

AD-A012 989

**CHEMICAL EQUILIBRIUM PROBLEMS TREATED BY GEOMETRIC AND
TRANSCENDENTAL PROGRAMMING**

Gideon Lidor

Stanford University

Prepared for:

**Office of Naval Research
National Institutes of Health
Atomic Energy Commission
National Science Foundation**

April 1975

DISTRIBUTED BY:

NTIS

**National Technical Information Service
U. S. DEPARTMENT OF COMMERCE**

**Best
Available
Copy**

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 75-8	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER AD-A012 989
4. TITLE (and Subtitle) CHEMICAL EQUILIBRIUM PROBLEMS TREATED BY GEOMETRIC AND TRANSCENDENTAL PROGRAMMING		5. TYPE OF REPORT & PERIOD COVERED Technical Report
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Gideon Lidor		8. CONTRACT OR GRANT NUMBER(s) N-00014-75-C-0267
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Operations Research Stanford University Stanford, CA 95305		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR-047-064
11. CONTROLLING OFFICE NAME AND ADDRESS Operations Research Program Code 434 Office of Naval Research Arlington, Virginia 22217		12. REPORT DATE April 1975
		13. NUMBER OF PAGES 215
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) This document has been approved for public release and sale; its distribution is unlimited.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) CHEMICAL EQUILIBRIUM GEOMETRIC PROGRAMMING DUALITY NONLINEAR PROGRAMMING FREE ENERGY TRANSCENDENTAL PROGRAMMING		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) SEE REVERSE SIDE		

PRICES SUBJECT TO CHANGE

DD FORM 1473
1 JAN 73

EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. Abstract, TR 75-8

The chemical equilibrium problem--finding the equilibrium composition of a multiphase, multicomponent system--is of interest in the study of chemical systems in general, with many potential applications in biochemistry and biomedicine. The problem can be posed as a nonlinear program, where a convex "free energy" function is minimized, subject to linear mass balance equations. There is an associated dual chemical problem, equivalent to a geometric program when the system is ideal.

This work studies the chemical duality and applies the existing theory of geometric programming to analyze and solve chemical problems.

Some general characteristics of free energy functions are developed and are used to analyze the properties of equilibrium solutions. Chemical duality is applied to formulate and solve a class of related problems which are of a different nature than the original chemical equilibrium problem.

A dual cutting-plane algorithm is adapted from a method developed for geometric programs and is tested and compared to a standard chemical equilibrium code. Geometric programming theory is extended to include forms having variables as exponents. The resulting "transcendental geometric programs" are shown to be a generalization of chemical problems, where the system is not ideal.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

320137

CHEMICAL EQUILIBRIUM PROBLEMS
TREATED BY GEOMETRIC AND TRANSCENDENTAL PROGRAMMING

BY

STEVEN LIDOR

TECHNICAL REPORT 75-8

APRIL 1975

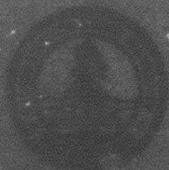
D D C
RECEIVED
JUL 30 1975
RECEIVED
B

Department of
Operations
Research

Reproduced by
NATIONAL TECHNICAL
INFORMATION SERVICE
U.S. Department of Commerce
Springfield, VA. 22151

Stanford
University

Stanford
California
94305



CHEMICAL EQUILIBRIUM PROBLEMS
TREATED BY GEOMETRIC AND TRANSCENDENTAL PROGRAMMING

by
Gideon Lidor

TECHNICAL REPORT 75-8

April 1975

DEPARTMENT OF OPERATIONS RESEARCH

Stanford University
Stanford, California

Research and reproduction of this report were partially supported by the National Institutes of Health Grant GM 14789; Office of Naval Research under Contract N00014-75-C-0267; National Science Foundation Grant MPS 71-03341 A03; and the U.S. Atomic Energy Commission Contract AT(04-3) 326 PA #18.

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

ACKNOWLEDGMENTS

I am deeply grateful to Professor Douglass Wilde, my thesis advisor, whose continuous encouragement, advice, and enthusiasm during my four years of association at Stanford were invaluable in making it all possible.

I am indebted to Professor George Dantzig for his generous support of my research and for inspiring many of the ideas in this work.

Special thanks are due to Rosemarie Stampfel for the superb typing of the manuscript and to Gail Lemmond who provided skillful secretarial assistance.

To my teachers and colleagues at Stanford, I am indebted for providing a pleasant and stimulating atmosphere.

Last but not least, to my dear parents, to whom this is the light at the end of a tunnel.

TABLE OF CONTENTS

CHAPTER		PAGE
1	INTRODUCTION	1
2	FORMULATION AND BASIC CONCEPTS	4
	2.1 General Notation Conventions	4
	2.2 The Chemical System	5
	2.3 The Mass Balance Constraints	9
	2.4 Free Energy Functions and Ideality	11
	2.5 The Chemical Equilibrium Problem.....	14
3	THERMODYNAMICS OF CHEMICAL EQUILIBRIA.....	17
	3.1 Reactions and Reaction Vectors	17
	3.2 The Mass Action Laws	21
	3.3 Systems, Properties and States	23
	3.4 Spontaneous Processes and the Gibbs Free Energy	24
	3.5 Equilibrium States	27
	3.6 Fundamental Equations	28
	3.7 Legendre Transforms and Homogeneous Functions	31
	3.8 Convexity of Free Energy Functions.....	37
4	GEOMETRIC PROGRAMMING AND CHEMICAL DUALITY	44
	4.1 Primal Geometric Programs	44
	4.2 Dual Geometric Programs	46
	4.3 Duality Theory of Geometric Programming.....	47
	4.4 Transformed Primal Programs	50
	4.5 Chemical Duality	51
	4.6 Interpreting the Dual Chemical Problem.....	54
	4.7 Equivalent Formulations of the Chemical System	55
	4.8 The Physical Significance of the Dual Chemical Problem	61

TABLE OF CONTENTS Cont.

CHAPTER		PAGE
5	PROPERTIES OF EQUILIBRIUM SOLUTIONS	65
	5.1 Introduction and Notation	65
	5.2 Some Equilibrium Set Characteristics.....	67
	5.3 Reduction of Degenerate Systems	72
	5.4 Uniqueness of Solutions	76
	5.5 Extensions to Nonideal Systems	81
6	APPLICATIONS OF CHEMICAL DUALITY	86
	6.1 Testing and Characterization of Equilibrium	86
	6.2 Verification of the Model	88
	6.3 Sensitivity Analysis	92
	6.4 Goaling Problems	99
7	A DUAL CHEMICAL ALGORITHM	108
	7.1 Introduction and Review of Existing Methods..	108
	7.2 The Geometric Inequality and Condensation....	110
	7.3 Linearization of Condensed Programs	113
	7.4 Convex Cutting Plane Algorithms	119
	7.5 Generation of Cutting Planes	121
	7.6 The MDA Algorithm	124
	7.7 Computational Aspects of the Algorithm.....	127
	7.8 Comparative Test Results	133
8	TRANSCENDENTAL GEOMETRIC PROGRAMMING	141
	8.1 Introduction and Formulation	141
	8.2 Necessary Conditions for Minima	146
	8.3 Duality in Transcendental Programs.....	151
	8.4 Some Properties of Transcendental Programs...	161
	8.5 Condensing Transcendental Programs.....	164
	8.6 Solving Transcendental Programs by Condensation	166
	8.7 Nonideal Systems and Transcendental Programming	168
APPENDIX A	TEST PROBLEMS AND SOLUTIONS	176
REFERENCES	203

CHAPTER 1

INTRODUCTION

Predicting the chemical composition of a multi-phase, multi-component system at equilibrium is of much practical, as well as theoretical, interest. This can be posed as a nonlinear programming problem in which a convex function, called the free energy, is minimized, subject to linear mass balance equations in nonnegative variables. There is an associated dual problem, equivalent to a geometric program. The main purpose of this work is to study this "chemical duality" and apply the existing theory of geometric programming to analyze and solve chemical problems.

Willard Gibbs [30], who developed in 1876 the fundamental concept of free energy, paved the way for a systematic study of chemical equilibrium. For many years the inability to handle large complex systems manually and lack of theory dealing with minimization problems, prevented the use of his powerful elegant ideas.

In the last thirty years, rapid developments of both optimization theory and computers proved the usefulness of Gibbs' ideas and their superiority over earlier methods involving solution of nonlinear

equations. Following the work of Brinkley [10], a serious effort began at the RAND Corporation in the fifties, contributing applications, theory, and computational methods. Notable among these are the works of Bigelow [8,9], Clasen [12-14], Dantzig [16-18], and Shapiro [44-47], whose theoretical developments form a framework to our work.

Geometric programming theory was developed in the sixties by Duffin, Peterson and Zener [25]. The relation of Geometric Programming to chemical equilibrium was shown by Avriel [4], and Passy and Wilde [40-41].

In addition to well known applications of chemical equilibrium to the study of chemical reactions in chemical process design, rocket propellants, and the oil industry, there are potentially many applications in biochemistry and medicine. Biological systems are in many cases closer to chemical "ideality" than are commercial inorganic processes, but their size and complexity were until recently beyond solution by available computational and analytical methods. Very little is known today about the intricate chemistry of many physiological functions, for example, the way medications react in the human body. This work is motivated in part by the need to analyze these and related systems.

Chapter 2 reviews the formulation of the problem. Chapter 3 studies thermodynamic concepts, deriving general properties of free energy functions from basic principles. Geometric programming and

the dual chemical problem are introduced in Chapter 4, with a new chemical interpretation of the dual variables. In Chapter 5 we apply the chemical duality and the results of Chapter 4 to study the nature of equilibrium solutions. It is shown that uniqueness of solutions is insured under regularity conditions relating to the dual problem. Geometric programming methods can also handle degenerate systems and degenerate solutions.

The theoretical discussion is followed, in Chapter 6, by several applications of duality in characterization and formulation of chemical problems. Among the results is a generalization of a "Goaling" method, in which side conditions are added to the problem and some of its mass balance equations are relaxed.

Chapter 7 presents a dual algorithm for solving the chemical equilibrium problem. The method, adapted from an algorithm proposed by Dembo [20], is shown to be a convex cutting plane algorithm. It solves a transformed geometric program through a sequence of "condensed" linearized programs. Comparative results of tests with this algorithm are presented, together with a small collection of test problems, which appear in an appendix.

The last chapter presents a new extension to geometric programming, which includes variables as exponents. This extension, called "transcendental geometric programming" was primarily developed because of its potential applications to engineering problems. However, a dual problem to the transcendental program is shown to be an interesting generalization of the chemical equilibrium problem, one particularly suitable for treating nonideal systems.

CHAPTER 2

FORMULATION AND BASIC CONCEPTS

This chapter introduces notation and terminology used throughout this work. The chemical system is described, together with the concepts required for mathematical formulation of the chemical equilibrium problem.

2.1 General Notation Conventions

We denote by \mathbb{R} the real line; \mathbb{R}^n an n -dimensional Euclidean space, and \mathbb{R}_+^n its positive orthant. There will be no distinction in notation between scalars and vectors since the dimensions will be clear from the context. Vectors are assumed to be column vectors. The transpose of a vector x is denoted x' . However, the prime will be omitted at times, when no ambiguity arises. Inner products will be written xy and sometimes $x' \cdot y$, where x and y are vectors. Matrices will always be denoted by capitals. If A is a matrix, A_j will denote its j -th column; A_i its i -th row. The symbol $\langle k \rangle$ denotes a set of consecutive integers belonging to the k -th partition of some integer set N .

References are enclosed in square brackets []. The material is divided into chapters; the chapters into sections. Theorems, lemmas, propositions and other major items are numbered sequentially within each chapter in the form k.n where k is the chapter number. Other items and equations are numbered sequentially within each section by a single number. Internal reference to equation i is by (i) if the equation is in the same section, (j.i) if it is in section j of the same chapter and (k.j.i) if it is in chapter k section j.

2.2 The Chemical System

The description below is based mostly on the notation of Shapiro and Shapley [47]. Some changes were introduced to adapt the formulation to Geometric Programming conventions.

The system under consideration is composed of a finite number (K) of homogeneous phases denoted $\phi_1, \phi_2, \dots, \phi_K$. We shall sometimes refer to phase ϕ_k simply as phase k. A phase is homogeneous when every part of it has the same pressure, temperature and chemical composition. By this definition, a homogeneous phase need not occupy a contiguous space in the system. For example, oil drops in water can be considered a homogeneous phase.

Each phase is a subset of a set of chemical species. The species are those chemical compounds which one can expect to find in the system. The question of which compounds should be considered

when constructing a model is best answered on the basis of experience and judgment. In general the choice of species is up to the designer of the model. The mathematical analysis of the system assumes a given set of species S_1, S_2, \dots, S_n , and the results may be mathematically valid even if the system is chemically unsound.

For our purposes the species will be partitioned in the phases, that is, each species belongs to one, and only one phase. Thus a chemical compound having the same chemical formula but appearing in two different phases is considered as two different species. We shall denote the fact that species S_j appears in phase Φ_k by $S_j \in \Phi_k$ or equivalently $j \in \langle k \rangle$, where

$$\langle k \rangle \equiv \{j | S_j \in \Phi_k\}$$

denotes the set of indices of species in Φ_k . The "phase to which S_j belongs" is denoted by $\Phi_{k(j)}$. Thus we can write $k(j) = \{k | j \in \langle k \rangle\}$. If $l \in \langle k(j) \rangle$ then S_l and S_j belong to the same phase.

The quantity of species S_j in the system is denoted by x_j , measured in moles. (x_j is thus proportional to the number of molecules of S_j in the system.) The values of x_j are usually unknown and part of the problem to be solved is to find the vector $x = (x_1, x_2, \dots, x_n)$ of chemical composition when the system is at equilibrium. The composition-vector x is partitioned into the phases in the same way as described for the corresponding species, so that we can describe the

composition of phase ϕ_k by the vector $x_{\langle k \rangle}$. For simplicity we will omit the brackets and write x_k . It will always be clear from the context whether the reference is to a vector or to a single component of a vector.

We define the integer sets

$$\begin{aligned}\bar{K} &\equiv \{1, 2, \dots, K\} && \text{(the phase indices)} \\ N &\equiv \{1, 2, \dots, n\} && \text{(the species indices)}\end{aligned}$$

The total moles of all species in a phase is denoted by

$$\bar{x}_k \equiv \sum_{j \in \langle k \rangle} x_j \quad k \in \bar{K} \quad (1)$$

and the mole fraction of species S_j (in $\phi_{k(j)}$) is denoted by

$$\hat{x}_j \equiv x_j / \bar{x}_{k(j)} \quad \text{when } j \in N \text{ and } \bar{x}_{k(j)} > 0 \quad (2)$$

The mole fraction serves as a dimensionless concentration. Note that the concentration x_j is defined only when $\bar{x}_{k(j)} > 0$, namely, when the phase actually exists.

Each of the species in the system is composed of a set of basic units called subspecies which are usually smaller chemical units like atoms, ions or radicals. In general, the subspecies are units

which do not decompose into any smaller units in the given system. The choice of subspecies, much like the selection of species to be present in the model, is a matter of experience and judgment. However, once the species are specified, the set of subspecies must provide for every species so that each species is composed of one or more of the given subspecies, denoted by B_1, B_2, \dots, B_m .

The composition of one mole of species S_j can now be described by the "equation"

$$\sum_{i=1}^m a_{ij} B_i = S_j \quad j \in N \quad (3)$$

here a_{ij} is the number of moles of B_i in one mole of S_j . The constants a_{ij} form a matrix $A \in \mathbb{R}^{m \times n}$. The elements of A need not all be nonnegative, although in simple cases they usually are. With each species S_j we can associate a column vector of A as indicated by (3). This vector A_j is called the formula vector of S_j .

With the notation developed so far we can now describe the chemical structure of the system by the triplet $\{B, S, \Phi\}$ where B is an m -vector of subspecies, S is an n -vector of species and Φ is a K -vector of phases which induces a partition of the set N .

2.3 The Mass Balance Constraints

Except where noted otherwise, we shall assume that the chemical system is closed, that is, there is no flow of matter in or out of the system. The contents of the system are thus determined by the initial input to it, usually measured in terms of moles b_i of the subspecies B_i .

Although the actual input is usually implemented by introducing certain species, nothing is lost by assuming that these were decomposed to their subspecies which were then separately introduced. Species can decompose or form via chemical reactions and there is no law of conservation for moles of species. Subspecies on the other hand, by their definition cannot form or decompose so their quantities must be conserved, regardless of the species in which they appear.

Given the initial amount b_i of B_i , $i \in M \equiv \{1, 2, \dots, m\}$ we can write for any composition vector x

$$\sum_{j=1}^n a_{ij} x_j = b_i \quad i \in M \quad (1)$$

These relations, written in matrix form

$$Ax = b$$

are called the mass balance constraints. The composition vector x cannot have any negative components so we require naturally that

$$x \geq 0 \quad (2)$$

Some remarks concerning the matrix A are in order here. Although a_{ij} is the "amount of B_i in S_j " it need not represent an actual physical quantity. It reflects only the structure of S_j with the given set of B_i 's. Sometimes it is convenient to define the B_i 's in a way which assigns a negative value to a_{lj} for some l and j . In these situations, the rest of the B_i 's are defined in such a way that B_l is implicitly contained in some of them, so that equation (3) is more of an accounting device than a meaningful physical relation. In some models additional constraints reflecting electroneutrality are included in the matrix A . This has the effect of introducing a new subspecies (electrical charge) which normally has a quantity of 0 on input (i.e., $b_i = 0$ for the charge).

The previous remarks notwithstanding, one can always generate completely arbitrary matrices A , and associate with them abstract chemical systems, as opposed to real chemical systems. We shall be interested mostly in the latter, but the former have useful applications since they are dual to geometric programs, as will be shown in Chapter 4.

2.4 Free Energy Functions and Ideality

Although most of our chemical and mathematical ingredients are at hand, nothing has been said so far about the mathematical formulation and characterization of equilibrium. Traditionally, equilibrium conditions were defined in terms of equilibrium constants and the "Mass Action Laws" to be described in the following section. Gibbs [30] defined a potential function F related to the chemical potential of a system and showed that equilibrium is attained when F reaches its minimum. The function F , called the free energy function, will be described in more detail in the next chapter. For the time being we note that in general, F has the form

$$F(T, P, x) = RT \sum_{j=1}^n x_j \mu_j(T, P, x) \quad (1)$$

Here

P is the total pressure

R is the gas constant

T is the absolute temperature

x is the composition vector

$\mu_j(T, P, x)$ is a function called the chemical potential of S_j .

