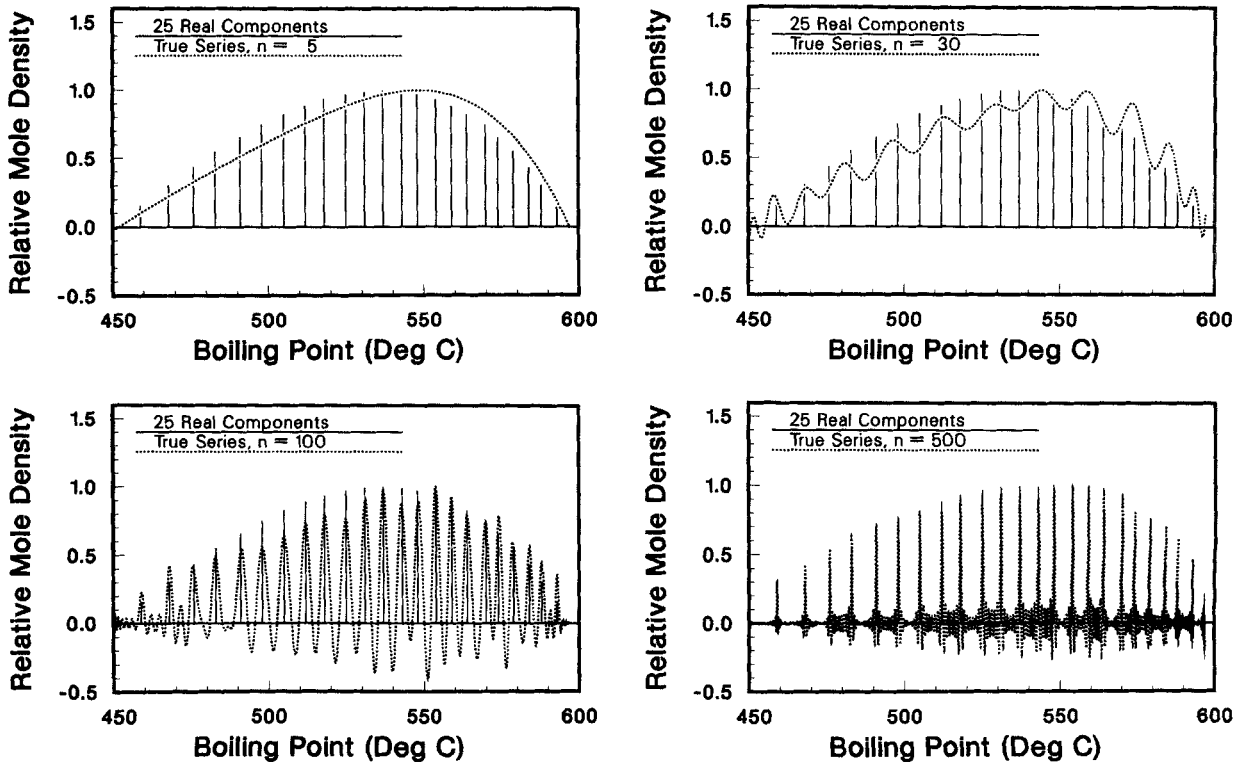

Figure 3 a. Series approximation of 25-component mixture, n = 5. b. Series approximation of 25-component mixture, n = 30. c. Series approximation of 25-component mixture, n = 100. d. Series approximation of 25-component mixture, n = 500.

index a_i are specified. For this example, we have chosen the component indices $\{a_i\}_{i=1,c}$ to represent the normal boiling point of the n-alkanes of carbon number 30 through 54 (Table 1). The normal boiling points were then converted to scaled values by mapping the boiling point domain (449.6°C, 597.0°C) which contained all 25 n-alkanes, to the standard Legendre polynomial domain, $(-1, 1)$. The use of this transformation eliminated any magnitude effects on the Legendre coefficients arising from the choice of units on the component index axis.

The effect of approximating a discrete distribution $x(u)$ with the truncated series $x_n(u)$ can be seen in Figures 3a–3d. The figures show both the discrete distribution $x(u)$ (solid line) and several Legendre polynomial expansions—$x_n(u)$, $n = 5, 30, 100, 500$—for a hypothetical 25 component n-alkane mixture (Table 1). The mole fractions for this mixture were generated using a quadratic expression in carbon number and are plotted as a discrete delta-function component distribution (scaled to maximum height 1.0) with the normal boiling point taken as the component index. The discrete distribution $x(u)$ was then expanded in series of 5, 30, 100, and 500 Legendre polynomials. The expansion coefficients are shown in Table 2.

Figure 3a shows the five-term Legendre expansion (scaled to unit height) of the discrete distribution $x(u)$. As can be seen in the figure, the five-term, "low-order," expansion effectively describes the distribution of material along the component index (boiling point) axis: that is, the amount of material with boiling point within a particular range is well represented by the integral of $x_5(u)$ over the range, provided that the range is not too narrow. The expansion $x_5(u)$ extracts this low-order information from the original distribution $x(u)$ but effectively ignores high-order information, such as the exact number of components, their exact boiling points, and their exact mole-fraction values. This "filtering" property of the series expansion $x_n(u)$ can often be used to advantage in engineering calculations involving complex mixtures, since the approximate distribution of the mixture components is usually of more interest than the exact composition or boiling point of each component.

