Thermodynamically Consistent Quasi-Newton Formulae

Newton and quasi-Newton methods have been used in chemical process design and optimization calculations for quite some time. They continue to be used today, both in the traditional sense and as part of the more recent hybrid method. While Newton-based fixed-point methods have been used to solve many different kinds of chemical process design and optimization problems, perhaps the point of single largest application has been that of multicomponent separation problems, especially equilibrium stage distillation.

Quite aside from this, classical thermodynamics provides us with certain fundamental mathematical and physical relationships governing the behavior of nonideal solutions, those being the homogeneity of partial molar excess properties and their derivatives and the Gibbs-Duhem equation.

In this work, we draw a connection between the class of Newton-based fixed-point methods and classical nonideal thermodynamics. That is, it is shown that, with the exception of Newton's method, none of the conventional Newton-like methods gives matrix approximations that are thermodynamically consistent. In other words, all existing nonsymmetric and symmetric quasi-Newton formulae generate matrix approximations that do not satisfy, for example, either the zero-degree homogeneity or Gibbs-Duhem equations.

In light of this, a new class of quasi-Newton formulae is presented for use in chemical process problems whose models include chemical and/or phase equilibrium. In particular, several new quasi-Newton formulae are presented that give Jacobian or Hessian approximations that satisfy the zero-degree homogeneity and/or Gibbs-Duhem equations, in addition to the usual secant and perhaps symmetry and sparsity conditions. This new class of formulae is called thermodynamically consistent quasi-Newton formulae. Some numerical results are presented that show that these thermodynamically consistent quasi-Newton formulae can provide improvements in reliability and computational efficiency when compared to existing Newton-like methods.

ANGELO LUCIA, D. C. MILLER
and ASHWANI KUMAR

Department of Chemical Engineering
Clarkson University
Potsdam, NY 13676

SCOPE

One common criticism of Newton's method as it applies to chemical process design and optimization is that it can be expensive to use on problems that involve complex nonideal thermodynamic calculations. This expense is due to the fact that the calculation of the derivatives of nonideal thermodynamic quantities can account for a very large percentage of the computer time needed to solve a problem. It remains, nevertheless, a highly reliable equation solving procedure.

The fact that Newton's method can frequently be a costly means of obtaining solutions to these problems has prompted both academic and industrial workers to seek alternatives. While there are many, for a long time quasi-Newton methods represented perhaps the best of these alternatives. They have been and continue to be more efficient than Newton's method in more than an insignificant number of cases. However, depending on the application, their reliability is frequently less than that of Newton's method. More recently, it has been shown that a hybrid approach which makes combined use of available analytical derivatives and a quasi-Newton formula, is superior to pure quasi-Netwon methods both from the standpoint of reliability and efficiency. Moreover, the robustness of this hybrid approach has been shown to be surprisingly close to that of Newton's method. The reliability is still not as good as Newton's method, but it is close. This has been clearly demonstrated for chemical process design problems involving nonideal solution behavior. Thus, in a limited sense, some real improvements have been made in equation solving capability.

Classical nonideal solution thermodynamics, on the other hand, provides us with certain fundamental mathematical and physical relationships that, to this date, have either been taken for granted or assumed to be completely divorced for certain aspects of the equation solving task. For example, any nonideal liquid mixture is subject to both the zero-degree homogeneity equations of $\ln \gamma$ and H^E as well as the Gibbs-Duhem equation.

In this work, we explore the relationship between equation solving by hybrid methods and thermodynamic consistency.

Correspondence concerning this paper should be addressed to Angelo Lucia.

CONCLUSIONS AND SIGNIFICANCE

All existing nonsymmetric and symmetric quasi-Newton formulae were shown to yield thermodynamically inconsistent Jacobian and Hessian matrix approximations, respectively. Accordingly, a new class of thermodynamically consistent quasi-Newton formulae was presented for use in multicomponent phase and chemical equilibrium calculations. Numerical results were presented that show that these thermodynamically consistent quasi-Newton formulae, when used in the context of a hybrid method, can result in improved computational efficiency and reliability.

INTRODUCTION

In a recent paper, Lucia (1985) has shown that the Schubert (1970) update, when used in the context of a hybrid fixed-point algorithm (Lucia and Macchietto, 1983) usually generates matrix approximations that are thermodynamically inconsistent. That is, when the matrix of terms involving nonideal thermodynamic derivatives is approximated by Schubert's method, that matrix does not satisfy either the zero degree homogeneity property of ln γ and H^E or the Gibbs-Duhem equation. Lucia has substantiated this using a set of single-stage, adiabatic flash problems involving a mixture of ethanol and n-hexane. Furthermore, in that same paper he has presented two new quasi-Newton formulae that exploit the zero degree homogeneity of ln γ and H^E. Thus, in a sense, they are thermodynamically consistent. Finally, using the same set of flash problems, Lucia has demonstrated that a thermodynamically consistent quasi-Newton formula can result in both improved reliability and greater computational efficiency when compared to Schubert's formula.

In this manuscript, we first extend the nonsymmetric quasi-Newton formulae presented in Lucia to the situation where zero-degree homogeneity equations must be independently satisfied in more than one phase, as in the cases of liquid-liquid and vapor-liquid-liquid equilibria. The updating formulae given in Lucia cannot do this. Second, we suggest a new symmetric quasi-Newton formula, for use in a hybrid optimization method for calculating chemical and/or phase equilibrium points by Gibbs free energy minimization. With this we study the effects of satisfying the zero-degree homogeneity and Gibbs-Duhem equations simultaneously. Next, we present some numerical results that show that these new thermodynamically consistent quasi-Newton formulae can result in improvements in reliability and/or computational efficiency when compared to existing Newton-like methods. Finally, we make some concluding remarks concerning the effects of thermodynamic consistency in equation solving and show how the Gibbs-Duhem equation can be satisfied in the nonsymmetric case by using partition symmetry.

