AD706020

COMPUTING EQUILIBRIUM COMPOSITIONS OF IDEAL CHEMICAL SYSTEMS

BY

JAMES H. BIGELOW

TECHNICAL REPORT NO. 70-4
MARCH 1970

DDC

This document has been approved for public release and sale; its distribution is unlimited.

OPERATIONS RESEARCH HOUSE

Stanford University
CALIFORNIA

16

COMPUTING EQUILIBRIUM COMPOSITIONS OF IDEAL CHEMICAL SYSTEMS

by

JAMES H. BIGELOW

TECHNICAL REPORT NO. 70-4

MARCH 1970

Department of Operations Research
Stanford University
Stanford, California

Research and reproduction of this report was partially supported by Office of Naval Research, Contract N-00014-67-A-0112-0011; U.S. Atomic Energy Commission, Contract AT(04-3) 326 PA#18; National Science Foundation, Grant GP 9329; U.S. Army Research Office, Contract DAHC04-67-0028; and National Institute of Health, Grant GM 14789-02.

Reproduction in whole or in part is permitted for any purpose of the United States Government. THis document has been approved for public release and sale; its distribution is unlimited.

CONTENTS

Section

I. INTRODUCTION 1

II. PRIMAL METHODS 5
1. The Augmented Form 5
2. The Reduced Form 8
3. The General Linear-Logarithmic Problem . 10
4. The Linear Approximation Method 11
5. The Quadratic Approximation Method . . . 14
6. The Basic Algorithm 16

III. A CLOSER LOOK AT THE REDUCED QUADRATIC METHODS 22
1. The Reduced Generalized Linear Program . 22
2. Lower Bounds 30
3. The Modified RQM Algorithm 35
4. The Other Methods 42

IV. DUAL METHODS 43
1. The Dual Augmented Linear Method 44
2. The Dual Formulation 47
3. Solving the Dual 54
4. Comparison of the Dual Methods 57
5. Degeneracy and the Dual Methods 60
6. Convergence of Dual Methods 62

Appendix. 63

REFERENCES. 70

PRECEDING PAGE BLANK

COMPUTING EQUILIBRIUM COMPOSITIONS OF IDEAL CHEMICAL SYSTEMS

James H. Bigelow*

The Rand Corporation, Santa Monica, California

I. INTRODUCTION

A single- or multi-phase chemical equilibrium problem may be expressed as a nonlinear programming problem. Thus to find the equilibrium composition of a chemical system one need only minimize a particular nonlinear function (the free energy) of composition subject to certain linear constraints (the mass-balance laws). The free energy is defined on the nonnegative orthant of n-space, where it is continuous, convex and homogeneous of degree one. In the interior of its domain it is infinitely differentiable; but at the boundary, the directional derivative may become infinite.

In this paper, the phrase "chemical equilibrium problem" refers only to a problem with a particular mathematical form. Problems of this form arise in many situations not classically denoted chemical equilibrium problems. For example, the dual to a geometric programming problem [1, 2] has this form. Also, steady-state problems, many of which arise naturally in the chemical laboratory, or in industry, can often be represented in this form.

The chemical equilibrium problem, then, is the problem

PRECEDING PAGE BLANK

*Any views expressed in this paper are those of the author. They should not be interpreted as reflecting the views of The Rand Corporation or the official opinion or policy of any of its governmental or private research sponsors. Papers are reproduced by The Rand Corporation as a courtesy to members of its staff.

of minimizing a function $F(x_1, x_2, \ldots, x_n)$, defined below, subject to the linear constraints

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

and the inequalities,

$$\text{(I.2)} \qquad x_j \geq 0, \quad j = 1, 2, \ldots, n.$$

The a_{ij} and b_i are given real constants. We assume that the m equations (I.1) are linearly independent and, so that the problem will be nontrivial, that $m < n$.

The n variables x_j are partitioned into p nonempty subsets called <u>compartments</u>, or <u>phases</u>. We denote the compartment containing the variable x_j by $\langle j \rangle$. We may indicate that x_j and x_k are in the same compartment by writing:

$$j \in \langle k \rangle \quad \text{or} \quad k \in \langle j \rangle \quad \text{or} \quad \langle j \rangle = \langle k \rangle .$$

Each compartment has associated with it a <u>sum</u>,

$$\text{(I.3)} \qquad \bar{x}_{\langle k \rangle} = \sum_{j \in \langle k \rangle} x_j .$$

Each variable has associated with it a <u>variable fraction</u>,

$$\text{(I.4)} \qquad f_j = \frac{x_j}{\bar{x}_{\langle j \rangle}} .$$

The objective function to be minimized is:

$$(I.5) \qquad F(x_1, \ldots, x_n) = \sum_{j=1}^{n} x_j(c_j + \log \hat{x}_j).$$

The $c_1, c_2, \ldots, c_n$ are given real constants.

When $x_j = 0$, either $\hat{x}_j = 0$ (if $\bar{x}_{<j>} > 0$) or $\hat{x}_j$ is undefined (if $\bar{x}_{<j>} = 0$). In either case, $\log \hat{x}_j$ is undefined; but to maintain the continuity of F at the boundary of the constraint set we define $x_j \log \hat{x}_j = 0$ whenever $x_j = 0$. ([3], p. 364)

It will be convenient to use matrix notation. Thus we let A be an m x n matrix whose $ij^{\underline{th}}$ element is a_{ij}. That is,

$$(I.6) \qquad A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}.$$

Similarly, we let b be the m-vector whose $i^{\underline{th}}$ component is b_i, or:

$$(I.7) \qquad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

In the same fashion we let x be the n-vector with components x_j; $\hat{x}$ the n-vector with components $\hat{x}_j$; c the n-vector with components c_j; and $\log \hat{x}$ the n-vector with components $\log \hat{x}_j$.

Using vector notation, we may write the chemical equilibrium problem in the following compact form:

$$\text{(I.8)} \qquad \begin{aligned} \text{Min } F(x) = \text{Min } & x \cdot (c + \log \hat{x}) \\ \text{s.t.} \quad & Ax = b \\ & x \geq 0 \ . \end{aligned}$$

The notation $x \cdot (c + \log \hat{x})$ denotes the inner product of the vectors x and $(c + \log x)$.

It will be convenient to define two sets.

$$\text{(I.9)} \qquad H(A, b = \{x \geq 0 | Ax = b\}$$

Clearly this is the set of all feasible points x.

$$\text{(I.10)} \qquad M(F|H) = \{x \in H | F(y) \geq F(x) \ \forall \ y \in H\}.$$

Evidently, for any function F defined on a set H, the set $M(F|H)$ is the set of all points in H where F achieves its minimum. In particular, if we take F to be the function defined by (I.5), and $H = H(A, b)$ from (I.9), then $M(F|H(A, b))$ is the set of points x which are solutions to the problem (I.8).

In previous papers* we have explored several ways in which problem (*.8) is related to physical phenomena, and some mathematical difficulties associated with solving it. In this paper we will discuss several methods for finding a solution to (I.8).

---

*This paper is the third in a series of three. The first was "Chemistry, Kinetics, and Thermodynamics;" the second, "Degeneracy in Ideal Chemical Equilibrium Problems."

II. PRIMAL METHODS

We separate methods for solving the chemical equilibrium problem into primal and dual methods, according to whether the the purpose of an iteration of the method is to find an improved composition vector (primal methods) or a better vector of Lagrange multipliers (dual methods). In an extension of Clasen's terminology [4, 5], the methods would be called, respectively, first- and second-order methods.

Further, we separate primal methods according to whether the set of variables is expanded to include the $\bar{x}_{<j>}$'s (the augmented form, due to Clasen [5]), or whether the $\bar{x}_{<j>}$'s are ignored during each iteration, and updated only at the end of the iteration (the reduced form, due to the author [6]). Both forms may be treated as special cases of a more general problem, the linear-logarithmic problem.

The first primal method, called here the Augmented Linear Approximation Method, was developed by Clasen [4, 5]. In his work, he assumed degeneracy away, and failed to present any means of coping with it. We will do the same in this section. In Sections III and IV we will deal with the problem of degeneracy.

1. The Augmented Form

The form in which the chemical equilibrium problem was originally presented, (I.8), is to find a vector x satisfying:

$$\text{Min } F(x) = \Sigma x_j(c_j + \log \hat{x}_j)$$

(II.1.1) $$\text{s.t.} \quad Ax = b$$
$$x \geq 0$$

It is more convenient, however, to change the form by expanding the Gibb's function. Remembering that $\hat{x}_j = x_j/\bar{x}_{<j>}$, we write:

(II.1.2) $$F(x) = \sum_j x_j(c_j + \log x_j) - \sum_{<j>} \bar{x}_{<j>} \log \bar{x}_{<j>}.$$

If we include the sums $\bar{x}_{<j>}$ in the vector of variables, and remember that $\bar{x}_{<j>} = \sum_{k\epsilon<j>} x_k$, we may solve, in place of (II.1.1),

(II.1.3) $$\begin{cases} \text{Min } F(x) = (\sum_j x_j(c_j + \log x_j) - \sum_{<j>} \bar{x}_{<j>} \log \bar{x}_{<j>}) \\ \text{s.t.} \qquad Ax = b \\ \qquad \sum_{k\epsilon<j>} x_k - \bar{x}_{<j>} = 0 \qquad \forall <j> \\ \qquad x_j \geq 0, \ \bar{x}_{<j>} \geq 0 \qquad \forall\ j\ ,\ <j>. \end{cases}$$

If there are n species and p compartments, there will be $N = n + p$ variables in problem (II.1.3). Similarly, if the equations $Ax = b$ are m in number, then there will be a total of $M = m + p$ equations in the constraint set of (II.1.3). When the chemical equilibrium problem has been cast in the form of (II.1.3), we say it is in <u>augmented form</u>.

Notice that we may define a matrix $\bar{A}$, called the augmented matrix, in such a way that it is the detached coefficient matrix for all M equations of (II.1.3). $\bar{A}$ is an MxN matrix. Define $\bar{d}$, an N-vector, by:

$$\text{(II.1.4)} \qquad \bar{d}_j = \begin{cases} 1 & \text{if } 1 \leq j \leq n \\ -1 & \text{if } n+1 \leq j \leq N \end{cases}$$

Define $\bar{c}$, an N-vector, by:

$$\text{(II.1.5)} \qquad \bar{c}_j = \begin{cases} c_j & \text{if } \quad 1 \leq j \leq n \\ 0 & \text{if } n+1 \leq j \leq N \end{cases}$$

Define an M-vector $\bar{b}$ by:

$$\text{(II.1.6)} \qquad \bar{b}_i = \begin{cases} b_i & \text{if } 1 \leq i \leq m \\ 0 & \text{if } m+1 \leq i \leq M. \end{cases}$$

Augment the vector x of unknowns by letting x_j be the appropriate $\bar{x}_{<j>}$ for $n+1 \leq j \leq N$. Then we may write (II.1.3) as:

$$\text{(II.1.7)} \qquad \begin{aligned} &\text{Min} \sum_{j=1}^{N} x_j(\bar{c}_j + \bar{d}_j \log x_j). \\ &\text{s.t.} \quad \bar{A}x = \bar{b} \\ &\qquad\quad x \geq 0. \end{aligned}$$

It can be shown ([7], Theorem II.1) that if (II.1.7) possesses a <u>strictly positive</u> solution x^*, then there exist Lagrange multipliers, M in number, such that x^* also satisfies the Kuhn-Tucker conditions [8] for problem (II.1.7).

2. The Reduced Form

Suppose the original problem (I.8) possesses at least one strictly positive optimal solution x^*. It need not be unique. For each j, $1 \leq j \leq n$ we define new values for the constants c_j to be:

$$c_j^* = c_j - 1 - \log \bar{x}^*_{<j>}. \tag{II.2.1}$$

Theorem II.2.1: Consider the following two problems:

$$\begin{aligned} \text{Min } F(x) &= \text{Min } \Sigma x_j (c_j + \log x_j) \\ \text{s.t. } Ax &= b \\ x &\geq 0 \end{aligned} \tag{II.2.2}$$

and a second problem,

$$\begin{aligned} \text{Min } W(x) &= \text{Min } \Sigma x_j (c_j^* + \log x_j) \\ \text{s.t. } Ax &= b \\ x &\geq 0 \end{aligned} \tag{II.2.3}$$

where c_j^* is defined by (II.2.1), and $x^* \in M(F|H)$, $x^* > 0$. Then x^* is the unique solution to (II.2.3).

Proof: First, note that the function $g(t) = t \log t$ is strictly convex for $t \geq 0$. (We define $g(0) = 0$ so that g will be continuous for $t \geq 0$.) To see this, note that for $t > 0$, $g''(t) = \frac{1}{t} > 0$. As the sum of a linear function

and functions such as $g(t)$, $W(x)$ is a strictly convex function on the nonnegative orthant of n-space. Therefore, (II.2.3) has at most a single solution.