Table 1. 25-Component n-Alkane Mixture Composition

Comp. No. i	Carbon No.	Mol. Wt.	Boil. Point (°C)	Mole Fraction x_i
1	30	422	449.6	0.00000000
2	31	436	459.0	0.01000000
3	32	450	468.0	0.01913043
4	33	464	476.0	0.02739131
5	34	478	483.0	0.03478261
6	35	492	491.0	0.04130435
7	36	506	498.0	0.04695652
8	37	520	505.0	0.05173913
9	38	534	512.0	0.05565218
10	39	548	518.0	0.05869565
11	40	562	525.0	0.06086957
12	41	576	531.0	0.06217391
13	42	590	537.0	0.06260870
14	43	604	543.0	0.06217391
15	44	618	548.0	0.06086957
16	45	632	554.0	0.05869565
17	46	646	559.0	0.05565218
18	47	660	564.0	0.05173913
19	48	674	570.0	0.04695652
20	49	688	574.0	0.04130435
21	50	702	579.0	0.03478261
22	51	716	584.0	0.02739131
23	52	730	588.0	0.01913043
24	53	744	593.0	0.01000000
25	54	758	597.0	0.00000000

Table 2. 25-Component Mixture Series Expansion Coefficients*

Funct. No. i	Series Coefficient $\bar{x}_i$	Lumping Index $G_x(i)$
1	0.70710680	1.0000
2	0.18408383	0.0635
3	−0.28766169	0.1342
4	−0.11167348	0.0198
5	−0.02853001	0.0013
6	−0.00420224	0.0000
7	−0.00056294	0.0000
8	−0.00446218	0.0000
9	−0.00378955	0.0000
10	0.00019569	0.0000
11	0.00262959	0.0000
12	0.00327808	0.0000
13	−0.00648256	0.0001
14	−0.01049675	0.0002
15	0.00833265	0.0001
16	−0.00539672	0.0000
17	0.00596742	0.0001
18	−0.00125803	0.0000
19	−0.00674853	0.0001
20	−0.00022739	0.0000
21	−0.01866262	0.0006
22	0.00072475	0.0000
23	0.01275526	0.0003
24	0.04206274	0.0028
25	0.02121477	0.0007
26	0.02196908	0.0008
27	0.02239637	0.0008
28	−0.06463284	0.0065
29	−0.01197364	0.0002
30	0.00747628	0.0001

*Component index, boiling point series expansion performed between 449.6°C and 597.0°C; function type used, orthonormal Legendre polynomials.

Close observation of Figure 3a reveals that the maximum of the series expansion $x_5(u)$ lies slightly to the right of the maximum in the original distribution $x(u)$ (labeled as '25 Real Components'). This phenomenon is due to the irregular spacing of the delta functions along the component index axis. As the carbon number increases, the difference in boiling point between successive components gradually becomes smaller; the delta functions are, therefore, spaced closer on the right of the figure. The low-order expansion $x_5(u)$, in effect, "recognizes" this fact, and the maximum is shifted accordingly. Thus, the low-order expansion reflects not only the amount of individual components within an interval, but also reflects the number of components in the interval.

As the number of terms in the series increases, the partial sums $x_n(u)$ converge, in the sense of generalized functions, to the original distribution $x(u)$. The convergence is illustrated in Figures 3b–3d. When 30 terms are included (Figure 3b), the expansion $x_{30}(u)$ begins to "remember" that the original $x(u)$ is composed of Dirac delta functions, and subsequently begins to exhibit some oscillatory character or "wiggle." As n increase to 100, the expansion $x_{100}(u)$ actually begins to approximate the individual Dirac delta functions of which $x(u)$ is composed. The ultimate result is well illustrated in Figure 3d, where 500 terms have been included. Figure 3d indicates that the series $x_{500}(u)$

produces a sharp "peak" for each delta function; the area (not necessarily the height) under each peak is proportional to the amount of that component in the mixture. The broad band of rapidly oscillating curve (or "fuzz") about the baseline clearly testifies to the fact that the partial sums $x_n(u)$ do not converge to $x(u)$ in the classical sense (Kaplan, 1973):

$$\lim_{n\to\infty} x_n(u) \neq x(u). \tag{15}$$

As $n \to \infty$, the fuzz about the baseline does not disappear. However, the partial sums $x_n(u)$ do converge to $x(u)$ in the sense of distributions, or generalized functions (Schwartz, 1966):

$$\lim_{n\to\infty} \int_\Omega x_n(u)\varphi(u)du = \int_\Omega x(u)\varphi(u)du, \tag{16}$$

where $\varphi(u)$ is an element from the set of infinitely-differentiable continuous functions of bounded support: $\varphi(u) \in C_0^\infty$. Thus, if the partial sum $x_n(u)$ is integrated with a smooth function $\varphi(u)$, as in Eq. 16, the positive and negative fluctuations will cancel and the integral of $x_n(u)$ will converge to the integral of $x(u)$.

As was shown above, an expansion $x_n(u)$ can be used to extract the low-order behavior from a discrete distribution $x(u)$; this process will be referred to herein as "smoothing." In order to smooth a discrete distribution $x(u)$, the number of terms, n, to be used in the partial sum must be known. If n is too small, the expansion $x_n(u)$ will fail to extract the entire low-order behavior of $x(u)$; if, on the other hand, too many terms are included, $x_n(u)$ will start to fluctuate due to the discrete influence of the delta functions (Figure 3b). A method is then needed to determine the optimum number of terms for the expansion $x_n(u)$.