For most of our analysis we assume that the system is maintained under fixed pressure and temperature, so that T and P are constants. For convenience we then eliminate the constant RT and redefine the dimensionless free energy at fixed T and P

$$F(x) = \sum_{j=1}^n x_j \mu_j(x) \quad (2)$$

The assumption of fixed temperature and pressure is usually made when the system is free to expand into the atmosphere surrounding it and is in contact with a large heat reservoir (e.g. the surrounding air). This is especially true of biological systems which maintain constant temperatures through internal mechanisms, and cannot tolerate large variations in either temperature or pressure.

From the form (2) of the free energy function it is clear that it can be written as the sum of free energies of the species in the system. Denoting the free energy of S_j by $f_j(x)$ we write

$$F(x) = \sum_{j=1}^n f_j(x)$$

and

$$f_j(x) = x_j \mu_j(x)$$

The function $\mu_j(x)$ can now be interpreted as the partial molar free energy of S_j . For $x_j > 0$ we can assume with no loss of generality that for any P, T and x

$$\mu_j(P, T, x) = c_j(P, T, x) + \log \hat{x}_j \quad (3)$$

The functions $c_j(P, T, x)$ can (and often do) have a rather complex form. However, much of our attention will be focused on ideal systems in which $c_j(P, T, x)$ is independent of x . Under the hypothesis of fixed P and T , the functions $c_j(\cdot)$ become constants c_j , and we have for $x > 0$

$$F(x) = \sum_{j=1}^n x_j (c_j + \log \hat{x}_j) \quad (4)$$

The assumption of ideality is by no means a mere theoretical construction. Many systems, especially those under fairly low pressures and moderate temperatures, exhibit ideal behavior over some range of P , T and x . Most biological systems satisfy these conditions and can therefore be treated as ideal systems. Dantzig et al. [17], who modelled the human respiratory system, obtained results which are in excellent agreement with experimental data by using an ideal free energy of the form (4).

Equation (3) is not the only form expressing the partial molar free energy [44]. In chemical thermodynamics the most widely used (dimensional) equivalent of (3) is

$$\bar{\mu}_j(P, T, x) = \bar{\mu}_j^0(P, T) + RT \log a_j \quad (5)$$

where a_j is the activity of S_j , which usually depends on P , T and x . The bar indicates that μ is a dimensional function.

Dividing (5) by RT , it is easy to verify that our previous definition of ideality implies

$$a_j = \gamma_j(P, T) \cdot \hat{x}_j \quad (6)$$

Here γ_j is called the activity coefficient. In the general (nonideal) case γ_j is a function of P , T and x . Some of the more fundamental properties of free energy functions will be developed and analyzed in the next chapter.

2.5 The Chemical Equilibrium Problem

We can now state the mathematical form of the chemical equilibrium problem, assuming that the system is ideal and under fixed pressure and temperature.

Problem CPI

$$\begin{aligned} \text{Minimize} \quad & F(x) \equiv x' \cdot (c + \log \hat{x}) & (1) \\ \text{subject to} \quad & Ax = b \\ & x \geq 0 \end{aligned}$$

Here $x \in \mathbb{R}^n$, $c \in \mathbb{R}^n$, $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$. The notation $\log \hat{x}$ represents a vector with components $\log \hat{x}_j$. When $x_j = 0$, $\log \hat{x}_j$ is undefined. To maintain continuity of F over the nonnegative orthant of \mathbb{R}^n , we define

$$x_j \log \hat{x}_j = 0 \quad \text{when } x_j = 0 .$$

With this definition F is well defined and "right continuous" over its domain. Note that \hat{x} implicitly partitions the species into the phases. The matrix A is partitioned into submatrices

$$A_{\langle k \rangle} \equiv (A_j | j \in \langle k \rangle) \quad k \in \bar{K} \quad .(2)$$

$A_{\langle k \rangle}$ contains the formula vectors of species in phase ϕ_k .

In the general case, under fixed temperature and pressure the problem is

Problem CPN

$$\begin{aligned} \text{minimize} \quad & F(x) \equiv x' \mu(x) & (3) \\ \text{subject to} \quad & Ax = b \\ & x \geq 0 \end{aligned}$$

where we define $x_j \mu_j(x) = 0$ when $x_j = 0$ to maintain continuity.

It is assumed that $\mu_j(x)$ is well defined for $x_j > 0$.

The problem of finding the equilibrium composition for a system $\{B, S, \phi\}$ can now be represented by the triplet $\{F, A, b\}$ where F is

the free energy function, A is a (partitioned) $m \times n$ matrix and b is an m -vector.

We define the set of possible solutions

$$X \equiv X(A, b) \equiv \{x \in \mathbb{R}^n \mid Ax = b, x \geq 0\}$$

A composition-vector x is said to be feasible if $x \in X$.

CHAPTER 3

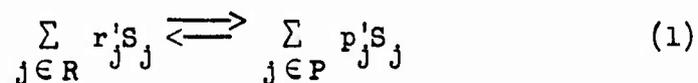
THERMODYNAMICS OF CHEMICAL EQUILIBRIA

There are two basic approaches to the study of chemical equilibrium; one dealing with processes, the other dealing with states. In this chapter we shall first review some notions of the classical "process" approach and then develop general properties of free energy functions based on the "state" approach.

3.1 Reactions and Reaction Vectors

Chemical species usually react to produce other species. In principle, all chemical reactions taking place in a system proceed simultaneously, each at some characteristic reaction rate which depends on temperature, pressure and concentrations of the species in the system. Equilibrium is reached when no observable change with time occurs in the system, in other words, when reaction rates are all zero.

A typical chemical reaction can be expressed schematically by



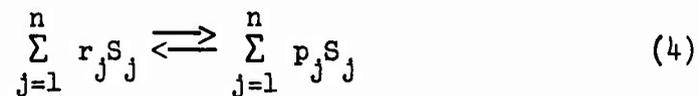
Here R and P are the sets of reactants and products respectively; r'_j and p'_j are the stoichiometric coefficients of the reactants and products respectively. The double arrow indicates a possible reaction in the reverse direction.

Letting

$$r_j = \begin{cases} r'_j & j \in R \\ 0 & j \notin R \end{cases} \quad (2)$$

$$p_j = \begin{cases} p'_j & j \in P \\ 0 & j \notin P \end{cases} \quad (3)$$

we can write



The relation above can be expressed as a formal equation by noting that S_j is represented by a column vector A_j via (2.2.3). Let $\theta_j = p_j - r_j$ for all $j \in N$. From (4) we get

$$\sum_{j=1}^n \theta_j s_j \rightleftharpoons 0$$

which can be converted to equations of mass conservation

$$\sum_{j=1}^n \theta_j A_j = 0$$

or for short

$$A\theta = 0 \tag{5}$$

The vectors $\theta \in \mathbb{R}^n$ satisfying (5) are called reaction vectors.

By this definition, the set of all possible reaction vectors is the null space of A.

Not all possible reactions implied by (5) can occur in a system with a given composition x . If $x_j = 0$, species S_j cannot be a reactant, so that only reaction vectors having $\theta_j \geq 0$ are possible. We are thus led to the definition of feasible reaction vectors at a given composition x , as those reaction vectors θ , for which $\theta_j \geq 0$ whenever $x_j = 0$.

For $x \in X$ and θ feasible at x we can find $\epsilon > 0$ sufficiently small so that for $|\theta| < \epsilon$

$$A(x + \theta) = b, \quad x + \theta \geq 0$$

Hence

$$x + \theta \in X$$

This observation identifies feasible reaction vectors as feasible directions for changes in composition. When a reaction characterized by θ takes place, x changes in the direction θ . A system for which there exists at least one nonzero vector θ satisfying (5) is said to be reactive. This work deals only with reactive systems.

As mentioned earlier, there is no loss of generality in assuming that the subspecies B_i are elements, in which case $A \geq 0$. (For the moment we ignore electroneutrality equations.) With this assumption we can state the following intuitive result.

Proposition 3.1: For a real chemical system, the set $X \equiv \{x \in \mathbb{R}^n \mid Ax = b, x \geq 0\}$ is bounded.

Proof: We may assume that X is nonempty, and that $A \geq 0$.

If X were unbounded, there would exist an $x^0 \in X$ and a nonzero reaction vector $\theta \geq 0$ such that

$$x^0 + \lambda\theta \in X \quad \text{for all} \quad \lambda \geq 0$$

Since the matrix A would have at least one positive element in each column, the equation $A\theta = 0$ would imply $\theta = 0$, a contradiction. \square

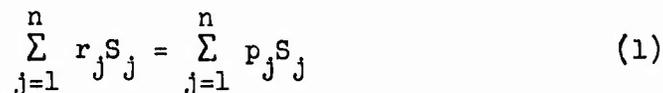
This result will be useful in the study of properties of equilibrium sets.

Corollary 3.2: Any nonzero reaction vector must have at least one positive and one negative component.

Intuitively, the corollary states that a species can form by reaction only if at least one other species decomposes.

3.2 The Mass Action Laws

The classical approach to chemical equilibrium considers reactions as "reversible" processes proceeding both forward and backward simultaneously [31]. For the reaction



the forward rate ρ_f is given by

$$\rho_f = k_1(r) \prod_{j=1}^n \hat{x}_j^{r_j} \quad (2)$$

Similarly, the backward rate is

$$\rho_b = k_2(p) \prod_{j=1}^n \hat{x}_j^{p_j} \quad (3)$$

Here $k_1(r)$ and $k_2(p)$ are the rate constants for the forward and backward reactions respectively, the system being assumed ideal.

Equilibrium is attained when the forward and backward rates are equal, so that no net change is observed with time. Therefore, at equilibrium we must have

$$k_1(r) \prod \hat{x}_j^{r_j} = k_2(p) \prod \hat{x}_j^{p_j}$$

Using the definition of θ from the previous section we have at equilibrium

$$\prod_{j=1}^n \hat{x}_j^{\theta_j} = \frac{k_2(p)}{k_1(r)} \equiv K(\theta) \quad (4)$$

Equation (4) is called the mass action law. It has to be satisfied for every feasible reaction vector θ at equilibrium. It can be shown [7,47] that (4) is a necessary, as well as sufficient condition for x to be an equilibrium composition of an ideal system when $x \in X$.

The mass action laws are closely related to the Kuhn-Tucker Conditions [36] for the minimum of $F(x)$ in problem CPI. We shall show in the next chapter that (4) appears in a geometric program associated with CPI.

When the matrix A has full rank--(m), the dimension of its null space is $n-m$. Thus there are no more than $n-m$ linearly independent reaction vectors.

Taking the logarithm of (4) we have

$$\theta \log \hat{x} = \log K(\theta) \quad (5)$$

where $\theta_j \log \hat{x}_j \equiv 0$ when $x_j = 0$.

Equation (5) indicates that $\log K(\theta)$ is a linear function of θ , and therefore (5) represents a linear system of equations in \hat{x} with coefficients θ from the null space of A . By the preceding remark, this system has at most $n-m$ independent equations, which together with the mass balance equations (2.3.1) determine x .

3.3 Systems, Properties, and States

Much of the terminology used in thermodynamics refers to observable physical entities and to measurable quantities. Thus many of the definitions are operational rather than conceptual.

The following definitions, though not completely precise mathematically, will serve our purposes.

- (i) A system is any collection of well defined physical entities.
- (ii) A property is any measure of the system.
- (iii) A state of a system is the collection of values of all its properties.

For a given chemical system the temperature and pressure are properties, as are the location, direction, and velocity of every molecule in the system. It would seem that a specification of the state would require an almost infinite amount of information.

Fortunately, only a limited number of properties are of interest in thermodynamic studies. Moreover, experience has shown that these properties are not independent of each other. For most purposes then, only a small set of independent properties need be specified in order to determine all other properties of interest and the state. These observations lead to formulating equations of state which relate independent and dependent properties. In what follows we occasionally use the term variables for properties. There are two types of variables:

- (i) Intensive variables (e.g. pressure P and temperature T) which are independent of the mass of the system.
- (ii) Extensive variables (e.g. volume V , energy U , composition x) which are mass-dependent. It will be shown later that this classification is significant in duality relations.

One point to be emphasized about properties and states is that they refer to points stationary in time. It is therefore appropriate to speak of properties as "state variables," in contrast to quantities measured during transition like heat flow into the system or work done by the system.

3.4 Spontaneous Processes and the Gibbs Free Energy

In the study of chemical processes, one often faces the question: "Will this process (or reaction) occur spontaneously under given conditions?". The second law of thermodynamics provides a quantitative

measure, the entropy function S , to answer this question. One consequence of this law is that every spontaneous process is accompanied by a net increase in entropy for the system and its surroundings. This criterion, though theoretically valid, is nevertheless of little practical use since it requires some knowledge of the surroundings. The free energy functions F and A introduced by Gibbs and Helmholtz resolve this difficulty. In terms of other thermodynamic properties, F is defined by

$$F \equiv U + PV - TS \quad (1)$$

Here U is the internal energy of the system, P , T , and V are the pressure, temperature, and volume respectively. Similarly,

$$A \equiv U - TS \quad (2)$$

Note that both F and A are extensive variables since U and S are.

We shall assume that the system is closed, under fixed temperature T and pressure P , and that it undergoes a process from state 1 (U_1, V_1, S_1) to state 2 (U_2, V_2, S_2). Following Denbigh [21], we let q be the heat absorbed by the system; w , the work done by it in the process. The first law of thermodynamics states that

$$U_2 - U_1 = q - w \quad (3)$$

The second law requires that for every spontaneous process

$$q \leq T(S_2 - S_1) \quad (4)$$

From (1), (3) we have

$$F_2 - F_1 = q - w + P(V_2 - V_1) - T(S_2 - S_1)$$

And by (4)

$$F_2 - F_1 \leq -w + P(V_2 - V_1) \quad (5)$$

The work w can be partitioned into pressure-volume work $P(V_2 - V_1)$ and other (chemical) work w' so that

$$w = P(V_2 - V_1) + w'$$

$$F_2 - F_1 \leq -w'$$

Since no external work is done on the system, $w' \geq 0$, whence

$$F_2 - F_1 \leq 0 \quad (6)$$

The inequality in (4) and (6) is strict when the process is irreversible as are all real processes. We conclude, therefore, that every spontaneous process in a closed system at fixed T and P is accompanied by a decrease in its free energy. (A similar analysis holds for A under fixed T and V .)

The free energy during the process can be described as a function of time t from the start of the process. $F(t)$ is thus a decreasing function. Taking the limit

$$\lim_{\Delta t \rightarrow 0} \frac{F(t + \Delta t) - F(t)}{\Delta t} = \frac{dF(t)}{dt}$$

we conclude from (6) that for a real, spontaneous process

$$\frac{dF}{dt} < 0 \quad (7)$$

This conclusion is extensively treated by Bigelow [7], who calls it the "global least action principle."

3.5 Equilibrium States

From a theoretical, as well as computational point of view, there are several advantages to the "state" approach over the "process" approach. Following Callen [11], we present here such a state approach and derive some new mathematical consequences of its postulates.

The first and second laws of thermodynamics will be replaced here by two postulates on simple systems. A simple system is a homogeneous, uncharged, and chemically-inert system which is not acted upon by any external fields like gravity or magnetic fields.

Postulate 1: There exist states, called equilibrium states, of simple systems in which the system is completely characterized macroscopically by its internal energy U , its volume V and its composition vector x .

For the second postulate we define a composite system to be a collection of two or more simple systems. A composite system is said to be closed if no flow of energy or matter can occur through its boundaries and if its volume is fixed. All restrictions on

flow of energy or matter between the simple systems in a closed composite system are called internal constraints.

Postulate 2: There exists a function $S = S(U, V, x)$, called the entropy of a composite system, defined for all equilibrium states and having the following properties:

- (i) S is differentiable
- (ii) S is monotone increasing in U
- (iii) S is additive over the constituent simple systems
- (iv) The values assumed by the extensive variables (U, V, x) in the absence of internal constraints, maximize S over the manifold of internally constrained equilibrium states.

The seemingly complicated formulation of the second postulate comes from the fact that S is only defined for equilibrium states, which in turn were defined for simple systems. For our purposes, each species (or subspecies) can be viewed as a simple system which is initially separated by walls from all other species. The composite system is then in an internally constrained equilibrium. When all internal constraints are removed the system will seek a new equilibrium state, the state of interest in this work.

3.6 Fundamental Equations

The second postulate asserts a functional relation between S and the extensive variables U, V, x . If this relation is known, one can, in principle, find all equilibrium properties of the system.

Such a functional relation is called a fundamental equation (or characteristic equation). Since the free energy is derived from fundamental equations, we shall now study some of their properties.

Proposition 3.3: Let $S = \Psi(U, V, x)$. Then Ψ is a homogeneous function of degree one.

Proof: By definition, a function $f : V \rightarrow W$ is homogeneous of degree p if $f(\lambda y) = \lambda^p f(y)$ for all $\lambda > 0$ and $y \in V$. Consider N identical simple systems in equilibrium. System i is characterized by S^i, U^i, V^i , and x^i . For the composite system formed by these N simple systems we have by postulate 2 (iii)

$$S = \Psi(U, V, x) = \Psi(NU^i, NV^i, Nx^i) = N\Psi(U^i, V^i, x^i)$$

Extending the above to any real N we obtain the result. \square

Proposition 3.4: Let $S = \Psi(U, V, x)$. Then Ψ is invertible with respect to U .

Proof: Clear from postulate 2, (i) (ii). \square

The above proposition offers another way to write fundamental equations, with U substituting for S as the dependent variable. Thus we can write $U = U(S, V, x)$.

It is easily verified that U is also a homogeneous function of degree one. This property of U (and as we shall see, also of the free energy F) plays a major role in most of the subsequent parts of this work. To simplify the terminology, we shall occasionally refer to "homogeneous functions of degree one" simply as 1-homogenous functions.

Proposition 3.5:

$$U = \left(\frac{\partial U}{\partial S} \right)_{V,x} S + \left(\frac{\partial U}{\partial V} \right)_{S,x} V + \sum_{j=1}^n \left(\frac{\partial U}{\partial x_j} \right)_{S,V,x_1} x_j$$

Proof: By the previous remarks $U = U(S,V,x)$ is 1-homogeneous and differentiable (over some domain) since S is. The result then follows directly from the well known Euler's Theorem which states that if $f(x)$ is differentiable and homogeneous of degree p then

$$p f(x) = x' \nabla f(x) \quad \square$$

This result and some of its implications were discussed by Dantzig [15,17].

The propositions presented here set the stage for a similar treatment of free energy functions. To keep the discussion in the most general terms, we shall derive free energy functions from fundamental equations by Legendre transforms.

