A LINEAR PROGRAMMING APPROACH TO THE CHEMICAL EQUILIBRIUM PROBLEM

George Dantzig, Selmer Johnson,

Wayne White

P-1060

Revised April 21, 1958

The RAND Corporation

1700 MAIN ST • SANTA MONICA • CALIFORNIA

SUMMARY

The well known chemical equilibrium problem is expressed in the form of minimizing the free energy of a mixture in order to compute the chemical composition at equilibrium.

By piece-wise linear approximations to the free energy function, the problem becomes a linear program which can be solved by a standard code on a computing machine. Successive approximations give any degree of accuracy.

A LINEAR PROGRAMMING APPROACH TO THE CHEMICAL EQUILIBRIUM PROBLEM

1. INTRODUCTION

The extension of linear programming methods into the non-linear area is recognized as one of the outstanding areas of research today. One way to treat a general convex objective function is to locally linearize it by taking partial derivatives. However the authors believe that researchers in the linear programming field are interested in ideas that convert a "near" separable convex function into a completely separable case where more efficient methods can then take over [1], [2].

Although the particular problem resolved has a non-management application, it should also be of interest to a rather large audience scheduling for the Petroleum Industry. These are for the most part chemical engineers who are familiar with the application to the Chemical Equilibrium Problem and who can take advantage of the solution for their respective companies. They will find the companion paper [3] discusses more of the chemical background to the problem and gives an alternative procedure. The present paper stresses the mathematical development of the linear programming approach to this problem.

The determination of the chemical composition of a complex mixture under chemical equilibrium conditions is a classic problem. There have been many computing techniques proposed and the constant appearance of new ones attests that none are entirely satisfactory. In our second paper [3], entitled "Chemical Equilibrium in Complex Mixtures", the formulation of two methods discovered in the course of our researches were reviewed, one of which is a steepest descent based on a quadratic fit to the free energy function to be minimized, and the other which reduces it to a linear programming problem. The purpose of this paper is to give a complete account of the latter. Our purpose is to show that there is an elegant way to transform the free energy function into a convex separable function which permits convenient piece-wise linear approximation and consequent solution by linear programming. Any desired degree of accuracy can be reached by successively improving the approximation.

2. THE PROBLEM

We consider an equilibrium mixture containing m different atom types. While in theory these will combine into all chemically possible molecular species, in practice only standard types are considered including the monotonic types

which are known to occur in measurable amounts.

Let b_i = the number of atomic weights of species i present in the mixture.

x_j = the number of moles of molecular species j present in the mixture where

(1) $$x_j \geq 0, \quad j = 1, 2, \ldots n.$$

$\bar{x}$ = the total number of moles of gas in the mixture, i.e.

(2) $$\bar{x} = \sum x_j.$$

a_{ij} = the number of atoms of species i in a molecule of species j.

Then the mass balance equations are

(3) $$\sum_{j=1}^{n} a_{ij} x_j = b_i \quad \text{for } i = 1, 2, \ldots m.$$

The determination of the equilibrium composition of a gaseous mixture is equivalent to the determination of the values of the mole numbers x_j that obey constraint (3) and minimize the total free energy of the mixture given by

(4) $$F(x_1, \ldots, x_n) = \sum_{j=1}^{n} c_j x_j + \sum_{j=1}^{n} x_j \ln (x_j/\bar{x})$$
$$= \sum_{j=1}^{n} c_j x_j + \bar{x} \sum_{j=1}^{n} (x_j/\bar{x}) \ln (x_j/\bar{x}).$$

which can be shown to be a convex function.* The values c_j

*[3] contains a direct proof; alternatively the fact that the chemical equilibrium problem can be reduced to a linear programming problem to any desired degree of accuracy also proves convexity.

are the modified Gibbs free energy function F°/RT of the atomic species at a given temperature plus the natural logarithm of the pressure in atmospheres.

Our problem is to minimize (4) subject to the linear equality and inequality constraints (1), (2), (3).

In order to apply linear programming, we make a piecewise linear approximation to each of the terms $(x_j/\bar{x})\ \ln\ (x_j/\bar{x})$ that appear in (4). If we set $\alpha = x_j/\bar{x}$ and $\beta = \alpha \ln \alpha$ then we shall replace each such curve by a broken line function $\beta = \beta(\alpha)$ such as the one below.

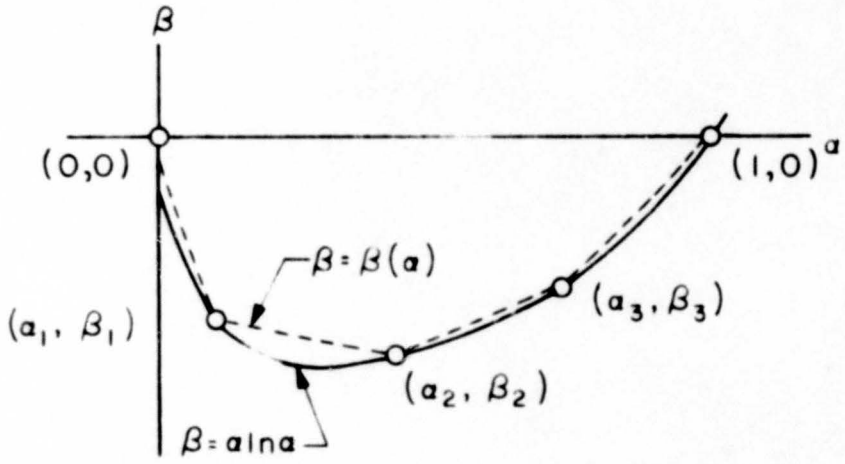

Fig. 1

The k points where the two curves agree are denoted by (α_i, β_i) where $i = 1, 2, \ldots, k$.