For the approximation $x_n(u)$ of the distribution $x(u)$ we may construct a measure of the approximation accuracy using the objective function $J(n)$:

$$J(n) = \|x(u) - x_n(u)\|_2^2 \overset{\text{def}}{=} \int_\Omega [x(u) - x_n(u)]^2\, du. \tag{17}$$

Substituting the expression for $x_n(u)$ given by Eq. 12 into Eq. 17 and simplifying results in the following expression for $J(n)$:

$$J(n) = \int_\Omega [x(u)]^2\, du - \sum_{i=1}^{n}\sum_{j=1}^{n} \bar{x}_i \bar{x}_j\, \bar{g}^{ij}, \tag{18}$$

where the $\{\bar{g}^{ij}\}_{i,j=1,n}$ are the contravariant components of the fundamental, or metric, function $\delta(u - w)$ obtained from the basis $\{\phi^i(u)\}_{i=1,n}$:

$$\begin{aligned}\bar{g}^{ij} &= \int_\Omega \int_\Omega \phi^i(w)\delta(w - u)\phi^j(u)\, du\, dw \\ &= \int_\Omega \phi^i(u)\phi^j(u)\, du.\end{aligned} \tag{19}$$

The first term on the righthand side of Eq. 18 is the square of the Euclidean function norm, $\|x(u)\|_2^2$, of the distribution $x(u)$. The value of $\|x(u)\|_2$ is fixed by $x(u)$ and does not depend on n; in fact, for discrete distributions (consisting of delta functions) the value of $\|x(u)\|_2$ is infinity.

The second term on the righthand side of Eq. 18 is the square of the Euclidean function norm $\|x_n(u)\|_2^2$ of the series approximation $x_n(u)$; it represents the reduction in the objective function $J(n)$ due to the approximation $x_n(u)$. When orthonormal Legendre polynomials are used, $\phi^i(u) = \phi_i(u)$; hence, $\bar{g}^{ij} = \delta_i^j$ and $\|x_n(u)\|_2^2$ becomes

$$\|x_n(u)\|_2^2 \overset{\text{def}}{=} \sum_{i=1}^{n}\sum_{j=1}^{n} \bar{x}_i \bar{x}_j \bar{g}^{ij} = \sum_{i=1}^{n} (\bar{x}_i)^2. \tag{20}$$

Since $\|x(u)\|_2 = \infty$, the effectiveness of $x_n(u)$ in reducing $J(n)$ is difficult to determine. However, the reduction in $J(n)$ due to the addition of more terms can also be evaluated:

$$\begin{aligned}J(n) - J(m) &= \|x_m(u)\|_2^2 - \|x_n(u)\|_2^2 \\ &= \sum_{i=1}^{m}\sum_{j=1}^{m} \bar{x}_i^{(m)}\bar{x}_j^{(m)}\bar{g}_{(m)}^{ij} - \sum_{i=1}^{n}\sum_{j=1}^{n} \bar{x}_i^{(n)}\bar{x}_j^{(n)}\bar{g}_{(n)}^{ij},\end{aligned} \tag{21}$$

where $m > n$. The (n) and (m) symbols have been used as subscripts and superscripts since the series coefficient $\bar{x}_i^{(n)}$ for an n-term expansion does not, in general, remain the same when additional terms are added. The reduction in $J(n)$ indicated in Eq. 21 cannot be compared to $J(n)$, since $J(n) = \infty$; however, $[J(n) - J(n+1)]$ can be compared to the total reduction in $J(n)$ due to all of the $n+1$ terms: the value $\|x_{n+1}(u)\|_2^2$. Thus, we may construct a statistic function $G_x(n)$ as a measure of the approximation effectiveness:

$$G_x(n+1) = \frac{\|x_{n+1}(u)\|_2^2 - \|x_n(u)\|_2^2}{\|x_{n+1}(u)\|_2^2}. \tag{22}$$

For an orthonormal basis set, such as Legender polynomials, the norms are given by Eq. 20. Furthermore, the value of $\bar{x}_i$ remains the same as new terms are added. Equation 22 can then be written as follows:

$$G_x(n+1) = \frac{(\bar{x}_{n+1})^2}{\sum_{i=1}^{n+1} (\bar{x}_i)^2} \quad \text{(Orthonormal Basis)}. \tag{23}$$

The behavior of the $G_x(n)$ statistic is illustrated in Figure 4 and summarized in Table 2. The value of $G_x(n)$ for $n = 1$ is, of course, 1, since the first (and only) term accounts for all of the reduction in $J(n)$ due to $x_1(u)$. As n is increased through 5 to 12, the values of $G_x(n)$ approach 0, indicating that the addition of the nth term causes little improvement or change to the approximation accuracy: a plot of $x_{10}(u)$ essentially coincides with the plot of $x_5(u)$ given in Figure 3a. As n increased past 12, $G_x(n)$ increases from zero to a sequence of small finite values.

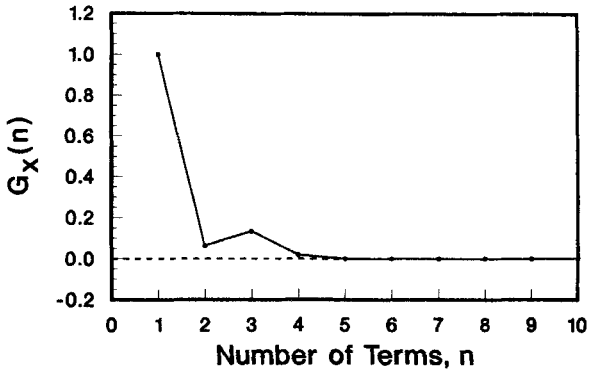

Figure 4. 25-component lumping effectiveness.

Although the incremental effect of these terms is small, the cumulative effect can become significant; these terms cause the fluctuations in the plot of $x_{30}(u)$ in Figure 3b. As n becomes large, the value of $G_x(n)$ can become quite significant; for example, $G_x(73) = 0.1166$.