To see that x^* is that point, we look at the Kuhn-Tucker optimality conditions [8]. There must exist π^* such that

$$\frac{\partial W(x^*)}{\partial x_j} - A_j^T \pi^* = 0. \qquad 1 \leq j \leq n$$

(We have equality since $x^* > 0$. Were one of the $x_j^* = 0$, we would replace the equality by ≥ 0.) Evaluating this, we find that x^* is the solution to (II.2.3) if and only if for some π^*,

$$c_j^* + 1 + \log x_j^* = A_j^T \pi^*, \quad 1 \leq j \leq n.$$

Substitute for c_j^* from equation (II.2.1), and we find that:

$$\text{(II.2.4)} \qquad c_j + \log x_j^* = A_j^T \pi^*, \quad 1 \leq j \leq n.$$

But equations (II.2.4) constitute the optimality conditions for the original problem (II.2.2). Since x^* solves (II.2.2), we know that the required π^* exists QED.

When the problem has been put in the form (II.2.3) we say it is in the <u>reduced</u> form. The chief difficulty with this form is that we don't know what the $\bar{x}_{<j>}^*$'s will be. In the course of using this form in computational methods, we shall use instead the $\bar{x}_{<j>}$'s currently available, those corresponding to our current solution.

3. The General Linear-Logarithmic Problem

Clearly both the augmented and reduced forms of the chemical equilibrium problem may be considered special cases of the following problem:

$$\text{(II.3.1)} \qquad \begin{aligned} \text{Min } F(x) &= \Sigma\, x_j\, (c_j + d_j \log x_j) \\ \text{s.t.} \quad Ax &= b \\ x &\geq 0 \end{aligned}$$

The matrix A and the vectors b, c, d are constant quantities, and what they are, as well as their dimensions, depends on which of the two forms we choose for the problem. We take A to have dimension $M \times N$.

Notice that unless for each j, d_j is nonnegative, F(x) is not convex. It may, however, be convex on the domain H(A, b), where $H(A, b) = \{x \geq 0 | Ax = b\}$, as it is in either the augmented or the reduced form of the chemical equilibrium problem.

4. The Linear Approximation Method

This method is due to Clasen [4, 5].

If we blindly apply the Kuhn–Tucker theorem to the general linear–logarithmic problem, (II.3.1), we may believe we are looking for a vector (x^*, π^*) satisfying:

$$\text{(II.4.1)} \qquad \frac{\partial L(x^*, \pi^*)}{\partial x_j} \geq 0; \quad \frac{\partial L(x^*, \pi^*)}{\partial \pi_i} = 0 \qquad 1 \leq j \leq N, \quad 1 \leq i \leq M.$$

where $\frac{\partial L(x^*, \pi^*)}{\partial x_j} > 0$ only if $x_j = 0$, and where :

$$\text{(II.4.2)} \qquad L(x, \pi) = F(x) - \pi^T(Ax - b).$$

Performing the partial differentiation indicated in (II.4.1), we find that (x^*, π^*) must satisfy:

$$\text{(II.4.3)} \qquad c_j + d_j \log x_j^* + d_j \geq A_j^T \pi^*, \text{ with equality if } x_j^* > 0.$$

$$\text{(II.4.4)} \qquad Ax^* = b.$$

We will assume that $x^* > 0$, so that (II.4.3) is satisfied as an equation for each j. Solving (II.4.3) for $\log x_j^*$ we find that:

$$\text{(II.4.5)} \qquad \log x_j^* = d_j^{-1} \left[A_j^T \pi^* - c_j - d_j \right].$$

Suppose we are given a starting solution y satisfying

$Ay = b$, $y > 0$. We may expand the function $(\log x_j)$ around the point y_j in a Taylors expansion. Then:

$$\text{(II.4.6)} \qquad \log x_j = \log y_j + \frac{x_j - y_j}{y_j} + \text{(higher-order terms)}.$$

Ignoring the higher-order terms, we substitute (II.4.6), evaluated at x_j^*, into (II.4.5). Then:

$$\text{(II.4.7)} \qquad x_j^* \approx x_j = y_j \, d_j^{-1} \left[A_j^T \pi^* - c_j - d_j \log y_j \right].$$

To find the value of the right-hand side of (II.4.7), we must first evaluate π^*. This is not, of course, possible, but we can find an approximation π to π^* by substituting (II.4.7) into (II.4.4), as is done below in (II.4.8).

Let D be a diagonal matrix whose $jj^{\underline{th}}$ element is d_j. Assume $d_j \neq 0$, $1 \leq j \leq N$.

Let Y be a diagonal matrix whose $jj^{\underline{th}}$ element is y_j, $1 \leq j \leq N$.

From (II.4.7) we write:

$$\text{(II.4.8)} \qquad b = Ax = (AYD^{-1}A^T)\pi - AYD^{-1}(c + d \log y).$$

Define:

$$\text{(II.4.9)} \qquad R = (AYD^{-1}A^T)$$

Then

$$\text{(II.4.10)} \qquad R\pi = b + AYD^{-1}(c + d \log y)$$

In the reduced case, it can be shown that R is nonsingular if and only if the matrix $\tilde{A}$, obtained from A by deleting those columns A_j with $y_j = 0$, has full rank m. A more complex comment applies to the matrix R obtained in the augmented case. Surely, if R is nonsingular, then equations (II.4.10) possess a solution π.

To find our new point x, we substitute π from equation (II.4.10) into (II.4.7) in place of π^*.

Unfortunately, we cannot insure that the new x will be positive, and so we cannot in general use x as our next iterate. Instead, we use a vector u satisfying:

$$u = \alpha x + (1 - \alpha)y \tag{II.4.11}$$

for some $0 < \alpha \leq 1$. We insist $u > 0$.

Section II.6 will give an algorithm using this development, and an indication of when it might work and when it might fail.

5. The Quadratic Approximation Method

To circumvent some possible convergence difficulties of Linear methods, one might try to insure that the new point x computed from y could be used as the next iterate. This suggests that we restrict x to be nonnegative.

The most straightforward way of doing this is to expand F(x) around the initial solution y and minimize the second degree Taylor's approximation, subject to the constraints $Ax = b$, $x \geq 0$. In place of problem (II.3.1) we solve the quadratic programming problem:

$$\text{Min } Q(x) = \tfrac{1}{2} \Sigma \frac{d_j(x_j - y_j)^2}{y_j} + \Sigma(x_j - y_j)(c_j + d_j \log y_j + d_j) + F(y)$$

$$\text{(II.5.1)} \qquad \text{s.t.} \qquad Ax = b$$

$$x \geq 0.$$

Since F(x) is convex on H(A, b), Q(x) will also be convex. Thus a vector x will solve (II.5.1) if and only if it satisfies the Kuhn-Tucker optimality conditions [8]. That is, x solves (II.5.1) if and only if there exists an M-vector π and an N-vector δ such that (x, π, δ) satisfy:

$$\text{(II.5.2)} \quad \begin{cases} \dfrac{d_j x_j}{y_j} - A_j^T \pi + (c_j + d_j \log y_j) = \delta_j & i \leq j \leq N \\ Ax - b = 0 \\ x \geq 0, \ \delta \geq 0, \ \Sigma x_j \delta_j = 0. \end{cases}$$

In matrix notation, the problem of finding (x, π, δ) to satisfy (II.5.2) may be written:

$$
\text{(II.5.3)} \quad
\left[\begin{array}{ccc|c}
\frac{d_1}{y_1} & & & -A_1^T \\
 & \ddots & & \vdots \\
 & & \frac{d_N}{y_N} & -A_N^T \\
\hline
A_1 & \cdots & A_N & 0
\end{array}\right]
\begin{bmatrix} x_1 \\ \vdots \\ x_N \\ \hline \pi_1 \\ \vdots \\ \pi_m \end{bmatrix}
+
\begin{bmatrix} c_1 + d_1 \log y_1 \\ \vdots \\ c_N + d_N \log y_N \\ \hline -b_1 \\ \vdots \\ -b_M \end{bmatrix}
=
\begin{bmatrix} \delta_1 \\ \vdots \\ \delta_N \\ \hline 0 \\ \vdots \\ 0 \end{bmatrix}
$$

$$x_j \geq 0, \ \delta_j \geq 0, \ x_j \delta_j = 0 \quad 1 \leq j \leq N.$$

This is recognizable as the sort of problem considered in complementary pivot theory [9]. One might also decide to solve problem (II.5.1) directly, using one of the many quadratic programming algorithms available (see, for example, [10]).

6. The Basic Algorithm

The Linear Method and the Quadratic Method are quite similar. We may in fact write that the new point x found using the Linear Method must satisfy:

$$(II.6.1)\quad \begin{cases} \dfrac{d_j x_j}{y_j} - A_j^T \pi + (c_j + d_j \log y_j & = 0 \\ Ax \qquad - b & = 0 \end{cases}$$

Compare these equations with equations (II.5.2) of the quadratic method. In the same way as for (II.5.2), satisfying equations (II.6.1) is equivalent to solving a quadratic program. The quadratic program to solve is just:

$$(II.6.2)\quad \text{Min } Q(x) = \left[\tfrac{1}{2} \sum \frac{d_j (x_j - y_j)^2}{y_j} + \Sigma (x_j - y_j)(c_j + d_j \log y_j + d_j) + F(y) \right]$$

$$\text{s.t.} \quad Ax = b$$

This is precisely the same as (II.5.1) except that the solution x is not constrained to be nonnegative.

Theorem II.6.1. Let F(x) be convex on H(A, b). Let $y > 0$, $y \in H(A, b)$ be the initial solution (either in the Linear or the quadratic method), and let x be a solution to the associated quadratic program ((II.5.1) in the quadratic case, (II.6.2) in the Linear case).

Let $\theta = x - y$.

Denote the derivative of F at y in the direction θ by $F'_\theta(y)$. Then:

$$F'_\theta(y) \leq 0$$

with equality if and only if y is a solution to the original problem (II.3.1).

<u>Proof</u>: It can be shown that since $y > 0$,

$$\text{(II.6.3)} \qquad F'_\theta(y) = \Sigma\, \theta_j(c_j + d_j \log y_j + d_j).$$

In the quadratic case, we may write (from (II.5.2)):

$$\text{(II.6.4)} \qquad \frac{d_j(x_j - y_j)}{y_j} = \delta_j + A_j^T\pi - (c_j + d_j \log y_j + d_j)$$

In the linear case, we have from (II.6.1) that equation (II.6.4) is still true with δ_j set equal to zero.

Multiply (II.6.4) by θ_j and sum over all $1 < j < N$. The result is:

$$\text{(II.6.5)} \qquad \Sigma\, \frac{d_j\theta_j^2}{y_j} = -\Sigma\, y_j\delta_j - F'_\theta(y)$$

since in both the linear and the quadratic methods, $\Sigma\, x_j\delta_j = 0$ (in the linear case because $\delta_j = 0$) and since $A\theta = 0$.

The left-hand side of (II.6.5) is just:

$$\text{(II.6.6)} \qquad \Sigma\, \frac{d_j\theta_j^2}{y_j} = \frac{d^2}{dt^2} F(y + t\theta)\Big|_{t=0} \geq 0$$

which is nonnegative because F is convex on H(A,b). Thus:

$$\text{(II.6.7)} \quad F'_\theta(y) = \begin{cases} -\Sigma \dfrac{d_j \theta_j^2}{y_j} - \tau\, y_j \delta_j & \text{(Quadratic)} \\ -\Sigma \dfrac{d_j \theta_j^2}{y_j} & \text{(Linear)} \end{cases} \leq 0$$

If y is not an optimal solution of (II.3.1), then there exists $y^* \in H(A,b)$ such that $F(y^*) < F(y)$. Let:

$$\theta^* = y^* - y.$$

Since F is convex on H(A,b), clearly:

$$F'_{\theta^*}(y) < 0.$$

If we let $x(\alpha) = y + \alpha\theta^*$ for some $0 < \alpha \leq 1$, then $x(\alpha) \in H(A, b)$. Further,

$$\text{(II.6.8)} \quad Q(x(\alpha)) = \frac{\alpha^2}{2} \tau \frac{d_j (\theta_j^*)^2}{y_j} + \alpha\tau\, \theta_j^*(c_j + d_j \log y_j + d_j) + F(y).$$

Since F is convex on H(A,b), surely $Q(x(\alpha))$ is a convex function of α, and

$$\frac{dQ}{d\alpha}\Big|_{\alpha=0} = F'_{\theta^*}(y) < 0.$$

Thus for some $\alpha > 0$,

$$\text{(II.6.9)} \quad Q(x(\alpha)) < Q(x(0)) = Q(y) = F(y).$$

That is, if y is not an optimal solution to (II.3.1), then the solution x^0 to (II.5.1) or (II.6.2) satisfies:

(II.6.10) $$Q(x^o) < Q(y) = F(y).$$

Now suppose x^o is a solution to (II.5.1), or (II.6.2) and $\theta = x^o - y$. Further, let:

$$F'_\theta(y) = 0.$$

From (II.6.7) we see that $\Sigma \frac{d_j(x_j^o - y_j)^2}{y_j} = 0$. Substituting this into $Q(x^o)$, we have:

(II.6.11) $$Q(x^o) = \tfrac{1}{2} \Sigma \frac{d_j(x_j^o - y_j)^2}{y_j} + \Sigma\theta_j(c_j + d_j \log y_j + d_j) + F(y)$$

$$= 0 + F'_\theta(y) + F(y) = F(y).$$

That is, $F'_\theta(y) = 0$ implies $Q(x^o) = F(y)$. But we showed that this in turn implies that y solves (II.3.1).

