

Mass Action and Polynomial Optimization

U. PASSY AND D. J. WILDE*

Israel Institute of Technology, Haifa, Israel

(Received March 1, 1969)

An analogy between the law of mass action and generalized polynomial optimization is proven, interpreted, and illustrated numerically. The example, which involves finding optimal temperatures in three chemical reactors in series, has been solved previously by other optimization methods. In this case the theory leads to insight into the structure of the optimum rather than to improved computational procedures.

1. Introduction

There is an instructive analogy between generalized polynomial optimization and the law of mass action describing chemical equilibrium. Generalized polynomial optimization [11], [16], [17], [20], applicable to many chemical engineering design situations, replaces a highly nonlinear constrained optimization problem with another "auxiliary" one whose variables satisfy more tractable linear equations. Only rarely, however, are there as many linear equations as auxiliary variables, and this article gives, for the first time, a set of additional equations sufficient to determine solutions. These new equations, although nonlinear, each have only one variable term, whose form resembles the equilibrium expression of the law of mass action. Since the linear equations can be interpreted as the results of a dimensional analysis, the new "equilibrium conditions" complete an analogy between the original optimization problem and laws of nature well known to chemical engineers.

Section 2 states the original generalized polynomial optimization problem, a form occurring widely in chemical engineering systems. The main result, together with statement of the auxiliary problem, is given without proof in the next section. Chemical and physical analogies are developed in section 4 and verified numerically in section 5, which applies the results to a well known reactor system serial optimization problem solved previously by dynamic programming [1], and variational methods [9], [12]. It is no doubt interesting to optimization theorists that this particular problem can be formulated with generalized polynomials, for it is a prototype of many chemical reactor design problems. However, this approach shows no computational advantage in this case. Section 6 shows how the original equality constraints of the problem were replaced by inequality constraints, a surprising procedure which happens to be advantageous in polynomial optimization. The formal proof of the equilibrium conditions, an extension of the first paper on polynomial optimization [16], is reserved until the final technical section so as not to impede the development of ideas of more direct interest to chemical engineers not specializing in optimization theory.

The equilibrium conditions given here extend a close analogy between geometric programming [2], [3], [11], [17] and the fact that free energy is minimum at chemical equilibrium [4], [5], [6], [7], [8], [10], [14], [18], [19]. Just as generalized polynomial optimization is an extension of geometric programming, so the law of mass action formulation is based on the minimum free energy concept.

2. The Primal Generalized Polynomial Problem

The notation follows [20, Ch. 4]. Consider $M+1$ generalized polynomial functions $g_m(\mathbf{x})$ ($m=0, 1, \dots, M$) of N real finite positive variables x_n ($n=1, \dots, N$). That is

* Chemical Engineering, Stanford University, Stanford, California 94305.

$$g_m(\mathbf{x}) \equiv \sum_{t=1}^{T_m} \sigma_{mt} c_{mt} \prod_{n=1}^N x_n^{a_{mtn}} \quad (1)$$

with $\infty > x_n > 0$ unknown, assembled into a vector $\mathbf{x} = (x_1, \dots, x_n)$, and the following real constants given: $c_{mt} > 0$, $\sigma_{mt} = \pm 1$, and a_{mtn} unrestricted in sign. The generalized polynomial optimization problem is to find \mathbf{x}^* minimizing g_0

$$g_0(\mathbf{x}^*) \equiv \min_{\mathbf{x}} g_0(\mathbf{x}) \quad (2)$$

among all \mathbf{x} satisfying the M constraints

$$\sigma_m [g_m(\mathbf{x})]^{\sigma_m} \leq 1, \quad m = 1, \dots, M \quad (3)$$

where

$$\sigma_m = \pm 1$$

is a known constant. For technical reasons discussed elsewhere, [16] it is assumed that the Kuhn-Tucker constraint qualification [13], [15] is satisfied at \mathbf{x}^* . When all σ_{mt} and σ_m are +1, the problem is called *geometric programming* [11], [13].

3. Main Result

The solution \mathbf{x}^* to the primal problem can be found by solving a different problem involving $N+1$ linear equations, and $T - (N+1)$ single term algebraic nonlinear equations in T unknowns. Here T , being the total number of terms in the primal problem, is given by

$$T \equiv \sum_{m=0}^M T_m. \quad (4)$$

Notice that since there are exactly as many equations as there are unknowns, no optimization is involved in the secondary "auxiliary" problem. In general the auxiliary dual problem has only a finite number of solutions, provided the equations are independent. As long as the primal problem has a finite minimum, the auxiliary problem will have at least one solution.