The behavior of $G_x(n)$ described above suggests a method by which the optimum number of terms can be chosen. First, the basis functions should be arranged in terms of increasing "order" (e.g., $1, x, x^2, x^3, \ldots$). The values of $G_x(1), G_x(2), \ldots$, etc., should then be calculated. When $G_x(n)$ reaches a consistent minimum, preferably zero, all of the low-order character of the original distribution $x(u)$ has been extracted.

The expansion of $x(u)$ in terms of a low-order expansion $x_n(u)$ in effect performs the process of lumping: a many-parameter mixture description, $x(u)$, is approximated by a few-parameter mixture description, $x_n(u)$. The approximate distribution $x_n(u)$ is also easier to interpret, since much of the confusing detail associated with a many-component, complex mixture is filtered out. Furthermore, the low-order coefficients $\{\bar{x}_i\}_{i=1,n}$ are often sufficient to provide accurate approximate solutions to VLE problems, as the following section illustrates.

Approximation of the Equilibrium Equations

The approximation of the series-coefficient form of the equilibrium equations (Eq. 11) will be illustrated for a bubble-point pressure calculation on an ideal mixture. First, however, the problem will be formulated in terms of the traditional discrete equilibrium equations.

For an ideal mixture and a perfect-gas vapor the equilibrium constants $\{K_i\}_{i=1,c}$ are given as follows:

$$K_i = \frac{P_i^\sigma}{P}, \tag{24}$$

where P_i^σ is the vapor-pressure of component A_i at the mixture temperature T, and P is the mixture pressure. In a bubble-point pressure calculation, T and $\{x_i\}_{i=1,c}$ are known and P and $\{y_i\}_{i=1,c}$ are to be calculated. The $\{y_i\}_{i=1,c}$ are given as follows:

$$y_i = K_i x_i = \frac{P_i^\sigma x_i}{P}. \tag{25}$$

The sum of the $\{y_i\}_{i=1,c}$ must equal 1; using this fact and then solving for P results in the familiar expression for Raoult's Law:

$$P = \sum_{i=1}^{c} x_i P_i^\sigma. \tag{26}$$

Equations 24–26 may be recast in function form. First, define a bivariate vapor-pressure function $V(w, u)$ as follows:

$$V(w, u) = PK(w, u). \tag{27}$$

The vapor-phase distribution $y(u)$ is given in terms of $x(u)$ and $K(w, u)$ by Eq. 4:

$$y(w) = \int_\Omega K(w, u)x(u)\, du. \tag{28}$$

The integral of $y(w)$ must equal 1.0. Using Eqs. 27 and 28 with this fact results in the following:

$$P = \int_\Omega \int_\Omega V(w, u)x, (u)\, dudw. \tag{29}$$

The functional equations given by Eqs. 28–29 are then expanded into series-coefficient form, truncated to n terms:

$$\bar{y}_i = \sum_{n=1}^{n} \bar{K}_i^j \bar{x}_i = \frac{1}{P_n} \sum_{i=1}^{n} \bar{V}_i^j \bar{x}_i, \tag{30}$$

$$P_n = \sum_{i=1}^{n} \sum_{j=1}^{n} \bar{l}^i \bar{V}_i^j \bar{x}_j, \tag{31}$$

where the subscript n on P_n denotes an n term approximation. The $\{\bar{l}^i\}_{i=1,n}$ values are given as follows:

$$\bar{l}^i = \int_\Omega \phi_i(u)\, du. \tag{32}$$

The $\{\bar{V}_i^j\}_{i,j=1,n}$ are similar to the $\{\bar{K}_i^j\}_{i,j=1,n}$; the values of $\{\bar{V}_i^j\}_{i,j=1,n}$ are given by equations similar to Eq. 9:

$$\bar{V}_i^j = \int_\Omega \phi_i(u)v(u)\phi^j(u)\, du, \tag{33}$$

where $v(u)$ is the vapor-pressure correlation function: $P_i^\sigma = v(a_i)$.

When the first basis function $\phi_i(u)$ equals a constant, C, then all the $\{\bar{l}^i\}_{i=1,n}$ will equal zero except for $\bar{l}^1$. For, from the definition of $\bar{l}^i$ (Eq. 32) and the duality condition (Eq. 10), we have

$$\bar{l}^i = \int_\Omega \phi^i(u)\, du = \frac{1}{C} \int_\Omega \phi_1(u)\phi^i(u)\, du = \frac{1}{C}\delta_1^i. \tag{34}$$

Equation 34 allows Eq. 31 to be simplified:

$$P_n = \bar{l}^1 \sum_{i=1}^{n} \bar{V}_1^i \bar{x}_i. \tag{35}$$

Thus, the evaluation of P_n requires a summation of n terms whereas the classical calculation (Eq. 26) requires a summation of c terms. If $n \ll c$, a substantial savings in computation time and effort may be achieved, as will be shown in an example calculation below.