3.7 Legendre Transforms and Homogeneous Functions

Let $G \equiv G(x)$ be a differentiable function on an open set $\Omega \subseteq \mathbb{R}^n$. Let $y \equiv y(x) \equiv \nabla G(x)$ for $x \in \Omega$. The function

$$\varphi(y) = G(x) - x'y \quad x \in \Omega, y = \nabla G(x) \quad (1)$$

is called the Legendre Transform of $G(x)$.

Let $H(x)$ be the matrix of second partial derivatives (assuming that they exist)

$$H_{ij}(x) = \frac{\partial y_i(x)}{\partial x_j}$$

By the implicit function theorem, if H is nonsingular there exist functions ϕ_i , $i = 1, 2, \dots, n$ such that

$$x_i = \phi_i(y)$$

In that case x can be eliminated from (1). By the same token we could write

$$\varphi(x) = G(x) - x'\nabla G(x) \quad (2)$$

or

$$\varphi(x,y) = G(x) - x'y \quad (3)$$

Legendre transforms (1) replace the independent variables x by the partial derivatives of the function G . The specific form of φ will be (1), (2) or (3) depending on whether y , x or

both, respectively, serve as independent variables. In the latter case only part of the x 's and the y 's are independent since we assume that G has at most n independent variables.

Lemma 3.6: Let $f(x)$ be differentiable on an open set $\Omega \subseteq \mathbb{R}^n$. If $f(x)$ is homogeneous of degree one then its Legendre transform is $\varphi(x) \equiv 0$.

Proof: $f(x) = x' \nabla f(x)$ by Euler's Theorem. The result follows. \square

Lemma 3.7: Let $f(x)$ be a homogeneous function (of degree one), differentiable on $\Omega \subseteq \mathbb{R}^n$. Then $y_1(x) \equiv \partial f(x) / \partial x_1$ is homogeneous of degree zero for $i = 1, 2, \dots, n$.

Proof:

$$f(\lambda x) = \lambda f(x) = \lambda \sum_{i=1}^n y_1(x) \cdot x_i \quad \text{for all } \lambda > 0$$

but

$$f(\lambda x) = \sum_{i=1}^n y_1(\lambda x) \cdot \lambda x_i = \lambda \sum_{i=1}^n y_1(\lambda x) \cdot x_i$$

We conclude

$$\sum_{i=1}^n y_1(\lambda x) \cdot x_i = \sum_{i=1}^n y_1(x) \cdot x_i \quad \text{for all } x \in \Omega, \lambda > 0$$

Thus

$$y_1(\lambda x) = y_1(x) = \lambda^0 y_1(x) . \quad \square$$

The last Lemma shows that Legendre transforms can be used to replace extensive variables like S, V (which are homogeneous of degree one since they are proportional to the mass of the system) by intensive ones which are homogeneous of degree zero and thus independent of the mass of the system.

We can thus define for $U = U(S, V, x)$

$$T \equiv \left(\frac{\partial U}{\partial S} \right)_{V, x} \quad (4)$$

$$P \equiv - \left(\frac{\partial U}{\partial V} \right)_{S, x} \quad (5)$$

$$\mu_i \equiv \left(\frac{\partial U}{\partial x_i} \right)_{S, V, x_j} \quad i = 1, 2, \dots, n \quad (6)$$

The intensive variables T and P defined above are the well known absolute temperature and pressure, while μ_i is the chemical potential of species S_i (see section 2.4).

The complete Legendre transform of U is, by Lemma 3.6,

$$\phi(T, P, \mu) = U - TS + PV - \sum_{i=1}^n \mu_i x_i = 0 \quad (7)$$

The quantity $U - TS + PV$ is a partial transform with respect to S and V . We shall now formalize this useful notion [11, p. 96].

Let $G(x)$ be a differentiable function on $\Omega \subseteq \mathbb{R}^n$. Let α be an integer subset of $\{1, 2, \dots, n\}$. The partial Legendre transform with respect to the variables $x_i, i \in \alpha$ is defined as

$$\varphi_{\alpha}(x, y) \equiv G(x) - \sum_{i \in \alpha} x_i y_i \quad (8)$$

where $y \in \Lambda(G), x \in \Omega$. Assume without loss of generality that $\alpha = \{1, 2, \dots, k\}$ and $\beta = \{k+1, k+2, \dots, n\}$ for some k . In most applications of (8) we choose as independent variables the sets $y_i (i \in \alpha)$ and $x_j (j \in \beta)$. We write

$$\varphi_{\alpha}(x, y) = \varphi_{\alpha}(x_{k+1}, x_{k+2}, \dots, x_n, y_1, y_2, \dots, y_k)$$

Free energy functions can now be defined in terms of partial Legendre transforms of $U = U(S, V, x)$. Let $\alpha = \{1, 2\}$ corresponding to S and V . Define

$$F(x, T, P) \equiv \varphi_{\alpha}(x, T, P) = U - TS + PV \quad (9)$$

It follows from (7) that

$$F(x, T, P) = \sum_{i=1}^n \mu_i x_i \quad (10)$$

Here

$$\mu_i = \mu_i(x, T, P) .$$

Equation (10) is identical to (2.4.1) which served as our original definition of the Gibbs free energy function. Similar to F we can define other free energy functions by suitable choice of α . Letting $\alpha = \{1\}$ we obtain the Helmholtz free energy A ,

$$A(V, x, T) \equiv \varphi_{\alpha}(V, x, T) = U - TS \quad (11)$$

and by (7), the Helmholtz free energy A becomes

$$A(V, x, T) = \sum_{i=1}^n x_i \mu_i(V, x, T) - V \cdot P(V, x, T)$$

The enthalpy H is defined by letting $\alpha = \{2\}$

$$H(S, x, P) = U + PV = T(S, x, P) \cdot S + \sum_{i=1}^n x_i \mu_i(S, x, P) \quad (12)$$

Our main goal is to derive general properties of free energy functions, based only on the postulates and Legendre transforms. Two basic properties are of interest in our discussion: homogeneity and convexity.

Theorem 3.8: Any partial Legendre transform of a fundamental equation is a homogeneous function of degree one in the extensive variables.

Proof: We have already shown that a fundamental equation is homogeneous of degree one in Proposition 3.3, so let the equation be in general $X_0 = G(X_1, X_2, \dots, X_n)$. By Lemma 3.6 we know that

$$\varphi(X, Y) = G(X) - X'Y = 0$$

where $X = (X_1, X_2, \dots, X_n) \in \Omega \subseteq \mathbb{R}^n$, $Y = \nabla G(x)$. Let $\alpha \subseteq \{1, 2, \dots, n\}$, $\beta = \{1, 2, \dots, n\} - \alpha$. Then

$$\varphi_\alpha(X) = G(X) - \sum_{i \in \alpha} X_i Y_i(X) = \sum_{j \in \beta} X_j Y_j(X)$$

where

$$Y_j(X) = \frac{\partial G(X)}{\partial X_j}$$

Therefore, for any $\lambda > 0$

$$\begin{aligned} \varphi_\alpha(\lambda X) &= \sum_{j \in \beta} (\lambda X_j) \cdot \frac{\partial G(\lambda X)}{\partial (\lambda X_j)} \\ &= \lambda \sum_{j \in \beta} X_j \frac{\lambda \partial G(X)}{\lambda \partial X_j} = \lambda \sum_{j \in \beta} X_j Y_j(X) = \lambda \varphi_\alpha(X). \quad \square \end{aligned}$$

This result is not entirely new. It was intuitively known and proven in the past (see for example [17]) for specific functions. Our approach is a generalization which applies to all thermodynamics free energy functions, without reference to their specific functional form.

The foregoing discussion illuminates the reason for using F for reactions under constant pressure and temperature. Under these conditions $\mu_i(x, P, T) = \mu_i(x)$ so that F is a function of composition only. Under constant V and T we use the function A (11), whence $\mu_i(V, x, T) = \mu_i(x)$ and $P(V, x, T) = P(x)$, and again, A is a function of composition only.

The homogeneity of F explains the reference to $\mu_i(x)$ as the "partial molar free energy." At fixed T and P we have

$$F(x) = \sum_{i=1}^n \mu_i(x) \cdot x_i = \sum_{i=1}^n \frac{\partial F(x)}{\partial x_i} x_i \quad (13)$$

Moreover, by Lemma 3.7, $\mu_i(x)$ is a homogeneous function of degree zero for $i = 1, 2, \dots, n$, which accounts for the fact that in most cases $\mu(x) = \mu(\hat{x})$.

Many difficulties in analysis and computation of free energy functions arise from their bad behavior for compositions which are not strictly positive, i.e., which have vanishing species. We observe that right continuity of F always requires that $F(0) = 0$ since

$$\lim_{\lambda \rightarrow 0} F(\lambda x) = \lim_{\lambda \rightarrow 0} \lambda F(x) = 0$$

Loomis and Steinberg [37] show that a 1-homogeneous function F is differentiable at 0 if, and only if, it is an affine function. Since the free energy is not linear, its boundary pathologies are in this sense a result of its 1-homogeneity.

3.8 Convexity of Free Energy Functions

The postulates of Section 3.5 were defined in terms of the entropy S . Tisza [50] and Callen [11] show that these can be described equally well in terms of the internal energy U , by

exchanging the roles of U and S in postulate 2 and by replacing the "maximize S " in (iv) with "minimize U ." Internally unconstrained equilibrium is then reached when the minimum of the internal energy is attained.

Based on postulate 2 we can now prove

Lemma 3.9: $U = U(S, V, x)$ is a convex function.

Proof: Let the subscripts c and u denote internally constrained and unconstrained systems respectively.

Consider two simple systems characterized by (S', V', x') and (S'', V'', x'') , respectively, which form a composite system. From postulate 2 (iii) we have

$$U_c(S', V', x') + U_c(S'', V'', x'') = U_c(S' + S'', V' + V'', x' + x'') .$$

By the energy version of (iv)

$$U_c(S' + S'', V' + V'', x' + x'') \geq U_u(S' + S'', V' + V'', x' + x'')$$

Since we are dealing only with equilibrium states we can now eliminate the subscripts and combine to get

$$U(S', V', x') + U(S'', V'', x'') \geq U(S' + S'', V' + V'', x' + x'')$$

Homogeneity of U requires

$$U(\lambda S, \lambda V, \lambda x) = \lambda U(S, V, x)$$

Therefore

$$\begin{aligned} & U(\lambda S', \lambda V', \lambda x') + U((1-\lambda)S'', (1-\lambda)V'', (1-\lambda)x'') \\ &= \lambda U(S', V', x') + (1-\lambda) U(S'', V'', x'') \\ &\geq U(\lambda S' + (1-\lambda)S'', \lambda V' + (1-\lambda)V'', \lambda x' + (1-\lambda)x'') \end{aligned}$$

This completes the proof. \square

Theorem 3.10: Let $\Omega \subseteq \mathbb{R}^n$, and let $G: \Omega \rightarrow \mathbb{R}$ be a twice continuously differentiable, homogeneous function of degree one on Ω . Let

$$\psi_{\alpha}(x) = G(x) - \sum_{i \in \alpha} y_i(x) \cdot x_i$$

be a partial Legendre transform of G with respect to x_i , $i \in \alpha$, where

$$y_i(x) = \frac{\partial G(x)}{\partial x_i} \quad i \in N \equiv \{1, 2, \dots, n\}$$

Under these conditions, $\psi_{\alpha}(x)$ is convex in the "untransformed" variables x_j , $j \notin \alpha$ for all α , if and only if, $G(x)$ is convex.

Proof: The "only if" is trivially proved by taking $\alpha = \emptyset$ (the empty set). Then $\varphi_{\emptyset}(x) = G(x)$ which is convex.

Now suppose G is convex, and let

$$Q_{ij}(x) = \frac{\partial y_i(x)}{\partial x_j} = \frac{\partial^2 G(x)}{\partial x_i \partial x_j}, \quad i, j \in N$$

The symmetric matrix $Q(x) = \{Q_{ij}(x)\}$ must be positive semidefinite (see Zangwill [55, p. 30]). Accordingly, all principal submatrices of Q are also positive semidefinite. Let $\alpha \subseteq N$, $\alpha \neq \emptyset$. By Lemma 3.6

$$\varphi_{\alpha}(x) = \sum_{i \in N-\alpha} y_i(x) \cdot x_i$$

From Theorem 3.8, $\varphi_{\alpha}(x)$ is homogeneous of degree one. Hence

$$y_i(x) = \frac{\partial \varphi_{\alpha}(x)}{\partial x_i}, \quad i \in N-\alpha$$

Therefore

$$\frac{\partial^2 \varphi_{\alpha}(x)}{\partial x_i \partial x_j} = \frac{\partial y_i(x)}{\partial x_j} = \frac{\partial^2 G(x)}{\partial x_i \partial x_j} = Q_{ij}(x) \quad \text{for } i, j \in N-\alpha$$

The second partial derivatives of $\varphi_{\alpha}(x)$ with respect to the untransformed variables form a principal submatrix of Q which is positive semidefinite. Therefore $\varphi_{\alpha}(x)$ is convex. \square

Theorem 3.10 proves that free energy functions, defined as partial transforms of fundamental equations, are convex in the extensive variables. In our case

$$F = F(T, P, x) \quad \text{is convex in } x.$$

Homogeneity, on the other hand, shows that F cannot be strictly convex, for let $x'' = \gamma x'$ for some positive scalar γ . Then

$$\begin{aligned} F(\lambda x' + (1-\lambda)x'') &= F([\lambda + (1-\lambda)\gamma]x') = [\lambda + (1-\lambda)\gamma] F(x') \\ &= \lambda F(x') + (1-\lambda) F(\gamma x') = \lambda F(x') + (1-\lambda) F(x'') \end{aligned}$$

Thus $F(x)$ is linear for $x = x' + \gamma x'$.

We have shown so far that convexity and homogeneity of U and F are outcomes of the basic postulates. We shall now prove the converse, namely, that any convex homogeneous function will satisfy the minimum energy principle.

Proposition 3.11: Let $G(x)$ be a convex homogeneous function defined on a convex set $\Omega \subseteq \mathbb{R}^n$. Let X' and X'' denote respectively the extensive variables of two simple systems, and let $G^*(X)$ denote the equilibrium value of $G(X)$. Then

$$G^*(X' + X'') \leq G^*(X') + G^*(X'').$$

Proof: By homogeneity we have for $\lambda \geq 0$

$$G^*(X' + \lambda X'') = (\lambda + 1) G^* \left(\frac{1}{\lambda + 1} X' + \frac{\lambda}{\lambda + 1} X'' \right)$$

Together with convexity, this leads to

$$\begin{aligned} G^*(X' + \lambda X'') &\leq (\lambda + 1) \left[\frac{1}{\lambda + 1} G^*(X') + \frac{\lambda}{\lambda + 1} G^*(X'') \right] \\ &= G^*(X') + \lambda G^*(X'') \end{aligned}$$

Letting $\lambda = 1$ we get the result. \square

We conclude this chapter by noting that a Legendre transform of a function can be geometrically interpreted as the envelope of tangent hyperplanes to the graph of the function. As such, Legendre transforms are closely related to "conjugate functions" which play a major role in Rockafellar's convex analysis [43]. When a function is replaced by its partial derivatives, a constant of integration must always be added when returning to the original function since differentiation "loses" some information. This is the reason for $\varphi(x)$ being defined the way it is and not by $x'y$. The last point is demonstrated here by defining the chemical equilibrium problem in terms of the intensive variables μ_1 , for an ideal system under fixed temperature and pressure.

We have

$$\mu_j = \mu_j(x) = c_j + \log(x_j/\bar{x}_k), \quad j \in \langle k \rangle$$

Hence

$$x_j = \bar{x}_k \cdot \exp(\mu_j - c_j), \quad j \in \langle k \rangle \quad (1)$$

$$F(\mu) = \sum_{j=1}^n \mu_j \bar{x}_{k(j)} \exp(\mu_j - c_j) \quad (2)$$

The mass balance constraints are

$$\sum_k \bar{x}_k \sum_{j \in \langle k \rangle} a_{ij} \exp(\mu_j - c_j) = b_i, \quad i = 1, 2, \dots, m \quad (3)$$

The variables \bar{x}_k are in essence the "integration" constants for the free energy of each phase.

CHAPTER 4

GEOMETRIC PROGRAMMING AND CHEMICAL DUALITY

The first part of this chapter reviews the duality theory of geometric programming due to Duffin, Peterson and Zener [25]. The review, including some important theorems, is followed by a dual formulation of the chemical equilibrium problem. We shall look at several forms of the dual chemical problem and then try to give a chemical interpretation relating these problems to well known chemical concepts and laws. In this chapter we shall at times sacrifice mathematical rigor for brevity.

4.1 Primal Geometric Programs

In the sections dealing with geometric programming we shall rely mostly on Duffin, Peterson and Zener's notation [25], with some changes to illuminate the relation to chemical problems. The theory of geometric programming was developed primarily for its engineering applications, which take the form of the following primal geometric program:

Program PGP: Find a vector $t \in \mathbb{R}^m$ so as to

$$\text{Minimize } g_0(t) \equiv \sum_{j \in \langle 0 \rangle} u_j(t) \quad (1)$$

$$\text{Subject to } g_k(t) \equiv \sum_{j \in \langle k \rangle} u_j(t) \leq 1, \quad k = 1, 2, \dots, K \quad (2)$$

and

$$t \equiv (t_1, t_2, \dots, t_m) > 0 \quad (3)$$

Here

$$u_j(t) \equiv C_j \prod_{i=1}^m t_i^{a_{ij}}, \quad j = 1, 2, \dots, n \quad (4)$$

C_j are positive real numbers, a_{ij} are real numbers, the sets $\langle k \rangle$ for $k = 0, 1, 2, \dots, K$ are integer sets partitioning the set $N \equiv \{1, 2, \dots, n\}$ as described in Section 2.2. For convenience we also define the integer sets

$$M \equiv \{1, 2, \dots, m\}$$

$$\bar{K} \equiv \{1, 2, \dots, K\}$$

The terms $u_j(t)$ resemble positive general polynomial terms and are therefore called posynomial terms. Similarly, the functions $g_k(t)$, $k \in \{0\} \cup \bar{K}$ are called posynomial functions or simply posynomials.

The matrix $A = \{a_{ij}\}$ is the exponents matrix, and the positive vector $C \equiv (C_1, C_2, \dots, C_n)$ is the coefficients vector. Problem PGP can therefore be completely characterized by

- (i) An $m \times n$ exponents matrix A
 - (ii) A positive n -vector of coefficients C
 - (iii) A partition of the set N defining the sets $\langle k \rangle$, $k = 0, 1, \dots, K$.
- We say that program PGP is consistent if there exists a vector t satisfying (2) and (3). The program is super-consistent if there is a $t > 0$ for which (2) are all satisfied with strict inequality.