AN ILLUSTRATIVE EXAMPLE

In the discussions that follow, it is convenient to consider the nonsymmetric and symmetric cases separately. We believe that it makes the presentation clearer.

Nonsymmetric Case.

Consider, for the sake of definiteness, a single-stage liquid-liquid flash drum which is assumed to operate at constant temperature and pressure. One set of equations that model this process are the equilibrium equations

$$K_i \ell_i^1 \Big/ \left(\sum_{j=1}^{n_c} \ell_j^1\right) - \ell_i^2 \Big/ \left(\sum_{j=1}^{n_c} \ell_j^2\right) = 0, \quad i = 1,2,\ldots,n_c \tag{1}$$

and the component mass balances

$$f_i^1 + f_i^2 - \ell_i^1 - \ell_i^2 = 0, \quad i = 1,2,\ldots,n_c \tag{2}$$

where ℓ_i denotes the ith component molal flow rate and f_i is the corresponding component molal feed flow rate. Also, K_i is the distribution ratio for the ith component. Finally, i is used to denote the component index, n_c is the total number of components in the mixture under consideration, and the superscripts 1 and 2 are used to simply distinguish the two dense phases.

Such a problem is typically solved by some Newton-based fixed-point algorithm. Therefore, in order to describe Newton's method and its many recent modifications, we let x denote a vector of unknown process variables (e.g., in the case of the illustrative example, ℓ_i^1 and ℓ_i^2, $i = 1,2,\ldots,n_c$). Moreover, we denote our collection of model equations by f and further suppose that the Jacobian matrix of f, $J(x)$, can be written as

$$J(x) = C(x) + A(x) \tag{3}$$

where $C(x)$, the computed part of the Jacobian, can be calculated from readily available derivative information and $A(x)$, the approximated part, contains terms involving derivative information that can be expensive to obtain (e.g., activity coefficient-component flow rate derivatives).

From the standpoint of classical thermodynamics, $C(x)$ is usually comprised of derivative information that looks very much like the ideal solution part of the model, while $A(x)$ contains terms that involve the derivatives of nonideal thermodynamic quantities. More specifically, for the illustrative example if the variables are ordered as $\ell_1^1, \ell_2^1, \ldots, \ell_{n_c}^1, \ell_1^2, \ell_2^2, \ldots, \ell_{n_c}^2$ and the equations are ordered as they appear, then $A(x)$, which is most important for our discussions, is given by

$$A(x) = \begin{bmatrix}
\frac{K_1\ell_1^1}{(\Sigma\ell_j^1)^2}\left(\frac{\partial \ln \gamma_1^1}{\partial \ell_1^1}\right) & \cdots & \frac{K_1\ell_1^1}{(\Sigma\ell_j^1)^2}\left(\frac{\partial \ln \gamma_1^1}{\partial \ell_{n_c}^1}\right) & \frac{K_1\ell_1^1}{(\Sigma\ell_j^1)^2}\left(\frac{\partial \ln \gamma_1^2}{\partial \ell_1^2}\right) & \cdots & \frac{K_1\ell_1^1}{(\Sigma\ell_j^1)^2}\left(\frac{\partial \ln \gamma_1^2}{\partial \ell_{n_c}^2}\right) \\
\vdots & & & \vdots & & \vdots \\
\frac{K_{n_c}\ell_{n_c}^1}{(\Sigma\ell_j^1)^2}\left(\frac{\partial \ln \gamma_{n_c}^1}{\partial \ell_1^1}\right) & \cdots & \frac{K_{n_c}\ell_{n_c}^1}{(\Sigma\ell_j^1)^2}\left(\frac{\partial \ln \gamma_{n_c}^1}{\partial \ell_{n_c}^1}\right) & \frac{K_{n_c}\ell_{n_c}^1}{(\Sigma\ell_j^1)^2}\left(\frac{\partial \ln \gamma_{n_c}^2}{\partial \ell_1^2}\right) & \cdots & \frac{K_{n_c}\ell_{n_c}^1}{(\Sigma\ell_j^1)^2}\left(\frac{\partial \ln \gamma_{n_c}^2}{\partial \ell_{n_c}^1}\right) \\
0 & \cdots & 0 & 0 & \cdots & 0 \\
\vdots & & \vdots & \vdots & & \vdots \\
0 & \cdots & 0 & 0 & \cdots & 0
\end{bmatrix} \tag{4}$$

where γ_i^1 and γ_i^2 denote the ith component activity coefficient in phases 1 and 2 respectively and $K_i = \gamma_i^1/\gamma_i^2$.

Symmetric Case.

We can also calculate the equilibrium point by minimizing the Gibbs free energy for the system subject to the mass balance constraints, instead of solving the model defined by Eqs. 1 and 2. Such a procedure is, certainly by now, a well-known method for computing points of chemical and/or phase equilibrium at constant temperature and pressure. Thus, in the case of the illustrative example the equivalent optimization problem that we wish to solve is given by

$$\min G = \sum_{m=1}^{2} \sum_{i=1}^{n_c} \ell_i^m (\overline{G}_i^{Im} + \overline{G}_i^{Em}) \tag{5}$$

subject to

$$f_i^1 + f_i^2 - \ell_i^1 - \ell_i^2 = 0, \quad i = 1,2,\ldots,n_c \tag{6}$$

and

$$\ell_i^m \geq 0, \quad i = 1,2,\ldots,n_c;\ m = 1,2 \tag{7}$$

where

$$\overline{G}_i^{Im} = \overline{G}_i^{0m} + RT \ln\left(\ell_i^m \Big/ \sum_j \ell_j^m\right) \tag{8}$$

and

$$\overline{G}_i^{Em} = RT \ln \gamma_i^m. \tag{9}$$

It should be clear to the reader that the solution to the above linearly constrained optimization problem by Newton's method requires, as in the case of the solution of the process model defined by Eqs. 1 and 2, first partial derivatives of $\ln \gamma_i^m$ with respect to ℓ_j^m and can again be computationally expensive.