If Eq. 35 is evaluated successively using an increasing number of terms, a sequence of pressures P_k will be obtained which will ultimately converge to the true pressure P:

$$P_1 P_2 \ldots P_{k-1} P_k P_{k+1} \ldots P. \tag{36}$$

The error $|P_k - P|$ of the approximation may be estimated using the remainder formula from the ratio test for convergence (Kaplan, 1973). The sequence P_k will first be converted to series form. Define

$$a_1 = P_1,$$
$$a_k = P_k - P_{k-1}, \; k = 2, \ldots, N, \tag{37}$$

where N is the total number of terms in the sequence. Each term of the sequence may then be expressed as a partial sum:

$$P_k = \sum_{i=1}^{k} a_i. \tag{38}$$

The remainder term may be estimated from the ratio test:

$$|P - P_k| \leq \frac{|a_{k+1}|}{1 - r}, \tag{39}$$

where

$$\left|\frac{a_{k+1}}{a_k}\right| \leq r < 1. \tag{40}$$

In order to utilize the ratio remainder equation the ratios $|a_{k+1}/a_k|$ must be less than 1.0. Furthermore, if the values of the ratios $|a_{k+1}/a_k|$ decrease as k increases, then the most recent value at $|a_{k+1}/a_k|$ may be used in place of r in Eq. 40.

When the simplified form for P_n given by Eq. 35 is used, the ratio test can be most effectively applied. The terms of the summation can be calculated incrementally and the ratio test applied as n is increased. When the ratio test indicates that $|P - P_n|$ is within the desired bounds the calculation is terminated. Such an application of the ratio test is illustrated below for an example bubble-point calculation utilizing Legendre polynomials.

The continuous-lumped method was used to perform a bubble-point pressure calculation and a flash calculation on a hypothetical 64-component mixture containing the n-alkanes from n-$C_{11}H_{24}$ to n-$C_{74}H_{150}$. The mole fraction of each component was calculated from

$$x_i = \frac{(n_i - 11)(74 - n_i)}{41{,}664}, \tag{41}$$

where n_i is the carbon number of component A_i.

The equilibrium ratio correlation $k(u)$ was obtained as follows. The Clausius-Clapeyron equation for vapor-pressure was combined with Trouton's rule in order to relate the ideal equilibrium constant to the normal boiling point of each component:

$$k(u) = \frac{1}{P}\exp\left[10.58\left(1 - \frac{u}{T}\right)\right], \tag{42}$$

where the component index u is taken as the normal boiling point. A detailed derivation of Eq. 42 is given by Anthony and Moore (1986).

The mole fractions given by Eq. 41 were then used to specify a discrete distribution function $x(u)$, as in Eq. 16. The distribution $x(u)$, together with the expression for $k(u)$ (Eq. 42), was then used to perform bubble-point pressure and flash calculations using the continuous-lumping method described above. The convergence of successive approximations using up to ten terms was evaluated and error bounds on the calculated results were obtained using the series ratio-remainder test.

An exact 64-component bubble-point pressure calculation was performed (Eq. 26) in order to provide a standard by which to judge the continuous-lumping method. A set of pseudocomponent calculations were also performed for comparison purposes. Each pseudocomponent consisted of a fixed number of neighboring components. The equilibrium ratio for each pseudocomponent was taken as the average of the equilibrium ratios of the contributing components. Traditional discrete equilibrium calculations were then performed, treating the pseudocomponents as real, physical components.

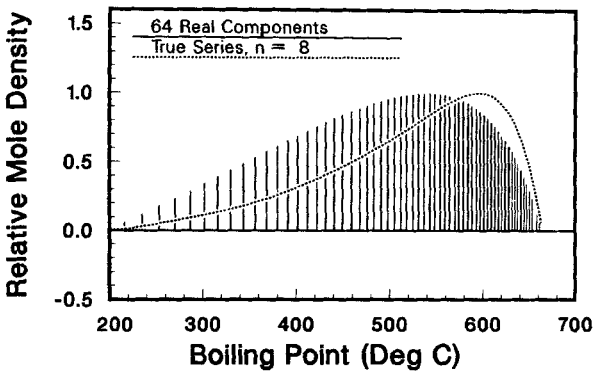

Figure 5. n-Alkane feed distribution.

The discrete distribution $x(u)$ of the 64-component n-alkane mixture has been plotted in Figure 5 (scaled to maximum height 1.0). Also shown in Figure 5 is the eight-term Legendre series expansion $x_8(u)$ (dashed line). Note that the maximum in the approximate distribution $x_8(u)$ occurs 60°C to the right of the maximum in the original distribution $x(u)$. This shift in the distribution maximum is due to the closer spacing of the Dirac delta functions on the right side of the figure.

The plot (Figure 6) of the lumping indicator function $G_x(n)$ for the 64-component mixture indicates that after seven Legendre terms are included in the series, the effect of adding additional terms becomes negligible; the eight-term approximation $x_8(u)$ contains essentially all of the low-order character of the original distribution $x(u)$. The series coefficients $\{\bar{x}_i\}_{i=1,n}$ and the $G_x(n)$ function values have been listed in Table 3.

Figure 7 illustrates the n-alkane feed composition and eight-term expansion $x_8(u)$ in terms of a cumulative mole fraction plot. Such plots are similar in nature to true-boiling-point (TBP) distillation curves. Whereas the eight-term pseudocomponent calculation as illustrated in Figure 8 gave only a crude approximation to the true cumulative mole-fraction curve, Figure 7 indicates that the cumulative plot of the eight-term expansion $x_8(u)$ is virtually identical to the original curve, with the excep-

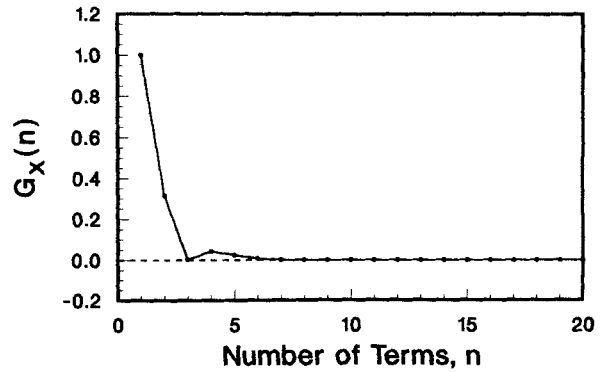

Figure 6. 64-Component lumping effectiveness.