Consider T non-negative auxiliary variables ω_{mt} , one for each term $\sigma_{mt} c_{mt} \sum_{n=1}^N x_n^{a_{mtn}}$ in the primal problem

$$\omega_{mt} \geq 0. \quad (5)$$

They are required to satisfy a single *normality condition* involving an unknown signum function $\sigma_0 (= \pm 1)$

$$\sigma_0 \sum_{t=1}^{T_0} \sigma_{0t} \omega_{0t} = 1 \quad (6)$$

as well as N linear *orthogonality conditions*

$$\sum_{m=0}^M \sum_{t=1}^{T_m} \sigma_{mt} a_{mtn} \omega_{mt} = 0, \quad \omega = 1, \dots, N. \quad (7)$$

The orthogonality conditions can be regarded as the results of the dimensional analysis of a certain function. Discussion of this interesting interpretation will, however, be deferred to a later section because it has been pointed out in other articles [3], [11].

In [16] it was proven that the following $M+1$ quantities, one for each primal generalized polynomial, must also be non-negative.

$$\omega_{00} \equiv 1 > 0 \quad (8)$$

$$\omega_{m0} \equiv \sigma_m \sum_{t=1}^{T_m} \sigma_{mt} \omega_{mt} \geq 0, \quad m = 1, \dots, M. \quad (9)$$

For reasons to be justified later, let these quantities be forbidden to vanish

$$\omega_{m0} > 0 \tag{10}$$

so that for future abbreviation the following *weights*, w_{mt} one for each primal term, can be defined.

$$w_{mt} \equiv \omega_{mt}/\omega_{m0}, \quad m = 0, 1, \dots, M. \tag{11}$$

They are called weights because they must sum to unity for each generalized polynomial.

$$\sigma_m \sum_{t=1}^{T_m} \sigma_{mt} w_{mt} = 1, \quad m = 1, 1, \dots, M. \tag{12}$$

These M conditions may be called the *weight normality conditions*, since they include the usual normality condition.

The linear normality and orthogonality conditions, assuming they are linearly independent equations, have $T - (N + 1)$ degrees of freedom, or “degrees of difficulty” as they have been termed in this context [11]. Hence the equations have $T - (N + 1)$ independent homogeneous solutions, each involving T numbers, one for each term. Let the final index $r = 1, \dots, T - (N + 1)$ distinguish between these solutions, and let the first two subscripts m and t of their individual components correspond to those of the weights w_{mt} . Then the components v_{mtr} are defined to satisfy, for every r , the homogeneous linear equations

$$\sum_{t=1}^{T_0} \sigma_{0t} v_{0tr} = 0 \tag{13}$$

$$\sum_{m=0}^M \sum_{t=1}^{T_m} \sigma_{mt} a_{mtn} v_{mtr} = 0, \quad n = 1, \dots, N. \tag{14}$$

The real numbers v_{mtr} , very easy to calculate, are combined with the primal coefficients c_{mt} to compute $T - (N + 1)$ real *equilibrium constants* K_r .

$$K_r \equiv \prod_{m=0}^M \prod_{t=1}^{T_m} c_{mt}^{\sigma_{mt} v_{mtr}}, \quad r = 1, \dots, T - (N + 1). \tag{15}$$

These constants are used to construct a third set of *equilibrium conditions* involving the weights which in principle completely determine numerical values of the auxiliary variables ω .

$$\prod_{m=0}^M \prod_{t=1}^{T_m} w_{mt}^{\sigma_{mt} v_{mtr}} = K_r, \quad r = 1, \dots, T - (N + 1). \tag{16}$$

The adjective “equilibrium” is suggested by the form of these equations, which resembles that describing the “law of mass action” for predicting equilibrium concentrations of chemicals reacting reversibly. The analogy will be discussed at length in the next section.