In this case, we suppose that the matrix of second derivatives (or Hessian matrix) can be written as

$$H(x) = C(x) + A(x). \tag{10}$$

For the liquid-liquid flash problem, the approximated part of the Hessian matrix, which in this case is the Hessian matrix of the Gibbs free energy function because the constraints are all linear, is given by

$$A(x) = RT \begin{bmatrix} \left(\dfrac{\partial \ln \gamma_1^1}{\partial \ell_1^1}\right) & \cdots & \left(\dfrac{\partial \ln \gamma_1^1}{\partial \ell_{n_c}^1}\right) & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ \left(\dfrac{\partial \ln \gamma_{n_c}^1}{\partial \ell_1^1}\right) & \cdots & \left(\dfrac{\partial \ln \gamma_{n_c}^1}{\partial \ell_{n_c}^1}\right) & 0 & \cdots & 0 \\ 0 & \cdots & 0 & \left(\dfrac{\partial \ln \gamma_1^2}{\partial \ell_1^2}\right) & \cdots & \left(\dfrac{\partial \ln \gamma_1^2}{\partial \ell_{n_c}^2}\right) \\ \vdots & & \vdots & \vdots & & \\ 0 & \cdots & 0 & \left(\dfrac{\partial \ln \gamma_{n_c}^2}{\partial \ell_1^2}\right) & \cdots & \left(\dfrac{\partial \ln \gamma_{n_c}^2}{\partial \ell_{n_c}^2}\right) \end{bmatrix} \tag{11}$$

THERMODYNAMIC CONSISTENCY

It is well known from classical thermodynamics that many of the important thermodynamic quantities are homogeneous functions of mole numbers, typically of degree zero or one. For example, the total Gibbs free energy and total excess Gibbs free energy are homogeneous functions of degree one, whereas the molar excess enthalpy and the function $\ln \gamma$ are homogeneous of degree zero.

Because these thermodynamic quantities are homogeneous functions, they satisfy Euler's theorem, which for an arbitrary homogeneous function, say $F(x)$, of n variables is given by

$$\sum_{j=1}^{n} \left(\frac{\partial F}{\partial x_j}\right) x_j = pF(x), \tag{12}$$

where p is the degree of homogeneity. Thus in the case of $\ln \gamma$, which is homogeneous of degree zero in mole numbers, we have that

$$\sum_{j=1}^{n_c} \left(\frac{\partial \ln \gamma_i^m}{\partial \ell_j^m}\right) \ell_j^m = 0, \quad i = 1,2,\ldots,n_c \quad m = 1,2,\ldots,n_p \tag{13}$$

where n_p denotes the number of phases. Moreover, a similar relationship holds for the molar excess enthalpy since H^E is also homogeneous of degree zero.

In addition to the zero-degree homogeneity property, we also know that $\ln \gamma$ must satisfy the Gibbs-Duhem equation, which is given by

$$\sum_{j=1}^{n_c} \left(\frac{\partial \ln \gamma_j^m}{\partial \ell_i^m}\right) \ell_j^m = 0, \quad i = 1,2,\ldots,n_c \quad m = 1,2,\ldots,n_p. \tag{14}$$

From these rather simple equations from classical thermodynamics we can, in our opinion, deduce a great deal about the approximated part as a linear operator.

Nonsymmetric Case.

In the nonsymmetric case, Eq. 13 implies that the approximated part of the Jacobian matrix has a nontrivial null space (or kernel) defined by the component molal flow rates. That is,

$$A(x)z^m = 0, \quad m = 1,2,\ldots,n_k \tag{15}$$

$$z_j^m = \begin{cases} \ell_j^m, & \text{for } j \text{ such that } (i,j) \epsilon I_m, 1 \leq i,j \leq n \\ 0, & \text{otherwise} \end{cases} \tag{16}$$

where I_m is the set of index pairs (i,j) that define the nonzero entries in the approximated part associated with the component molal flow rates in phase m and where n_k is the dimension of the null space (for example, for a multistage process, in general, $n_k = n_p n_s$ where n_s is the number of stages). Also, observe that for the problems in which we are interested the vectors z^m, $m = 1,2,\ldots,n_k$, are always mutually orthogonal.

Symmetric Case.

Together, the zero-degree homogeneity and Gibbs-Duhem equations imply that the approximated part of the Hessian matrix is symmetric and that it has this nontrivial null space of dimension n_k that is defined by the component molal flow rates. More specifically, Eqs. 13 and 14 imply Eq. 15 and the condition

$$A(x) = A^T(x). \tag{17}$$

EXISTING NEWTONLIKE METHODS AND THERMODYNAMIC CONSISTENCY

Nonsymmetric Case.

For the nonsymmetric case, we have some collection of nonlinear process model equations, $f(x)$, that we wish to solve for a fixed point, x^*. We make the choice that we will use some Newton-based fixed-point method to solve our nonlinear process model equations. Consequently, we are concerned with an iterative procedure of the form

$$B_k(x_{k+1} - x_k) = -f_k, \tag{18}$$

where B_k represents some approximation to the Jacobian matrix of f, $f_k = f(x_k)$, and k is the iteration number. We further suppose that the Jacobian matrix of f can be written in the form defined by Eq. 3.