Let us imagine that the values of $\bar{x}$ and x_1 are fixed for the moment and that the values of $x_{11} \geq 0,\ x_{12} \geq 0, \ldots, x_{1n} \geq 0$ are chosen so to satisfy

(5) $$\bar{x} = x_{11} + x_{12} + \ldots + x_{1k}$$

(6) $$x_1 = \alpha_1 x_{11} + \alpha_2 x_{12} + \ldots + \alpha_k x_{1k}$$

and to minimize z_1 where

(7) $$z_1 = \beta_1 x_{11} + \beta_2 x_{12} + \ldots + \beta_k x_{1k}.$$

We shall now prove that $\beta = \beta(\alpha)$ is convex, that

(8) $$\text{Min } z_1 = \bar{x}\,\beta(x_1/\bar{x}) \doteq x_1 \ln(x_1/\bar{x}).$$

Proof: It is clear that $\beta = \beta(\alpha)$ will be convex if $\beta = \alpha\ln\alpha$ is convex and this follows by noting that

(9) $$\frac{d\beta}{d\alpha} = 1 + \ln \alpha$$

is monotonically increasing. Next let us substitute

(10) $$x_{11} = \lambda_1\bar{x}, \quad x_{12} = \lambda_2\bar{x}, \ \ldots, \ x_{1k} = \lambda_k\bar{x}$$

then (5), (6), and (7) may be rewritten

(11) $$1 = \lambda_1 + \lambda_2 + \ldots + \lambda_k$$

(12) $$(x_1/\bar{x}) = \alpha_1\lambda_1 + \alpha_2\lambda_2 + \ldots + \alpha_k\lambda_k$$

(13) $$(z_1/\bar{x}) = \beta_1\lambda_1 + \beta_2\lambda_2 + \ldots + \beta_k\lambda_k$$

and the problem is equivalent to finding $\lambda_1 \geq 0$, $\lambda_2 \geq 0, \ldots, \lambda_k \geq 0$ satisfying (11), (12) for fixed x_1 and $\bar{x}$ minimizing $(z_1/\bar{x})$. If we interpret $\lambda_1 \geq 0$ as the weights assigned to the points (α_1,β_1), then $\sum_1^k \alpha_i\lambda_i$ and $\sum_1^k \beta_i\lambda_i$ are coordinates of the <u>center of gravity</u> of the points. Hence we are seeking weights to assign to the points such that the abscissa of the center of gravity is $(x_1/\bar{x})$, see (12), and the ordinate $(z_1/\bar{x})$, see (13), is as small as possible. Obviously this smallest value, for any convex curve $\beta = \beta(\alpha)$, is $\beta(x_1/\bar{x})$, and this value is ob-

tained for a broken line function by assigning $\lambda_i = 0$ to all points except the two points on either side of $(x_1/\bar{x})$ and weighting up these two points appropriately.

To solve the chemical equilibrium problem by linear programming, consider the problem of finding $x_j \geq 0$, $x_{jl} \geq 0$ and minimum z satisfying

$$
\begin{array}{l}
a_{11}x_1 + \ldots + a_{1n}x_n = b_1 \\
a_{21}x_1 + \ldots + a_{2n}x_n = b_2 \\
\hdashline
a_{m1}x_1 + \ldots + a_{mn}x_n = b_m \\
\hline
x_1 + \ldots + x_n - \bar{x} = 0 \\
\hline
-\bar{x} + \sum_{l=1}^{k} x_{1l} = 0 \\
-\bar{x} + \sum_{l=1}^{k} x_{2l} = 0 \\
\hdashline
-\bar{x} + \sum_{l=1}^{k} x_{nl} = 0 \\
\hline
-x_1 + \sum_{l=1}^{k} \alpha_l x_{1l} = 0 \\
-x_2 + \sum_{l=1}^{k} \alpha_l x_{2l} = 0 \\
\hdashline
-x_n + \sum_{l=1}^{k} \alpha_l x_{nl} = 0 \\
\hline
c_1x_1 + \ldots + c_nx_n + \sum_{l=1}^{k} \beta_l x_{1l} + \sum_{l=1}^{k} \beta_l x_{2l} + \ldots + \sum_{l=1}^{k} \beta_l x_{nl} = z
\end{array}
\qquad (14)
$$