When a set of positive ω_{mt} and w_{mt} satisfying the normality, orthogonality, and equilibrium conditions (eqs. 12, 7 and 16) have been found, together with the corresponding value of the signum function σ , they may be used to evaluate the *auxiliary function*, given by:

$$d(\omega) \equiv \sigma \left[\prod_{m=0}^M \prod_{t=1}^{T_m} (c_{mt}/w_{mt})^{\sigma_{mt} \omega_{mt}} \right]^\sigma. \tag{17}$$

At such a point ω the auxiliary function has the same value as the primal objective function $g_0(\mathbf{x})$ at a primal point $\mathbf{x} > \mathbf{0}$ whose coordinates are related to the weights. That is,

$$g_0(\mathbf{x}) = d(\omega) \tag{18}$$

where x and w are related by

$$w_{0r} = c_{0r} \prod_{n=1}^N x_n^{a_{0rn}} / d(\omega) \quad (19)$$

and

$$w_{mt} = c_{mt} \prod_{n=1}^N x_n^{a_{mnt}}, \quad m = 1, \dots, M. \quad (20)$$

The x_n may be found from the known w_{mt} by direct multiplication and division of appropriate powers of eqs. (19) and (20). They may also be found by taking logarithms of eqs. (19) and (20), solving the resulting equations *linear* in the variables $\log x_n$, and then taking antilogarithms.

Such points x are interesting because it has been proven [16] that one of them gives the minimum value of g_0 , if there is a minimum, and hence solves the primal problem. When there are multiple solutions, one need evaluate x only where the corresponding dual solution ω gives the smallest value of $d(\omega)$.

Eqs. (19) and (20) show why the weights can be assumed not to vanish; the positivity and finiteness of the primal variables forbids this.

$$w_{mt} > 0. \quad (21)$$

This avoids difficulty with the equilibrium conditions, which could not be satisfied if any weights were zero.

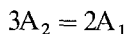
In [16] it is proven that eq. (10), which assumes the ω_{m0} do not vanish, is equivalent to requiring that the corresponding primal constraint eqs. (3) are tight, i.e., satisfied as equalities. This situation is tolerable when, as in many engineering problems, it is known in advance which constraints will be tight at the optimum. If the tightness of a constraint is doubtful, it may be deleted from the constraint set, and dual solutions computed. Any dual solution based on deleting primal constraints must have its primal point checked for feasibility. While in principle all combinations of tight and loose constraints should be examined, this is rarely necessary in practice. In the example to be described, all constraints were known to be tight, and the auxiliary equations had only one solution.

4. Physical and Chemical Analogies

The equilibrium conditions of eq. (16) lend themselves to instructive analogies with the law of mass action in chemical equilibrium. Let the weights w_{mt} be regarded as concentrations of a chemical in a phase m . The exponents v_{mtr} can be viewed as analogs of stoichiometric coefficients for a chemical reaction involving the weights w_{mtr} , with K_r as the equilibrium coefficient. Thus for example the expression

$$w_{11}^2 w_{12}^{-3} = 100$$

in which the exponents are $v_{111} = 2$ and $v_{121} = -3$ and $K_r = 100$ is mathematically analogous to the reversible chemical reaction



with A_1 and A_2 representing chemical species whose concentrations are respectively w_{11} and w_{12} . The high equilibrium constant $K_1 = 100$ would indicate that equilibrium would favor large concentration of A_1 and small concentrations of A_2 at equilibrium. In this way the equilibrium conditions, nonlinear though they are as formulated, may yield insight into the relative values of the weights and produce reasonable starting guesses for numerical procedures.

The analogy is not complete, however. For one thing, the weights may be greater than unity and hence do not resemble concentrations as expressed in mol fractions. Moreover, the expo-

nents v_{mtr} may not be integers, as required in chemistry by the law of definite proportions. Finally, all of the exponents v_{mtr} may have the same sign, which would correspond to the impossible situation of all reactants and no products, or vice versa.

The N orthogonality conditions of eq. (7) can be interpreted physically as the results of a dimensional analysis, for they restrict values of the ω to those which make the following expression, formed from the primal terms

$$c_{mt} \prod_{n=1}^N x_n^{a_{mnt}}$$

reduce to the auxiliary function $d(\omega)$ in which the x have cancelled out.

$$\sigma \left[\prod_{m=0}^M \prod_{t=1}^{T_m} \left(c_{mt} \prod_{n=1}^N x_n^{a_{mnt}/w_{mt}} \right)^{\sigma_{mt}\omega_{mt}} \right]^\sigma = \sigma \left[\prod_{m=0}^M \prod_{t=1}^{T_m} (c_{mt}/w_{mt})^{\sigma_{mt}\omega_{mt}} \right]^\sigma = d(\omega).$$

This fact, which may be verified by direct application of eq. (7), is too well known to warrant further discussion here [11], [16], [20].