Within this general framework, it is possible to define all of the existing nonsymmetric Newtonlike methods. For example, when $C(x)$ is evaluated analytically and $A(x)$ is calculated analytically or by finite difference, we have an implementation of Newton's method. If $C(x)$ is again evaluated analytically but $A(x)$ is instead calculated by

$$\overline{A} = A + \sum_{i=1}^{n} (s_i^T s_i)^+ e_i^T [y\text{-}\overline{C}s\text{-}As] e_i s_i^T, \tag{19}$$

then we have the hybrid method of Lucia and Macchietto (1983). Here A and $\overline{A}$ and successive approximations to $A(x)$, $s = x_{k+1} - x_k$, $y = f(x_{k+1}) - f(x_k)$, and $\overline{C} = C(x_{k+1})$. Moreover, s_i is the vector formed from s such that it has zero everywhere the ith row of $A(x)$ does and the superscripts $+$ and T denote the generalized inverse and matrix transposition respectively. Observe that Eq. 19 is simply a generalized representation of the Schubert update. Finally, we note that it is also quite easy to represent the pure quasi-Newton methods of Broyden (1965, 1971) and Schubert (1970) within this Newton-based framework.

When the approximated part of the Jacobian matrix is calculated by finite differences or analytically, it has this correct null space. In other words, it satisfies the zero-degree homogeneity equations of $\ln \gamma$ and H^E. On the other hand, when the Schubert update (i.e., Eq. 19) is used to build approximations to $A(x)$, such approximations are usually thermodynamically inconsistent. This is because the variational calculus problem from which the Schubert update is derived contains no provision (i.e., no constraint equation) for forcing A, the approximation to $A(x)$, to satisfy the zero-degree homogeneity equations.

To partially circumvent this, Lucia (1985) suggested the following two quasi-Newton formulae for use in the context of a hybrid method for chemical process design:

$$\begin{aligned}\overline{A} = A &- \sum_{i=1}^{n} (z_i^T z_i)^+ (e_i^T A z) e_i z_i^T \\ &+ \sum_{i=1}^{n} (w_i^T s_i)^+ e_i^T [y\text{-}\overline{C}s\text{-}As] e_i w_i^T \\ &+ \sum_{i=1}^{n} (s_i^T z_i)(z_i^T z_i)^+ (w_i^T s_i)^+ (e_i^T A z) e_i w_i^T,\end{aligned} \tag{20}$$

and the tandem equations

$$\hat{A} = A + \sum_{i=1}^{n} (z_i^T s_i)^+ e_i^T [y\text{-}\overline{C}s\text{-}As] e_i s_i^T \tag{21}$$

and

$$\overline{A} = \hat{A} + \sum_{i=1}^{n} (z_i^T z_i)^+ e_i^T \hat{A} z e_i z_i^T, \tag{22}$$

where $z = \sum_{m=1}^{nk} z^m$, $w_i = s_i - (s_i^T z_i / z_i^T z_i) z_i$, and s_i and z_i are the vectors formed from s and z, respectively, such that they have zeros in every position that the ith row of $A(x)$ does. Observe that the quasi-Newton formula defined by Eq. (20) satisfies the usual secant and sparsity conditions as well as the zero-degree homogeneity equations for $\ln \gamma$ and H^E at each iteration of the calculations. In contrast, the update defined by Eqs. 21 and 22, which is simply two successive applications of the Schubert update, satisfies the sparsity and zero-degree homogeneity conditions at each iteration but only satisfies the secant condition in the limit. However, neither update can be applied to situations in which zero-degree homogeneity equations are required to be satisfied independently in several phases in equilibrium with each other, such as in liquid-liquid and vapor-liquid-liquid equilibria. Rather, the updating formulae given by Lucia apply to the case where only one phase is subject to the zero-degree homogeneity equations, which is often a good approximation in low-to-moderate pressure vapor-liquid equilibrium calculations where the vapor phase can often be assumed ideal.

Symmetric Case.

For symmetric problems, the analysis is much the same; only here we are concerned with building iterative representations of the approximated part of the Hessian matrix $H(x)$. We assume, as in the nonsymmetric case, that $H(x)$ can be written as in Eq. 10.

For the class of problems that calculate phase and/or chemical equilibrium points for nonideal mixtures by Gibbs free energy minimization at constant temperature and pressure, the approximated part of the Hessian is subject to the zero-degree homogeneity equations for $\ln \gamma$ and the Gibbs-Duhem equation.

As in the nonsymmetric case, when $A(x)$ is calculated analytically or by finite differences, it is thermodynamically consistent. On the other hand, when any symmetric quasi-Newton formula is used to build iterative representations of $A(x)$, these approximations do not satisfy either the zero-degree homogeneity of $\ln \gamma$ or the Gibbs-Duhem equation. More specifically, the approximation to $A(x)$ calculated by the Powell-symmetric-Broyden (PSB) formula (Powell, 1970), which is defined by

$$\overline{A} = A + \frac{(y\text{-}\overline{C}s\text{-}As)s^T + s(y\text{-}\overline{C}s\text{-}As)^T}{s^T s} - \frac{(y\text{-}\overline{C}s\text{-}As)^T s}{(s^T s)^2} s s^T \tag{23}$$

where s is defined as before and where y represents the difference in the gradient of the Gibbs function at successive iterates, is usually thermodynamically inconsistent. This is again because there are no constraints contained in the variational calculus problem from which the PSB formula is derived that force the zero-degree homogeneity of $\ln \gamma$ and the Gibbs-Duhem equation to be satisfied. Furthermore, the same is true of the Broyden-Fletcher-Goldfarb-Shanno (BFGS) update (Goldfarb, 1970), which is given by

$$\overline{A} = A + \frac{(y\text{-}\overline{C}s)(y\text{-}\overline{C}s)^T}{(y\text{-}\overline{C}s)^T s} - \frac{A s s^T A}{s^T A s}, \tag{24}$$

and the sparse PSB update (Toint, 1977; Marwil, 1978).

Thermodynamically Consistent Quasi-Newton Formulae

By a thermodynamically consistent quasi-Newton formula, we mean any quasi-Newton update that satisfies the zero-degree homogeneity and/or Gibbs-Duhem equations at each iteration of the calculations. Thus, we include in this class the updating formulae given by Eq. 20 and Eqs. 21 and 22.