4.2 Dual Geometric Programs

With each primal geometric program there is an associated dual geometric program.

Program DGP: Find vectors $\delta \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}^K$ so as to

$$\text{Maximize } v(\delta) = \left[\prod_{j=1}^n (c_j / \delta_j)^{\delta_j} \right] \left[\prod_{k=1}^K \lambda_k \right] \quad (1)$$

$$\text{Subject to } \sum_{j \in \langle 0 \rangle} \delta_j = 1 \quad (\text{Normality condition}) \quad (2)$$

$$A\delta = 0 \quad (\text{Orthogonality conditions}) \quad (3)$$

$$\delta = (\delta_1, \delta_2, \dots, \delta_n) \geq 0 \quad (\text{Nonnegativity}) \quad (4)$$

and

$$\lambda_k = \sum_{j \in \langle k \rangle} \delta_j, \quad k \in \bar{K} \quad (5)$$

Here the coefficients C_j , the matrix A and the partitions $\langle k \rangle$ of N are the same as defined for PGP. To maintain continuity of $v(\delta)$ on the nonnegative orthant of \mathbb{R}^n we define

$$\delta^\delta = \delta^{-\delta} \equiv 0 \quad \text{when} \quad \delta = 0$$

Note that the dual program can be completely characterized by the same information characterizing PGP. Program DGP is said to be consistent if there exists a $\delta \geq 0$ satisfying (2) and (3).

4.3 Duality Theory of Geometric Programming

The main lemma of geometric programming states the relation between the primal and dual programs.

Lemma 4.1: (Duffin et al. [25, p. 114]). If t satisfies the constraints of primal program PGP, and δ satisfies the constraints of its dual program DGP then

$$g_0(t) \geq v(\delta) \tag{1}$$

Moreover, under the same conditions $g_0(t) = v(\delta)$ if, and only if,

$$\delta_j = \begin{cases} u_j(t)/g_0(t) & j \in \langle 0 \rangle \\ \lambda_k(\delta) u_j(t) & j \in \langle k \rangle, k \in \bar{K} \end{cases} \tag{2}$$

The lemma serves as a basis for the following duality theorem:

Theorem 4.2: ([25, p. 80]). Suppose primal program PGP is superconsistent, and $g_0(t)$ attains a minimum value at a feasible point t^* . Then

- (i) Dual program DGP is consistent and attains its maximum at a dual feasible point δ^* .
- (ii) $v(\delta^*) = g_0(t^*)$
- (iii) There exist nonnegative Lagrange multipliers η_k^* , $k \in \bar{K}$ such that the Lagrange function

$$L(t, \eta) = g_0(t) + \sum_{k=1}^K \eta_k [g_k(t) - 1] \quad (3)$$

has the property

$$L(t^*, \eta) \leq g_0(t^*) = L(t^*, \eta^*) \leq L(t, \eta^*) \quad (4)$$

for arbitrary $t > 0$ and arbitrary $\eta \geq 0$. Moreover, there exists a maximizing vector δ^* for DGP such that

$$\delta_j^* = \begin{cases} u_j(t^*)/g_0(t^*), & j \in \langle 0 \rangle \\ \eta_k^* u_j(t^*)/g_0(t^*), & j \in \langle k \rangle, k \in \bar{K} \end{cases} \quad (5)$$

Furthermore,

$$\lambda_k(\delta^*) = \eta_k^*/g_0(t^*) \quad k \in \bar{K} \quad (6)$$

(iv) If δ^* is a maximizing point for dual program DGP, each minimizing point t^* for PGP satisfies the system of equations

$$u_j(t^*) = \begin{cases} \delta_j^* v(\delta^*) & j \in \langle 0 \rangle \\ \delta_j^* / \lambda_k(\delta^*) & j \in \langle k \rangle \end{cases} \quad (7)$$

where k ranges over the integers for which $\lambda_k(\delta^*) > 0$.

A dual program is said to be canonical if there exists a positive vector $\delta > 0$ satisfying the dual constraints (2.2) and (2.3). Otherwise the program is said to be degenerate. We shall say that a primal program is canonical or degenerate when the corresponding dual program is.

Theorem 4.3: ([25, p. 169]). Suppose dual program DGP and its corresponding PGP are canonical. Then DGP is always consistent, but PGP is consistent if, and only if, DGP has a finite positive maximum. Moreover, under these conditions the minimum of PGP is attained at some finite primal feasible point $t^* > 0$ and is equal to the maximum of DGP.

4.4 Transformed Primal Programs

For theoretical and practical reasons, it is sometimes advantageous to consider a transformed version of PGP. Let

$$z_i = \log t_i \quad i = 1, 2, \dots, m$$

The transformed (primal) geometric program is

Program TGP

$$\text{Minimize } g_0(z) \equiv \sum_{j \in \{0\}} u_j(z) \quad (1)$$

$$\text{Subject to } g_k(z) \equiv \sum_{j \in \{k\}} u_j(z) \leq 1, \quad k \in \mathcal{K} \quad (2)$$

where

$$u_j(z) = \exp\{-c_j + \sum_{i=1}^m a_{ij} z_i\}, \quad j \in N \quad (3)$$

$$c_j = -\log C_j$$

We used identical notation for PGP. It will always be clear from context whether we are considering program PGP or TGP.

Two points should be noted about program TGP

- (i) The variables z_i are unrestricted in sign.
- (ii) The program is convex since $u_j(z)$ are convex functions.

4.5 Chemical Duality

Following the development of the optimality conditions by Kuhn and Tucker [36], it was shown by White et al. [54] and later by Dorn [23] that the Lagrange multipliers of the chemical equilibrium problem CPI give rise to a new problem called a dual chemical problem. To avoid confusion between geometric programming duality and chemical duality we shall refer to problems CPI and CPN, i.e., to the chemical equilibrium problems, as primal chemical problems. Dual chemical problems are those problems involving the multipliers of problems CPI or CPN of Section 2.5.

White's and Dorn's dual chemical programs are not "pure" duals in that they contain, in addition to multipliers, also primal chemical variables, namely, composition variables. Avriel [4], and Passy and Wilde [40], showed that geometric programming can be applied to generate a pure dual chemical problem, in fact, they showed that the chemical equilibrium problem is equivalent to the dual of a geometric program (DGP). It is then a straightforward task to show that a dual chemical problem equivalent to PGP can be formulated.

To show that problem CPI is equivalent to DGP we define

$$a_{i0} = -b_i \quad i = 1, 2, \dots, m$$
$$\delta_0 = \lambda_0 = 1 \quad \text{and} \quad \langle 0 \rangle = \{0\}$$

We then identify

$$\begin{aligned} \delta_j & \text{ with } \hat{x}_j & j \in N \\ \lambda_k & \text{ with } \bar{x}_k & k \in \bar{K} \\ C_j = \exp(-c_j) & & j \in N \end{aligned}$$

so that

$$x_j = \delta_j \lambda_k(j)$$

Rather than minimize $F(x)$ in (2.5.1), we can maximize $\exp(-F(x))$, which after the above transformations is seen to be $v(\delta)$. The orthogonality conditions $A\delta = 0$ are satisfied for the augmented matrix A which has the added column $A_0 = -b$.

There is no difficulty in showing the converse--that DGP can be formulated as a CPI so that the relation is indeed two-sided. We remark here that the general DGP is equivalent to an abstract chemical problem. The special properties of matrix A in chemical problems (for example, Proposition 3.1) do not hold for a general DGP.

Our interest will be focused mostly on the dual chemical problem. In light of the transformations above, it has the following form:

Problem DCP (Dual Chemical Problem)

$$\begin{aligned} \text{Minimize } & -\log g_0(t) \equiv -b' \log t \equiv -\sum_i b_i \log t_i \\ \text{Subject to } & g_k(t) \equiv \sum_{j \in \langle k \rangle} u_j(t) \leq 1 \quad k \in \bar{K} \\ & t_1 > 0, t_2 > 0, \dots, t_m > 0 \end{aligned}$$

Here

$$u_j(t) = C_j \prod_{i=1}^m t_i^{a_{ij}} \quad j \in N$$

$$C_j = \exp(-c_j) \quad j \in N$$

where c_j are the free energy coefficients, $A = \{a_{ij}\}$ and $b = (b_1, b_2, \dots, b_m)$ are the matrix and right hand side of the mass balance equations.

$$\log t \equiv (\log t_1, \log t_2, \dots, \log t_m)$$

Another formulation of the chemical dual is based on the transformed geometric program. This form is essentially the one used by Dorn [23] and Bigelow [7], and it simply replaces $\log t_i$ by z_i so that we

$$\text{Minimize } -\log g_0(z) = -b'z$$

$$\text{Subject to } g_k(z) = \sum_{j \in \langle k \rangle} u_j(z) \leq 1 \quad k \in \bar{K}$$

where

$$u_j(z) = \exp(-c_j + \sum_{i=1}^m a_{ij} z_i) \quad j \in N$$

4.6 Interpreting the Dual Chemical Problem

Most of the implications and applications of chemical duality, especially those of equivalence to geometric programs will be treated in the following chapters. In the rest of this chapter we shall study the physical meaning of the dual chemical problem. Specifically, we shall interpret the dual chemical variables and constraints.

Dorn [23] and White et al. [54], who realized the existence of duality in chemical equilibrium, also attempted to interpret the dual variables. It is easily seen that dual chemical variables t_i or z_i relate to the subspecies and their mass balance equations. The form of the objective function in DCP (which equals $F(x)$ at an optimal t^* and x^*) suggests that $\log t_i$ is some energy measure of subspecies B_i . Indeed all the interpretations, including a later one by Duffin and Zener [26] which considered entropy functions, followed this line.

We find two major defects in all these attempts. First none of the above approaches accounts for the fact that the dual chemical variables for a given chemical equilibrium problem are not unique. Second, all are quite unintuitive and do not readily relate to known chemical concepts.

Dorn suggests that "the dual variables are dimensionless energies, contributed by the elements (subspecies)." White et al. state that " π_i (the multiplier of $[b_i - \sum_{j=1}^n a_{ij}x_j]$) is the free energy contribution due to the presence of one mole of atom (subspecies) i ." The set of subspecies for a given system need not be

unique (see Section 2.3). Moreover, the free energy function itself is only a relative measure. Its value depends on the definition of certain standard states and energies. We shall deal first with these states and then return to interpret the dual chemical variables.

4.7 Equivalent Formulations of the Chemical System

Consider the problem

Problem CP1

Minimize $F(x)$

Subject to $Ax = b, \quad x \geq 0$

Where $x \in \mathbb{R}^n, A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$.

Let $D \in \mathbb{R}^{m \times m}$ be a real nonsingular matrix and let $y \in \mathbb{R}^m$ be a given vector. Let $\bar{A} = DA, \bar{b} = Db$, and consider the following problems:

Problem CP2

Minimize $F(x)$

Subject to $\bar{A}x = \bar{b}, \quad x \geq 0$

Problem CP3

Minimize $\bar{F}(x) \equiv F(x) - y'Ax$

Subject to $Ax = b, \quad x \geq 0$

It is easy to see that problems CP1, CP2, and CP3 are equivalent in the sense that x is a solution to one of them if, and only if, it is a solution to all of them. The equivalence of CP1 and CP3 was used by Dantzig et al. [18] to obtain more convenient coefficients in $F(x)$. In the ideal case

$$F(x) = x' \cdot (c + \log \hat{x}) \quad (1)$$

Hence, the objective function of CP3 becomes

$$\bar{F}(x) = x' \cdot (c - A'y + \log \hat{x}) \quad (2)$$

Looking at the dual chemical problems associated with CP1, CP2 and CP3 we see that for CP1 (in the ideal case) the dual chemical problem is DCP.

For CP2 we have a similar dual, with \bar{a}_{ij} replacing a_{ij} and \bar{b}_i replacing b_i . For CP3 we obtain the dual chemical problem by replacing c_j of DCP with

$$\bar{c}_j \equiv \exp\{-c_j + y'A_j\} \quad (3)$$

Obviously the solutions t^* or z^* to these dual chemical problems are not the same.

In a more general context, let

$$F(x) = \sum_{j=1}^n \mu_j(x) \cdot x_j \quad (4)$$

which is the most general form, as shown in (2.4.2), (3.7.13).

Kuhn Tucker Conditions applied to these problems then state that if F is differentiable and a constraint qualification holds, a necessary condition for x^* to be a solution to CP1, CP2 or CP3, respectively, is that there exist vectors $z^* \in \mathbb{R}^m$ such that

$$\text{For CP1} \quad z^* A = \mu'(x^*) \quad (5)$$

$$\text{For CP2} \quad z^* \bar{A} = \mu'(x^*) \quad (6)$$

$$\text{For CP3} \quad z^* A = \mu'(x^*) - y'A \quad (7)$$

Since the values of z^* in (5), (6) and (7) are clearly not the same, the question arises as to what is the nature of the differences between the original chemical problem CP1 and the two equivalent problems CP2 and CP3.

Problem CP2 amounts to a redefinition of the subspecies. The only restrictions on the subspecies being that they be independent and sufficient to describe all the species uniquely, one can apply non-singular linear transformations on the subspecies and maintain the conditions above (although the resulting set of new subspecies may appear strange to a chemist). We demonstrate such a transformation by the following example.

Example 4.4. Consider a single liquid phase system with the set S of species containing CO_2 , H^+ , OH^- , H_2O , Ac^- , HAc , H_2CO_3 , $\text{CO}_3^{=}$ where Ac^- represents acetic ion. The set B of subspecies is composed

of CO_2 , H^+ , OH^- and Ac^- . The matrix A for this system is

B \ S	CO_2	H^+	OH^-	H_2O	Ac^-	HAc	H_2CO_3	$\text{CO}_3^{=}$
CO_2	1	0	0	0	0	0	1	1
H^+	0	1	0	1	0	1	1	-1
OH^-	0	0	1	1	0	0	1	1
Ac^-	0	0	0	0	1	1	0	0

Let

$$D = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & -1 \\ -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

The new system, obtained by premultiplying A by D is

B \ S	CO_2	H^+	OH^-	H_2O	Ac^-	HAc	H_2CO_3	$\text{CO}_3^{=}$
$\text{CO}_3^{=}$	1	0	0	0	0	0	1	1
H^+	1	1	0	1	-1	0	2	0
OH^-	-1	0	1	1	0	0	0	0
HAc	0	0	0	0	1	1	0	0

Here the new column of subspecies is obtained from the previous one by premultiplying it by $(D^{-1})'$.

$$\begin{bmatrix} 1 & -1 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} \text{CO}_2 \\ \text{H}^+ \\ \text{OH}^- \\ \text{Ac}^- \end{bmatrix} = \begin{bmatrix} \text{CO}_2 - \text{H}^+ + \text{OH}^- \\ \text{H}^+ \\ \text{OH}^- \\ \text{H}^+ + \text{Ac}^- \end{bmatrix} = \begin{bmatrix} \text{CO}_3^{=} \\ \text{H}^+ \\ \text{OH}^- \\ \text{HAc} \end{bmatrix}$$

Thus, linear transformations of A amount to changing the set of subspecies B . Notice that the free energy coefficients were not changed. How can one change the subspecies and maintain the same free energy function? The answer is found in the definition of the free energy coefficients c_j . The same question arises with regard to CP3. In this case we can write

$$\bar{F}(x) = x(\bar{c}(y) + \log \hat{x})$$

where

$$\bar{c}(y) = c - A'y \quad (8)$$

For a given fixed vector y , problem CP3 amounts to changing only the free energy coefficients.

To see the meaning of these transformations, we review the derivation of the vector c [19, 21]. To compute all c_j with a common reference, it is assumed that each species S_j is formed from

the subspecies B_i via the formation reaction

$$\sum_{i=1}^m a_{ij} B_i = S_j \quad (9)$$

where each of the subspecies is in its reference state. A reference state is usually chosen to be the state where the pure subspecies is in its most stable form at 25°C. At any rate, each subspecies has some reference state and some associated reference free energy f_i . The reaction (9) has some associated change in free energy, since it can be viewed as an internally unconstrained equilibrium (see Section 3.4). Let this change be ΔF_j , then

$$c_j = \Delta F_j + \sum_{i=1}^m a_{ij} f_i \quad (10)$$

where c_j represents the dimensionless free energy of one (pure) mole of S_j . (ΔF_j and f_i are assumed dimensionless.) Now suppose that the reference states are redefined so that the new reference free energy of B_i is $f_i + y_i$. The new coefficient is

$$\bar{c}_j \equiv \Delta F_j + \sum_{i=1}^m a_{ij} (f_i + y_i) = c_j + \sum_{i=1}^m a_{ij} y_i$$

Thus

$$\bar{c} = c + A'y$$

Problem CP3 can thus be interpreted as a system with different reference states, which have new reference free energies. Similarly, the fact that c was unchanged in CP2 is explained by an implicit change in reference states, associated with the new set of subspecies.

4.8 The Physical Significance of the Dual Chemical Problem

The relation of the dual chemical problem to geometric programming and our observations about reference states will now be employed to interpret the variables and constraints of problem DCP. Let t^* be a solution to DCP (Section 4.5). Let the primal chemical problem be CPI with $F(x)$ given by (7.1) (this is problem CPI of Chapter 2--the standard ideal chemical equilibrium problem). Let x^* be a solution, namely, x^* is an equilibrium composition vector.

Applying the equivalence to geometric programming and the duality theorem 4.2, we have by (4.3.7)

$$x_j^* = \bar{x}_k^* u_j(t^*) = \bar{x}_k^* e^{-c_j} \prod_{i=1}^m (t_i^*)^{a_{ij}}$$

where

$$\bar{x}_k^* = \sum_{i \in \langle k \rangle} x_i^*, \quad k = k(j)$$

In other words, when $\bar{x}_{k(j)}^* > 0$

$$\hat{x}_j^* \prod_{i=1}^m (t_i^*)^{-a_{ij}} = e^{-c_j} \quad (1)$$

This relation has exactly the form of the mass action law (Section 3.2) for the formation reaction of S_j (Equation (7.9)), where $K(\theta) \equiv K_j = e^{-c_j}$. Indeed we know from classical thermodynamics (Denbigh [21, p. 140]) that the (dimensional) free energy of S_j is given by $-RT \log K_j$. From (7.10) we then have

$$RT \ln c_j = -RT \log K_j \quad (2)$$

which then reduces (1) to

$$\hat{x}_j^* \prod_{i=1}^m (t_i^*)^{-a_{ij}} = K_j \quad (3)$$

The mass action law (3) stated here relates concentrations (mole fractions) of reactants and products in the formation reaction to the equilibrium constant of the reaction, where the reactants (subspecies) are in their reference states. We conclude therefore that the dual chemical variables t_i represent concentrations of the subspecies in their reference states.