QED.

The algorithm based on either of these methods may be stated as follows:

1. Given $y > 0$, $y \in H(A,b)$, find (x,π) satisfying (II.5.3) in the quadratic case, or (II.6.1) in the linear case. Let $\theta = x - y$.

2. If $\Sigma \frac{d_j\theta_j^2}{y_j} + \Sigma\delta_j y_j$ ($\delta_j = 0$ in the linear case) is sufficiently small, terminate. If not, go to step 3.

3. Let $u = y + \alpha\theta$ for some $0 < \alpha \leq 1$. The vector u must satisfy $u > 0$ and $F(u) < F(y)$.

4. Replace y by u in step 1, and continue the cycle from there.

The termination criterion of step 2 does not say how small $\Sigma \frac{d_j \Delta_j^2}{y_j} + \Sigma \delta_j y_j$ must be. This is a matter to be decided by the requirements of the problem.

Combining the two forms of the chemical equilibrium problem, augmented and reduced, with the two methods of solving the linear-logarithmic problem, linear and quadratic, yields four different methods for solving, or at least trying to solve, the chemical equilibrium problem. We name these methods the Augmented Linear Method (ALM), the Augmented Quadratic Method (AQM), the Reduced Linear Method (RLM), and the Reduced Quadratic Method (RQM).

Reduced methods have the advantage that the number of variables, and hence the size of the problem, is smaller in the reduced case than in the augmented. Linear methods hold an advantage over quadratic methods in that each iteration requires less time, or at least no more time. That is, the finding of a new direction in a linear method is never a more lengthy process, and is sometimes a shorter one, than in a quadratic method. Whether this advantage is counterbalanced by a possible saving in the number of iterations required by the quadratic methods, the author cannot say.

The linear methods, however, need not converge to a solution y^* of (II.3.1). If (II.3.1) has a bounded solution set, it can be shown that the sequence of iterates $\{y^{(k)}\}$ generated by the method has a limit point. But it may happen that the successive step sizes $\alpha^{(k)}$ converge to zero, allowing the quantity $\sum \frac{d_j \theta_j^2}{y_j}$ to remain bounded away from zero. This might occur if, for example, one of the limit points y^o of $\{y^{(k)}\}$ has some component $y_j^o = 0$.

With some modifications, it can be shown (and will be shown) that the quadratic method applied either to the reduced or the augmented form yields a solution to a slightly modified problem.

The author recommends that whatever primal method is used, slacks be included in the formulation (see [7]). To include slacks in the problem is equivalent to including in each compartment a tiny amount of a substance that will not diffuse to any other compartment and will not participate in any reaction. This can be done by replacing $\bar{y}_{<j>}$ by $\bar{y}_{<j>} + s_{<j>}$, and $\hat{y}_j$ by $\frac{y_j}{\bar{y}_{<j>}+s_{<j>}}$ wherever they occur in the formulas, where each $s_{<j>}$ is some small positive number. This will insure that if an answer exists, it will be unique, and so that if a solution method generates a sequence $\{y^{(k)}\}$ of iterates with some y^* as a limit point, and if $\{y^*\} = M(F|H)$, then the entire sequence $\{y^{(k)}\}$ converges to y^*.

III. A CLOSER LOOK AT THE REDUCED QUADRATIC METHOD

This section deals with a method, first proposed by the author in [6], for solving the chemical equilibrium problem. Because of its derivation, one might call it the Reduced Generalized Linear Programming Method. However, it turns out in the end to be the same as RQM. We retain the derivation, even though it leads to nothing new, because it is both interesting in its own right and because it leads us to suspect that RQM might be more powerful than supposed.

1. The Reduced Generalized Linear Program

If the objective function of a chemical equilibrium problem were linear, we would have a ready-made, efficient method for solving the problem, the simplex method. Since it is not linear, we try to make it so. We will find it convenient to use the reduced form (II.2.3) in place of the original form (II.2.2).

Instead of solving (II.2.3) directly, let us "linearize" it by introducing n new variables α_j, $1 \leq j \leq n$, and solving the problem:

$$\text{(III.1.1)} \quad \begin{array}{lll} \text{Min } \phi(y,\alpha) = \text{Min } \Sigma\, y_j(c_j^* - \log \alpha_j) & & \\ \text{s.t.} \quad Ay = b & & ;\pi \\ \qquad \alpha_j y_j \leq 1 & & ;(-x_j) \\ \qquad y \geq 0 & & \end{array}$$

Each α_j may be chosen independently of every y_k.

Clearly, given any $y \in H(A,b)$, we can find a corresponding feasible solution (y,α) to (III.1.1) by choosing:

$$\alpha_j = \begin{cases} 1/y_j & \text{if } y_j > 0 \\ 1 & \text{if } y_j = 0 \end{cases} \qquad 1 \leq j < n.$$

Alternatively, given a feasible solution (y, α) to (III.1.1), the vector y is a feasible solution to the original problem (II.2.3) — i.e., $y \in H(A,b)$.

Further, if (y, α) is a feasible solution to (III.1.1), it may easily be improved unless

$$\text{(III.1.2)} \qquad y_j\alpha_j = 1 \qquad \forall j \ni y_j > 0.$$

We may consider (III.1.1) to be a generalized linear program ([11], p. 434). Thus we wish to minimize a linear form subject to linear constraints, where the vector of coefficients of a variable y_j is not constant, but rather can vary over some set.

The quantities π and $(-x_j)$ appearing to the right of the constraints of (III.1.1) are Lagrange multipliers. The multipliers π corresponding to equality constraints are unrestricted in sign, but the multipliers $(-x_j)$ corresponding to inequalities must satisfy:

$$\text{(III.1.3)} \qquad x_j \geq 0 \qquad 1 \leq j \leq n.$$

Suppose we have an initial, feasible solution (y, α)

to (III.1.1) which is strictly positive. Let (III.1.2) be satisfied for each j.

Consider the α_j as constants. For the given α_j, assume that the solution y is the optimal solution to (III.1.1). (This will surely be the case if H(A, b) is bounded. The reader may easily check that in this case y is the only solution.) Under this assumption there must exist multipliers (π, x) such that (y, π, x) satisfies the Kuhn-Tucker conditions [8], namely:

$$\text{(III.1.4)} \quad f_j(\alpha_j) = c_j^* - \log \alpha_j - A_j^T\pi + \alpha_j x_j = 0 \qquad 1 \leq j \leq n.$$

To improve the solution (y, α), we must find new columns of coefficients from the sets mentioned before which it would be profitable to introduce--i.e., which "price out" negatively. That is, we must find new values α_j' for the variables α_j satisfying:

$$\text{(III.1.5)} \qquad f_j(\alpha_j') \leq 0, \qquad 1 \leq j \leq n,$$

with strict inequality for at least one j.

The function $f_j(\alpha_j)$ is defined for $\alpha_j > 0$ and is convex, a fact easy to check. To satisfy (III.1.5) we would be well-advised to minimize the function f_j, subject to $\alpha_j > 0$. The minimum occurs where the derivative vanishes.

$$\text{(III.1.6)} \qquad \frac{df_j}{d\alpha_j} = 0 = -\frac{1}{\alpha_j'} + x_j.$$

If $x_j > 0$, we should therefore choose $\alpha_j' = \frac{1}{x_j}$. We

will deal with the case that $x_j = 0$ later.

We have so far only asserted the existence of multipliers (π, x). Conditions (III.1.3-4) are in general too few in number to determine their values. But equations (III.1.6) suggest that the multipliers x_j be interpreted as composition variables. Thus we demand that $x \in H(A, b)$.

This latest condition, with (III.1.4) requires that (π, x) satisfy (substituting $\frac{1}{y_j}$ for α_j and $c_j - 1 - \log \bar{y}_{<j>}$ for c_j^*):

$$\text{(III.1.7)} \quad \begin{cases} \frac{x_j}{y_j} - A_j^T\pi + (c_j + \log \hat{y}_j - 1) = 0 & 1 \leq j \leq n \\ Ax \qquad - b \qquad\qquad = 0 \\ x \geq 0 \end{cases}$$

We have gone from one extreme to another. The equations of (III.1.7) are those of the Reduced Linear Method. It can be shown that if the matrix A has full rank there is a unique solution x to these equations. However, this solution need not be nonnegative.

To insure that $x \geq 0$, we relax the first n equations as follows:

$$\text{(III.1.8)} \quad \begin{cases} \frac{x_j}{y_j} - A_j^T\pi + (c_j + \log \hat{y}_j - 1) = \delta_j \\ Ax \qquad - b \qquad\qquad = 0 \\ x \geq 0,\ \delta \geq 0,\ x^T\delta = 0. \end{cases}$$

We find that the reduced generalized L. P. method leads back to the Reduced Quadratic Method.

This suggests, however, that as in a generalized L. P., the new solution x will itself be strictly better than the old. In fact, this is the case.

To prove our theorem, we will need the following lemma.

<u>Lemma III.1.1.</u> Let $g(r) = r - 1 - \log r$ be defined for $r \geq 0$, with $g(0) = \infty$. Then $g(r) \geq 0$ for all r, and $g(r) = 0$ if and only if $r = 1$.

<u>Proof</u>: It is well known that g(r) is strictly convex in r, and differentiable for $r > 0$. Its minimum occurs where the derivative vanishes. Thus:

$$g'(r) = 1 - \frac{1}{r} = 0 \tag{III.1.9}$$

implies the minimum occurs at $r = 1$, and nowhere else (by strict convexity). Substituting, we find for each $r \geq 0$,

$$g(r) \geq g(1) = 1 - 1 - \log 1 = 0 \tag{III.1.10}$$

with equality if and only if $r = 1$.

QED.

<u>Theorem III.1.2.</u> Let $y \in H(A, b)$, $y > 0$ be an initial point, and let (x, π, δ) satisfy (III.1.8). Define $\theta = x - y$, and $u(\alpha) = y + \alpha\theta$. Then:

(i) $F(u(\alpha)) \leq F(y) \quad \forall\ 0 \leq \alpha \leq 1$

and:

(ii) $F(u(\alpha)) = F(y)$ if and only if <u>either</u> $\alpha = 0$ <u>or</u> y is a solution to the original problem (I.8) (or II.2.2). In this case, $x = y$.

In this theorem, F refers to the original Gibbs function. not the objective of the linear-logarithmic problem.

<u>Proof:</u> x satisfies:

$$\text{(III.1.11)} \qquad \frac{x_j}{y_j} - A_j^T\pi + (c_j + \log \hat{y}_j) - 1 = \delta_j$$

where $\delta_j \geq 0$ and $x_j\delta_j = 0$ for $1 \leq j \leq n$. Multiply (III.1.11) by $(y_j - u_j(\alpha))$ and sum over $1 \leq j \leq n$. After rearranging terms this gives us:

$$\text{(III.1.12)} \qquad F(y) - F(u(\alpha)) = \alpha \sum_j \delta_j y_j + \sum_j u_j\left(\frac{u_j}{y_j} - 1 - \log \frac{u_j}{y_j}\right)$$
$$+ \sum_{<j>} \bar{u}_{<j>}\left(\frac{\bar{y}_{<j>}}{\bar{u}_{<j>}} - 1 - \log \frac{\bar{y}_{<j>}}{\bar{u}_{<j>}}\right)$$
$$+ \alpha(1 - \alpha) \sum_j \frac{\theta_j^2}{y_j} .$$

Equation (III.1.12) holds for all $0 \leq \alpha < 1$.