Nonsymmetric Case.

There are, as stated earlier, situations where one might want to satisfy independent zero-degree homogeneity equations in several phases simultaneously (for example, in liquid-liquid, vapor-liquid-liquid or, for that matter, vapor-liquid equilibrium calculations). The quasi-Newton formula defined by Eqs. 20 through 22 cannot do this. To see this, note that these updates were derived to satisfy the condition

$$\bar{A}z = 0 \tag{25}$$

where $z = \Sigma_{m=1}^{n_k} z^m$ (Lucia, 1985). When several zero-degree homogeneity equations must be satisfied, the constraints are given by

$$\bar{A}z^m = 0, \quad m = 1,2,\ldots,n_k. \tag{26}$$

Clearly, Eq. 25 does not necessarily imply Eq. 26.

To extend the updating formula defined by Eq. 20 to the case of multiple zero-degree homogeneity conditions, we simply solve the variational problem

$$\min \|\bar{A} - A\|_F \tag{27}$$

subject to

$$\bar{A}s = y\text{-}\bar{C}s \tag{28}$$

$$\bar{A}z^m = 0, \quad m = 1,2,\ldots,n_k \tag{29}$$

and

$$\bar{a}_{ij} = 0, \quad (i,j)\epsilon I \tag{30}$$

where I is the set of index pairs defining the zero elements of the approximated part of the Jacobian and where it is assumed that the vector $z^1, z^2, \ldots, z^{n_k}$ are mutually orthogonal.

The solution to this problem is given by

$$\bar{A} = A + \sum_{i=1}^{n} (w_i^T s_i)^+ e_i^T[y\text{-}\bar{C}s\text{-}As]e_i w_i^T + \sum_{m=1}^{n_k} \sum_{i=1}^{n} \frac{e_i^T A z^m}{(z_i^m)^T(z_i^m)} \left[\left(\frac{s_i^T z_i^m}{w_i^T s_i}\right) e_i w_i^T - e_i (z_i^m)^T\right] \tag{31}$$

where

$$w_i = s_i - \sum_{m=1}^{n_k} \left(\frac{s_i^T z_i^m}{(z_i^m)^T(z_i^m)}\right) z_i^m.$$

The extended form of the updating formula defined by Eqs. 21 and 22, which is an iterated projection (Dennis and Schnabel, 1979), is given by

$$\hat{A} = A + \sum_{i=1}^{n} (s_i^T s_i)^+ e_i^T[y\text{-}\bar{C}s\text{-}As]e_i s_i^T \tag{32}$$

and

$$\bar{A} = \hat{A} + \sum_{m=1}^{n_k} \sum_{i=1}^{n} [(z_i^m)^T(z_i^m)]^+ e_i^T \hat{A} z^m e_i (z_i^m)^T. \tag{33}$$

Observe that the updating formula defined by Eqs. 32 and 33 corresponds to $n_k + 1$ Schubert updates and will only satisfy the secant condition in the limit, unlike the one defined by Eq. 31, which satisfies the secant condition at each iteration. Both updates, however, satisfy all zero-degree homogeneity and sparsity conditions at each iteration.

Symmetric Case.

In the symmetric case it is relatively easy to satisfy any number of independent zero-degree homogeneity and Gibbs-Duhem equations at each iteration of the calculations. Furthermore, because of the block diagonal structure of minimum Gibbs free energy problems, it is more convenient to update the approximated part of the Hessian blockwise. For this purpose we suggest the updating formula defined by

$$\bar{A} = \sum_{m=1}^{n_k} \bar{A}_m, \tag{34}$$

where

$$\hat{A}_m = A_m + \frac{(y_m - \bar{C}_m s_m - A_m s_m)s_m^T + s_m(y_m - \bar{C}_m s_m - A_m s_m)^T}{s_m^T s_m} - \frac{(y_m - \bar{C}_m s_m - A_m s_m)^T s_m}{(s_m^T s_m)^2} s_m s_m^T \tag{35}$$

$$\bar{A}_m = \left[I - \frac{z_m z_m^T}{z_m^T z_m}\right] \hat{A}_m \left[I - \frac{z_m z_m^T}{z_m^T z_m}\right]. \tag{36}$$

Also, $\bar{C}_m$ is the associated diagonal block of the computed part of the Hessian, $z_m = z^m$ (which is defined by Eq. 16) and s_m and y_m are the appropriate projections of s and y, respectively, that are associated with the mth phase.

Observe that the update defined by Eqs. 34 through 36 is simply comprised of successive PSB corrections to each diagonal block of the approximated part of the Hessian. Because of Eq. 36, $\bar{A}$ is necessarily thermodynamically consistent at each iteration of the calculations.

NUMERICAL RESULTS

In this section, we present some numerical results for the thermodynamically consistent quasi-Newton formulae given in the previous section. In particular, we solve various single-stage, isothermal flash problems involving different multicomponent mixtures. We do this for both the nonsymmetric and symmetric cases and in each case compare the results to those obtained when $A(x)$ is approximated by finite-difference and by existing quasi-Newton formulae. Finally, in all examples we used the UNIQUAC equation (Prausnitz et al., 1980) to model the liquid phase activity coefficient, while the vapor phase was assumed to be ideal. Thus, in all cases, the dimension of the null space, n_k, was one.

Nonsymmetric Case.

The example problems here were as follows:

1. A feed mixture of 1.2404×10^{-2} kmol/s of methanol, 7.1404×10^{-3} kmol/s of n-heptane, and 1.3215×10^{-2} kmol/s of benzene was flashed at 337.36 K and 1.013×10^5 Pa.
2. A mixture of ethanol, acetone, and water was flashed at 341.46 K and 1.013×10^5 Pa. The feed component flow rates for ethanol, acetone, and water were 2.300×10^{-3} kmol/s, 1.8358×10^{-2} kmol/s, and 1.2101×10^{-2} kmol/s, respectively.
3. A mixture of 1.197×10^{-6} kmol/s of ethanol, 1.2536×10^{-2} kmol/s of n-propanol, 1.1946×10^{-2} kmol/s of cyclohexane, and 8.265×10^{-3} kmol/s of benzene was flashed at 382.87 K and 3.039×10^5 Pa.