Table 3. Bubble-Point Pressure Results Continuous-Lumped Method*

Funct. No. i	Feed Coefficients		Vapor Coefficients: True		Vapor Coefficients: Calculated	
	$\bar{x}_i$	$G_x(i)$	$\bar{y}_i$	$G_y(i)$	$\bar{y}_i^*$	$G_y^*(i)$
1	0.70710678	1.0000	0.70710677	1.0000	0.70710678	1.0000
2	0.47904247	0.3146	−0.31685999	0.1672	−0.31691753	0.1673
3	−0.05027740	0.0035	−0.18969356	0.0565	−0.18940064	0.0564
4	−0.18160690	0.0431	0.18573713	0.0514	0.18461101	0.0508
5	−0.13806961	0.0243	−0.09987969	0.0147	−0.09619639	0.0136
6	−0.07649182	0.0074	0.03218314	0.0015	0.02229899	0.0007
7	−0.03498079	0.0015	−0.01315921	0.0003	0.00770445	0.0001
8	−0.01375080	0.0002	0.00973279	0.0001	−0.02288933	0.0008

*Mixture temperature, 500°C; mixture pressure, 209.48 kPa; fraction vaporized, 0.0.
Component index, boiling point; series expansion performed between 195.93°C and 663.0°C; function type used, orthonormal Legendre polynomials; number of terms, 8.

tion that the $x_8(u)$ curve does not exhibit the stair-step discontinuities of the original curve, which are indicative of the discrete nature of the original mixture. Thus, an exact 64-parameter mixture description has been replaced with an eight-parameter description which describes the mixture as a whole equally well.

Table 3 summarizes the results of an eight-term bubble-point pressure calculation using the continuous-lumped method. The calculated pressure equals the true pressure, 209.48 kPa. The calculated vapor coefficients $\{\bar{y}_i^*\}_{i=1,n}$ agree fairly well with the true coefficients $\{\bar{y}_i\}_{i=1,n}$ through the fifth term, after which the agreement deteriorates. However, the 6th, 7th and 8th terms contribute little to the total objective function $J(8)$, as evidenced by the values for $G_y(6)$, $G_y(7)$, and $G_y(8)$. The lumping index values for the calculated coefficients, $G_y^*(6)$, $G_y^*(7)$, and $G_y^*(8)$ also indicate that the corresponding series coefficients—$\bar{y}_6^*$, $\bar{y}_7^*$, and $\bar{y}_8^*$— are of little consequence. Note, however, the sudden increase in the value of $G_y^*(n)$ from $n = 7$ to $n = 8$, a change from 0.0001 to 0.0008. Such a sudden increase so soon after $G_y^*(n)$ reaches a minimum is inconsistent with the behavior usually observed for a low-order expansion of a discrete distribution. The sudden increase may therefore be taken as indicative of a large error in $\bar{y}_8^*$, an observation which is verified by comparing $\bar{y}_8^*$ to the true coefficient $\bar{y}_8^*$.

Figure 9 shows the vapor-phase cumulative mole fraction plots for the discrete mixture, the true series expansion, and the calculated series expansion. As for the feed mixture (Figure 7), the true series expansion curve lies on top of the stair-step curve for the 64 real components. Furthermore, the calculated series curve is indistinguishable from the true series curve, a further verification that the last few terms of the series are relatively insignificant.

The good fit exhibited by the eight-term series approximation in Figure 9 is a vast improvement over the eight-term pseudocomponent approximation, exhibited in Figure 10. First, the series method provides a closer approximation to the true discrete curve (stair-step curve). Furthermore, there is little bias in the calculated series curve; the pseudocomponent bubble-point calculation significantly overpredicted the mole fractions of the less-volatile pseudocomponents (as much as 23% over the true value), which resulted in the pseudocomponent curve lying beneath the curve for the discrete mixture. Figure 9 indicates that no such bias is present, with both the true and calculated series curves lying on top of the discrete curve.

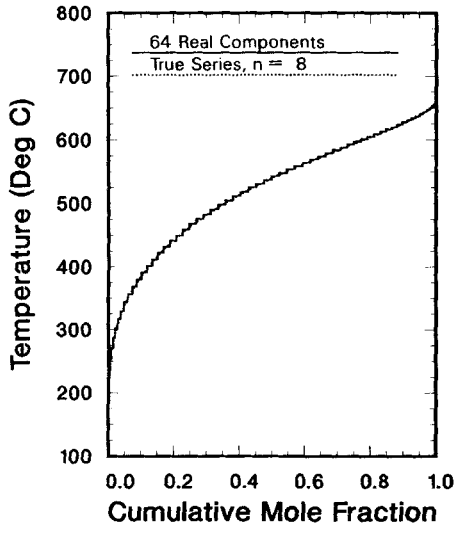

Figure 7. *n*-Alkane feed composition Legendre series expansion.

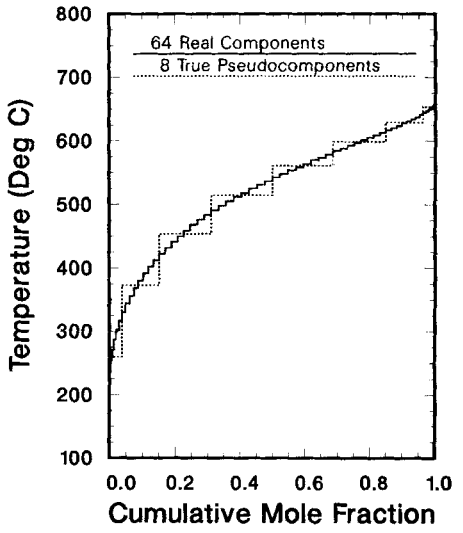

Figure 8. *n*-Alkane feed composition pseudocomponent approximation.