Taking the right-hand side of (III.1.12) term by term, we see that:

$$\text{(III.1.13)} \qquad \alpha \sum_j \delta_j y_j \geq 0$$

$$\text{(III.1.14)} \qquad \sum_j u_j\left(\frac{u_j}{y_j} - 1 - \log \frac{u_j}{y_j}\right) = \sum_j u_j \; g\left(\frac{u_j}{y_j}\right) \geq 0$$

where g is defined in Lemma III.1.1.

$$\text{(III.1.15)} \quad \sum_{<j>} \bar{u}_{<j>}\left(\frac{\bar{y}_{<j>}}{\bar{u}_{<j>}} - 1 - \log \frac{\bar{y}_{<j>}}{\bar{u}_{<j>}}\right) = \sum_{<j>} \bar{u}_{<j>}\, g\left(\frac{\bar{y}_{<j>}}{\bar{u}_{<j>}}\right) \geq 0.$$

This term must be modified if $\alpha = 1$. For $\alpha < 1$, note that $\bar{u}_{<j>}(\alpha) > 0$ for every phase <j>. But at $\alpha = 1$, $u = x$, and some $\bar{x}_{<j>}$ may be zero. Thus we write (III.1.15), for $\alpha = 1$, as:

$$\text{(III.1.16)} \quad \sum_{<j>}\left(\bar{y}_{<j>} - \bar{x}_{<j>} - \bar{x}_{<j>} \log \frac{\bar{y}_{<j>}}{\bar{x}_{<j>}}\right) \geq 0 \text{ if } \alpha = 1.$$

For each <j> such that $\bar{x}_{<j>} > 0$, this is the same as for $\alpha < 1$. If $\bar{x}_{<j>} = 0$, then the only nonzero term of the sum (III.1.16) equals $\bar{y}_{<j>} > 0$.

Finally,

$$\text{(III.1.17)} \qquad \alpha(1 - \alpha) \sum_j \frac{\theta_j^2}{y_j} \geq 0.$$

Combining (III.1.13-17) we have:

$$\text{(III.1.18)} \qquad F(y) - F(u(\alpha)) \geq 0 \qquad 0 \leq \alpha \leq 1$$

proving statement (i).

To show statement (ii), note that if $\alpha = 0$, then $u(\alpha) = y$, so that surely $F(u(\alpha)) = F(y)$.

Suppose, on the other hand, that $\alpha > 0$, and $F(u(\alpha)) = F(y)$. We first take $\alpha = 1$. From (III.1.14) we see that $F(y) - F(x) = 0$ only if for each j,

(III.1.19) $$x_j \, g\left(\frac{x_j}{y_j}\right) = 0$$

which is true if and only if either $x_j = 0$, or $x_j = y_j$, by Lemma III.2.1. If $\theta \neq 0$, then for at least one j, $x_j = 0$.

But this implies that $\bar{x}_{<j>} \neq \bar{y}_{<j>}$. Hence the inequalities (III.1.15) or (III.1.16) apply, proving that:

(III.1.20) $$F(x) < F(y) \qquad \text{if } x \neq y.$$

For $0 < \alpha < 1$, we need only note that F is convex, so that, for every α, $0 < \alpha \leq 1$, $F(u(\alpha)) < F(y)$ if $x \neq y$.

We have shown that $F(u(\alpha)) = F(y)$ if and only if $\alpha = 0$ or $x = y$. To complete the proof, we must show that in the latter case, y solves (I.8).

But if $x = y$, then $\theta = 0$, so that $F'_\theta(y) = 0$. Thus by Theorem II.4.1, y solves the original problem. QED.

2. Lower Bounds

We may separate feasible solutions y into "bad" points and "good" points. "Bad" points are points $y \in H(A, b)$ which are not interior to the nonnegative orthant. other feasible points are "good".

The reason for this terminology, particularly with regard to RQM, is easy to see upon examining equations (III.1.8). If we have as our initial point y a "bad" point, and we try to generate a new point x from it, then x must satisfy:

$$\text{(III.2.1)} \qquad \frac{x_j}{y_j} - A_j^T \pi + (c_j + \log \hat{y}_j - 1) \quad {}_j, \qquad 1 \leq j \leq n.$$

If $y_j = 0$, this is undefined. Large values of y_j present their own difficulties. Similar reasoning points out that bad points are bad for any of our primal methods.

As it happens, the only "bad" points that ordinarily concern us are those satisfying $y_j = 0$ for some j. This is a consequence of the following result.

Lemma III.2.1: Let $y \in H(A, b)$, and define:

$$\text{(III.2.2)} \qquad S(y) = \{x \in H(A, b) \mid F(x) \leq F(y)\}.$$

Then S(y) is bounded if and only if the solution set $M(F|H)$ of the original problem (II.2.2) is bounded and non-empty.

Proof: We know that since F is convex, $S(y)$ must be convex. If $S(y)$ is unbounded, then by a well-known result (see, for example, [12], Lemma 3), there exists a vector $v \neq 0$ such that for all $x \in S(y)$ and all $t \geq 0$,

$$x + tv \in S(y). \tag{III.2.3}$$

Notice that this implies in particular, since $x + tv \in H(A, b)$, that:

$$\begin{aligned} Av &= 0 \\ v &\geq 0. \end{aligned} \tag{III.2.4}$$

Furthermore, for every $t \geq 0$,

$$F(y + tv) \leq F(y).$$

By the homogeneity of F

$$F(v + \frac{1}{t} y) - F(\frac{1}{t} y) \leq 0.$$

Letting $t \to \infty$, we have, since $F(0) = 0$ and F is continuous,

$$F(v) \leq 0. \tag{III.2.5}$$

By [12], Theorems 4 and 5, it is implied by Equations (III.2.4-5) that $M(F|H)$ is either unbounded or empty.

Conversely, it is clear that $M(F|H) \subset S(y)$. (This does not rule out the chance that $M(F|H)$ is empty.) Thus instead of solving the original problem (II.2.2), we could replace it with:

$$\text{Min } F(x)$$
$$\text{s.t. } x \in S(y).$$

If $S(y)$ is bounded, then it is compact (since by the continuity of F it must be closed). And a continuous function always achieves its minimum on a compact set. Thus $M(F|S(y))$ is nonempty, and (since $S(y)$ is bounded) $M(F|S(y))$ is also bounded.

Clearly, $M(F|S(y)) = M(F|H)$. QED.

The set $S(y)$ has for us the significance that if we begin the algorithm at an initial point $y^{(0)} \in H(A, b)$, $y^{(0)} > 0$, then every iterate $y^{(k)}$ must also be in the set $S(y^{(0)})$. If $M(F|H)$ were bounded and nonempty, then the sequence of iterates $\{y^{(k)}\}$ generated by any of the primal algorithms would be a bounded set.

Furthermore, even if $M(F|H)$ is either empty or unbounded, we may continue generating new points until some $y_j^{(k)}$ grows too large to be handled. This will merely be evidence that the solution we are chasing is a will-o-the-wisp. In addition, we may rule out the possibility that $M(F|H)$ is unbounded by introducing slacks into the problem, as suggested in [7].

On the other hand, that some $y_j^{(k)} = 0$ does not suggest that the problem is without a solution. More than this, in a practical sense $y_j^{(k)}$ need not actually be zero to cause trouble; it need only be sufficiently small that the electronic

computer cannot work with it. We must have some means to insure that no iteration has a component $y_j^{(k)}$ too close to zero.

The obvious way to accomplish this is to modify the problem by placing lower bounds on each variable. Thus we might solve:

$$\text{(III.2.6)} \qquad \begin{aligned} &\text{Min } F(x) = \Sigma\, x_j(c_j + \log \hat{x}_j) \\ &\text{s.t.} \quad Ax = b \\ &\qquad\qquad x \geq \ell > 0 \end{aligned}$$

where ℓ is an n-vector of lower bounds, each ℓ_j, $1 \leq j \leq n$, being a small, positive number.

By expanding the number of variables and the number of linear constraints, we may cast (III.2.6) into the form of a classical chemical equilibrium problem.

$$\text{(III.2.7)} \qquad \begin{aligned} &\text{Min } F(x) = \Sigma\, x_j(c_j + \log \hat{x}_j) + \Sigma\, s_j\,(c_{j+n} + \log \hat{s}_j) \\ &\text{s.t.} \quad Ax = b \\ &\qquad x_j - s_j = \ell_j \\ &\qquad x \geq 0,\ s \geq 0. \end{aligned}$$

Each variable s_j is taken to be in a new compartment, and is the only species there. Its free-energy constant c_{j+n} is set to zero.

It is not necessary, however, to expand the problem in this way. Lower bounds can be included without increasing the size of the problem at all. For the RQM, in place of equations (III.1.8), we would let:

(III.2.8) $\qquad x' = x - \ell$

and require x' to satisfy:

$$\text{(III.2.9)} \qquad \begin{bmatrix} \frac{1}{y_1} & & & -A_1^T \\ & \ddots & & \vdots \\ & & \frac{1}{y_n} & -A_n^T \\ A_1 & \cdots & A_n & 0 \end{bmatrix} \begin{bmatrix} x_1' \\ \vdots \\ x_n' \\ \pi_1 \\ \vdots \\ \pi_m \end{bmatrix} + \begin{bmatrix} c_1 + \log \hat{y}_1 - 1 + \frac{\ell_1}{y_1} \\ \vdots \\ c_n + \log \hat{y}_n - 1 + \frac{\ell_n}{y_n} \\ -b_1 + \sum a_{1j}\ell_j \\ \vdots \\ -b_m + \sum a_{mj}\ell_j \end{bmatrix} = \begin{bmatrix} \delta_1 \\ \vdots \\ \delta_n \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$x_j' \geq 0, \ \delta_j \geq 0, \ \delta_j \cdot x_j' = 0, \ 1 \leq j \leq n.$$

It is easy to show that if x' satisfies (III.2.9), then $x = x' + \ell$ (from (III.2.8)) solves the quadratic program (II.5.1) for the reduced quadratic case, with the non-negativity conditions $x \geq 0$ replaced by $x \geq \ell$. It is also easy to show that (x, s) will be the point generated by RQM applied to the expanded problem (III.2.7), where $s = x' = x - \ell$ are the new values for the additional variables.

This observation insures that all the theorems that have been proven for the problem without lower bounds still are true for the problem with lower bounds.

3. The Modified RQM Algorithm

The algorithm for the RQM may be modified as a result of Theorem III.1.2, and to include lower bounds. These two modifications, as we shall see, insure that the algorithm will yield a solution y^* which is optimal for the lower-bounded problem (III.2.7)

1. Given $y \geq \ell$, $y \in H(A, b)$, solve (III.2.9) for x', and let $x = x' + \ell$. We may use complimentary pivot methods, or we may choose to solve this as a quadratic program.
2. Let $\theta = x - y$. If $\Sigma \frac{\theta_j^2}{y_j} + \Sigma \delta_j(y_j - \ell_j)$ is sufficiently small, terminate. If not, go to step 3.
3. Let $\underline{\alpha}$ and $\overline{\alpha}$ be prechosen numbers. Pick as a step size any α satisfying:

$$0 < \underline{\alpha} \leq \alpha \leq \overline{\alpha} \leq 1.$$

   That is, we bound α away from zero and if we wish, from one. A perfectly adequate choice is $\underline{\alpha} = \overline{\alpha}$, and hence a constant step size.
4. Let $u = y + \alpha\theta$. Replace y by u, and return to step 1.

The convergence criterion of step 2 has been modified to take account of the lower bounds.

What follows is a proof that the algorithm finds a solution to (III.2.7). Throughout, we will use the following notation:

$\{y^{(k)}\}$ is the sequence of iterates generated by the modified algorithm.

$\{x^{(k)}\}$ is the sequence such that $x(k)-\ell$ satisfies (III.2.9) if y is replaced by $y^{(k)}$.

$\delta^{(k)} = x^{(k)} - y^{(k)}$ for each k.

α_k is the kth step size, satisfying $0 < \underline{\alpha} \leq \alpha_k \leq \overline{\alpha} \leq 1$ for all k, so that:

$y^{(k+1)} = y^{(k)} + \alpha_k \delta^{(k)}$.

We will also assume that the solution set $M_\ell(F|H)$ of (II.2.7) is bounded and nonempty. The "subscript" ℓ denotes lower bounding. It is easy to show, using [12], Theorems 4 & 5, that $M_\ell(F|H)$ is bounded and nonempty if and only if $M(F|H)$, the solution set of the non-lower-bounded problem, is bounded and nonempty.

<u>Lemma III.3.1</u>: $\lim_{k\to\infty} \delta^{(k)} = 0$.

<u>Proof</u>: Since $M_\ell(F|H)$ is nonempty, $F(x)$ is bounded below on $H_\ell(A, b) = \{x \geq \ell | Ax = b\}$. Thus the monotone decreasing sequence $\{F(y^{(k)})\}$ must have a limit.