We solved each of these example problems to an accuracy of $\|f\|_2 \leq 10^{-5}$ by Newton's method with finite difference approximations to $A(x)$, the hybrid approach of Lucia and Macchietto (1983), the two thermodynamically consistent hybrid (TCH1 and TCH2) methods, in which $A(x)$ was defined by Eq. 31 and Eqs. 32 and 33, respectively, and Schubert's method. The results are shown in Table 1.

For the examples presented here, as well as others, the calculations were initiated by setting

$$\ell_i = (1 - \alpha)\left(\sum_{j=1}^{n_c} f_j\right) x_i$$

TABLE 1. RESULTS FOR ISOTHERMAL FLASH CALCULATIONS

Example	Newton	Hybrid	TCH1	TCH2	Schubert
1	7(21)	32(35)	14(17)	10(13)	20(23)
2	9(27)	14(17)	15(18)	10(13)	23(26)
3	6(24)	27(31)	9(13)	15(18)	72(75)

Figure pairs in each column are: Iterations (Rigorous thermodynamic calculations). In all cases $B_o = J(x_o)$.

and

$$v_i = \alpha\left(\sum_{j=1}^{n_c} f_j\right) y_i,$$

where x_i and y_i represent the incipient liquid and vapor compositions for dew point and bubble point calculations performed on the feed. The vaporization fraction, α, was set equal to 0.5 for all examples and the dogleg strategy of Powell (1970) was used as a convergence safeguard. We also investigated other initialization procedures and, barring a few exceptions, the results remained qualitatively the same.

It is important to note that when a liquid component flow rate tended to zero, which happened reasonably frequently, the quasi-Newton correction predicted by either Eq. 31 or Eqs. 32 and 33 was usually inordinately large. In part, we believe that this is due to the fact that the scalar quantity $[(z_i^m)^T(z_i^m)]$ can become small and because the elements in the matrix A corresponding to this zero component flow rate can take on any value and still be consistent with the zero-degree homogeneity condition. To avoid this we reset the corresponding row of A to zero, which is what Newton's method does under these circumstances.

With regard to the numerical results, perhaps the most we can say at this point of our study is that nonsymmetric quasi-Newton formulae that account for the zero-degree homogeneity of partial molar excess properties can provide improvements in computational efficiency. We have observed this in many cases; however, it is not uniformly true. Sometimes these thermodynamically consistent quasi-Newton formulae do not provide any improvements. They take the same number of iterations as Schubert's formula to reach the desired accuracy. Rarely, they require more iterations. We believe this is due to the fact that we have not, as yet, made use of the Gibbs-Duhem equation in the nonsymmetric case.

Symmetric Case.

For the symmetric case, which is a very natural setting for investigating the use of both the zero-degree homogeneity and Gibbs-Duhem equations in quasi-Newton approximations, the results are very encouraging. To illustrate this, we present numerical results for the following isothermal flash problems solved by minimizing the Gibbs free energy of the system:

1. A feed mixture of 1×10^{-2} kmol/s of ethanol and 1×10^{-2} kmol/s of n-hexane was flashed at 335 K and 1.013×10^5 Pa.
2. An equimolar mixture of methanol, ethanol, acetone, and water was flashed at 343 K and 1.013×10^5 Pa.
3. A mixture of carbon tetrachloride, chloroform, methanol, ethanol, acetone, and benzene was flashed at 335 K and 1.013×10^5 Pa. The feed component flow rates for carbon tetrachloride, chloroform, and benzene were each 1.260×10^{-2} kmol/s, while those for methanol, ethanol, and acetone were each 6.300×10^{-2} kmol/s.

We solved these example problems by successive quadratic programming (Powell, 1978) using various methods for approximating both the entire Hessian of the Gibbs function and the approximated part of the Hessian. In particular, we approximated the complete Hessian by the full BFGS, full PSB, and sparse PSB updates, and approximated $A(x)$ by finite difference, the sparse PSB formula, and the thermodynamically consistent updating formula (TC update) defined by Eqs. 34–36. For each problem, the calculations were terminated when the Kuhn-Tucker conditions were satisfied to an accuracy of 10^{-5}. The results are shown in Tables 2, 3 and 4.

The three starting points were chosen as follows:

1. $\ell_i = (1-\alpha)(\sum_{j=1}^{n_c} f_j)x_i$ and $v_i = \alpha(\sum_{j=1}^{n_c} f_j)x_i$ where x_i represents the composition of the liquid phase in equilibrium with vapor feed at the specified T and p; $\alpha = 0.55, 0.50, 0.60$ for examples 1, 2, and 3, respectively.
2. Same as initialization 1 with x_i replaced by y_i, where y_i is the composition of the vapor phase in equilibrium with liquid feed at the specified T and p; $\alpha = 0.75, 0.50, 0.55$ for examples 1, 2, and 3, respectively.
3. $\ell_i = v_i = f_i/2$.

We investigated other starting points. Again, the results were qualitatively unchanged.

For the symmetric case, the results indicate that a hybrid approach with the approximated part of the Hessian built by a thermodynamically consistent quasi-Newton formula can compete favorably with Newton's method and the full BFGS formula on phase equilibrium problems solved by Gibbs free energy minimization. This approach is clearly superior to both quasi-Newton methods for constrained optimization when either the full or sparse PSB update is used to build the Hessians of the Gibbs function and a hybrid approach that uses only secant information to update the approximated part of the Hessian. (Also, BFGS updates to $A(x)$ did not work very well, most likely because the approximated part of the Hessian is not sign-definite).