The physical interpretation of the normality conditions (eq. 12) is obvious; they permit viewing the w as weights. Thus all of the auxiliary conditions have physical or chemical analogies now that the nonlinear equations, one for each degree of difficulty, have been interpreted as equilibrium conditions.

In most engineering problems, the exponents a_{mnt} are fixed by physical or geometric considerations and rarely change, whereas the coefficients c_{mt} are often derived from such transitory economic data as unit costs or prices. Hence all of the auxiliary information concerning the varying c_{mt} is concentrated in the equilibrium constants K_p , (eq. 15). In this way the equilibrium conditions, by explicitly showing how the optimal weights depend on the c_{mt} , give insight into the sensitivity of the solution.

5. Example: Three Reactors in Series

The theory will now be illustrated by application to a well known chemical engineering problem—the choice of temperatures in three continuous serial tank reactors which maximize chemical yield from the third. The desired material B is produced from A by the following sequence of first order reactions, which lead ultimately to a degradation product D .



The rates for the two reactions are respectively

$$r_1 = k_1 \exp(-E_1/RT) \tag{22}$$

$$r_2 = k_2 \exp(-E_2/RT). \tag{23}$$

These reactions occur in a series of three stirred tank reactors having the same volume. It is desired to maximize the yield of material B leaving the third reactor when the initial flow rate of materials A and B are given and the flow rate of A leaving the last reactor is fixed.

This problem, first solved by Aris [1], was chosen because it has been used frequently for demonstrating various optimization methods [9], [12], and is a prototype of many involving chemical reaction system optimization. Hence it is interesting that the same problem can be formulated with generalized polynomials and solved using the normality, orthogonality, and equilibrium conditions. Also the equality constraints of this problem can be replaced by inequalities, as required by generalized polynomial optimization theory, and still yield the proper solution. Further, through a simple transformation one can guarantee that no auxiliary weights ever exceed unity. The example, while providing a numerical check on the theory, does not, however, demonstrate any present computational advantage in the generalized polynomial formulation.

To emphasize the chemical structure of the example, the abstract use of x_n for primal problem variables will be sacrificed in favor of A_n, B_n ($n=1, 2, 3$) respectively for mol flow of materials

A and B (mols/min) and T_n for the reactor temperatures. Let the feed rates be $A_0 = 10$ mols/min and $B_0 = 0.1$ mols/min, and suppose the fixed production rate A_3 is 0.1 mols/min. Then the problem is constrained by the following mass balance equations, one for each reactor $n = 1, 2, 3$.

$$A_{n-1} = A_n + A_n k_1 \exp(-E_1/RT_n) \quad (24)$$

$$B_{n-1} = B_n + B_n k_2 \exp(-E_2/RT_n) - A_n k_1 \exp(-E_1/RT_n). \quad (25)$$

To express these transcendental functions as generalized polynomials, define, for $n = 1, 2, 3$:

$$\theta_n \equiv \exp(-E_1/RT_n). \quad (26)$$

To obtain a numerical solution, suppose that $k_1 = 10$, $k_2 = 1$, and $E_2/E_1 = 2$. A further special feature of this reactor optimization problem is that one can guarantee that no weight in the auxiliary problem will exceed unity, by introducing the following variables for $n = 1, 2, 3$:

$$C_n \equiv B_{n-1}/(B_{n-1} + A_n k_1 \theta_n). \quad (27)$$

Then the optimization problem can be written as to maximize the yield rate B_3 , or equivalently to minimize its reciprocal

$$\min B_3^{-1} \equiv \min g_0 \quad (28)$$

subject to positivity constraints obviously satisfied automatically for physical reasons

$$A_n B_n C_n \theta_n > 0 \quad (29)$$

and to the following generalized polynomial constraints derived from eqs. (24–27). Note that all these constraints must be satisfied as equalities, since they are derived from equations. Justification for replacing these equations by inequalities of the senses given is discussed in the next section. The constraints are:

$$\left. \begin{aligned} g_1 &= 0.1A_1 + A_1\theta_1 && \leq 1 \\ g_2 &= 10B_1C_1 + 10B_1C_1\theta_1^2 && \leq 1 \\ -g_3^{-1} &= -(-C_1 - 100A_1C_1\theta_1)^{-1} && \leq 1 \end{aligned} \right\} \text{Reactor 1}$$