Using the function g from Lemma III.1.1 and equations (III.1.12-17) from Theorem III.1.2, we have that:

$$\text{(III.3.1)} \quad F(y^{(k)}) - F(y^{(k+1)}) \geq \sum_j y_j^{(k+1)} g\left(\frac{y_j^{(k+1)}}{y_j^{(k)}}\right) \geq 0.$$

Substituting for g, and taking the limit as $k \to \infty$, we find that for each $1 \leq j \leq n$, by Lemma III.1.1,

$$\text{(III.3.2)} \quad \left[\lim_{k\to\infty} y_j^{(k+1)}\left(\frac{y_j^{(k+1)}}{y_j^{(k)}} - 1 - \log \frac{y_j^{(k+1)}}{y_j^{(k)}}\right)\right] = 0.$$

Since for every k, $y_j^{(k)} \geq \ell_j > 0$, we see that by Lemma III.1.1,

$$\lim_{k\to\infty} \frac{y_j^{(k+1)}}{y_j^{(k)}} = 1,$$

or:

$$\text{(III.3.3)} \quad \lim_{k\to\infty} \frac{y_j^{(k+1)} - y_j^{(k)}}{y_j^{(k)}} = \lim \alpha_k \frac{\theta_j^{(k)}}{y_j^{(k)}} = 0.$$

But $y^{(k)} \in S(y^{(0)})$, a bounded set, and the α_k we uniformly bounded away from zero by $\underline{\alpha}$. The conclusion follows.

Q.E.D.

Notice that since $M_1(F|H)$ is bounded and nonempty, $S(y^{(0)})$ is bounded. Thus the set of iterates $\{y^{(k)}\}$ is a bounded, infinite set, and hence must have a limit point.

<u>Theorem III.3.2:</u> Let N be a subsequence of integers (1, 2, ...) such that $\{y^{(k)}\}_{k\in N}$ converges, and let:

$$\text{(III.3.4)} \qquad y^* = \lim_{k\in N} y^{(k)}.$$

Assume there exists $v > 0$ such that $Av = b$, and suppose the matrix A has full rank m. Then,

$$y^* \in M_\ell(F|H).$$

<u>Note:</u> It is no hardship to assume the existence of $v > \ell$ such that $Av = b$. The appendix shows a method for finding an initial, strictly positive solution $y^{(0)}$, if one exists. We may always set the lower bounds ℓ to satisfy $y^{(0)} > \ell$.

<u>Proof:</u> Surely y^* is feasible. That is, $y^* \geq \ell$ and $Ay^* = b$.

From (III.2.9) we write that

$$\text{(III.3.5)} \qquad c_j + \log \hat{y}_j^{(k)} = A_j^T \pi^{(k)} + \lambda_j^{(k)} + \frac{y_j^{(k)} - x_j^{(k)}}{y_j^{(k)}}.$$

Since $y_j^{(k)} \geq \ell_j > 0$, and $y^{(k)} \in S(y^{(0)})$ (a bounded set) implies that $\bar{y}_{<j>}^{(k)}$ is bounded, we have:

$$\text{(III.3.6)} \qquad \hat{y}_k^* = \frac{y_j^*}{\bar{y}_{<j>}^*} = \lim_{k \in N} \hat{y}_j^{(k)} > 0. \qquad 1 \leq j \leq n.$$

Partition the indices into two sets, $I = \{j | y_j^* = \ell_j\}$, $J = \{j | y_j^* > \ell_j\}$, and let $v > \ell$, $Av = b$. Define a reaction vector r by:

$$\text{(III.3.7)} \qquad r = v - y^*.$$

Clearly, $r_j > 0$ for $j \in I$.

Multiply each equation j of (III.3.5) by r_j and sum over $1 \leq j \leq n$. The result is:

$$\text{(III.3.8)} \quad \sum_j r_j(c_j + \log \vartheta_j^{(k)}) = \Sigma\, r_j\, \delta_j^{(k)} + \Sigma\, r_j\left(\frac{y_j^{(k)} - x_j^{(k)}}{y_j^{(k)}}\right)$$

Notice that for $j \in J$, we must have for $k \in N$ sufficiently large that $y_j^{(k)} > \ell_j$, and since for each j, $\vartheta_j^{(k)} \to 0$, we will also have, for $k \in N$ sufficiently large and $j \in J$, $x_j^{(k)} > \ell_j$. But from (III.2.8) and (III.2.9) we see that for every k,

$$\text{(III.3.9)} \qquad (x_j^{(k)} - \ell_j)\delta_j^{(k)} = 0.$$

Thus for $k \in N$ large enough and $j \in J$ we have $\delta_j^{(k)} = 0$, so we can write:

$$\text{(III.3.10)} \qquad \lim_{k \in N} \delta_j^{(k)} = 0 \qquad j \in J.$$

From Lemma III.3.1 we have that:

$$\text{(III.3.11)} \qquad \lim_{k \in N} \sum_j r_j \left(\frac{y_j^{(k)} - x_j^{(k)}}{y_j^{(k)}}\right) = 0.$$

And from (III.3.6) we know that there exists a bound B such that for every $k = 1, 2, \ldots,$

$$\text{(III.3.12)} \qquad \lim_{k \in N} \sum_j r_j(c_j + \log \vartheta_j^{(k)}) \leq B.$$

Take the limit of each side of (III.3.8) and, using (III.3.10-12) we find that:

$$\text{(III.3.13)} \qquad \lim_{k \in N} \sum_{j \in I} r_j\, \delta_j^{(k)} \leq B.$$

But for each $j \in I$ we have $r_j > 0$, and certainly each $\delta_j^{(k)} \geq 0$. Thus for each $j \in I$, the sequence $\{\delta_j^{(k)}\}_{k \in N}$ is a bounded sequence, and must have a limit point. Let N_1 be a subsequence of N such that for each $j \in I$ the sequence $\{\delta_j^{(k)}\}_{k \in N_1}$ converges. Let:

$$\delta_j^* = \lim_{k \in N_1} \delta_j^{(k)} \quad \forall\, j \in I \tag{III.3.14}$$

Take the limit of each side of (III.3.5) as $k \in N_1$ tends to ∞. Note that $\{A_j^T \pi^{(k)}\}_{k \in N_1}$ must have a limit for each j, since every other term has a limit. Thus, since A has full rank m, $\{\pi^{(k)}\}_{k \in N_1}$ must have a limit.

Let:

$$\pi^* = \lim_{k \in N_1} \pi^{(k)} \tag{III.3.15}$$

Then:

$$c + \log \hat{y}^* = A^T \pi^* + \delta^*. \tag{III.3.16}$$

Looking at the bounded problem (III.2.6) we see that (III.3.16) is just the optimality condition, where $-\delta_j^*$ is the multiplier on the lower bound inequality $x_j \geq \ell_j$.

Alternatively, if we look at the expanded chemical equilibrium problem (III.2.7), and taking $-\delta_j^*$ to be the multiplier on the equation $x_j - s_j = \ell_j$ we see that (III.3.16), together with the equations:

$$\log s_j^* = -\delta_j^* \qquad 1 \leq j \leq n \tag{III.3.17}$$

form the optimality conditions for (III.2.7). Note that for $y_j^* > \ell_j$, we have $s_j^* > 0$, so $\hat{s}_j^* = 1$. In this case, (III.3.17) is satisfied, since $\delta_j^* = 0$. If $y_j^* = \ell_j$, then $s_j^* = 0$, so by the virtual mole fractions theorem II.4 of [7], we must have $\hat{s}_j^* \leq 1$. But from (III.3.17), $\hat{s}_j^* = \exp(-\delta_j^*)$, and since $\delta_j^* \geq 0$, clearly, $\hat{s}_j^* \leq 1$.

Thus $y^* \in M_\ell(F|H)$. QED.

<u>Corollary III.3.3:</u> If δ^* from Theorem III.3.2 satisfies $\delta^* = 0$, then $y^* \in M(F|H)$. That is y^* solves the original problem. In particular this will be the case if $y^* > \ell$.

<u>Proof</u>: From (III.3.16), if $\delta^* = 0$, then

$$c + \log y^* = A^T \pi^*. \tag{III.3.18}$$

This shows $y^* \in M(F|H)$. Clearly $\delta^* = 0$ if $y^* > \ell$. QED.

<u>Corollary III.3.4:</u> If $M_\ell(F|H)$ contains exactly one point x^*, then $\{y^{(k)}\}$ is a convergent sequence, and:

$$x^* = \lim_{k\to\infty} y^{(k)}.$$

<u>Proof</u>: Obvious.

4. The Other Methods

One may use lower bounds in the Augmented Quadratic Method (AQM) in a fashion similar to the way they are used in RQM, and indeed, if the successive step sizes α_k are intelligently chosen, the method will then converge. This is relatively easy to prove, because on the set $\{x \geq \ell\}$ the second partial derivatives of $F(x) = \Sigma\, x_j(c_j + d_j \log x_j)$ are bounded.

Linear methods, recall, had trouble with the original lower bounds of zero. They can have the same difficulties with nonzero lower bounds. While it might be possible to modify the direction found by a linear method so that it would not violate the bounds, this would make the method much the same as a quadratic method. There seems to be no convenient way to use lower bounds in the linear methods.

IV. DUAL METHODS

We will present two dual methods. The first, called here the Dual Augmented Linear Method (DALM), was developed by Clasen [4, 5], as a method for refining the solution obtained from his primal, or first order method, ALM.

The second dual method solves a form of the dual of the chemical equilibrium problem. The author had believed that this dual formulation was new. A closer study of Rockafellar's work in duality [13, 14, 15, 16], of the work of Duffin, et al [1] on Geometric Programming, and of a paper of Hámala and Milan [2] reveals a problem of the form of the chemical equilibrium problem may be considered to be the dual of a geometric program. Because the relation between a problem and its dual is reciprocal —i.e., the dual of the dual is the primal—so may a geometric program be considered the dual of the chemical equilibrium problem. The dual formulation presented here, for the problem without slacks, may, by a simple change of variables be shown to be a geometric program.

This dual problem requires one to maximize a linear function subject to nonlinear constraints. Experience has shown that problems with nonlinear constraints are in general more difficult and less efficient to solve than problems whose sole nonlinearity is in the objective function. Therefore, we suggest that a primal method be used first to obtain an approximate solution to a chemical equilibrium

problem perturbed by slacks and (in the case of quadratic methods) lower bounds. With the resulting values of the dual variables as a starting solution, a method may be applied to the dual problem of the original chemical equilibrium problem—unperturbed by slacks or lower bounds—to refine the solution obtained by a primal method.

1. The Dual Augmented Linear Method

From Section II.1, we know that $x > 0$, $x \in K(A, b)$ solves the chemical equilibrium problem in augmented form if and only if there exists a vector π such that:

$$\text{(IV.1.1)} \qquad \bar{c}_j + \bar{d}_j \log x_j + \bar{d}_j = \bar{A}_j^T \pi \qquad 1 \leq j \leq N.$$

For definitions of $\bar{A}$, $\bar{b}$, $\bar{c}$ and $\bar{d}$, see Section II.1

Solving (IV.1.1) for x_j, we find that:

$$\text{(IV.1.2)} \qquad x_j = \exp\left[\bar{d}_j^{-1}(\bar{A}_j^T \pi - \bar{c}_j - \bar{d}_j)\right].$$

Suppose we have initial estimates of the numbers π_i, $1 \leq i \leq M$, perhaps from the final solution from a primal method. Surely, for any π at all, the vector x computed from (IV.1.2) will satisfy the optimality conditions (IV.1.1). But such an x may not be feasible. The nature of the exponential function insures $x > 0$, but perhaps:

$$\text{(IV.1.3)} \qquad \bar{b} - \bar{A}x = \bar{g} \neq 0.$$

Substituting (IV.1.2) into (IV.1.3) shows that $\bar{g}$ may be considered a function of π. Then we wish to find π^* so that:

$$\text{(IV.1.4)} \qquad \bar{g}(\pi^*) = 0.$$

If our initial guess π is close to π^*, then, taking a Taylor's expansion of $\bar{g}$ around π to first order, it is <u>approximately</u> true that:

$$\text{(IV.1.5)} \qquad 0 = \bar{g}(\pi^*) \approx \bar{g}(\pi) + \frac{\partial \bar{g}(\pi)}{\partial \pi} (\pi^* - \pi)$$

where $\frac{\partial \bar{g}}{\partial \pi}$ is an MxM matrix whose ij^{th} element is $\frac{\partial \bar{g}_i}{\partial \pi_j}$, $1 \leq i, j \leq M$. The matrix $\frac{\partial \bar{g}}{\partial \pi}$ is computed in [5], p. 24, to be:

$$\text{(IV.1.6)} \qquad \frac{\partial \bar{g}}{\partial \pi} = -(\bar{A}XD^{-1}\bar{A}^T) = -\bar{R}$$

Substitute into (IV.1.5) to get:

$$\text{(IV.1.7)} \qquad \bar{R}\Delta\pi = \bar{g}(\pi).$$

$\bar{R}$ is the matrix that occurs in the primal linear method applied to the augmented form. See Section II.4.

The algorithm based on DALM is as follows:

0. Let the problem be in augmented form, and let π be an initial estimate of π^*.
1. Evaluate x_j by means of (IV.1.2), $\bar{g}(\pi)$ by means of (IV.1.3), and $\bar{R}$ by means of (IV.1.6).