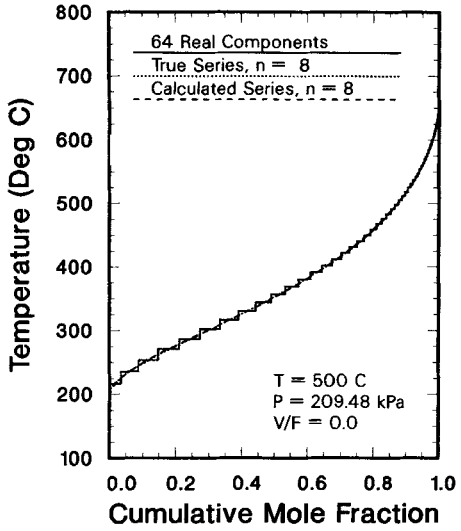

Figure 9. Bubble-point vapor composition continuous-lumping method.

The effect of the number of terms upon the calculated pressure is illustrated in Figure 11 and summarized in Table 4. As Figure 11 indicates, the pressure essentially converges to the true value after only six terms. By contrast, the pseudocomponent method converges slowly, needing 16 to 32 pseudocomponents to obtain an accurate value for the calculated pressure.

Table 4 also illustrates the use of the series ratio-remainder test to obtain error bounds for the calculated pressures. Provided that the increment ratio remains less than 1.0, the ratio test usually provides both a bound on and an estimate for the magnitude of the error. The only exception in Table 4 is for term eight, where the error, −0.0365, exceeds the error bound, 0.0011. This violation of the ratio test is probably due to errors in the numerical evaluation of the integrals used in calculating the vapor pressure coefficients $\{\overline{V}_i^j\}_{i,j=1,n}$. This line of reasoning is consistent with the poor value obtained for $\overline{y}_8^*$. Thus, the ratio test provides a convenient means of determining the number of terms

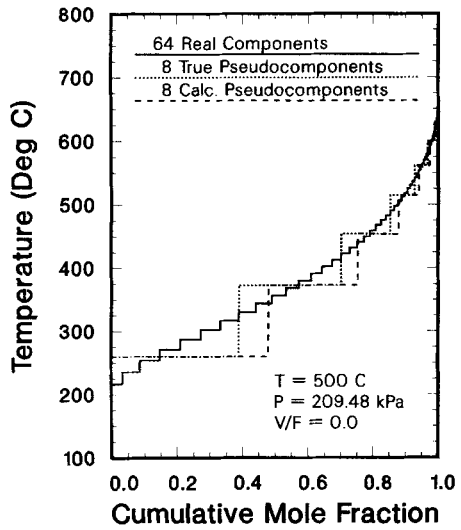

Figure 10. Bubble-point vapor composition pseudocomponent method.

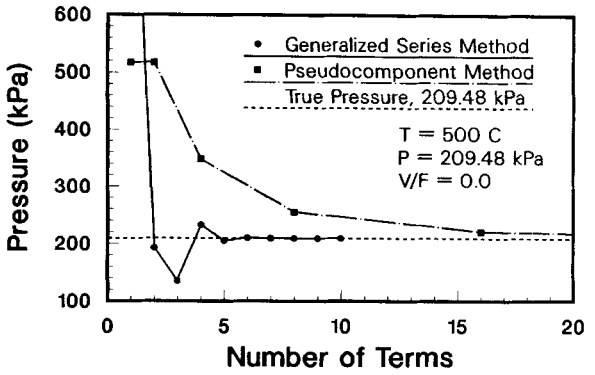

Figure 11. Bubble-point pressure comparison.

required to obtain a specified accuracy; for example, the ratio test indicates that the value for the pressure obtained using seven terms, $P_7 = 209.438$ kPa, is within 0.02% of the true value, $P = 209.48$ kPa. Thus, a high degree of precision was obtained before the ratio test failed due to numerical difficulties.

Conclusions

The continuous-lumping method extends the application of continuous thermodynamics to mixtures which do not satisfy the continuous-mixture hypothesis. Continuous-lumping is useful especially for mixtures described by discontinuous component distributions, such as the diesel fuel chromatogram in Figure 2. Furthermore, the method provides an effective means of reducing the number of parameters required to adequately describe the mixture. The detail provided by an n-term series expansion is vastly superior to the description provided by n pseudocomponents. In addition to parameterizing the component distribution, the continuous-lumping method provides a mathematical means, whereby the governing equilibrium equations may be converted to the new parameterized form. Finally, since the method is based on a mathematical approximation instead of the continuous-mixture assumption, the errors involved in the approximation can be evaluated using standard mathematical techniques such as the series ratio-remainder formula.