$$\left. \begin{aligned} g_4 &= -(-A_1^{-1}A_2 - 10A_1^{-1}A_2\theta_2)^{-1} \leq 1 \\ g_5 &= B_1^{-1}B_2C_2 + B_1^{-1}B_2C_2\theta_2^2 \leq 1 \\ -g_6^{-1} &= -(-C_2 - 100A_2B_1^{-1}C_2\theta_2)^{-1} \leq 1 \end{aligned} \right\} \text{Reactor 2}$$

$$\left. \begin{aligned} g_7 &= -(-0.1A_2^{-1} - A_2^{-1}\theta_3)^{-1} \leq 1 \\ g_8 &= B_2^{-1}B_3C_3 + B_2^{-1}B_3C_3\theta_3^2 \leq 1 \\ -g_9^{-1} &= -(C_3 - B_2^{-1}C_3\theta_3)^{-1} \leq 1 \end{aligned} \right\} \text{Reactor 3}$$

The signum functions for the constraints g_{3n} ($n = 1, 2, 3$) are negative ($\sigma_{3n} = -1$), while all the rest are positive. Since the yield B_3 can only be positive, it is known in advance that $\sigma_0 = 1$, i.e., that the optimal value of the objective, and hence of the auxiliary function, must be positive. Since the signs of each of the terms are also given by eqs. (30), this completes the formulation of the primal problem.

For each of the 19 terms in eqs. (28) and (30) define an auxiliary variable ω_{mt} where in this example $m = 0, 1, \dots, 9$, while $t = 1$ for $m = 0$, and $t = 1, 2$ for $m = 1, \dots, 9$. The normality condition (eq. 6) is simply

$$\omega_{01} = 1. \quad (31)$$

The orthogonality conditions are, for the 11 primal variables listed to the left of each equation, given by eq. (7) as:

$$\begin{array}{rcl}
 A_1: & \omega_{11} + \omega_{12} & -\omega_{32} + \omega_{41} + \omega_{42} & = 0 \\
 B_1: & & \omega_{21} + \omega_{22} & -\omega_{51} - \omega_{52} & + \omega_{62} & = 0 \\
 C_1: & & \omega_{21} + \omega_{22} - \omega_{31} - \omega_{32} & & & = 0 \\
 \theta_1: & & \omega_{12} & + 2\omega_{22} & - \omega_{32} & = 0 \\
 A_2: & \omega_{71} + \omega_{72} & & -\omega_{41} - \omega_{42} & & -\omega_{62} & = 0 \\
 B_2: & & -\omega_{81} - \omega_{82} & + \omega_{92} & & -\omega_{51} + \omega_{52} & = 0 \\
 C_2: & & & & & \omega_{51} + \omega_{52} - \omega_{61} - \omega_{62} & = 0 \\
 \theta_2: & & & & -\omega_{42} & 2\omega_{52} & -\omega_{62} & = 0 \\
 B_3: & -\omega_{01} & + \omega_{81} + \omega_{82} & & & & & = 0 \\
 C_3: & & \omega_{81} + \omega_{82} - \omega_{91} - \omega_{92} & & & & & = 0 \\
 \theta_3: & -\omega_{72} & + 2\omega_{82} & - \omega_{92} & & & & = 0
 \end{array} \tag{32}$$

There are $19 - (11 + 1) = 7$ degrees of difficulty. The fact that $\sigma_m = \sigma_{m_t}$ for all $m = 1, \dots, 9$ and $t = 1, 2$ together with eqs. (8-12) imply that the weights w_{m_t} all sum to unity

$$\sum_{t=1}^{T_m} w_{m_t} = 1, \quad m = 0, 1, \dots, 9 \tag{33}$$

and that therefore they never exceed unity individually.

$$0 \leq w_{m_t} \leq 1. \tag{34}$$

Below are exhibited linearly independent column vectors v_r ($r = 1, \dots, 7$) all orthogonal to the coefficients of the normality and orthogonality conditions as required by eqs. (13) and (14). To save space, each vector is given beneath its vector symbol.