This interpretation is not merely an abstract construction. When one or more phases of the system do actually correspond to a reference state of some subspecies, say phase ϕ_k is the reference state of B_i , then the concentration t_i^* will be the actual equilibrium concentration of B_i in that phase. Note that we can always assume without loss of generality that B_i is included also as a species in each phase. In practice (depending on the reference states chosen), t_i^* is usually a very small number, indicating that B_i does not exist free in the system in any significant amount. (See Section 6.2.)

Equation (1) yields an easy interpretation to the polynomial terms $u_j(t)$ of the dual chemical problem DCP. We already saw their relation to mass action laws of formation reactions

$$u_j(t^*) = e^{-c_j} \prod_{i=1}^m (t_i^*)^{a_{ij}} = \hat{x}_j^* \quad (4)$$

Hence, at equilibrium each posynomial term represents the equilibrium concentration (mole fraction) of an associated species. A point t which is feasible but not optimal represents a state in which the formation reactions are not completed so that $u_j(t) \leq \hat{x}_j^*$. Finally, the dual chemical constraints

$$g_k(t) = \sum_{j \in \langle k \rangle} u_j(t) \leq 1$$

are a restatement of the fact that mole fractions sum to unity for each existing phase at equilibrium.

Equations (3.7) imply that $g_k(t^*) = 1$ whenever $\bar{x}_k^* > 0$. Duffin and Zener [26] noted that for a gas phase the dual chemical constraint can be reformulated

$$\sum_{j \in \langle k \rangle} P_j(t) \leq P$$

This is Dalton's Law which states that the partial pressures P_j of ideal gases at equilibrium sum to the total pressure P when the gas phase exists.

The basic difference between the primal chemical variables x and the dual variables t (or z) is that while the former are extensive variables whose values depend on the mass of the system,

the latter are intensive variables and can thus be viewed as the outcome of a Legendre transform of an extensive function. The hidden "integration constants" required to recover extensive variables from intensive ones are the multipliers η_k (of the dual chemical constraints), which are needed to compute δ_j^* (or x_j^*) in equation (3.5).

CHAPTER 5

PROPERTIES OF EQUILIBRIUM SOLUTIONS

5.1 Introduction and Notation

In this chapter we shall study the solutions to the chemical equilibrium problem. The main issues are existence, uniqueness, bounds, common properties of solution points, and general properties of solution sets, henceforth called equilibrium sets. These points will be reviewed in light of previous results relating to Legendre transforms and geometric programming duality theory.

Credit for much of the original work on this subject is due to Shapiro and Shapley [47], and to Bigelow [7]. Some of our results are not new, and were elaborately proved in [47]. We shall apply geometric programming duality to obtain simpler proofs which are direct consequences of the theory developed by Duffin, Peterson and Zener [25]. Other results here generalize previous results in two ways: first, by considering more general free energy functions, and second, by obtaining results under somewhat weaker conditions. This chapter provides a theoretical basis to the study of equilibrium sets and for bridging the gap between the mathematical implications of the model and the behavior of real chemical systems.

Unless stated otherwise, we assume a closed system under constant temperature and pressure. The system is defined by the triplet $\{F, A, b\}$ as described in Section 2.5, where F is taken to be

$$F(x) = \sum_{j=1}^n \mu_j(x) \cdot x_j \quad (1)$$

When F has the form

$$F(x) = \sum_{j=1}^n x_j (c_j + \log \hat{x}_j) \quad (2)$$

we refer to it (and to the system) as ideal. F is assumed differentiable on the positive orthant of \mathbb{R}^n . The feasible set of the system is denoted by X

$$X = X(A, b) \equiv \{x \in \mathbb{R}^n \mid Ax = b, x \geq 0\} \quad (3)$$

Its intersection with the positive orthant \mathbb{R}_+^n is

$$X^+ \equiv \{x \in X \mid x > 0\} .$$

We say that the system is canonical when $X^+ \neq \emptyset$ where \emptyset denotes the empty set. Otherwise the system is degenerate. (See Section 4.3.) The equilibrium set (solution set) of a given system is denoted

$$D(F|X) = \{y \in X \mid F(y) \leq F(x) \text{ for all } x \in X\} \quad (4)$$

5.2 Some Equilibrium Set Characteristics

Theorem 5.1: For a real chemical system (F, A, b) , the solution set $\mathcal{D}(F|X)$ is bounded.

Proof: By Proposition 3.1, X is bounded so the result is immediate. \square

F was shown to be convex on its domain of definition, which is normally \mathbb{R}_+^n . When possible, the domain is extended to the non-negative orthant. The following simple lemma insures that in these cases convexity is maintained on the extended domain.

Lemma 5.2: Let $F : \Omega \rightarrow \mathbb{R}$ be a convex function on the open set $\Omega \subseteq \mathbb{R}^n$.

Let $\bar{\Omega}$ denote the closure of \mathbb{R}^n . Suppose that

$\lim_{t \rightarrow 0^+} F(x + t\theta)$ exists for all x and θ in \mathbb{R}^n for which

$$x + t\theta \in \Omega \quad \text{for all} \quad 0 < t \leq t_0$$

Then F is convex on $\bar{\Omega}$

Proof: Let $x, y \in \mathbb{R}^n$. Let $\theta > 0$, so that $x + t\theta \in \Omega$ for all $t > 0$. Then, for any $\lambda \in [0, 1]$ we have

$$F[\lambda(x + t\theta) + (1-\lambda)(y + t\theta)] \leq \lambda F(x + t\theta) + (1-\lambda) F(y + t\theta)$$

Hence

$$F[\lambda x + (1-\lambda)y + t\theta] \leq \lambda F(x + t\theta) + (1-\lambda) F(y + t\theta)$$

Taking the limit as $t \rightarrow 0^+$ on both sides, we obtain the result. \square

Theorem 5.3: If F satisfies the conditions of the preceding lemma, then $\mathcal{D}(F|X)$ is convex.

Proof: \mathcal{D} is simply the minimum set of a convex program. \square

This is Lemma 9.3 in [47], applied here to a general free energy function. Since convexity on \mathbb{E}_+^n is assured by Theorem 3.10, only the right continuity of F on X is required to satisfy Lemma 5.2.

Two important characteristics of solutions which were developed via "quasi dependence" in [47] are shown here to be direct results of geometric programming duality.

Theorem 5.4: Let the system $\{F, A, b\}$ be canonical, where F is ideal.

Let $x^* \in \mathcal{D}(F|X)$. Then

$$x_j^* = 0 \quad \text{if and only if} \quad \bar{x}_{k(j)}^* = 0 .$$

Proof: Let $k \equiv k(j)$.

If $\bar{x}_k^* = 0$ clearly $x_j^* = 0$ by definition of \bar{x}_k . Now suppose $x_j^* = 0$ but $\bar{x}_k^* > 0$.

By Theorem 4.3 there exists a finite positive t^* solving the dual chemical problem DCP. According to Theorem 4.2

$$u_j(t^*) = e^{-c_j} \prod_{i=1}^m (t_i^*)^{a_{ij}} = x_j^*/\bar{x}_k^* = 0$$

contrary to the positivity and finiteness of t^* . \square

This well-known result has far reaching chemical implications. It says that at equilibrium, either a whole phase vanishes, or each of the species in the phase must be present in some positive amount, when the system is canonical. This justifies our remark about equilibrium concentrations of subspecies. These can always be assumed to exist if a reference phase exists at equilibrium (Section 4.8).

Theorem 5.5: Suppose (F, A, b) is canonical and F ideal.

Let $x \in \mathcal{D}(F|X)$ and $y \in \mathcal{D}(F|X)$ be two distinct equilibrium solutions. Then

$$\hat{x}_j = \hat{y}_j$$

whenever both are defined.

Proof: Theorem 4.3 guarantees the existence of finite and positive vectors $t(x)$ and $t(y)$ solving the dual chemical problem DCP. Furthermore, every solution t to DCP must satisfy condition (iv) of Theorem 4.2, namely

$$u_j(t) = x_j/\bar{x}_{k(j)} \quad \text{when } \bar{x}_{k(j)} > 0$$

similarly

$$u_j(t) = y_j/\bar{y}_{k(j)} \quad \text{when } \bar{y}_{k(j)} > 0$$

Both equations must hold for every t , therefore

$$x_j/\bar{x}_{k(j)} = y_j/\bar{y}_{k(j)}$$

whenever both are defined. \square

To the chemist, this result comes as no surprise. It indicates a certain concentration invariance, i.e., some uniqueness of concentrations in equilibrium solutions. It may be harder, though, to **accept** the fact that the same phase may vanish in one solution but not in another solution to the same system! To realize such a situation, observe that the mathematical formulation does not prevent us from "splitting" a phase in our model into two identical phases and treating these as two completely separate phases.

Suppose $x_{\langle k \rangle}$ is a solution for ($j \in k$) when ϕ_k is treated as a single phase, that is, x_j ($j \in k$) are equilibrium values. Assume $\bar{x}_k > 0$ so that there are equilibrium mole fractions \hat{x}_j , $j \in k$.

We now separate ϕ_k into ϕ_{k_1} and ϕ_{k_2} and construct a solution to the new system by letting $\alpha \in [0,1]$ and

$$\begin{aligned} \bar{x}_k^1 &= \alpha \bar{x}_k & \bar{x}_k^2 &= [1-\alpha] \bar{x}_k \\ x_j^1 &= \alpha x_j & x_j^2 &= [1-\alpha] x_j \end{aligned}$$

Clearly the new solution satisfies the constraints. Let \bar{F} be the new free energy. Since $\hat{x}_j = \hat{x}_j^1 = \hat{x}_j^2$, we obtain

$$\begin{aligned} F(x) - \bar{F}(x) &= \sum_{j \in \langle k \rangle} x_j (c_j + \log \hat{x}_j) - \left\{ \sum_{j \in \langle k \rangle} x_j^1 (c_j + \log \hat{x}_j^1) + \sum_{j \in \langle k \rangle} x_j^2 (c_j + \log \hat{x}_j^2) \right\} \\ &= \sum_{j \in \langle k \rangle} [x_j - (x_j^1 + x_j^2)] (c_j + \log \hat{x}_j) = 0 \end{aligned}$$

This implies that the new system is also in equilibrium, because its free energy cannot decrease below $F(x)$, the new system being "internally constrained" by the separation of Φ_k . Convexity of \mathcal{D} then implies that for all $\lambda \in [0,1]$

$$y_j = \lambda x_j^1 + (1-\lambda)x_j^2 \quad \text{for all } j \in k$$

is also a solution, since α was arbitrary. In particular, $\lambda = 0$ and $\lambda = 1$ cause Φ_{k_2} and Φ_{k_1} respectively to vanish. In a well formulated system such a situation should not arise. Conditions which insure that phases are not "split" in this way are part of the regularity condition discussed in the following sections.

5.3 Reduction of Degenerate Systems

A system $\{F, A, b\}$ may be canonical and yet the solution need not be strictly positive as was just shown. A solution x is degenerate if

$$x \in \mathcal{D}(F|X) \text{ but } x \notin X^+$$

This section deals both with system degeneracy and solution degeneracy. In both cases it is shown that the problem can be reduced by eliminating vanishing species.

Lemma 5.6: A feasible system is degenerate if and only if there exists an index set $I \subseteq N \equiv \{1, 2, \dots, n\}$, $I \neq \emptyset$, such that

- (i) $j \in I$ implies $x_j = 0$ for all $x \in X$
- (ii) there exists an $x \in X$ with

$$x_j > 0 \text{ for all } j \in N - I.$$

Proof: The "if" is trivially true since (i) implies degeneracy.

Suppose on the contrary that (i) is not true, i.e., for each $j \in N$ there exists at least one vector $x^{(j)} \in X$ with $x_j^{(j)} > 0$. We can choose a set of positive numbers λ_j such that $\sum_{j=1}^n \lambda_j = 1$ and form the linear combination

$$x = \sum_{j=1}^n \lambda_j x^{(j)} \quad \text{with} \quad x_j = \sum_{j=1}^n \lambda_j x_j^{(j)} > 0$$

Therefore $x > 0$ and $x \in X$ (since X is convex), a contradiction. Let \bar{I} be the maximal set satisfying (i). \bar{I} exists since it can be taken as the union of all sets I satisfying (i). If (ii) is not satisfied then the set $N - \bar{I}$ is degenerate, and by the argument used above there must exist a set $J \subseteq N - \bar{I}$ such that $j \in J$ implies $x_j = 0$ for all $x \in X$, contrary to the assumption that \bar{I} is maximal. \square

The lemma shows that in a degenerate system there is a set of species which must identically vanish, while the rest of the species always have a nondegenerate feasible composition. Intuitively, one tends to remove the vanishing species from the system and consider a canonical reduced system with species S_i , $i \in N - \bar{I}$. The strong duality theory of geometric programming justifies this reduction when F is ideal.

Let the reduced integer set be

$$\Gamma = N - \bar{I}$$

The problem formed by removing all species which are not in Γ is called the reduced problem or Γ CPI.

Problem Γ CPI

$$\text{Minimize } F(x) = \sum_{j \in \Gamma} x_j (c_j + \log \hat{x}_j)$$

$$\text{Subject to } \sum_{j \in \Gamma} a_{ij} x_j = b_i, \quad i \in M$$

$$x_j \geq 0, \quad j \in \Gamma$$

For convenience we shall use Γ as an operator, and write

ΓA for the reduced matrix,

ΓF for the reduced objective function,

Γx for the reduced composition vector, etc.

Theorem 5.7: Suppose problem CPI (Section 2.5) is feasible and degenerate.

Then the reduced problem Γ CPI is canonical. Furthermore,

$$F(\text{CPI}) = F(\Gamma\text{CPI})$$

where $F(\text{CPI})$ and $F(\Gamma\text{CPI})$ denote the equilibrium free energies for the two problems, respectively.

Proof: The boundedness of X (Proposition 3.1) and Theorems 3 and 4 of Duffin et al. [25] establish the result. \square

Theorem 5.7 reduces the study of problem CPI to a study of canonical programs. Canonicity can be tested and the reduction performed using a linear programming method due to Clasen [13]. His technique finds a positive composition $x \in X^+$ if one exists. It is discussed further in Chapter 7.

A more difficult question is that of degenerate solutions. Although in this case one cannot perform an a priori reduction, one can speak of a reduced solution and reduced solution integer set in a way analogous to reduced systems. Given $x^* \in \mathcal{D}(F|X)$, the reduced solution integer set is $\Lambda(x^*) = \{j \in \Gamma | x_j^* > 0\}$.

The Λ -problem ($\Lambda \equiv \Lambda(x^*)$) is the problem formed by ignoring all species S_j , $j \in N - \Lambda(x^*)$. We can easily prove

Lemma 5.8: If $x^* \in \mathcal{D}(F|X)$ then

$$\Lambda x^* \in \mathcal{D}(\Delta F|\Delta X)$$

where Λx^* , ΔF , ΔX represent "reduced" quantities and $\Lambda = \Lambda(x^*)$.

Proof: $x^* \in X$ implies $\Lambda x^* \in \Delta X$ since only zero valued components of x^* were deleted. Suppose there exists a $y \in \Delta X$ with

$$\Delta F(y) < \Delta F(\Lambda x^*)$$

Define a vector $y^* \in \mathbb{R}^n$

$$y_j^* = \begin{cases} y_j & j \in \Lambda(x^*) \\ 0 & j \in N - \Lambda(x^*) \end{cases}$$

Then $y^* \in X$

$$\begin{aligned} F(y^*) &= \sum_{j=1}^n y_j^* (c_j + \log \hat{y}_j^*) \\ &= \sum_{j \in \Lambda(x^*)} y_j^* (c_j + \log \hat{y}_j^*) + 0 \\ &= \sum_{j \in \Lambda(x^*)} y_j (c_j + \log \hat{y}_j) = \Delta F(y) \end{aligned}$$

Similarly

$$F(x^*) = \Lambda F(\Lambda x^*)$$

We conclude

$$F(y^*) < F(x^*)$$

a contradiction, since $x \in \mathcal{D}(F|X)$. \square

Unfortunately Λ is not known in advance and therefore the lemma cannot be applied to computations. Furthermore $\Lambda(\cdot)$ need not be unique. The preceding result is true only for ideal F .

5.4 Uniqueness of Solutions

In this section F is always ideal.

Theorem 5.9: Let $\{F, A, b\}$ be canonical, $A \in \mathbb{R}^{m \times n}$, $n \geq m$.

Let $I = \{i \mid \text{there exists an } x \in \mathcal{D}(F|X) \text{ with } x_i > 0\}$.

Then the solution to the dual problem DCP is unique if, and only if, $\text{rank}(IA) = m$. Here $IA \equiv [A_j \mid j \in I]$ is a reduced matrix.

Proof: Let $|I|$ denote the number of elements in the set I . First we show that if $|I| = \ell$ there exists an $x^* \in \mathcal{D}(F|X)$ with $|\Lambda(x^*)| = \ell$. The argument is by taking convex combinations, similar to the proof of Lemma 5.6. Therefore we can use Λ rather than I and require that there exist an $x^* \in \mathcal{D}(F|X)$ such that $\text{rank}(\Lambda A) = m$ where $\Lambda = \Lambda(x^*)$.

Theorem 4.3 requires that if t^* solves DCP then

$$u_j(t^*) = \hat{x}_j^* \quad \text{whenever } j \in \Lambda(x^*)$$

Taking logarithms and writing $u_j(t^*)$ explicitly we find

$$\sum_{i=1}^m a_{ij} \log t_i^* = c_j + \log x_j^*, \quad j \in \Lambda(x^*)$$

This linear system in $\log t$ has a unique solution if, and only if, the rank of the coefficient matrix is m . Note that a solution t^* exists by Theorem 4.3 and that the above linear system is always consistent by Theorem 4.2. This completes the proof. \square

When the conditions of the theorem are satisfied, namely, the solution to DCP is unique, the same linear equations in $\log t$ imply the existence of a unique positive vector $y \in \mathbb{R}_+^n$ such that

$$A^i \log t = c + \log \hat{y}.$$

Although \hat{y} is strictly positive, a nonpositive solution x to CPI may still exist. Bigelow [7] calls y the vector of "virtual mole fractions." Of course, if $|I| = n$, there exists an $x \in \mathcal{D}(F|X^+)$, i.e., $x > 0$, and then \hat{x} and \hat{y} are identical.

As mentioned earlier, the type of uniqueness implied by Theorem 5.5 does not exclude identical phases and is therefore somewhat unrealistic for the chemist. As we shall subsequently show,

under some regularity conditions one can obtain a stronger uniqueness than was obtained in the past. The regularity conditions (on the dual chemical problem), while mathematically restrictive, do not pose much difficulty in real systems.