It will now be easy to see that the values $x_j = x_j^o$, that form part of the optimal solution to (14), give the optimal solution to (1), (2), (3) if (4) is replaced by the approximation

$$(15) \qquad F + \bar{F} = \sum_{j=1}^{n} c_j x_j + \sum_{j=1}^{n} \bar{x}\, \beta(x_j/\bar{x}).$$

Proof: Consider an optimal solution to (14) and let the values of $x_j = x_j^o$. For these fixed values of x_j, it is clear from the structure of (14) that the optimal choice of values $x_{11}, x_{12}, \ldots, x_{1k}$ for example must satisfy (5) and (6) and minimize z_1 given by (7). Moreover by (8) the value of $z_1 = \bar{x}^o \beta(x_j^o/\bar{x}^o)$ and

$$\text{Min } z = \sum_{j=1}^{n} c_j x_j^o + \sum_{j=1}^{n} x^o \beta(x_j^o/\bar{x}^o) = \bar{F}^o$$

Conversely, consider an optimal solution to (1), (2), (3), (15) with values $x_j = x_j^*$, we can always find values $x_{j\ell} = x_{j\ell}^* \geq 0$ such that

$$\bar{x}^* = \sum_{\ell=1}^{k} x_{j\ell}^*$$

$$x_j^* = \sum_{\ell=1}^{k} \alpha_\ell x_{j\ell}^*$$

$$\beta(x_j^*/\bar{x}^*) = \sum_{\ell=1}^{k} \beta_\ell x_{j\ell}^* .$$

As we saw in the discussion following (13), these x_j^*, $x_{j\ell}^*$ are a solution to (14) whose value $z = z^*$ is the same as the value Min $\bar{F}$ obtained by substituting $x_j = x_j^*$ in (15); hence Min $\bar{F} = z^*$. From the relations Min $\bar{F} = z^* \geq$ Min z and Min $\bar{F} \leq F^o$ = Min z follows Min $\bar{F}$ = Min z and our proof is complete.

This approach has several advantages over previous methods.

(1) It uses a standardized code of the simplex method of linear programming.

(2) Only one curve $\alpha \log \alpha$ for $0 \leq \alpha \leq 1$, has to be approximated and this can be done as accurately as desired since the number of equations is not increased.

(3) There is no necessity of deciding in advance which are the so called major components and which are the minor ones as in the usual case for other methods.

The evaluation of $\alpha_\ell \log \alpha_\ell$ can be included in the coding so that the successive approximations could be carried out automatically to any degree of accuracy.

In order not to tax the memory of the computing machine and also to reduce the number of computations of $\beta_\ell = \alpha_\ell \log \alpha_\ell$, the following system of "screening" should be added to the code.

First compute $\beta_\ell = \alpha_\ell \log \alpha_\ell$ at $\alpha_\ell = 0, .5, 1$ with grid size .5. Then after the first linear program approximation has been solved, for each term $(x_j/\bar{x}) \ln (x_j/\bar{x})$, halve the grid size and compute only those new values which are adjacent to the current value of $\alpha = x_j/\bar{x}$. Thus on the second piecewise approximation, if some $x_j = 0$, compute $\beta_\ell = \alpha_\ell \log \alpha_\ell$ at $\alpha_\ell = .25$, discarding the value at $\alpha_\ell = 1$; if $x_j/\bar{x} = .5$, compute $\alpha_\ell \log \alpha_\ell$ at .25 and .75 and discard the values at 0 and 1. If $x_j/\bar{x}$ is a weighted average of two grid points 0 and .5, then include a grid value of α_ℓ at .25 and discard the value at 1, etc. In this way for each new piecewise approximation we have at most three values of α_ℓ and $\alpha_\ell \log \alpha_\ell$ for each $j = 1, 2, \ldots, n$ such that the range of α_j

values is halved each time. It has been observed empirically that successive values of $x_j/\bar{x}$ have stayed within the ranges set up for all previous cycles i.e. the accuracy has doubled with each successive approximation.

After a sufficient number of cycles have been carried out, any x_j which is still at zero level in the approximate solution could be adjusted to a trace level by standard methods used to get the order of magnitude of such components when the non-trace level components have been determined very closely.

When there are a large number of problems to be run, it is recommended that the simplex code be modified to generate internally in the machine the columns associated with the variables x_{jl}. If this is done it is probably best to determine the best choice of α_l, β_l analytically for each j rather than by the above grid technique. The present RAND code is being modified along these lines.

There is also another improvement possible which takes further advantage of the structure by partitioning the basis into two parts—the first part associated with the first m rows and the other associated with the remaining rows. This results in an essential $m \times m$ subbasis whose inverse is needed—the remainder of the computation is then carried out implicitly.

BIBLIOGRAPHY

1. Dantzig, G.B., Recent Advances in Linear Programming Management Science, January 1956, pp. 131-143.

2. Charnes, A. and Lamke, C.E., "Minimization of Non-Linear Separable Convex Functions." Naval Research Logistics Quarterly VI, pp. 301-312 (1954).

3. White, W.B., Johnson, S.M., and Dantzig, G.B., Chemical Equilibrium in Complex Mixtures, The RAND Corporation Paper P-1059, April 15, 1957 (To appear in the Journal of Chemical Physics).