2. Find $\Delta\pi$ satisfying (IV.1.7).

3. Replace π by $\pi + \Delta\pi$ and return to step 1.

There are two possible termination criteria. We might stop if $\bar{g}(\pi)$ computed in step 1 were sufficiently small. Alternately, we could stop if $\Delta\pi$ computed in step 2 were small enough.

This method does not in general converge. In the first place, there must exist an $x > 0$ solving the original problem, or some of the variables π must become infinite. In the second place, even when this condition is satisfied, if the initial π is too far from the optimal π^*, the method may diverge, just as any Newton's method may diverge.

2. The Dual Formulation

Since the reduced form of the problem is also written as a linear-logarithmic problem, one might think that we could develop a Dual Reduced Linear Method. However, if we substitute the appropriate quantities for $\bar{d}_j$ and $\bar{c}_j$, we discover that in the reduced case (IV.1.2) becomes:

$$\text{(IV.2.1)} \qquad x_j = \bar{x}_{<j>} \exp\,(A_j^T\pi - c_j) \qquad 1 \le j \le n.$$

In this case we have no way of defining the x_j's in terms of the π_i's, but only of expressing the x_j's as functions of π.

This does, however, lead to a dual method. In fact, it suggests a new form for the dual problem.

From [7], Theorem II.4, we know that $x \in H(A,b)$ solves the original problem (I.8) if and only if there exists a vector of virtual mole fractions $\xi > 0$ satisfying:

$$\text{(IV.2.2)} \quad \begin{cases} \text{(a)} \quad \xi_j = \hat{x}_j & \forall\ j \text{ such that } \bar{x}_{<j>} > 0 \\ \text{(b)} \quad \bar{\xi}_{<j>} = \sum\limits_{k\in<j>} \xi_k \le 1 & \forall \text{ compartments } <j> \\ \text{(c)} \quad c + \log \xi = A^T\pi & \text{for some m-vector } \pi. \end{cases}$$

Of course, this depends on the existence of $y \in H(A,b)$, $y > 0$.

Compare (IV.2.1) and (IV.2.2) (a) and (c). This suggests we consider ξ as a function of π. That is, we let:

$$\text{(IV.2.3)} \qquad \xi_j(\pi) = \exp\,(A_j^T\pi - c_j)$$

Then $\bar{\xi}_{<j>}$ defined by (IV.2.2) (b) is a function of π, and we have the constraints on π that:

$$\text{(IV.2.4)} \qquad \bar{\xi}_{<j>}(\pi) \leq 1 \qquad \forall <j>.$$

If we compute derivatives of $\xi_{<j>}(\pi)$ we find that:

$$\text{(IV.2.5)} \qquad \frac{\partial \bar{\xi}_{<j>}(\pi)}{\partial \pi} = \sum_{k \in <j>} A_k \, \xi_k(\pi) \equiv \hat{\delta}_{<j>}(\pi)$$

where the symbol $\equiv$ indicates a definition. Also,

$$\text{(IV.1.6)} \qquad \nabla^2 \bar{\xi}_{<j>}(\pi) = \sum_{k \in <j>} A_k \, \xi_k(\pi) A_k^T.$$

Since $\xi_j(\pi) > 0$ for every π, $\nabla^2 \bar{\xi}_{<j>}(\pi)$ is positive semi-definite. Thus the function $\bar{\xi}_{<j>}(\pi)$ is convex, and the constraint set defined by (IV.2.4) is a convex set.

The dual may be formulated for the problem with or without slacks (see [7], or the comment at the end of Section II.6 above). We will do it both ways, and show that the dual without slacks is a limiting case of the dual with slacks. In what follows, we assume that there exists $y \in H(A,b)$ satisfying $y > 0$, and that the solution set $M(F|H)$ of the problem without slacks is bounded and nonempty.

<u>Theorem IV.2.1</u>: Let the problem have slack $s_{<j>} > 0$ in each phase $<j>$. Let π^* solve:

$$\text{(IV.2.7)} \quad \begin{cases} \text{Max } w(\pi) = [b^T\pi + \sum_{<j>} s_{<j>} \log (1 - \bar{\xi}_{<j>}(\pi))] \\ \text{s.t.} \qquad \bar{\xi}_{<j>}(\pi) \leq 1 \quad \forall <j> \end{cases}$$

We take $\log 0 = -\infty$.

Define $x_j^* = \dfrac{s_{<j>}\,\xi_j(\pi^*)}{1 - \bar{\xi}_{<j>}(\pi^*)}$ for each j, $1 \le j \le n$.

Then x^* is the unique solution to the chemical equilibrium problem with slacks $s_{<j>}$.

<u>Proof</u>: Since $M(F|H)$ is bounded and nonempty, and since there exists $y > 0$, $y \in H(A,b)$, [7], Theorems III.1 and III.2 us that there exists a unique, strictly positive solution y^* to the problem with slacks, and an associated vector of multipliers η^*. Clearly,

$$\text{(IV.2.8)} \qquad \bar{\xi}_{<j>}(\eta^*) = \sum_{k\in<j>} \frac{y_j^*}{\bar{y}^*_{<j>} + s_{<j>}} = \frac{\bar{y}^*_{<j>}}{\bar{y}^*_{<j>} + s_{<j>}} < 1.$$

It is easy to see, then, that π^* must also satisfy $\bar{\xi}_{<j>}(\pi^*) < 1$, or else $w(\pi^*) = -\infty$.

[7], Lemma III.2.1 shows that $w(\pi)$ is a concave function of π. Thus, since π^* is surely in the interior of its constraint set, we know that:

$$\text{(IV.2.9)} \qquad \frac{\partial w(\pi^*)}{\partial \pi} = b - \sum_j \frac{A_j s_{<j>}\xi_j(\pi^*)}{1 - \bar{\xi}_{<j>}(\pi^*)} = b - Ax^* = 0.$$

Also, $x_j^* > 0$ and finite.

Finally,

$$\text{(IV.2.10)} \qquad \hat{x}_j = \frac{x_j^*}{\bar{x}^*_{<j>} + s_{<j>}} = \left(\frac{s_{<j>}\,\xi_j(\pi^*)}{1 - \bar{\xi}_{<j>}(\pi^*)}\right) \Big/ \left(\frac{s_{<j>}}{(1 - \bar{\xi}_{<j>}(\pi^*))}\right) = \xi_j(\pi^*).$$

Thus, by the definition of $\xi_j(\pi)$ (Equation (IV.2.3)) we know that:

$$c_j + \log \hat{x}_j^* = A_j^T \pi^* \qquad 1 \leq j \leq n. \tag{IV.2.11}$$

Therefore, by (IV.2.9) and (IV.2.11), x^* solves the problems with slacks. QED.

Problem (IV.2.7) is the dual problem with slacks. To show it is truly a dual, we must show that $w(\pi^*)$ is equal to the optimal value of the primal objective function.

<u>Theorem IV.2.2</u>: Let π^* solve (IV.2.7), and x^* solve the primal problem with slacks. Then:

$$F(x^*, s) = \Sigma\, x_j^*(c_j + \log \hat{x}_j^*) + \Sigma s_{<j>} \log \hat{s}_{<j>} = w(\pi^*). \tag{IV.2.12}$$

<u>Proof</u>: The quantity $\hat{s}_{<j>} = 1 - \sum_{k \in <j>} \hat{x}_k^*$

since in each compartment, the mole fractions (including that of the slack) must sum to one. Thus by (IV.2.10):

$$\hat{s}_{<j>} = 1 - \Xi_{<j>}(\pi^*). \tag{IV.2.13}$$

Now substitute (IV.2.11) and (IV.2.13) into (IV.2.12) to find that

$$F(x^*, s) = w(\pi^*). \quad \text{QED.}$$

Note that for any $x \in H(A,b)$, $F(x,s) \geq F(x^*, s)$, and for any π satisfying $\Xi_{<j>}(\pi) \leq 1 \quad \forall <j>$, $w(\pi) \leq w(\pi^*)$.

Thus weak duality is also satisfied. The theorem proves strong duality.

The dual problem without slacks is similar.

<u>Theorem IV.2.3</u>: Let π^* solve:

$$\text{(IV.2.14)} \qquad \text{Max } b^T\pi$$
$$\text{s.t. } \bar{\xi}_{<j>}(\pi) \leq 1 \; : \; \lambda_{<j>} \qquad \forall <j>.$$

Let $\lambda^*_{<j>}$ be the value of the multiplier $\lambda_{<j>}$ at the optimum.

Define $x^*_j = \lambda^*_{<j>} \, \xi_j(\pi^*)$.

Then:

$$x^* \in M(F|H).$$

<u>Proof</u>: Note that since the constraints are inequalities of the sense ($\leq$), and we are maximizing, $\lambda^*_{<j>} \geq 0$, and $\lambda^*_{<j>} = 0$ if $\bar{\xi}_{<j>}(\pi^*) < 1$.

That (π^*, λ^*) solves (IV.2.14) requires that π^* solve the unconstrained Lagrangian problem:

$$\text{(IV.2.15)} \qquad \text{Max } [b^T\pi - \sum_{<j>} \lambda^*_{<j>}(\bar{\xi}_{<j>}(\pi) - 1)].$$

Clearly, this is a concave, differentiable function that achieves its maximum where its gradient vanishes. Thus (from (IV.2.5))

$$\text{(IV.2.16)} \qquad b - \sum_{<j>} \lambda^*_{<j>}\hat{\xi}_{<j>}(\pi^*) = b - Ax^* = 0.$$

Since $\lambda^*_{<j>} \geq 0$, clearly $x^* \geq 0$. Since if $\lambda^*_{<j>} > 0$, $\bar{\xi}_{<j>}(\pi^*) = 1$, we may make the identification $\bar{x}^*_{<j>} = \lambda^*_{<j>}$, and be assured that if $\bar{x}^*_{<j>} > 0$, then:

$$\text{(IV.2.17)} \qquad \hat{x}^*_j = \xi_j(\pi^*).$$

That is, $x^* \in H(A,b)$ (because of (IV.2.16) and $x^* \geq 0$), and $\xi(\pi^*)$ is a vector of virtual mole fractions satisfying (IV.2.2) for $x = x^*$. Thus by [7], Theorem II.4, $x^* \in M(F|H)$. QED.

It is easy to show that both strong and weak duality hold in the case without slacks. The proof is much the same as for Theorem IV.2.2.

Problem (IV.2.14) is the dual problem without slacks. We now wish to show that it is a limiting case of the dual with slacks.

<u>Theorem IV.2.4</u>: Let A have full rank m. Let $\{s^{(k)}\}$ be a sequence of slack vectors satisfying $s^{(k)} > 0$, $\lim_{k \to 0} s^{(k)} = 0$. Let $\pi^{(k)}$ solve (IV.2.7), the dual with slacks $s^{(k)}$. Then

(i) $\pi^{(k)}$ is the unique solution to (IV.2.7) with slacks $s^{(k)}$

(ii) The sequence $\{\pi^{(k)}\}$ has at least one limit point (ie subsequence converging to) π^*.

(iii) π^* solves (IV.2.14), the dual without slacks.

<u>Proof</u>: Since $M(F|H)$ is bounded and nonempty, and there exists $y > 0$, $y \in H(A,b)$, [7], Theorems III.1.1 and III.1.2 show

that the primal problem with slacks $s^{(k)}$ possesses a unique, strictly positive solution $x^{(k)}$. By Theorem IV.2.1, clearly

$$\xi(\pi^{(k)}) = \hat{x}^{(k)} > 0.$$

By [7], Lemma II.11, $\xi(\pi^{(k)})$ is unique. But by definition

$$c + \log \xi(\pi^{(k)}) = A^T\pi^{(k)}. \tag{IV.2.18}$$

Since A has full rank, and $\xi(\pi^{(k)})$ is unique, $\pi^{(k)}$ is unique, proving (i).

By [7], Theorem III.1.5, there exists a subsequence N of (1, 2, ...) such that $\xi^* = \lim_{k \in N} \xi(\pi^{(k)})$ is a vector of virtual mole fractions for the problem without slacks. Since $\xi^* > 0$, we take limits as $k \to \infty$, $k \in N$, of (IV.2.18). Clearly then, $\lim_{k \in N} A^T\pi^{(k)}$ exists, and since A has full rank, $\pi^* = \lim \pi^{(k)}$ also exists, proving (ii).

Finally, by continuity,

$$\xi^* = \xi(\pi^*).$$

Since ξ^* is a vector of virtual mole fractions for the original problem, by Theorem IV.2.3, π^* solves (IV.2.14). Q.E.D.