Consider the problem

$$\begin{aligned} & \text{Minimize } f(t) \\ & \text{Subject to } g_k(t) \leq 0, \quad k = 1, 2, \dots, K \end{aligned}$$

where $f(t)$ and $g_k(t)$ are all differentiable functions on some set $\Omega \subseteq \mathbb{R}^n$. This problem is said to be regular [5] at t if

- (i) $g_k(t) \leq 0, k = 1, 2, \dots, K$
- (ii) At the point t , the gradient vectors $\nabla g_k(t)$ of the active constraints ($k \in \bar{K}(t) \equiv \{k: g_k(t) = 0\}$) are linearly independent.
- (iii) $\nabla f(t)$ is interior to the cone generated by $\nabla g_k(t), k \in \bar{K}(t)$.

Conditions (ii) and (iii) imply that all active constraints are "strongly binding," that is, none of them can be relaxed without affecting the solution.

Theorem 5.10: Let $\{F, A, b\}$ be canonical and F ideal.

Let $x \in \mathcal{D}(F|X)$ and let $\Lambda = \Lambda(x)$. Suppose

- (i) $\text{rank}(\Lambda A) = m$
- (ii) The corresponding dual problem DCP is regular at an optimal solution $t = t(x)$. Then

x is the unique solution to $\{F, A, b\}$.

Proof: By Theorem 5.9, if $\text{rank}(\Lambda A) = m$, t is the unique solution to DCP. For the dual chemical problem, the regularity assumption means: $\eta_k > 0$ if, and only if, $g_k(t) = 1$. Here η_k , $k = 1, 2, \dots, K$ are the Lagrange multipliers in Theorem 4.2 (iii). From this theorem we also find that if $x \in \mathcal{D}(F|X)$

$$\bar{x}_k = \lambda_k = \eta_k / g_0(t) = \eta_k \exp[F(x)] > 0 \quad (1)$$

for all $k \in \bar{K}(t)$ and $\bar{x}_k = 0$ for $k \notin \bar{K}(t)$. Thus, according to Theorem 5.4, $x_j = 0$ for all j such that $k(j) \notin \bar{K}(t)$. This is true for all $x \in \mathcal{D}(F|X)$. By Theorem 5.5 for j such that $k(j) \in \bar{K}(t)$ we have

$$\hat{x}_j = \hat{y}_j \quad \text{whenever } x, y \in \mathcal{D}(F|X) .$$

But regularity assures a unique vector η of multipliers. Thus \bar{x}_k , $k \in \bar{K}(t)$ are also unique by (1). Hence $\bar{x}_k = \bar{y}_k$ and thus, for all $x, y \in \mathcal{D}(F|X)$

$$x_j = \hat{x}_j \bar{x}_{k(j)} = \hat{y}_j \bar{y}_{k(j)} = y_j \quad \square$$

In general we may assume without loss of generality that $\text{rank}(A) = m$. We then have

Corollary 5.11: If there exists an $x \in \mathcal{D}(F|X^+)$ and the dual problem DCP is regular at a solution $t(x)$, then x is the unique solution to CPI.

Corollary 5.12: If $K = 1$ (single phase) and the system is canonical, it has a unique solution $x = \mathcal{D}(F|X^+)$.

Proof: We assume that $\text{rank}(A) = m$. Canonicity and boundedness of X imply $b \neq 0$ so that $x \in X$ implies $x \neq 0$. In particular $x \in \mathcal{D}(F|X)$ is nonzero, so $\bar{x} > 0$ and by Theorem 5.4, $x > 0$. The result then follows from Theorem 5.10, even without requiring regularity. \square

We have defined regularity in terms of the dual chemical problem. It is interesting to see what these conditions mean in chemical terms. Suppose that at some $x \in \mathcal{D}(F|X)$ and $t = t(x)$ solving DCP, the problem is not regular. In particular, let $g_\ell(t) = 1$ but $\eta_\ell = 0$. Then $\bar{x}_\ell = 0$. Let $\epsilon > 0$ be a small number and consider a change in some c_s , $s \in \langle \ell \rangle$ such that

$$\tilde{c}_s = c_s - \epsilon.$$

Evaluating the perturbed constraint we have

$$\begin{aligned} \tilde{g}_\ell(t, \epsilon) &= \sum_{\substack{j \in \langle \ell \rangle \\ j \neq s}} \left[e^{-c_j} \prod_{i=1}^m t_i^{a_{ij}} \right] + e^{\epsilon - c_s} \prod_{i=1}^m t_i^{a_{is}} \\ &= g_\ell(t) + (e^\epsilon - 1) e^{-c_s} \prod_{i=1}^m t_i^{a_{is}} > 1 \end{aligned}$$

Therefore the constraint ℓ becomes binding and a solution of the perturbed DCP will yield $\mu_\ell(\epsilon) > 0$ implying $\bar{x}_\ell(\epsilon) > 0$. Phase Φ_ℓ

which did not exist, appears after the perturbation for any $\epsilon > 0$. This indicates instability, or rather, discontinuity in the chemical system, **reflected by discontinuity in concentrations.**

As the c_j 's are usually empirical numbers with finite accuracy, the probability of such instability in practice is zero, **so chemists can be content** knowing that chemical systems are regular and thus are likely to have a unique equilibrium composition.

5.5 Extensions to Nonideal Systems

Geometric programming duality serves well in analyzing the ideal case. Unfortunately, no parallel theory was developed for the more general nonideal case. Moreover, there is no guarantee that a meaningful duality even exists. In this section we shall assume only homogeneity of degree one and convexity, as postulated in Chapter 3. We shall examine some conditions leading to dual problems, particularly "pure" duals.

A major complication in subsequent analysis is that $F(x)$ may be undefined and surely is not differentiable when $x \not> 0$. We shall bypass these obstacles by sacrificing some **generality and dealing** only with positive compositions. Our basic problem is CPN (Section 2.5).

Theorem 5.13: Let F be differentiable on \mathbb{E}_+^n , convex and homogeneous of degree one. Consider problem CPN. Then there exists an $x^* \in \mathcal{D}(F|X^+)$ if, and only if, there exists a vector $z^* \in \mathbb{E}^m$ such that z^* solves the problem:

$$\begin{aligned} & \text{Maximize } z'b \\ & \text{Subject to } z'A \leq \mu'(x^*) \end{aligned}$$

Furthermore,

$$z^*b = F(x^*) \quad \text{and} \quad A'z^* = \mu(x^*)$$

Here

$$F(x) = \mu'(x) \cdot x$$

$\mu(x)$ is the vector of chemical potentials.

Proof: By direct application of Kuhn-Tucker conditions [36] to the convex, linearly constrained problem CPN, the necessary and sufficient conditions based on the Lagrangean

$$L(x, z) = F(x) - z'(Ax - b)$$

are

- (i) $Ax^* = b$
- (ii) $\nabla_x L(x^*, z^*) \geq 0$, i.e., $\mu(x^*) - z^*A \geq 0$
- (iii) $z^*Ax^* = \mu(x^*) \cdot x^* = F(x^*)$ (complementary slackness).

From (i), (ii)

$$\mu(x^*) \cdot x^* = F(x^*) \geq z^*Ax^* = z^*b$$

From (i), (iii) and the condition $x^* > 0$

$$z^*Ax^* = z^*b = F(x^*), \quad A'z^* = \mu(x^*)$$

The rest is clear. \square

The dual problem implied by the theorem is not "pure" for it involves primal variables x . Eisenberg [28] studied the existence of "pure" dual problems to "homogeneous programs" (where $F(x)$ is homogeneous of degree one), but his assumptions are not valid in our case.

Let $F(x) = \mu'(x) \cdot x$ and suppose

- (1) $\mu_j(x) = \mu_j(\hat{x}_j)$, $j \in \langle k \rangle$, $k = 1, 2, \dots, K$
- (2) $\mu_j(\hat{x})$ is well defined for $\hat{x}_j > 0$, $j \in N$
- (3) $\mu_j(\hat{x}_j)$ is strictly monotone in \hat{x}_j and $\mu_j^{-1}(y)$ exists for all y in the row space of A .

Then the constraints of Theorem 5.13, namely,

$$A'z \leq \mu(\hat{x}^*)$$

can be inverted

$$\mu_j^{-1}(A_j'z) \leq \hat{x}_j^*$$

Summation over $j \in \langle k \rangle$ yields

$$g_k(z) = \sum_{j \in \langle k \rangle} \mu_j^{-1}(A_j'z) \leq 1$$

Therefore, if (1), (2) and (3) hold, a pure dual resembling a geometric program can be formed. The hardest condition to accept is probably condition (3). Condition (1), although quite restrictive, is nevertheless common in ideal or nearly ideal systems. It is natural to expect that $\mu(x) = \mu(\hat{x})$ since μ is homogeneous of degree zero.

In practical nonideal systems, $\mu(x)$ is based in many cases on empirical relations (see Section 8.7 for example) which are valid in some range of x but need not preserve either homogeneity or convexity. The best one can hope to do in the general situation is to be able to handle such problems computationally.

We conclude our short treatment of the general case with

Proposition 5.14: Let $F(x) = \mu(x) \cdot x$ be differentiable on \mathbb{R}_+^n , convex and homogeneous of degree one. Consider problem CPN. Then

$$x \in \mathcal{D}(F|X^+)$$

if and only if

$$\theta' \mu(x) = 0$$

for every reaction vector θ .

Proof: By Theorem 5.13 there is a z such that

$$z'A = \mu(x)$$

Consequently

$$z'A\theta = \mu(x) \cdot \theta$$

but by definition of reaction vectors $A\theta = 0$. Thus

$$0 = \mu(x) \cdot \theta \quad \square$$

The relation $\mu(x) \cdot \theta = 0$ is the analog to the logarithm of the mass action laws (Section 3.2). Restoring the usual form, we can write the generalized mass action laws as

$$\prod_{j=1}^n e^{\theta_j \mu_j(x)} = 1$$

In the ideal case: $\mu_j(x) = c_j + \log \hat{x}_j$. We have

$$\prod_{j=1}^n \hat{x}_j^{\theta_j} = \exp(-c' \cdot \theta) \equiv K(\theta)$$

as was shown in Chapter 3.

CHAPTER 6

APPLICATIONS OF CHEMICAL DUALITY

Several applications of the duality presented in Chapter 4 will be demonstrated in this chapter. A major application, computational algorithms based on duality, will be discussed in the next chapter.

The chemical system is $\{B, S, \bullet\}$, characterized mathematically by $\{F, A, b\}$ as described in Chapter 2. The applications presented are

- a. Testing for equilibrium, given a feasible composition.
- b. Verification of the model--upper bounds on concentrations of trace components.
- c. Sensitivity analysis--changes of the solution with variation in the free energy parameters.
- d. Goaling techniques--solution of problems with side conditions ("goals") on the equilibrium composition.

6.1 Testing and Characterization of Equilibrium

Chemical duality can be applied in a simple and straightforward way to test whether a given composition-vector $x \in X$ is an equilibrium solution for an ideal system. The results of the

previous Chapter, especially Theorem 5.4, indicate that if $x \in \mathcal{D}(F|X)$ (the "equilibrium set") then $x_j = 0$ implies $\bar{x}_{k(j)} = 0$, so we assume that this condition is satisfied. For an ideal system, geometric programming duality Theorem 4.2 and the form of the dual chemical problem DCP indicate that $x \in \mathcal{D}(F|X)$ if and only if there exists $z \in \mathbb{R}^m$ such that

$$\sum_{i=1}^m a_{ij} z_i = c_j + \log x_j \quad \text{for } j \in N \text{ such that } \bar{x}_{k(j)} > 0 \quad (1)$$

This equation can be easily obtained by applying the Kuhn-Tucker Conditions (see [7]). Thus, an optimality test will consist of solving the linear system (1) for z . In any nontrivial case, the number of equations will exceed the number of variables, so that this is a test for the consistency of system (1). A measure of how far x is from the optimum can be obtained by finding the "least squares" solution to the system. Let

$$\begin{aligned} \Lambda &= \Lambda(x) \equiv \{j | x_j > 0\} \\ \Lambda A &= \{A_j | j \in \Lambda\} \\ \Lambda \hat{x} &= \{\hat{x}_j | j \in \Lambda\} \\ \Lambda c &= \{c_j | j \in \Lambda\} \end{aligned}$$

Then the least squares solution z is given by

$$z = [\Lambda A (\Lambda A)']^{-1} \cdot \Lambda A \cdot [\log(\Lambda \hat{x}) + \Lambda c]$$

where ΛA is assumed to be of full rank m . If x is an equilibrium solution then the error ϵ is

$$\epsilon = (\Lambda A)' z - \log(\Lambda \hat{x}) - \Lambda c = 0$$

Otherwise, ϵ^2 is an appropriate measure of deviation from equilibrium. Note that ϵ^2 is actually a measure of the deviation from a solution which satisfies the mass action laws.

6.2 Verification of the Model

This application was first noted by White et al. [54]. We expand their result by adding bounds on concentrations of added species. A model $\{B, S, \Phi\}$ is based on the hypothesis that S is indeed the right set of species and that all other possible species may appear only in negligible amounts. The purpose here is to test this hypothesis.

It is assumed that an equilibrium composition x for the model is known, together with the dual chemical variables z computed by (1.1). We wish to test whether the exclusion of some species $S_q \notin S$ is justified. We assume that $k(q) \in \Phi$ and $\bar{x}_{k(q)} > 0$. Let $A_q = \{a_{iq}\}$ be the formula-vector of S_q . By (1.1) we must have at equilibrium

$$\sum_{i=1}^m a_{iq} z_i - c_q = \log(\hat{x}_q) \quad (1)$$

Thus, the approximate concentration of S_q , had it been included in the model, would be \hat{x}_q^0

$$\hat{x}_q^0 = \exp\left[\sum_{i=1}^m a_{iq} z_i - c_q\right] \quad (2)$$

If \hat{x}_q^0 is small (say $\hat{x}_q^0 < 10^{-5}$), one can usually justify the exclusion of S_q . The true value of \hat{x}_q , had S_q been included, is of course different since with S_q in the model, z is no longer optimal. In many situations, one can obtain an upper bound on any \hat{x}_j (whether $S_j \in S$ or not) without ever solving the problem, by noting that since $\hat{x}_j \leq 1$, we always have

$$\sum_{i=1}^m a_{ij} z_i \leq c_j, \quad j = 1, 2, \dots, n \quad (3)$$

when z solves DCP. To find a bound on \hat{x}_q we solve the linear program

$$\text{Maximize } y_q = \sum_{i=1}^m a_{iq} z_i$$

Subject to $A'z \leq c$

Letting y_q^* be the maximum, we obtain by (2)

$$\hat{x}_q \leq \exp[y_q^* - c_q] \quad (4)$$

The right hand side may be greater than unity, in which case the bound is useless. This happens when $y_q^* > c_q$. Solving a linear program to obtain a single bound is of course highly inefficient. In many practical situations, upper bounds can be found by inspection. Recall that there is no loss of generality in assuming that $A \geq 0$ for a real system (ignoring electroneutrality), and that each subspecies B_i

is also included as a species $S_{j(i)}$, where $j(i)$ is the index of the "species" B_i . Using this idea in (3) we have

$$z_i \leq c_{j(i)} \quad (5)$$

because

$$a_{lj(i)} = \begin{cases} 1 & l = i \\ 0 & \text{otherwise} \end{cases}$$

If $A \geq 0$ then for any j , and any z solving DCP

$$\sum_{i=1}^m a_{ij} z_i \leq y_j \equiv \sum_{i=1}^m a_{ij} c_{j(i)} \quad (6)$$

Similar to equation (4), we arrive at

$$\hat{x}_q \leq \exp\{y_q - c_q\} \quad (7)$$

Again, if $y_q > c_q$, the bound is useless since the right hand side is greater than one. However, we found this technique very useful in obtaining bounds and initial approximations with the dual algorithm of the next chapter.

Example 6.1. Suppose one wishes to test whether ozone O_3 is justly excluded from the hydrazine model (Appendix A.3). It is assumed that this could be species S_{11} with $c_{11} \simeq -14$ (approximated for 3500° from data in [38]). The formula vector for O_3 is $(0,3,0)$, corresponding to subspecies (H,O,N) . Now observe that $S_8 = \text{oxygen } (0)$

and $c_8 = -10.708$. Thus, $z_2 \leq -10.708$ from equation (5). Then, by (6)

$$y_{11} = 3 \cdot c_8 = -32.124$$

Equation (7) leads to

$$\hat{x}_{11} \leq \exp[-32.124 - (-14)]$$

$$\hat{x}_{11} \leq e^{-18.124} < 10^{-7}$$

The small concentration indicates that O_3 is only a trace element and can be excluded without affecting the model.

Up until now it was assumed that the phase $\Phi_{k(q)}$ of the excluded species S_q , exists at equilibrium. Difficulties arise when the phase does not exist, i.e., $\bar{x}_{k(q)} = 0$ at equilibrium. We wish to test whether inclusion of S_q in the model would have caused the phase to appear.

Assume that the dual chemical problem DCP is regular at a dual solution point z (Section 5.4). For the dual chemical problem DCP (without S_q) we have

$$g_{k(q)}^0 = \sum_{j \in \langle k(q) \rangle} \exp\left[\sum_{i=1}^m a_{ij} z_i - c_j\right] < 1.$$

After inclusion of S_q we obtain

$$g_{k(q)} = g_{k(q)}^0 + \exp\left[\sum_i a_{iq} z_i - c_q\right] \quad (8)$$

If $g_{k(q)} < 1$ then $\bar{x}_{k(q)} = 0$ and the phase $\Phi_{k(q)}$ still vanishes. Otherwise, the phase appears, in which case the concentrations of species S_j , $j \in \Phi_{k(q)}$ can be approximated by $u_j(z)/g_{k(q)}$, i.e.,

$$\hat{x}_j = \frac{1}{g_{k(q)}} \cdot \exp\left[\sum_{i=1}^m a_{ij} z_i - c_j\right], \quad j \in \langle k(q) \rangle$$

The vector z is no longer the exact optimal solution to the dual DCP of the expanded problem with S_q . If $\bar{x}_{k(q)}$ is not small, the mass balance equations are violated and one has to resolve the problem with S_q included in the model.

6.3 Sensitivity Analysis

When the parameters of a given system are changed, its equilibrium composition may also change. Bigelow and Shapiro [9] analyzed the variation of the equilibrium composition x with small changes of A , b , and c by using geometric programming arguments. A limited earlier version appeared in Duffin et al. [25]. Later analysis by Theil [49] was concerned with variations in primal geometric variables, the variables of the dual chemical problem.