We attribute some of the good performance of the hybrid approach with the thermodynamically consistent quasi-Newton update to the fact that $C(x)$ is positive-semidefinite on the mass balance constraint manifold. This has a stabilizing effect on the calculations. However, it cannot account for all of the good be-

TABLE 2. RESULTS FOR MINIMUM GIBBS FREE ENERGY CALCULATIONS FOR MIXTURES OF ETHANOL, n-HEXANE

	Method*					
	Approximation to $H(x)$			Approximation to $A(x)$		
Starting Point	Full BFGS	Full PSB	Sparse PSB	Finite-Difference	Sparse PSB	TC Update
1	22(22)	F	F	7(21)	24(24)	15(15)
2	9(9)	16(16)	16(16)	6(18)	10(10)	8(8)
3	23(23)	F	F	8(24)	F	20(20)

* All methods, except finite-difference, used $H(x_0) = C(x_0)$.

Figure pairs in Approximation columns are: Function and gradient calls (Rigorous thermodynamic calculations).

F = failure.

TABLE 3. RESULTS FOR MINIMUM GIBBS FREE ENERGY CALCULATIONS FOR MIXTURE OF METHANOL, ETHANOL, ACETONE, AND WATER

Starting Point	Method* Approximation to $H(x)$ Full BFGS	Full PSB	Sparse PSB	Approximation to $A(x)$ Finite-Difference	Sparse PSB	TC Update
1	12(12)	12(12)	12(12)	6(30)	F	12(12)
2	10(10)	18(18)	18(18)	6(30)	8(8)	8(8)
3	24(24)	10(10)	10(10)	6(30)	6(6)	6(6)

* All methods, except finite-difference, used $H(x_0) = C(x_0)$.
Figure pairs in Approximation columns are: Function and gradient calls (Rigorous thermodynamic calculations).
F = failure.

havior. When $A(x)$ is updated by quasi-Newton methods that only use secant information, the numerical performance can deteriorate. Because of this and the fact that the thermodynamically consistent update performs rather well, we believe that the second derivative information that comes from the zero-degree homogeneity and Gibbs-Duhem equations can be of value in the equation solving task.

CONCLUDING REMARKS

In summary, we have argued that when Newton's method is applied to chemical process problems whose models include the chemical and/or phase equilibria of a nonideal mixture, it generates matrix approximations that satisfy various thermodynamic constraints such as the zero-degree homogeneity of $\ln \gamma$ and H^E and the Gibbs-Duhem equation. All existing quasi-Newton formulae, on the other hand, do not. To circumvent this, we presented several new thermodynamically consistent quasi-Newton formulae. In the nonsymmetric case, we suggested two new quasi-Newton formulae that satisfy any number of linearly independent zero-degree homogeneity equations simultaneously. These updates were tested on single-stage, isothermal flash calculations involving a number of different multicomponent mixtures. The numerical results show that these thermodynamically consistent quasi-Newton formulae can give improvements in computational efficiency when compared to existing quasi-Newton methods. Unfortunately, the results were not observed to be uniformly in favor of the thermodynamically consistent updates.

For the symmetric case, we studied the effects of satisfying both the zero-degree homogeneity equations of $\ln \gamma$ and the Gibbs-Duhem equation. Single-stage, isothermal flash problems were again solved for several multicomponent mixtures, this time by minimizing the total Gibbs free energy of the system subject to the mass balance constraints. Here the numerical results were very encouraging and showed that a symmetric and thermodynamically consistent quasi-Newton formula, used in the context of a hybrid optimization algorithm, can compete favorably with the BFGS formula and Newton's method.

We believe that the results for the nonsymmetric case can be improved, perhaps markedly, by developing nonsymmetric quasi-Newton formulae that satisfy both the zero-degree homogeneity and Gibbs-Duhem equations simultaneously. The results for the minimum Gibbs free energy calculations seem to support this. However, to do this, we must split the Jacobian matrix differently and use the concept of partition symmetry. In particular, if we define

$$J(x) = C_1(x) + C_2(x)A(x),$$

where $C_2(x)$ is a diagonal matrix and easily computed in analytical form, then for problems involving chemical and/or phase equilibria $A(x)$ turns out to be partition-symmetric. To see this, consider the illustrative example. If $C_2(x)$ is defined by

$$C_2(x) = \begin{bmatrix} \dfrac{\ell_1^1 K_1}{\sum_j \ell_j^1} & & & & & \\ & \cdot & & & & \\ & & \cdot & & & \\ & & & \dfrac{\ell_{nc}^1 K_{n_c}}{\sum_j \ell_j^1} & & \\ & & & & 1 & \\ & & & & & \cdot \\ & & & & & & \cdot \\ & & & & & & & \cdot \\ & & & & & & & & 1 \end{bmatrix} \tag{37}$$

TABLE 4. RESULTS FOR MIMINUM GIBBS FREE ENERGY CALCULATIONS FOR MIXTURE OF CARBON TETRACHLORIDE, CHLOROFORM, METHANOL, ETHANOL, ACETONE, AND BENZENE

Starting Point	Method* Approximation to $H(x)$ Full BFGS	Full PSB	Sparse PSB	Approximation to $A(x)$ Finite-Difference	Sparse PSB	TC Update
1	10(10)	12(12)	F	7(49)	14(14)	12(12)
2	23(23)	23(23)	23(23)	12(84)	28(28)	18(18)
3	25(25)	F	F	F	12(12)	11(11)

* All methods, except finite-difference, used $H(x_0) = C(x_0)$.
Figures in Approximation columns are: Function and gradient calls (Rigorous thermodynamic calculations).
F = failure.