3. Solving The Dual

A pair of vectors (π^*, λ^*) solves the dual:

$$\text{(IV.3.1)} \qquad \text{Max } b^T\pi$$
$$\text{s.t. } \bar{\xi}_{<j>}(\pi) \leq 1 \; : \; \lambda_{<j>} \qquad \forall <j>$$

if and only if they satisfy the Kuhn-Tucker conditions, derived in the usual way from the Lagrangian.

$$\text{(IV.3.2)} \quad g(\pi^*, \lambda^*) = b - \sum_{<j>} \lambda^*_{<j>} \sum_{k\epsilon<j>} A_k \xi_k(\pi^*) = 0$$

$$\text{(IV.3.3)} \quad \bar{\xi}_{<j>}(\pi^*) \leq 1 \qquad \forall <j>$$

$$\text{(IV.3.4)} \quad \lambda^*_{<j>} \geq 0, \text{ and } \lambda^*_{<j>}(1 - \bar{\xi}_{<j>}(\pi^*)) = 0 \qquad \forall <j>.$$

If instead we have an initial solution (π, λ) which may not satisfy (IV.3.2-4), we may compute corrections of (π, λ) by expanding $g(\pi, \lambda)$ and $\bar{\xi}_{<j>}(\pi)$ around the known initial point. Using always linear approximations, and making the associations:

$$\text{(IV.3.5)} \qquad x_j = \lambda_{<j>} \xi_j(\pi) \qquad 1 \leq j \leq n$$

we have from (IV.2.5) that $\Delta\pi$ and $\Delta\lambda$ must satisfy:

$$\text{(IV.3.6)} \qquad \begin{aligned} R\Delta\pi + \Sigma\hat{\beta}_{<j>}\Delta\lambda_{<j>} &= g(\pi, \lambda) \\ -\hat{\beta}^T_{<j>}\Delta\pi &\geq (\bar{\xi}_{<j>}(\pi) - 1) \end{aligned}$$

where $R = AXA^T$.

Noticing that $g(\pi, \lambda) = b - \Sigma\, \lambda_{<j>}\hat{\beta}_{<j>}$, we may let $\lambda'_{<j>}$ be the new value of the multiplier, and write (IV.3.6) as:

$$
\text{(IV.3.7)} \quad \begin{aligned}
R\Delta\pi + \Sigma\, \hat{\beta}_{<j>}\lambda'_{<j>} - b &= 0 \\
-\hat{\beta}_{<j>}\Delta\pi + (1 - \xi_{<j>}(\pi)) &= \sigma_{<j>} \\
\lambda'_{<j>} \geq 0,\ \sigma_{<j>} \geq 0,\ \sigma_{<j>}\lambda'_{<j>} &= 0 \qquad \forall <j>
\end{aligned}
$$

where we have included the complementarity conditions (IV.3.4).

Notice that (IV.3.7) is equivalent to the quadratic program,

$$
\text{(IV.3.8)} \qquad \begin{aligned}
&\text{Min } \tfrac{1}{2}((\Delta\pi)^T R\Delta\pi) - b^T\Delta\pi \\
&\text{s.t. } -\hat{\beta}^T_{<j>}\Delta\pi \geq (\xi_{<j>}(\pi) - 1) : \lambda'_{<j>} \qquad \forall <j>.
\end{aligned}
$$

Notice also that the matrix R is exactly that found in RLM. For these reasons, we call the method based on this development the Dual Reduced Quadratic Method, or DRQM.

<u>Theorem IV.3.1</u>: If $M(F|H)$, the solution set to the original problem, is nonempty, there always exists $\Delta\pi$ such that $-\hat{\beta}^T_{<j>}\Delta\pi \geq \xi_{<j>}(\pi) - 1$, for every $<j>$. That is, the quadratic program (IV.3.8) is always feasible.

<u>Proof</u>: Let $x^* \in M(F|H)$, and π^* be an associated vector of multipliers. Then clearly,

$$\xi_{<j>}(\pi^*) \leq 1 \qquad \forall <j>.$$

Since $\xi_{<j>}(\pi)$ is a convex function, we know that (from (IV.2.5)).

$$\text{(IV.3.9)} \qquad 1 \geq \xi_{<j>}(\pi^*) \geq \xi_{<j>}(\pi) + \hat{\beta}^T_{<j>}(\pi^* - \pi).$$

Thus $\Delta\pi = \pi^* - \pi$ is feasible for every π. QED.

Unfortunately, it is not always true that (IV.3.8) has a solution. The quadratic form, though always positive semi-definite, need not be strictly convex (ie positive definite). In this case, (IV.3.8) may be unbounded.

The easiest method for coping with problems of unboundedness would be to arbitrarily bound $\Delta\pi$. Other methods may suggest themselves to the reader.

The algorithm for DRQM is as follows:

0. Find an initial (π, λ). π could be, for example, the final multipliers found by a primal reduced method. $\lambda_{<j>}$ could be set equal to $\bar{x}_{<j>}$.
1. Compute $x_j = \lambda_{<j>}\xi_j(\pi)$, $\hat{\beta}_{<j>}(\pi)$, and R.
2. Solve the complementary pivot problem (IV.3.7) for $\Delta\pi$ and λ'.
3. If $\Delta\pi$ is sufficiently small or if $R\Delta\pi$ is sufficiently small, terminate. Otherwise replace λ by λ' and π by $\pi' = \pi + \Delta\pi$. Then return to step 1.

The convergence criterion "if $\Delta\pi$ is sufficiently small" is clear. Note that $R\Delta\pi = b - \Sigma\, \lambda'_{<j>}\hat{\beta}_{<j>}(\pi)$. If this is small, then $x_j = \lambda'_{<j>}\, \xi_j(\pi)$ is nearly feasible. The other quantity that one should check is $\bar{\xi}_{<j>}$. But this cannot much exceed 1 if $\Delta\pi$ is small.

4. Comparison of the Dual Methods

Simple algebra shows that the matrix $\bar{R}$, defined by (IV.1.6) can be partitioned as follows:

$$\text{(IV.4.1)} \quad \bar{R} = \left[\begin{array}{c|ccc} R & \beta_{<1>} & \cdots & \beta_{<p>} \\ \hline \beta^T_{<1>} & & & \\ \vdots & & 0 & \\ \beta^T_{<p>} & & & \end{array}\right]$$

where R is the matrix AXA^T of Section IV.3,

and each $\beta_{<j>}$ is defined, one for each compartment, by:

$$\text{(IV.4.2)} \qquad \beta_{<j>} = \sum_{k\epsilon<j>} A_k x_k.$$

If we rename the variables π_i, $m + 1 \leq i \leq M$ in DALM, calling them $\eta_{<j>}$, we have from (IV.1.7), (IV.1.3) and (IV.4.1) that we wish to solve:

$$\text{(IV.4.3)} \qquad \begin{aligned} R\Delta\pi + \sum_{<j>} \beta_{<j>}\Delta\eta_{<j>} &= b = Ax \\ \beta_{<j>}\Delta\pi &= \bar{x}_{<j>} - \sum_{k\epsilon<j>} x_k. \end{aligned}$$

Clasen [5] suggests that the right-hand sides of equations $m + 1, \ldots, m + p = M$ be replaced by zero, but the author sees no reason for doing so.

In our new notation, given multipliers (π, η) we may compute from (IV.1.2) and the definition of $\bar{A}$ (see Section II.1)

(IV.4.4) $$x_j = \exp\,(A_j^T\pi - c_j - 1 + \eta_{<j>}) \qquad 1 \le j \le n$$

(IV.4.5) $$\bar{x}_{<j>} = \exp\,(\eta_{<j>} - 1) \qquad \forall\ <j>$$

$\eta_{<j>}$ is the multiplier in the equation $\sum_{k\in<j>} x_k - \bar{x}_{<j>} = 0$.

Combining (IV.4.4) and (IV.4.5), we find that computing $\hat{x}_j$ from (IV.1.2) and computing $\xi_j(\pi)$ from (IV.2.3) yield the same result. That is,

(IV.4.6) $$\hat{x}_j = \frac{x_j}{\bar{x}_{<j>}} = \exp\,(A_j^T\pi - c_j) = \xi_j(\pi).$$

That is, comparing (IX.4.2), the definition of $\beta_{<j>}$, and (IV.2.5), the definition of $\hat{\beta}_{<j>}$, we see that:

(IV.4.7) $$\beta_{<j>} = \bar{x}_{<j>}\hat{\beta}_{<j>}(\pi).$$

Let us then make the following changes in (IV.4.3):

(i) Let $\lambda_{<j>} = \bar{x}_{<j>}$,

(ii) Let $\Delta\lambda_{<j>} = \bar{x}_{<j>}\Delta\eta_{<j>}$, and $\lambda'_{<j>} = \lambda_{<j>} + \Delta\lambda_{<j>}$,

(iii) Divide each of the last p equations of (IV.4.3) by the appropriate $-\bar{x}_{<j>}$.

Then $(\Delta\pi, \lambda')$ should satisfy:

(IV.4.8) $$R\Delta\pi + \sum_{<j>} \hat{\beta}_{<j>}\lambda'_{<j>} = b$$

$$-\hat{\beta}_{<j>}\Delta\pi \qquad\qquad = \bar{\xi}_{<j>} - 1.$$

Compare (IV.4.8) with (IV.3.7). The matrix of detached coefficients and all the constant terms are the same. The difference is that in DALM, the variable $\lambda'_{<j>}$ may

become negative, whereas in DRQM, it must be nonnegative; and that the conditions that were inequalities in DRQM are equations in DALM.

The reason for these differences lies in the fact that DATM and DRQM treat the variables $\lambda'_{<j>}$, from which the new sums $\bar{x}'_{<j>}$ are computed, differently. Making the substitutions from (i)-(iii) above, we see that in DALM,

$$\text{(IV.4.9)} \quad \bar{x}'_{<j>} = \bar{x}_{<j>} \exp(\Delta\eta_{<j>}) = \lambda_{<j>} \exp \frac{\Delta\lambda_{<j>}}{\lambda_{<j>}} \; .$$

In DRQM, on the other hand,

$$\text{(IV.4.10)} \quad \bar{x}'_{<j>} = \bar{x}_{<j>}(1 + \Delta\eta_{<j>}) = \lambda'_{<j>} .$$

That the two methods are closely related is already evident. Additional evidence of this is the fact that (IV.4.10) may be considered a first-order Taylor's approximation of (IV.4.9).

5. Degeneracy and the Dual Methods

We found that none of the primal methods could cope with a degenerate problem in a natural way. It was necessary to insure that the solution set $M(F|H)$ contained at least one strictly positive point, either by using lower bounds or slacks. It was at least helpful to insure that if the problem had a solution at all, it was unique. This was done by adding slacks.

DALM has similar difficulties. By Equation (IV.1.2) every x_j, including the sums $\bar{x}_{<j>}$, must necessarily be positive. They can approach zero only if the multipliers approach infinite values. Thus if a problem has only degenerate solutions, or if all solutions are too nearly degenerate, this method will fail.

DRQM, on the other hand, or any method used for solving the problem in its dual form, can accomodate problems with degenerate solutions. It is only necessary to insure that no vector of virtual mole fractions the method arrives at is degenerate, or nearly so. If ξ is the vector in question, then for some π,

$$c + \log \xi = A^T \pi.$$

If $\xi_j = 0$ or $\xi_j < \epsilon$ for some small $\epsilon > 0$, then one or more of the π_i must be large. Of course, ξ_j must be very small indeed before $\log \xi_j$ is a large negative number. This problem rarely occurs, and when it does it is usually

evidence that some species j has been included that has for all practical purposes no importance in the system in question.

The author suggests that the problem first be solved using one of the primal methods with slacks, and if necessary, bounds. If the slacks used are large enough, and the primal solution sufficiently accurate, the multipliers π corresponding to the final iterate of the primal procedure will satisfy:

$$\xi_{<j>}(\pi) \leq 1 \qquad \forall <j>.$$

A method for solving the dual problem may therefore start with a feasible point.

Starting with the point π described above, a method for solving the dual can find not only a solution to the original problem (i.e., without slacks or bounds), but can express all the solutions. That is, if in problem (IV.3.7), we find that $\Delta\pi = 0$, then every solution may be recovered as a solution of:

$$\sum \hat{\beta}_{<j>} \bar{x}_{<j>} = b$$

$$\bar{x}_{<j>} \geq 0$$

$$\bar{x}_{<j>}(1 - \xi_{<j>}(\pi)) = 0$$

See [7] Lemma III.2.

6. Convergence of Dual Methods

As they stand, neither dual method is assured of converging to an optimal solution π^*, although if the starting π is sufficiently close to π^* each method will generally find an optimal solution.

There are, however, other methods for solving the dual problem (IV.3.1), some of which are sure of converging, whatever the initial point. Examples of such methods include Rosen [17], Kalfon, et al. [18], Daniel [19], Fletcher & Powell [20, 21], Zoutendijk [22], several methods found in [10], particularly in a chapter by Wolfe, and Uzawa [23].