For most practical purposes, the changes of interest occur primarily in the free energy coefficients c_j when the system is ideal. A special interesting case occurs when the vector c changes

with a single parameter, the temperature T for example. In biochemical systems, where changes are usually small due to internal stabilizing mechanisms, sensitivity analysis can analyze corresponding changes in composition without solving the problem afresh.

We present here an independent approach equivalent to that of [9].

Consider the ideal system $\{F, A, b\}$ where

$$F(x) = x \cdot (c + \log \hat{x})$$

Let the corresponding dual chemical problem be (Section 4.5):

Problem DCP

$$\text{Minimize } \bar{g}_0(z) \equiv -\log[g_0(z)] = -b'z$$

$$\text{Subject to } g_k(z) \equiv \sum_{j \in (k)} u_j(z) \leq 1, \quad k = 1, 2, \dots, K$$

where

$$u_j(z) \equiv \exp\left[\sum_{i=1}^m a_{ij} z_i - c_j\right]$$

For z to be a solution to DCP, the following necessary conditions must hold for some vector $\eta \in \mathbb{R}^K$

$$\nabla_z [\bar{g}_0(z) - \sum_{k=1}^K \eta_k (1 - g_k(z))] = 0 \quad (1)$$

$$g_k(z) \leq 1 \quad (2)$$

$$k = 1, 2, \dots, K$$

$$\eta_k \geq 0 \quad (3)$$

$$\eta_k (1 - g_k(z)) = 0 \quad (4)$$

Computing (1), we arrive at

$$-b_i + \sum_{k=1}^K \eta_k \left[\sum_{j \in \langle k \rangle} a_{ij} u_j(z) \right] = 0, \quad i = 1, 2, \dots, m \quad (5)$$

Let $c + \alpha \tilde{c}$ be the new coefficient vector with \tilde{c} being a direction vector and α a parameter. The corresponding changes in η , $u(z)$, and z are $\tilde{\alpha}\eta$, $\tilde{\alpha}u$, $\tilde{\alpha}z$, respectively. α can be ignored since our interest is only in directions of change. We assume that these directional derivatives exist, at the solution point z . (Existence requires that the conditions of the implicit function theorem be satisfied for problem DCP. This topic is thoroughly studied in [9].) Conditions (1)-(4) must hold also for $c + \alpha \tilde{c}$, $\eta + \alpha \tilde{\eta}$, $u + \alpha \tilde{u}$, $z + \alpha \tilde{z}$, so that (5) leads to

$$\sum_k \tilde{\eta}_k \cdot \left[\sum_{j \in \langle k \rangle} a_{ij} u_j(z) \right] + \sum_k \eta_k \cdot \left[\sum_{j \in \langle k \rangle} a_{ij} \tilde{u}_j(z) \right] = 0 \quad (6)$$

From the definition of u (which is considered a constraint)

$$\begin{aligned} \tilde{u}_j(z) &= \left[\sum_i a_{ij} \tilde{z}_i \right] \cdot u_j(z) - \tilde{c}_j u_j(z) \\ &= u_j(z) \cdot \left[\sum_i a_{ij} \tilde{z}_i - \tilde{c}_j \right] \end{aligned} \quad (7)$$

If $g_k(z) = 1$, we can write by (2)

$$\tilde{g}_k(z) \leq 0 \quad (8)$$

From (3)

$$\tilde{\eta}_k \geq 0 \quad \text{if} \quad \eta_k = 0 \quad (9)$$

and

$$\tilde{\eta}_k \cdot (1 - g_k(z)) - \eta_k \cdot \tilde{g}_k(z) = 0 \quad (10)$$

Assume that the problem DCP is regular at z . Then

$$\eta_k = 0 \quad \text{implies} \quad g_k(z) < 1$$

In this case, if $g_k(z) = 1$ then (10) implies $\tilde{g}_k(z) = 0$ so that (8) is redundant. Similarly, if $\eta_k(z) = 0$ then (10) implies $\tilde{\eta}_k(z) = 0$ so that (9) is redundant.

For the regular case, therefore, (6), (7), and (10) cover all the conditions.

Let

$$h_{ik} \equiv \sum_{j \in \langle k \rangle} a_{ij} u_j(z)$$

and let

$$d_j \equiv \eta_{k(j)} \cdot u_j(z)$$

Combining (6) and (7), changing order of summation, and inserting the definitions of h and d leads, after some algebraic manipulations, to

$$\sum_k \tilde{\eta}_k \cdot h_{ik} + \sum_j a_{ij} d_j \cdot \left(\sum_j a_{ij} \right) \cdot \tilde{z}_i = \sum_j a_{ij} d_j \tilde{c}_j \quad (11)$$

for $i = 1, 2, \dots, m$

Equation (10) leads to

$$\tilde{\eta}_k - \sum_{j \in \langle k \rangle} \tilde{\eta}_k \cdot u_j(z) - \eta_k \sum_i h_{ik} \cdot \tilde{z}_i + \sum_{j \in \langle k \rangle} d_j \tilde{c}_j = 0$$

$$\text{for } k = 1, 2, \dots, K \quad (12)$$

Let

$$A = \{a_{ij}\}, \quad A \in \mathbb{F}^{m \times n}$$

$$D = \text{diag}(d_1, d_2, \dots, d_n)$$

$$H = \{h_{ik}\}, \quad H \in \mathbb{F}^{m \times K}$$

$$\tilde{\eta} = \{\tilde{\eta}_k\}, \quad \tilde{\eta} \in \mathbb{F}^K$$

$$\Lambda = \text{diag}(\eta_1, \eta_2, \dots, \eta_K)$$

$$1_k = (1, 1, \dots, 1) \in \mathbb{F}^K$$

$$G(z) = \text{diagonal matrix in } \mathbb{F}^{K \times K}, \quad G_{kk}(z) = 1 - g_k(z)$$

$$E = \{e_{kj}\}, \quad E \in \mathbb{F}^{K \times n}$$

and

$$e_{kj} = \begin{cases} 1 & j \in \langle k \rangle \\ 0 & \text{otherwise.} \end{cases}$$

With these definitions and some additional algebra we arrive at

$$\begin{bmatrix} ADA' & | & AHA' \\ \hline & & \\ -\Lambda H & | & G(z) \end{bmatrix} \cdot \begin{bmatrix} \tilde{z} \\ \tilde{\eta} \end{bmatrix} = \begin{bmatrix} A \\ \hline -E \end{bmatrix} \cdot D\tilde{c} \quad (13)$$

This is a linear system of $m+K$ equations in $m+K$ unknowns-- \tilde{z} and $\tilde{\eta}$. If this system is nonsingular (which we assume, of course), \tilde{z} and $\tilde{\eta}$ can be computed explicitly for any given \tilde{c} (provided that the solution at c , namely z and η , is known). To recover the composition-vector x , our main interest, recall that $u_j(z) = \hat{x}_j$ whenever $\bar{x}_{k(j)} > 0$. Applying equation (5) we obtain the mass balance equations

$$\sum_j a_{ij} \cdot [u_j(z) \cdot \eta_{k(j)}] = b_i, \quad i = 1, 2, \dots, m$$

whence

$$u_j \cdot \eta_{k(j)} = x_j = \hat{x}_j \cdot \bar{x}_{k(j)} \quad (14)$$

Therefore

$$\tilde{x}_j = \tilde{u}_j \cdot \eta_{k(j)} + u_j \cdot \tilde{\eta}_{k(j)} \quad (15)$$

\tilde{u}_j can be computed from \tilde{z} via (7). $\tilde{\eta}$ is found directly from (13).

In conclusion, we have found the directional change \tilde{x} due to a change \tilde{c} in the coefficients.

We shall now examine a special case where c changes with the temperature, i.e., $c = c(T)$ and

$$\tilde{c} = \frac{dc(T)}{dT}$$

Recall that by equations (2.4.5), (2.4.6) in the ideal case

$$c_j(P,T) = \frac{\mu_j^0(P,T)}{RT} + \log[\gamma_j(P,T)] \quad (16)$$

We assume that $\gamma_j(P,T)$ is independent of T and that P is fixed so that $\gamma_j(P,T) = \gamma_j$ (constant). From elementary thermodynamics [21, p. 143]

$$\frac{d[\mu_j(x,P,T)/RT]}{dT} = - \frac{\Delta H_j(T)}{RT^2} \quad (17)$$

Here $\Delta H_j(T)$ is the enthalpy of formation of S_j at T . R is the gas constant. Furthermore

$$\frac{d[\Delta H_j(T)]}{dT} = C_{pj}(T) \approx \alpha_j + \beta_j T + \gamma_j T^2 \quad (18)$$

where C_{pj} is the heat capacity, α , β , and γ are empirical coefficients. We can integrate the last relation, adding an integration constant \bar{H}_j . Substituting in (17), integrating again, and identifying μ_j with c_j (for the pure species), we obtain

$$c_j^1 - c_j^0 = \frac{1}{R} \left[\bar{H}_j \cdot \left(\frac{1}{T_1} - \frac{1}{T_0} \right) - \alpha_j \cdot \log\left(\frac{T_1}{T_0}\right) - \frac{\beta_j}{4} \cdot (T_1^2 - T_0^2) - \frac{\gamma_j}{9} \cdot (T_1^3 - T_0^3) \right]$$

where the integration is from initial absolute temperature T_0 to final temperature T_1 and $c_j^0 = c_j(T_0)$. For small changes of T ,

like those expected in biological systems, the first term is usually dominant. Let $\bar{H} = (\bar{H}_1, \bar{H}_2, \dots, \bar{H}_n)$ and taking the (finite) difference $c^1 - c^0$ as direction we obtain the well known formula

$$\tilde{c} = \frac{\bar{H}}{R} \left(\frac{1}{T_1} - \frac{1}{T_0} \right)$$

Values of \bar{H} are typically tabulated in handbooks of chemistry for the temperature $T = 298.16^\circ\text{K}$. Changes in temperature can thus be directly translated to changes in composition.

6.4 Goaling Problems

Goaling problems are variations of the classical chemical equilibrium problem. They differ from the standard problem in that some predetermined conditions ("goals") are imposed on the equilibrium composition. An attempt to solve the problem by simply adding the goals to the mass balance constraints will violate the principle of internally unconstrained equilibrium, discussed in Section 3.5. To compensate for the loss of internal "degrees of freedom" due to the goals, one normally relaxes some of the "external" mass balance constraints. This is done by allowing an open system, in which some species can be freely added externally to achieve an equilibrium where the goals are satisfied. Although goaling problems do not resemble our familiar chemical equilibrium problem, duality can in some instances modify the problem to put it in a standard chemical

equilibrium format. Our main result here is a generalization of a procedure for concentration-goaling by DeHaven [19]. The problem is:

Problem CPG

Given an ideal system (F, A, b) , find the equilibrium composition and the amount of species S_r to be added (or removed) externally, such that the equilibrium concentration of S_r is some given value \hat{x}_r . DeHaven [19] and the RAND chemical composition code [48] deal only with the case where the species appears also as a subspecies. We shall relax this assumption. Furthermore, we show how to extend the procedure to several simultaneously goaled species. It is assumed that the problem is canonical and has a solution. The formula-vector of S_r is A_r .

Theorem 6.2: Problem CPG is equivalent to the chemical equilibrium problem $CP^\#$ defined by the triplet $(F^\#, A^\#, b^\#)$ where

$$F^\# = F + [-(c_r + \log \hat{x}_r) + \log \hat{x}_{n+1}]x_{n+1}$$

$$A^\# = (a_{ij}^\#) = \begin{cases} a_{ij} & 1 \leq j \leq n \\ -a_{ir} & j = n+1 \end{cases}$$

$$b^\# = (b_i^\#) = b_i - \theta \cdot a_{ir}$$

$$\theta = \min_i (b_i / a_{ir} | a_{ir} > 0)$$

Proof: First, note that $CP^\#$ was constructed from CPG by adding a species S_{n+1} which is in a separate phase Φ_{K+1} . $CP^\#$ is canonical since, by hypothesis, the underlying ungoal program CP is. To see this let $x \in X^+(A, b)$. For any $\epsilon > 0$, let

$$x_j^\# = \begin{cases} x_r + \epsilon & j = r \\ \epsilon + \theta & j = n+1 \\ x_j & \text{otherwise} \end{cases}$$

Then $x^\# > 0$ and

$$\begin{aligned} A^\# \cdot x^\# &= Ax + A_r \epsilon - A_r \cdot (\epsilon + \theta) \\ &= Ax - A_r \theta = b - A_r \theta = b^\# \end{aligned}$$

Thus

$$x^\# \in X^{\#+}(A^\#, b^\#).$$

Next, observe that if $l = \arg \min_i (b_i/a_{ir} | a_{ir} > 0)$, then x_{n+1}/a_{1l} equals the total number of moles of subspecies l , which is unknown due to the open ended nature of the problem.

If problem CPG has a solution x^* by adding x_{n+1}^* moles of S_r , then there is a solution to problem CP^* defined by (F, A, b^*)

where $b^* = b + A_r \cdot x_{n+1}^*$. The solution x^* satisfies the goal:

$$\hat{x}_r^* = \hat{x}_r.$$

Suppose x_{n+1}^* were known, then by duality, there exists a vector z such that

$$z'A_j = c_j + \log(\hat{x}_j^*) \quad \text{for all } j \text{ such that } \hat{x}_{k(j)}^* > 0$$

In particular

$$z'A_r = c_r + \log(\hat{x}_r^*)$$

From the solution x^* to CP^* we construct a solution $x^\#$ to $CP^\#$ as follows

$$x_j^\# = \begin{cases} x_j^* & 1 \leq j \leq n \\ \theta + x_{n+1}^* & j = n+1 \end{cases}$$

Indeed

$$\begin{aligned} A^\# \cdot x^\# &= Ax^* - A_r(x_{n+1}^* + \theta) \\ &= b + A_r x_{n+1}^* - A_r \theta - A_r x_{n+1}^* = b^\# \end{aligned}$$

Clearly, if x^* solves CP^* then z above will satisfy

$-z'A_r = -(c_r + \log x_r^*) + \log 1$. Therefore $x^\#$ is a solution to $CP^\#$. Conversely, any solution $x^\#$ to $CP^\#$ must have a dual chemical solution z such that

$$-z'A_r = -(c_r + \log(\hat{x}_r^\#))$$

Hence

$$\log x_r^\# = -c_r + z'A_r = \log \hat{x}_r$$

It is clear from the construction that $x^\#$ will satisfy the mass balance constraints and $x_{n+1}^\# \neq 0$. If $x_{n+1}^\# = 0$ all species with $a_{1l} > 0$ must vanish, but we can assume without loss of generality that at least one such species exists in each phase. Theorem 5.4 then leads to $x_j = 0$ for all j , which is contrary to the hypothesis that a solution exists. This completes the proof. \square

Corollary 6.3: The procedure indicated by the theorem can be extended to simultaneous goaling of several species.

Proof: One can proceed inductively, goaling on species S_r generates problem $CP^\#$. Starting with $CP^\#$ and goaling on S_s generates $(CP^\#)^\#$ and so on. \square

At most m independent species can be goaled this way, but even a smaller number may lead to inconsistencies.

We are unable to find a similar method for the more general problem of goaling S_r by adding or removing a different species S_s . It is possible, to cast this problem in the form of a parametric chemical equilibrium problem with a single parameter α .

Let

$$l \equiv \arg \min_i \{b_i/a_{is} \mid a_{is} > 0\}$$

and let

$$\theta \equiv b_l/a_{ls}$$

θ is interpreted as the total amount of S_s currently in the input.

Set $b^* \equiv b - \theta A_s$ where A_s is the formula vector of S_s . Let

$\alpha \geq 0$ be a parameter and set

$$b^{\#} \equiv b^{\#}(\alpha) \equiv b^* + \alpha A_s$$

The problem is now characterized by $\{F, A, b^{\#}(\alpha)\}$. To solve it, one has to find α , and the equilibrium solution $x(\alpha)$, such that $\hat{x}_r(\alpha)$ will have the goaled value. The amount of S_s to be added (or removed) is $\alpha - \theta$, assuming that θ is the original input of S_s .

The class of goaling problems can be generalized further by requiring the equilibrium composition x to minimize a given function $\varphi(x)$. Again, this goal is achieved by allowing an open system with varying input. The problem can thus be formulated as follows

Problem CP(Ω)

Let Ω be a subset of \mathbb{R}^m .

Find $y \in \Omega$ and a composition-vector $x^* \equiv x^*(y)$

$$x^* \in X(A, y)$$

such that

$$x^* \in \mathcal{D}(F|X(A, y))$$

and $\varphi(x^*) \leq \varphi(x)$ for all $w \in \Omega$ and x , such that

$$x \in \mathcal{D}(F|X(A, w)).$$

As a simple example, consider the case where

$$\Omega = \{y \in \mathbb{R}^m \mid y \geq 0, \|y\| = 1\}$$

and

$$\varphi(x) = \hat{x}_r^{-1}.$$

In chemical terms, we wish to find the input vector y (replacing the usual fixed b) such that \hat{x}_r , the concentration of S_r , is maximized. The solution is trivial--an input of pure S_r . To see this, notice that in the dual chemical problem $DCP(\Omega)$ associated with $CP(\Omega)$, nothing changes except the objective function which now reads

$$\text{minimize } -y'z.$$

\hat{x}_r is identified with the term $u_r(z)$

$$u_r(z) = \exp[\sum_i a_{ir} z_i - c_r]$$

To maximize \hat{x}_r we can maximize $u_r(z)$ in the dual chemical problem. The problem is then: find $y \in \Omega$ such that the solution $z(y)$ to $DCP(\Omega)$ will also be a solution to $\max_y \{u_r(z)\}$ where $z = z(y)$. The goaling condition (in the dual) can be written

$$\text{minimize } \psi(z) \equiv - \sum_i a_{ir} z_i.$$

By choosing $y_i^* = \frac{a_{ir}}{\|A_r\|}$,

one can easily verify that $y^* \in \Omega$ and that any solution z to the dual problem of $CP(y^*)$ will also minimize $\psi(z)$. In this case y^* was found even without solving $CP(\Omega)$. The equilibrium composition can now be found by solving the problem $CP(y^*)$, characterized by (F, A, y^*) .

A useful example of this type is when

$$\varphi(x) = \hat{x}_r / \hat{x}_s$$

Equivalently, we wish to maximize \hat{x}_s / \hat{x}_r . This particular application is of interest when the equilibrium products have to be separated, for instance, when the solution yields some undesired byproduct. The cost of separation depends significantly on the concentration ratio of the species involved. Naturally, one wishes to maximize the ratio of desired product to undesired product.