then

$$A(x) = \begin{bmatrix} \left(\frac{\partial \ln \gamma_1^1}{\partial \ell_1^1}\right) & \cdots & \left(\frac{\partial \ln \gamma_1^1}{\partial \ell_{n_c}^1}\right) & \left(\frac{\partial \ln \gamma_1^2}{\partial \ell_1^2}\right) & & \left(\frac{\partial \ln \gamma_1^2}{\partial \ell_{n_c}^2}\right) \\ \vdots & & \vdots & \vdots & & \\ \left(\frac{\partial \ln \gamma_{n_c}^1}{\partial \ell_1^1}\right) & \cdots & \left(\frac{\partial \ln \gamma_{n_c}^1}{\partial \ell_{n_c}^1}\right) & \left(\frac{\partial \ln \gamma_{n_c}^2}{\partial \ell_1^2}\right) & \cdots & \left(\frac{\partial \ln \gamma_{n_c}^2}{\partial \ell_{n_c}^2}\right) \\ 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix} \tag{38}$$

Now in order to satisfy both the zero-degree homogeneity and Gibbs-Duhem equations in each phase, we can construct an iterated projection between the space of secant, partition-symmetric matrices, and the space of thermodynamically consistent, partition-symmetric matrices.

Finally, we note that the extension of this idea of thermodynamically consistent quasi-Newton formulae for problems involving nonideal mixtures carries over to the more general setting of chemical process optimization since partial derivatives of homogeneous functions are themselves homogeneous. More specifically, in this context we are interested in approximating, among other things, matrices of $(\partial^2 \ln \gamma_i^m / \partial \ell_j^m \partial \ell_k^m)$. Because the functions $(\partial \ln \gamma_i^m / \partial \ell_j^m)$ are homogeneous of degree -1 in mole numbers, it follows that the thermodynamic constraints are of the form

$$\sum_{k=1}^{n_c} \left(\frac{\partial^2 \ln \gamma_i^m}{\partial \ell_j^m \partial \ell_k^m}\right) \ell_k^m = -\left(\frac{\partial \ln \gamma_i^m}{\partial \ell_j^m}\right), \; i = 1,2,\ldots,n_c \qquad j = 1,2,\ldots,n_c \qquad m = 1,2,\ldots,n_p. \tag{39}$$

Observe that, although more complicated, these constraints as well as the vectors z^m are the same form as those studied in this paper. The only term that is changed is the righthand side of Euler's theorem (i.e., Eq. 12).

ACKNOWLEDGMENT

This material is partially based upon work supported by the National Science Foundation under Grant No. CPE-8204005 and Grant No. CPE-8216857.

NOTATION

A = approximated part of the Jacobian or Hessian matrix
B = Jacobian approximation
C = computed part of the Jacobian or Hessian matrix
e_i = ith unit vector
f, f_i^m = vector function of model equations, molal flow rate for the ith component of the feed in the mth phase
G, G^0, G^I, G^E = Gibbs free energy, standard state Gibbs free energy, ideal solution Gibbs free energy, excess Gibbs free energy
H, H^E = Hessian matrix, excess enthalpy
J = Jacobian matrix of f
K_i = distribution ratio for the ith component
s = change in the independent variables
T = temperature
w = vector orthogonal to the thermodynamic constraints
x, x_k = vector of unknown variables, at the kth iteration
y = change in function or gradient from one iteration to the next
z, z^m = vector contained in the iterative subspace defined by the thermodynamic constraints, mth basis vector for the thermodynamic constraint manifold

Script

ℓ_i = liquid phase molal flow rate for the ith component

Greek Letters

γ_i = ith component liquid phase activity coefficient

Subscripts

i,j = component number
k = iteration counter

Superscripts

E = excess enthalpy
I = ideal solution
m = phase number or local coordinate number
T = transpose
$+$ = generalized inverse
o = standard state

LITERATURE CITED

Broyden, C. G., "A Class of Methods for Solving Nonlinear Simultaneous Equations," *Math. Comp.*, **19**, 577 (1965).

———, "The Convergence of an Algorithm for Solving Sparse Nonlinear Systems," *Math. Comp.*, **25**, 285 (1971).

Dennis, J. E., and R. B. Schnabel, "Least-Change Secant Updates for Quasi-Newton Methods," *SIAM Rev.*, **21**, 443 (1979).

Goldfarb, D., "A Family of Variable Metric Methods Derived by Variation Means," *Math. Comp.*, **24**, 23 (1970).

Lucia, A., "Partial Molar Excess Properties, Null Spaces and a New Update for the Hybrid Method of Chemical Process Design," *AIChE J.*, **31**(4), 558 (1985).

Lucia, A., and S. Macchietto, "New Approach to Approximation of Quantities Involving Physical Properties Derivatives in Equation-Oriented Process Design," *AIChE J.*, **29**, 705 (1983).

Marwil, E. S., "Exploiting Sparsity in Newton-like Methods," Ph.D. Thesis, Cornell Univ., Ithaca, NY (1978).

Powell, M. J. D., "A New Algorithm for Unconstrained Optimization," *Nonlinear Programming* J. B. Rosen, O. L. Mangasarian, and K. B. Ritter, Eds., Academic Press, New York, 31 (1970).

———, "A Fast Algorithm for Nonlinearly Constrained Optimization Calculations," *Numerical Analysis*, G. A. Watson, Ed., *Lecture Notes in Mathematics*, Springer, New York, 630 (1978).

Prausnitz, J., et al., *Computer Calculations for Multicomponent Vapor-Liquid and Liquid-Liquid Equilibria*, Prentice-Hall, Englewood Cliffs, NJ (1980).

Schubert, L. K., "Modification of a Quasi-Newton Method for Nonlinear Equations with a Sparse Jacobian," *Math. Comp.*, **24**, 27 (1970).

Toint, Ph. L., "On Sparse and Symmetric Matrix Updating Subject to a Linear Equation," *Math. Comp.*, **31**, 954 (1977).

Manuscript received Jan. 25, 1984; revision received Nov. 14 and accepted Nov. 24, 1984.