v_r	v_1	v_2	v_3	v_4	v_5	v_6	v_7
v_{01r}	0	0	0	0	0	0	0
v_{11r}	2	-1	-2	1	0	0	0
v_{12r}	-2	0	0	0	1	0	0
v_{21r}	-1	0	0	-1	0	1	0
v_{22r}	1	0	0	0	0	0	0
v_{31r}	0	0	0	-1	-1	1	0
v_{32r}	0	0	0	0	1	0	0
v_{41r}	0	1	0	0	0	0	0
v_{42r}	0	0	2	-1	0	0	0
v_{51r}	0	0	-1	0	0	1	0
v_{52r}	0	0	1	0	0	0	0
v_{61r}	0	0	0	-1	0	1	0
v_{62r}	0	0	0	1	0	0	0
v_{71r}	0	1	2	0	0	-1	-2
v_{72r}	0	0	0	0	0	1	2
v_{81r}	0	0	0	0	0	0	-1
v_{82r}	0	0	0	0	0	0	1
v_{91r}	0	0	0	0	0	1	0
v_{92r}	0	0	0	0	0	-1	0

Eqs. (15) and (16) can now be used to obtain the 7 equilibrium conditions.

$$\begin{aligned}
 w_{11}^2 w_{12}^{-2} w_{21}^{-1} w_{22} &= 0.01 \\
 w_{11}^{-1} w_{41} w_{71}^{-1} &= 100 \\
 w_{11}^{-2} w_{42}^2 w_{51}^{-1} w_{52} w_{71}^{-2} &= 100 \\
 w_{11} w_{21}^{-1} w_{31} w_{42}^{-1} w_{61} w_{62}^{-1} &= 0.01 \\
 w_{12} w_{31} w_{32}^{-1} &= 0.01 \\
 w_{21} w_{31}^{-1} w_{51} w_{61}^{-1} w_{71}^{-1} w_{72} w_{91}^{-1} w_{92} &= 1 \\
 w_{71}^{-2} w_{72}^2 w_{81}^{-1} w_{82} &= 0.01
 \end{aligned} \tag{36}$$

As a numerical check on the theory, eqs. (31), (33) and (36) were solved by the Newton-Raphson method on the Stanford B-5500 computer. No computational advantage over other optimization techniques is expected or claimed for this particular problem, which, having only two degrees of freedom in its primal form, is probably most speedily solved by a direct method ([20], ch. 7). Clasen [8] has developed a program for solving such equations, which he calls

TABLE 1

Optimal auxiliary variables

j	$\omega_{3j-2,1}$	$\omega_{3j-2,2}$	$\omega_{3j-1,1}$	$\omega_{3j-1,2}$	$\omega_{3j,1}$	$\omega_{3j,2}$
1	0.1491552	0.569085	0.6773674	0.0986109	0.0096715	0.7663067
2	0.0104917	0.0375745	0.8471265	0.1087228	0.7759782	0.1798711
3	0.0502989	0.1776384	0.8891054	0.1108945	0.9558492	0.0441507

TABLE 2

Optimum weights

j	$w_{3j-2,1}$	$w_{3j-2,2}$	$w_{3j-1,1}$	$w_{3j-1,2}$	$w_{3j,1}$	$w_{3j,2}$
1	0.207668	0.792332	0.872920	0.127079	0.0124635	0.9875364
2	0.218276	0.781723	0.886255	0.113744	0.811820	0.188179
3	0.220670	0.779330	0.889105	0.110894	0.955844	0.044150

“linear-logarithmic” system. This special problem has a unique solution given in Tables 1 and 2. Although only the first two decimal places are physically significant in view of the limited precision of kinetic data, seven decimal places have been given in case others wish to verify the computations.

Substitution of these values into eq. (17) gives a value for the auxiliary function $d=0.1406$, which can be used to generate eqs. (19) and (20). The solution of these equations is the set of

TABLE 3

Optimal primal variables

j	A_j^*	B_j^*	θ_j^*
1	2.0766	7.0037	0.3815
2	0.4531	7.6459	0.3582
3	0.1000	7.1120	0.3531

optimal primal variables given in Table 3. From these the primal objective function is computed to be 0.1406 at the optimum as predicted by eq. (18). Hence the optimum yield is $(0.1406)^{-1}$ mols/min.

6. Constraint Signs

This section proves that the constraint signum functions used in the preceding section are correct. The problem analyzed is a good example of how equations can be replaced by inequalities in many physical situations. Six out of nine signs can be obtained from the orthogonality conditions (eq. 32); finding the others requires a physical argument.