APPENDIX

The Initial Solution

Section

1. The Generalized Projection Method
2. The Linear Programming Method
3. Degenerate Cases
4. Remarks

Appendix: The Initial Solution

The primal algorithms presented in the body of this work all require a starting composition x satisfying:

$$Ax = b$$
$$x > 0.$$

Clasen [5] has developed two methods for finding such an x, the projection method and the linear programming method.

1. The Generalized Projection Method

It is actually not true that the primal algorithms require a starting composition which is feasible and strictly positive, although all of them must start with a vector $y > 0$. One may simply forget that $Ay \neq b$ and use any primal method to find a new composition x. Clasen ([5], p. 5) suggests that y be "... the exact [optimal] solution of another problem which differs from the one being considered in relatively minor ways."

If an initial infeasible but strictly positive y is used in either of the linear methods (ALM or RLM), it may be that the x obtained, although it will satisfy $Ax = b$, will not be positive. One may either admit defeat and use the linear programming method, or use a new starting point $u = \alpha y + (1 - \alpha)x$, for some $0 < \alpha < 1$ such that $u > 0$. The projection method of Clasen [5] is equivalent to using the vector $y > 0$, $Ay \neq b$ as the starting point in ALM.

The author thinks it better to use $y \geq \ell$, $Ay \neq b$ as a starting solution in one of the quadratic methods with lower bounds. The new point x that is generated will satisfy $Ax = b$ and $x \geq \ell$ so long as such an x exists.

2. The Linear Programming Method

This method can be found in Clasen [5], pp. 9–14.

If an initial guess y is not available for the projection method, or if projection has failed, one may use the L. P. method instead.

We wish to find a point x satisfying:

$$\text{(A.2.1)} \qquad \begin{aligned} Ax &= b \\ x &> 0. \end{aligned}$$

But $x > 0$ if and only if its smallest component x_j is also positive. Thus for any point x, we let:

$$v \leq \min_j x_j$$

and define new variables y by:

$$y_j = x_j - v.$$

Substituting in (A.2.1), we wish to find (y, v) such that:

$$\text{(A.2.2)} \qquad \begin{aligned} \sum_j A_j y_j + \left(\sum_{k=1}^{n} A_k\right) v &= b \\ y \geq 0, \quad v &> 0. \end{aligned}$$

Letting $Q = \sum_{k=1}^{n} A_k$, the obvious thing to do is to find the optimal solution (y^*, v^*) to:

$$\text{(A.2.3)} \qquad \begin{aligned} &\text{Max } v, \\ \text{s.t.} \quad & Ay + Qv = b \\ & y \geq 0. \end{aligned}$$

If $v^* > 0$, then the feasible composition x, defined by:

$$\text{(A.2.4)} \qquad x_j = y_j^* + v^*$$

satisfies (A.2.1).

3. Degenerate Cases:

The advantage of the L. P. method over the projection method using the quadratic algorithms is that while both will find a strictly positive feasible x if one exists, the L. P. method also discovers (i) if the matrix A has full rank (and if not which rows are linear combinations of the others), (ii) if the equations $Ax = b$, $x \geq 0$ can be satisfied at all, and (iii) if there is a feasible x but no strictly positive feasible x, which variables x_j are constrained to be zero. A discussion of these points may be found in [6], pp. 23–25.

(i) The Rank of A.

To solve the linear program (A.2.3), one must first find a basis, i.e., a set of m columns of the matrix [A, Q] which form a nonsingular matrix. It is easy to show that

since Q is a linear combination of the columns of A, [A, Q] possesses a basis if and only if A does. The matrix A has rank m if and only if it possesses a basis.

(ii) Infeasibility

If a basis can be found, then the equations $Ay + Qv = b$ are always solvable. Furthermore, there is a solution with $y \geq 0$. (This is because if $Ay = b$, we may let $v = \min_j y_j$, and let $y_j' = y_j - v$. Then $y' \geq 0$, and $Ay' + Qv = b$.)

Let (y^*, v^*) be an optimal solution of (A.2.3). If $v^* < 0$, then there is no feasible solution to the problem. That is, there is no composition $x \geq 0$ satisfying $Ax = b$. If there were, of course, we would have $y = x$, $v = 0$ a feasible solution to (A.2.3), and $v > v^*$.

(iii) Positivity

If $v^* = 0$, we must find at least one x_j constrained to be zero by the conditions

$$\text{(A.3.1)} \qquad \begin{aligned} Ax &= b \\ x &\geq 0. \end{aligned}$$

To find such an equation let λ^* be the vector of optimal Lagrange multipliers associated with the optimal solution (y^*, v^*) of (A.2.3).

From the duality theorem of linear programming,

$$\text{(A.3.2)} \qquad b^T\lambda^* = v^* = 0.$$

Since we are maximizing, each column should 'price out' negative. Thus:

$$1 - Q^T\lambda^* \leq 0 \tag{A.3.3}$$

and for each other column:

$$-A_j^T\lambda^* \leq 0. \tag{A.3.4}$$

Because $Q = \sum_{j=1}^{n} A_j$, we see from (A.3.3) that:

$$Q^T\lambda^* \geq 1 \Rightarrow A_j^T\lambda^* > 0 \tag{A.3.5}$$

for at least one j.

Let us form a new equation by taking the linear combination λ^* of the original equations (A.3.1).

$$\Sigma\, \alpha_j x_j = 0 \tag{A.3.6}$$

where $\alpha_j = A_j^T\lambda^*$, and the right-hand side comes from (A.3.2).

By (A.3.4), $\alpha_j \geq 0$ for all $1 \leq j \leq n$, and (A.3.5) insures that at least one $\alpha_j > 0$. Thus (A.3.6) shows that at least one x_j must be zero.

We may delete each column A_j from the problem such that $\alpha_j > 0$, and solve (A.2.3) again. While there is no guarantee that the new v^* for the reduced problem will be positive, we know that v^* cannot be negative. We may continue solving and deleting either until $v^* > 0$ or the number of columns is exactly m. At this point we know that (A.3.1) has a unique solution x.

The procedure never allows us to strike out a basic column, since each basic column "prices out" to zero. Thus there will always be at least m columns remaining, and the rank of the matrix can never diminish.

4. Remarks

Usually, if a strictly positive feasible solution x cannot be found, it is evidence that the problem is incorrectly formulated. The same is true if the equations $Ax = b$ are redundant, so that a basis in the L. P. method cannot be found.

The author suggests that if any of the degeneracies of Section 3 occur, the problem solver stop and examine his data, instead of dropping redundant equations or variables constrained to be zero.

REFERENCES

1. Duffen, R. J., E. L. Peterson, and C. Zener, Geometric Programming, John Wiley & Sons, 1967.

2. Hamala and Milan, "Geometric Programming in Terms of Conjugate Functions," Disc. Paper #6811, International Center for Management Science, Universite Catholique d Louvain, Belgium.

3. Shapiro, N. Z., and L. S. Shapley, "Mass Action Laws and the Gibbs Free-Energy Function," J. Soc. Indust. Appl. Math., Vol. 13, No. 2, 1965.

4. Clasen, R. J., The Linear-Logarithmic Programming Problem, The Rand Corporation, RM-3707-PR, 1963.

5. -----, The Numerical Solution of the Chemical Equilibrium Problem, The Rand Corporation, RM-4345-PR, 1965.

6. Bigelow, J. H., The Chemical Equilibrium Problem, Tech. Rep. #68-1, Operations Research House, Stanford University, 1968.

7. -----, Degeneracy in Ideal Chemical Equilibrium Problems, The Rand Corporation, P-4349, 1970.

8. Kuhn, H. W., and A. W. Tucker, "Nonlinear Programming," Proc. 2nd Berkeley Symp. on Math. Stat. & Prob., J. Newman (ed), Univ. Calif. Press, Berkeley, Calif., 1950, pp. 481-492.

9. Dantzig, G. B., and R. W. Cottle, Positive (Semi-) Definite Matrices and Mathematical Programming, Operations Research Center, Univ. of Calif., Berkeley, ORC 63-18, 1963.

10. Abadie, J. (ed.), Nonlinear Programming, North-Holland Pub. Co., 1967.

11. Dantzig, G. B., Linear Programming and Extensions, Princeton Press, 1963.

12. Bigelow, J. H., J. C. DeHaven and N. Z. Shapiro, Chemical Equilibrium Problems with Unbounded Constraint Sets, The Rand Corporation, RM-5952-PR, 1970. Also to appear in SIAM J. of Appl. Math.

13. Rockafellar, R. T., Convex Functions and Dual Extremum Problems, Ph.D. Thesis, Harvard, 1963.

14. -----, "Duality Theorems for Convex Functions," Bull. Amer. Math. Soc., Vol. 70, 1964, pp. 189-192.

15. -----, "Duality and Stability in Extremum Problems Involving Convex Functions," Pac. J. Math, Vol. 21, No. 1, 1967.

16. -----, Duality in Nonlinear Programming, presented at the Amer. Math. Soc. Seminar on the Mathematics of the Decision Sciences, Stanford Univ., 1967.

[illegible], [illegible], "The Gradient Projection Method for Nonlinear Programming. Part II. Nonlinear Constraints," J. Soc. Indust. Appl. Math., Vol. 9, No. 4, 1961.

[illegible] Falfon, P., G. Ribiere, J. C. Sogno, A Method of Feasible Directions Using Projection Operators, Institute Blaise Pascal, 23 Rue du Maroc, Paris XIX eme, France, 1968.

[illegible] Daniel, J. W., The Conjugate Gradient Method for Linear and Nonlinear Operator Equations, Ph.D. Thesis, Stanford, 1965.

[illegible]0. Fletcher, P., and M. J. D. Powell, "A Rapidly Convergent Descent Method for Minimization," Comp. J., Vol. 6, 1963.

[illegible] ------, "Function Minimization by Conjugate Gradients," Comp. J., Vol. 7, 1964.

[illegible] Zoutendijk, G., Methods of Feasible Directions, Elsevier Pub. Co., 1960.

[illegible] [illegible], "Iterative Methods for Concave Programming," Studies in Linear and Nonlinear Programming, K. Arrow, [illegible] and H. Uzawa (eds.) Stanford Press, 1958.

BEST
AVAILABLE COPY

UNCLASSIFIED
Security Classification

DOCUMENT CONTROL DATA - R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) Department of Operations Research Stanford University Stanford, Calif. 94305	2a. REPORT SECURITY CLASSIFICATION 2b. GROUP

3. REPORT TITLE

COMPUTING EQUILIBRIUM COMPOSITIONS OF IDEAL CHEMICAL SYSTEMS

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Technical Report

5. AUTHOR(S) (Last name, first name, initial)

BIGELOW, James H.

6. REPORT DATE March, 1970	7a. TOTAL NO. OF PAGES 71	7b. NO. OF REFS 23
8a. CONTRACT OR GRANT NO. N-00014-67-A-0112-0011 b. PROJECT NO. NR-047-064 c. d.	9a. ORIGINATOR'S REPORT NUMBER(S) Technical Report No. 70-4 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report) –	

10. AVAILABILITY/LIMITATION NOTICES

This document has been approved for public release and sale. Its Distribution is unlimited.

11. SUPPLEMENTARY NOTES none	12. SPONSORING MILITARY ACTIVITY Operations Research Program (code 434) Office of Naval Research Washington, D. c. 20360

13. ABSTRACT

A single- or multi-phase chemical equilibrium problem may be expressed as a nonlinear programming problem. Thus to find the quilibrium composition of of chemical system one need only minimize a particular nonlinear function (the free energy) subject to certain linear equality constraints (the mass-balance laws) and nonnegativity conditions. The free energy is defined and continuous on the nonnegative orthant of n-space. In the interior of its domain it is infinitely differentiable; but at the boundary, the directional derivative may become infinite.

The phrase "chemical equilibrium problem" refers only to a problem with a particular mathematical form. Such problems arise in many chemical problems not classically denoted as equilibrium problems. In addition, the dual to a geometric programming problem has this form.

We present six iterative methods for solving the chemical equilibrium problem, four primal and two dual. In chemical terms, each composition produced by a primal method satisfies the mass-balance laws while successive iterates more nearly satisfy the mass-action laws. Dual methods do the reverse.

We present two formulations of the chemical equilibrium problem as a more general linear-logarithmic problem, and two methods for solving the general problem. Of the four resulting primal methods, two (the Linear methods) need not converge to an optimal solution. The other two (the Quadratic methods) if applied to an appropriately modified chemical equilibrium problem, will certainly converge.

Neither dual method need converge to the optimum, unless the starting point is sufficiently close to the optimal solution. Nor has the author been able to modify the Linear dual method so that it is sure to converge. However, the Quadratic dual method is derived from a formulation of the dual problem to the chemical equilibrium problem as a geometric programming problem. Many standard methods of nonlinear programming will certainly converge when applied to this dual. Any such method allows one to ignore the possibility that the original problem has a nonunique or a degenerate solution.

DD FORM 1 JAN 64 1473