The applications of the continuous-lumping method illustrated herein have been kept simple in order to emphasize the basis for the method. However, the method is readily applicable

Table 4. Bubble-Point Pressure Convergence Summary Continuous-Lumped Method*

No. of Terms	Calculated Pressure (kPa)	Error in Pressure (kPa)	Incremental Pressure (kPa)	Ratio of Increments	Ratio Bound (kPa)
1	1014.9626	805.4881	1014.9626	—	4331.6
2	192.6738	−16.8007	822.2887	0.810167	60.984
3	135.8973	−73.5772	56.7765	0.0690	—
4	232.4015	22.9270	96.5041	1.6997	37.344
5	205.4762	−3.9983	26.9253	0.2790	5.3666
6	209.9509	0.4764	4.4747	0.1662	0.5793
7	209.4380	−0.0365	0.5129	0.1146	0.0436
8	209.4781	0.0036	0.0401	0.0783	0.0011
9	209.4770	0.0025	0.0011	0.0277	—
10	209.4743	−0.0002	0.0027	2.4545	—

*Mixture temperature, 500°C; mixture pressure, 209.48 kPa; true fraction vaporized, 0.0.

to virtually any type of phase equilibrium calculations involving either ideal or nonideal mixtures. For example, the continuous-lumping method has been incorporated into an equation of state flash routine by Moore (1988).

Thus, the development of continuous-lumping establishes numerous practical and fundamental advantages over both pseudocomponent lumping and continuous thermodynamics. We have removed the theory of continuous thermodynamics from the limiting concept of a continuous mixture and have placed it on a mathematical foundation. We believe that the methodology developed herein will allow the practical application of continuous-modeling concepts to complex mixtures of industrial significance.

Acknowledgments

The authors thank Dr. G. D. Allen of the Mathematics Department for his advice and recommendations concerning the theory of generalized functions (distributions).

The financial support in the form of graduate fellowships from Phillips Petroleum Company, Texaco Oil Company, Gulf Oil Company, and Texas Eastman Company is greatly appreciated. The financial support of the U. S. Department of Energy (Project Number DE-AC18-83FC10601), the Texas A&M University Center for Energy and Mineral Resources, and the Texas A&M Engineering Experiment Station is gratefully acknowledged. The cooperation extended by the Energy Research Center at the University of North Dakota is also greatly appreciated.

Notation

A_i = ith real component of the mixture
a_i = ith term of a mathematical series; component-index of ith component
c = number of real components
C_0^∞ = set of infinitely-differentiable functions of bounded support
$\bar{g}^{ij}$ = ijth series coefficient of the metric function $\delta(u - w)$ relative to $\phi_i(w)\phi_j(u)$
$G_x(n)$ = lumping-effectiveness function for $x_n(u)$
$J(n)$ = least-squares objective function
$k(u)$ = vapor-liquid equilibrium-ratio correlation
K_i = vapor-liquid equilibrium ratio
$K(w, u)$ = bivariate vapor-liquid equilibrium function
$\bar{K}_i^j$ = ijth series coefficient of $K(w, u)$ relative to $\phi^i(w)\phi_j(u)$
$\bar{l}^i$ = ith series-coefficient of the unit function relative to $\phi_i(u)$
n = number of pseudocomponents in a lumping scheme; number of terms in a series approximation
n_i = carbon number of component A_i
P = total mixture pressure
P_i° = vapor-pressure of component A_i
P_n = calculated pressure from the n-term series approximation
r = series coefficient ratio from the ratio convergence test
T = mixture temperature
T_i° = normal boiling-point of component A_i
u = component-index variable
$v(u)$ = vapor-pressure correlation
$V(w, u)$ = bivariate vapor-pressure function
$\bar{V}_i^j$ = ijth series coefficient of $V(w, u)$ relative to $\phi^i(w)\phi_j(u)$
w = component-index dummy variable
x_i = liquid-phase mole fraction of component A_i
$x(u)$ = liquid-phase distribution function
$x_n(u)$ = n-term series approximation to $x(u)$
$\bar{x}_i$ = ith series coefficient of $x(u)$ relative to $\phi^i(u)$
y_i = vapor-phase mole fraction of component A_i
$y(u)$ = vapor-phase distribution function
$y_n(u)$ = n-term series approximation to $y(u)$
$\bar{y}_i$ = ith series coefficient of $y(u)$ relative to $\phi^i(u)$

Greek letters

δ_j^i = Kronecker delta
$\delta(u)$ = Dirac delta-function
$\phi_i(u)$ = alternate basis function
$\phi^i(u)$ = alternate dual basis function
$\varphi(u)$ = infinitely differentiable function of bounded support: $\varphi(u) \in C_0^\infty$
Ω = domain of component index variable u

Mathematical symbols

$\|\cdot\|_2$ = Euclidean norm

Subscripts

n = n-term series approximation
(n) = n-term lumping scheme
(m) = m-term lumping scheme

Superscripts

(n) = n-term lumping scheme
(m) = m-term lumping scheme
$*$ = calculated or predicted parameter
$\bar{}$ = parameter taken relative to alternate basis set
σ = vapor-liquid interface value

Literature Cited

Anthony, R. G., and P. K. Moore, "Kinetic Model Development for Low-Rank Coal Liquefaction," Quarterly Report, DOE/FC/10601-12 (Apr. 1, 1986–June 30, 1986).

Cotterman, R. L., R. Bender, and J. M. Prausnitz, "Phase Equilibria for Mixtures Containing Very Many Components: Development and Application of Continuous Thermodynamics for Chemical Process Design," *Ind. Eng. Chem. Process Design and Dev.*, **24**, 194 (1985).

Kaplan, W., *Advanced Calculus*, 2nd ed., McGraw-Hill, New York (1973).

Moore, P. K., "A Continuous-Lumped Phase-Equilibrium Model for Complex Multiphase Reactions," PhD Diss., Texas A&M Univ. (Aug., 1988).

Schwartz, L., *Mathematics for the Physical Sciences*, Addison-Wesley, Reading, MA (1966).

Manuscript received Feb. 16, 1989, and revision received Apr. 17, 1989.