The normality condition (eq. 31), the orthogonality condition for B_3 , and the definition of ω_{80} in (eq. 9) give

$$0 \leq \omega_{80} = \sigma_8(\omega_{81} + \omega_{82}) = \sigma_8 \omega_{01} = \sigma_8 .$$

Hence $\sigma_8 = 1$. Condition C_3 gives

$$\begin{aligned} 0 \leq \omega_{90} &= \sigma_9(-\omega_{91} - \omega_{92}) \\ &= -\sigma_9(\omega_{81} + \omega_{82}) = -\sigma_9 . \end{aligned}$$

Whence $\sigma_9 = -1$. Next condition B_2 gives

$$\begin{aligned} 0 \leq \omega_{50} &= \sigma_5(\omega_{51} + \omega_{52}) = \sigma_5(\omega_{81} + \omega_{82} - \omega_{92}) \\ &\sigma_5(1 - \omega_{92}) = \sigma_5 \omega_{91} \end{aligned}$$

since $\omega_{91} < 0$, it follows that $\sigma_5 = 1$. Similar analysis, involving successive application of orthogonality conditions for C_2, B_1 , and C_1 give respectively that $\sigma_6 = -1, \sigma_2 = 1$, and $\sigma_3 = -1$ as assumed in the previous section.

The remaining signs $\sigma_1 = 1, \sigma_4 = \sigma_7 = -1$, are obtained by replacing the material balance eq. (24) for A by the inequality:

$$A_{n-1} \geq A_n(1 + k_1 \theta_n) . \tag{37}$$

To justify this, regard eq. (37) as giving the freedom to remove, from the stream entering reactor n , material A at any non-negative rate

$$\alpha_{n-1} \geq 0 .$$

Then eq. (37) becomes

$$A_{n-1} - \alpha_{n-1} = A_n(1 + k_1 \theta_1) .$$

Since the yield of B_3 can only be decreased by permitting α_{n-1} to be positive, the optimum value of α_{n-1} must be zero. Hence eq. (37) can only be satisfied as an equality at the optimum, and no generality is lost in replacing eq. (24) by eq. (37). If, on the other hand, the equality were reversed, it would never be tight at the optimum, since this sense of the inequality corresponds to being able to add arbitrary amounts of A at any reactor, increasing the yield without limit. Hence reversing the inequality is the same as deleting it entirely, which would obviously be absurd. Incidentally, this type of argument, when applied to the material balances on B , leads to $\sigma_2 = \sigma_5 = \sigma_8 = 1$ as proven in the preceding paragraph. Without the preliminary analysis described in this section, it would in principle have been necessary to examine all 2^9 possible combinations of the nine signum functions.

7. Proof of the Equilibrium Conditions

Reference [16] proves that if x^* is a primal point where $y(x)$ is minimum, then there exists an auxiliary point $\omega^* > 0$ where $d(\omega^*) = y(x^*)$ and whose coordinates satisfy the following (see eqs. 46, 47, 61, 62 and of ref. [16], for every $t = 1, \dots, T_m$ and $m = 1, \dots, M$

$$\begin{aligned} \sigma_{0t} \log(w_{0t}/c_{0t}) - \log \sigma_0 g_0 - \sum_{n=1}^N a_{0tn} \mu_n &= 0 \\ \sigma_{mt} \log(w_{mt}/c_{mt}) - \log \sigma_m g_m - \sum_{n=1}^N a_{mtn} \mu_n &= 0 \end{aligned}$$

where μ_n ($n=1, \dots, N$) are real constants that are not relevant here. Let each equation be multiplied by v_{mtr} ($r=1, \dots, T-(M+1)$), the $T-(M+1)$ solutions to eqs. (13) and (14), and let these equations be summed over all m and t . Then for all r ,

$$\begin{aligned} & \sum_{m=0}^M \sum_{t=1}^{T_m} \sigma_{mt} v_{mtr} \log(w_{mt}/c_{mt}) - \log \sigma_0 g_0 \sum_{t=1}^{T_0} \sigma_{0t} v_{0tr} \\ & - \sum_{m=1}^M \log \sigma_m g_m \sum_{t=1}^{T_m} \sigma_{mt} v_{mtr} - \sum_{m=0}^M \sum_{t=1}^{T_m} \sum_{n=1}^N \sigma_{mt} a_{mtn} v_{mtr} \mu_n = \\ & \sum_{m=0}^M \sum_{t=1}^{T_m} \sigma_{mt} v_{mtr} \log(w_{mt}/c_{mt}) = 0. \end{aligned} \quad (38)$$

Thus only the first sum remains; the second vanishes by eq. (13); the third, because the tightness of the primal constraints makes $\sigma_m g_m$ unity; and the fourth, because of eq. (14). The equilibrium conditions (Eqs. 15 and 16) follow from eq. (38) by exponentiation and reengagement.

8. Conclusions

The auxiliary problem of generalized polynomial optimization now can be interpreted by analogies with dimensional analysis and the chemical law of mass action. An example showed that an important class of highly nonlinear chemical reactor design problems can be formulated and solved using this theory. By itself, the theory would seem to require examination of all combinations, either of tight and loose inequality constraints, or of positive and negative signum functions for equality constraints. However, the example demonstrated how additional analysis in practice can determine the few combinations having physical relevance. In the example, such reasoning eliminated all but one of the 2^9 possible combinations of nine constraint signum functions. For the single case remaining, the auxiliary equations were solved without difficulty using a standard Newton-Raphson procedure.

9. Acknowledgements

The example was suggested by Dr. Samuel J. Wajc of the University of Brussels, Belgium. The research was supported by National Science Foundation grant GK-651 and by National Aeronautics and Space Administration grant NgR 05-020-073.

REFERENCES

- [1] R. Aris, *The Optimal Design of Chemical Reactors*, Academic Press, New York, 1961.
- [2] M. Avriel, *Topics in Optimization: Block Search; Applied Geometric Programming and Stochastic Geometric Programming*, Dissertation Stanford University, Stanford, California (June 1966).
- [3] M. Avriel and D. J. Wilde, Optimal Condenser Design by Geometric Programming, *Ind. Engng. Chem. Des. Dev.* (April 1967).
- [4] G. S. Bahn and E. E. Zukoski, *Kinetics, Equilibria and Performance of High Temperature Systems*, Butterworths, London (1960).
- [5] R. E. Bellman and S. E. Dreyfus, *Applied Dynamic Programming*, Princeton University Press, Princeton, New Jersey, 1962.
- [6] S. R. Brinkley, Jr., Calculation of the Equilibrium Composition of Systems of Many Constituents, *J. Chem. Phys.* 15, pp. 107-110, (Feb. 1967).
- [7] R. J. Clasen, *The Numerical Solution of the Chemical Equilibrium Problem*, the RAND Corporation Report RM-4345-PR (January 1965).
- [8] R. J. Clasen, *The Linear Logarithmic Programming Problem*, The RAND Corporation Report RM-3707-PR (June 1963).
- [9] M. M. Denn and R. Aris, Green's Functions and Optimal Systems, *I.E.C. Fundamentals*, 4, 1, pp. 7-16, 1965.
- [10] W. S. Dorn, Variational Principles for Chemical Equilibrium, *J. Chem. Phys.*, pp. 1490-1492 (1960).
- [11] R. J. Duffin, E. L. Peterson and C. Zener, *Geometric Programming*, John Wiley and Sons, Inc., New York, 1967.
- [12] L. T. Fan and C. S. Wang, *The Discrete Maximum Principle*, John Wiley and Sons, Inc., New York, 1965.

- [13] M. A. Hanson, A Duality Theorem in Nonlinear Programming with Nonlinear Constraints, *Aust. J. Stat.* 3, pp. 64–72 (1961).
- [14] J. H. Kandiner and S. R. Brinkley, Jr., Calculation of Complex Equilibrium Reactions, *I.E.C.* 42, pp. 850–855, May 1950.
- [15] H. W. Kuhn and A. W. Tucker, Nonlinear Programming, in J. Neyman (Ed.) *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley and Los Angeles, University of California Press, pp. 481–493, 1951.
- [16] U. Passy and D. J. Wilde, Generalized Polynomial Optimization, *SIAM J. Appl. Math.*, 15, 5 (Sept. 1967).
- [17] U. Passy, *Generalization of Geometric Programming: Partial Control of Linear Inventory Systems*, Dissertation Stanford University, Stanford, California (June 1966).
- [18] M. Planck, *Treatise on Thermodynamics*, Dover Publications, Inc., Third Edition.
- [19] W. B. White, S. M. Johnson and G. B. Dantzig, Chemical Equilibrium in Complex Mixtures, *J. Chem. Phys.*, 28, 751–755 (1958).
- [20] D. J. Wilde and C. S. Beightler, *Foundations of Optimization*, (Englewood Cliffs, N. J., Prentice-Hall, Inc. 1967).