In an analogous way to the method described before, the problem is: minimize $\psi(z) = \sum_i (a_{ir} - a_{is})z_i$ over all $y \in \Omega$ such that z solves the dual problem $DCP(y)$, i.e., over all $y \in \Omega$ such that z solves

$$\begin{aligned} &\text{Minimize } -z'y \\ &\text{Subject to } g_k(z) \leq 1, \quad k = 1, 2, \dots, K. \end{aligned}$$

Unfortunately, here the solution is not immediate, thus some iterative procedure is needed. Goaling problems are closely related to a problem treated by Dantzig et al. [16] and to the so-called "pooling problem."

In general one can apply the Kuhn-Tucker conditions to create a mixed primal-dual problem, which for problem $CP(\Omega)$ with goal function $\varphi(x)$ is

$$\begin{aligned} & \text{Minimize } \varphi(x) \\ & \text{Subject to } Ax = y \\ & \quad x \geq 0 \\ & \quad y \in \Omega \\ & \quad A'z = c + \log \bar{x} \end{aligned}$$

This form is not very attractive, as it resembles neither the primal nor the dual problems. Most of the useful structure of either primal or dual problems is lost. Finally we note that with goaling, the assumption of an open system implies that X is no longer bounded. Most of our previous results, especially those which assume existence of a finite $x \in \mathcal{D}(F|X)$ are still in force for the unbounded case. Unbounded problems were treated by Bigelow et al. [8], and by Kortanek et al. [35].

CHAPTER 7

A DUAL CHEMICAL ALGORITHM

7.1 Introduction and Review of Existing Methods

Most work on the chemical equilibrium problem has been directed toward development of computational algorithms. With increasing dimensions and complexity of the problems treated, neither ad hoc techniques for specific problems, nor the time-honored approach of solving a linear-nonlinear system of mass balance and mass action equations are adequate any more. Van Zeggeren and Storey [53] present a comprehensive review of techniques, divided into two broad classes, one using the classical approach of solving the mass action equations, the other using optimization techniques. Notable in the first class is the pioneering work of Brinkley [10].

Among the optimization techniques we mention Newton-type methods based on quadratic approximations by White et al. [54] and Dluzniewski et al. [22], the separable programming approach of Dantzig et al. [18], Dantzig's generalized programming method [15], Clasen's linear-logarithmic algorithms [12-14], and Bigelow's linear, quadratic, and dual algorithms [7]. All these methods can be classified as primal chemical algorithms since they all work directly with the

chemical problem, improving the composition vector at each iteration. Some of the primal chemical algorithms, especially Dantzig's [15], Clasen's [12-14] and Bigelow's methods [7] use dual multipliers within each iteration, but are nevertheless primal oriented. Passy and Wilde [41] were the first to implement a pure dual chemical method, but their approach is valid only for single-phase problems. Bigelow presented two "dual" methods and suggested some of their advantages, without reporting any computational experience.

After the relation of chemical equilibrium to geometric programming was realized, geometric programming problems were solved using chemical equilibrium codes, especially the highly efficient program developed at RAND [48]. In a sense, the algorithm presented here does just the opposite--it uses a geometric programming method to solve chemical problems.

Two points motivated the study of this dual chemical algorithm:

- (i) The dimensions of the problem are significantly reduced, from n variables and m constraints in the primal chemical problem, to m variables and K (nonlinear) constraints in the dual. This reduction can be advantageous both in computing time, and certainly in the storage requirements.
- (ii) The inability of the RAND code, and of most other primal chemical methods to handle degeneracies in the sense of Chapter 5. Problems with vanishing phases and linear dependencies due to ill-formulated problems are the two main causes of such degeneracies.

As our basic method, we have chosen an algorithm suggested by R. Dembo [20], originally developed to solve geometric programming problems with both positive and negative coefficients. Since the dual chemical problem involves only posynomials, we shall be interested only in the standard geometric programming part of his technique. In its original form the algorithm was found unsuitable for chemical problems due to the large coefficient values in the dual. After presenting the basic features of the algorithm we shall present a modified version of it, shown to be equivalent to Zangwill's concave cutting plane method [55]. The theory is followed by some computational results and a discussion of computational aspects.

7.2 The Geometric Inequality and Condensation

A general form of the well known inequality between the arithmetic and geometric weighted means of positive numbers states that

Theorem 7.1: Let u_1, u_2, \dots, u_n be positive numbers and let w_1, w_2, \dots, w_n be nonnegative numbers such that $\sum_{i=1}^n w_i = 1$.

Then

$$\sum_{i=1}^n w_i u_i \geq \prod_{i=1}^n u_i^{w_i} \quad (1)$$

Furthermore, equality holds if, and only if, there is a constant K such that $u_i = K$ for all i . A proof can be found in Hardy, Littlewood, and Polya [33].

This inequality called henceforth the "geometric inequality" plays a key role in geometric programming. It will now be put in a more general form, where the u_i 's and w_i 's are positive and nonnegative functions respectively.

Theorem 7.2. Let $\mathbb{R}_+ \equiv \{x \in \mathbb{R} | x > 0\}$, $\mathbb{R}^\# \equiv \{x \in \mathbb{R} | x \geq 0\}$. Let $u_i: \Omega_1 \subseteq \mathbb{R}^m \rightarrow \mathbb{R}_+$, $w_i: \Omega_2 \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^\#$, $i = 1, 2, \dots, n$, be functions, such that $\sum_{i=1}^n w_i(x) = 1$ for some $x \in \Omega_2$.

When $w_i(x) = 0$ we define $[w_i(x)]^{w_i(x)} = 1$. Then

$$\sum_{i=1}^n u_i(t) \geq \prod_{i=1}^n \left[\frac{u_i(t)}{w_i(x)} \right]^{w_i(x)} \quad \text{for all } t \in \Omega_1 \quad (2)$$

Furthermore, equality holds if and only if there is a constant K such that

$$\frac{u_i(t)}{w_i(x)} = K \quad \text{for all } i$$

This theorem is the basis for condensation of posynomials.

Let $t \in \Omega \subseteq \mathbb{R}^m$

$$g_k(t) = \sum_{j \in \langle k \rangle} u_j(t)$$

$$u_j(t) = c_j \prod_{i=1}^m t_i^{a_{ij}}, \quad j \in \langle k \rangle$$

From Chapter 4 we recall that $g_k(t)$ is called a "posynomial" and $u_j(t)$ a "posynomial term."

Let $t^0 > 0$, $t^0 \in \Omega$ and define

$$w_j(t^0) = \frac{u_j(t^0)}{g_k(t^0)} \quad (3)$$

The theorem can now be applied

$$\sum_{j \in \langle k \rangle} u_j(t) \geq \prod_{j \in \langle k \rangle} \left[\frac{u_j(t)}{w_j(t^0)} \right]^{w_j(t^0)} \equiv \tilde{g}_k(t, t^0) \quad (4)$$

The function $\tilde{g}_k(t, t^0)$, defined as the right hand side of the inequality, is called the condensation of $g_k(t)$ at t^0 , the term coined by Duffin [24].

The condensation of a posynomial function approximates it by a single term posynomial. The approximation has all the properties of a first order Taylor's expansion of a convex function, as shown by the following lemma.

Lemma 7.3: Let $\tilde{g}_k(t, t^0)$ be as defined in (4). Then

- (i) $\tilde{g}_k(t^0, t^0) = g_k(t^0)$
- (ii) $\nabla \tilde{g}_k(t^0, t^0) = \nabla g_k(t^0)$
- (iii) $\tilde{g}_k(t, t^0) \leq g_k(t)$ for all $t \in \Omega$.

Proof:

$$(i) \frac{u_j(t^0)}{w_j(t^0)} = \frac{u_j(t^0) \cdot g_k(t^0)}{u_j(t^0)} = g_k(t^0) \quad \text{for all } j \in \langle k \rangle$$

Thus, the condition for equality in Theorem 7.2 holds.

Substituting t^0 for t in (4) we obtain the result.

$$(ii) \frac{\partial g_k(t^0)}{\partial t_i} = \frac{1}{t_i^0} \sum_{j \in \langle k \rangle} a_{ij} u_j(t^0)$$

$$\frac{\partial \tilde{g}_k(t^0, t^0)}{\partial t_i} = \frac{1}{t_i^0} \left[\prod_{j \in \langle k \rangle} \left[\frac{u_j(t^0)}{w_j(t^0)} \right]^{w_j(t^0)} \cdot \sum_{j \in \langle k \rangle} a_{ij} w_j(t^0) \right]$$

and from (i)

$$\begin{aligned} \frac{\partial g_k(t^0, t^0)}{\partial t_i} &= \frac{1}{t_i^0} g_k(t^0) \cdot \sum_{j \in \langle k \rangle} a_{ij} w_j(t^0) \\ &= \frac{1}{t_i^0} \sum_{j \in \langle k \rangle} a_{ij} u_j(t^0) = \frac{\partial g_k(t^0)}{\partial t_i} \end{aligned}$$

(iii) This is just a restatement of (4). \square

Figure 1 shows the relation between $g(t)$ and $\tilde{g}(t, t^0)$.

7.3 Linearization of Condensed Programs

The dual chemical problem, stated in Chapter 4 is

Problem DCP

$$\text{Minimize } \{-\log[g_0(t)]\} \equiv - \sum_{i=1}^m b_i \cdot \log t_i \quad (1)$$

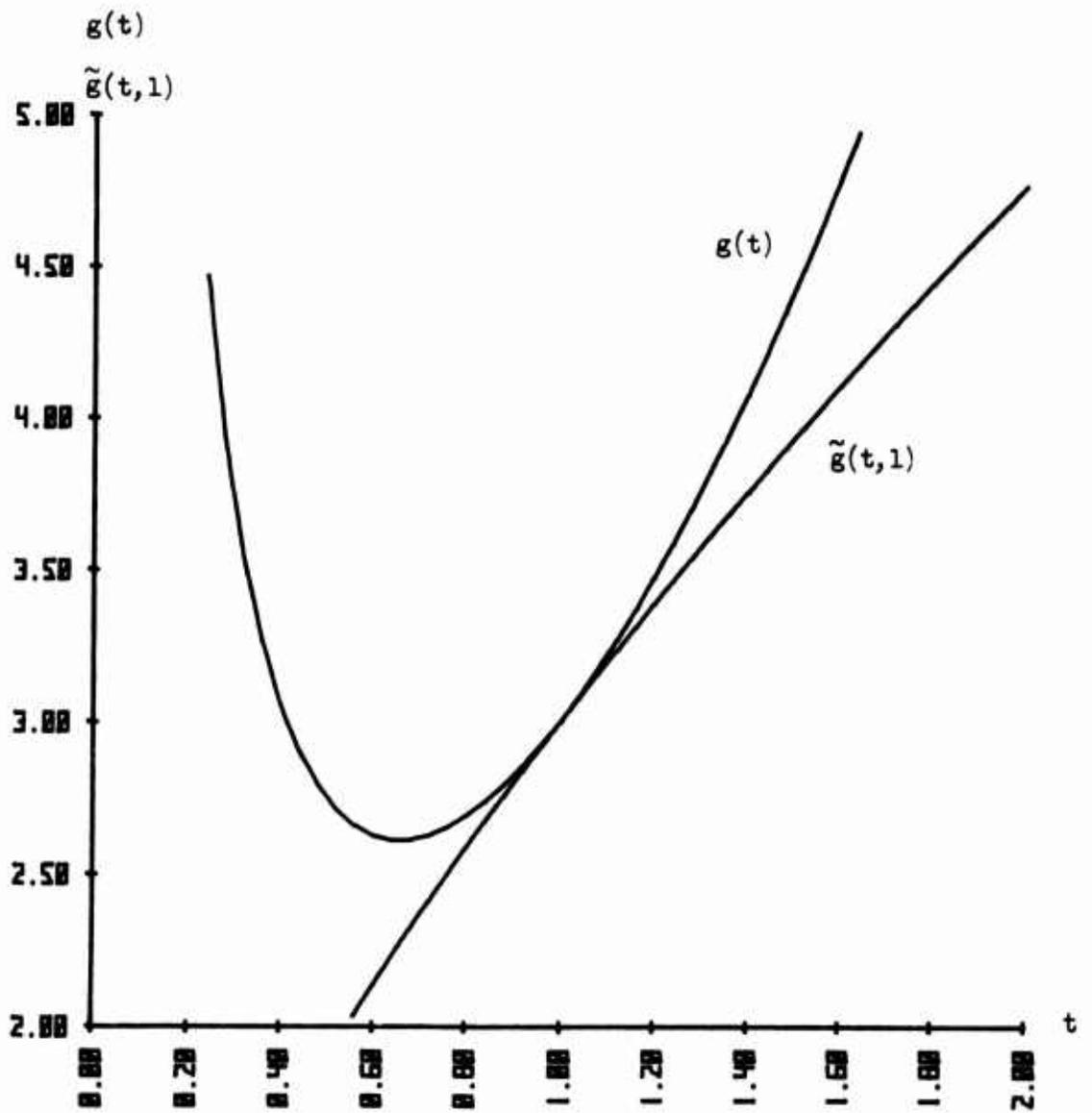


Figure 1
Posynomial Condensation

$$g(t) = 1/t + t + t^2$$

$$\tilde{g}(t,1) = 3 \cdot t^{2/3}$$

$$\text{Subject to } g_k(t) \equiv \sum_{j \in \langle k \rangle} u_j(t) \leq 1, \quad k = 1, 2, \dots, K \quad (2)$$

and

$$t \in \mathbb{R}_+^m$$

where $u_j(t)$ is defined in the preceding section and

$$c_j = \exp[-c_j], \quad j = 1, 2, \dots, n$$

Define the feasible set G of problem DCP by

$$G \equiv \{t \in \mathbb{R}_+^m \mid g_k(t) \leq 1, k = 1, 2, \dots, K\} \quad (3)$$

Let $t^0 \in \mathbb{R}_+^m$ and consider the totally condensed dual chemical problem.

Problem $\tilde{\text{DCP}}(t^0)$

$$\text{Minimize } \{-\log[g_0(t)] = -\sum_{i=1}^m b_i \cdot \log t_i\} \quad (4)$$

$$\text{Subject to } \tilde{g}_k(t, t^0) \leq 1, \quad k = 1, 2, \dots, K \quad (5)$$

$$t \in \mathbb{R}_+^m$$

Define the feasible set for the condensed problem

$$\tilde{G}(t^0) \equiv \{t \in \mathbb{R}_+^m \mid \tilde{g}_k(t, t^0) \leq 1, k = 1, 2, \dots, K\} \quad (6)$$

Proposition 7.4: $G \subseteq \tilde{G}(t^0)$ for all $t^0 \in \mathbb{R}_+^m$.

Proof: $t \in G$ implies $g_k(t) \leq 1$, $k = 1, 2, \dots, K$, which implies $\tilde{g}_k(t, t^0) \leq 1$ by Lemma 7.3. Hence $t \in \tilde{G}(t^0)$. \square

A totally condensed program such as $\tilde{DCP}(t^0)$ can be linearized simply by taking the logarithms of all the condensed functions \tilde{g}_k .

Let

$$\alpha_{ik}(t^0) \equiv \sum_{j \in \langle k \rangle} a_{ij} w_j(t^0) \quad (7)$$

$$\tilde{c}_k(t^0) \equiv \prod_{j \in \langle k \rangle} \left[\frac{c_j}{w_j(t^0)} \right]^{w_j(t^0)} \quad (8)$$

Here $w(t^0)$ is defined by (2.3). Note that

$$\tilde{g}_k(t, t^0) \equiv \tilde{c}_k(t^0) \prod_{i=1}^m t_i^{\alpha_{ik}(t^0)}, \quad k = 1, 2, \dots, K \quad (9)$$

Now, letting

$$z_i = \log t_i$$

$$z_i^0 = \log t_i^0$$

and taking the logarithms of the constraint functions $\tilde{g}_k(t, t^0)$

in (2.5), we arrive at

$$\log \tilde{g}_k(z, z^0) \equiv \log \tilde{c}_k(z^0) + \sum_i \alpha_{ik}(z^0) z_i \leq 0, \quad k = 1, 2, \dots, K \quad (10)$$

We use the somewhat ambiguous notation $g(t) = g(z)$. It will always be clear from context which function is used.

Note that in the case of a general geometric program, total condensation also requires condensing the objective function. For the dual chemical problem DCP, the objective function is already a single term posynomial and need not be condensed.

With the transformations described above, the problem becomes a linear program:

Problem $\tilde{LP}(z^0)$

$$\begin{aligned} &\text{Minimize } -bz \\ &\text{Subject to } \tilde{A}(z^0) \cdot z \leq \tilde{c}(z^0) \end{aligned} \quad (11)$$

where

$$\tilde{A}_{ik}(z^0) = \alpha_{ik}(z^0)$$

$$\tilde{c}_k(z^0) = -\log \tilde{C}_k(z^0)$$

Notice that z is unrestricted in sign.

Problem $\tilde{LP}(z^0)$, approximating DCP, has m variables and K constraints. For a typical chemical system, m is considerably smaller than the number of species n . Typically $m/n \approx 1/2$ or even smaller. K is usually very small--less than 10 even for large problems. The matrix $\tilde{A}(z^0)$ is thus much smaller than A , a definite advantage in storage and computations.

Instead of working with t , we could initially develop the condensation with $z = \log t$, that is, condense the transformed

geometric program TGP in Section 4.4 and its corresponding dual chemical program DCP. For the transformed program, define the transformed feasible set

$$Z \equiv \{z \in \mathbb{R}^m \mid g_k(z) \leq 1, k = 1, 2, \dots, K\} \quad (12)$$

Similarly, for the condensed program we have, based on (11)

$$\tilde{Z}(z^0) \equiv \{z \in \mathbb{R}^m \mid \tilde{A}(z^0) \cdot z \leq \tilde{c}(z^0)\} \quad (13)$$

Applying the strict monotonicity of the logarithmic function, it is easy to prove the analog of Proposition 7.4.

Proposition 7.5: $Z \subseteq \tilde{Z}(z^0)$ for all $z^0 \in \mathbb{R}^m$.

An immediate corollary is

Corollary 7.6:

$$\inf\{b'z \mid z \in \tilde{Z}(z^0)\} \leq \inf\{b'z \mid z \in Z\} \quad \text{for all } z \in \mathbb{R}^m.$$

Notice that all the functions in TGP and $\tilde{LP}(z^0)$ are well defined and that Z and $\tilde{Z}(z^0)$ are closed, in contrast to PGP (Section 4) and G . Unfortunately none of the feasible sets need be bounded. Even if G or Z are bounded, $\tilde{G}(t^0)$ or $\tilde{Z}(z^0)$ may still be unbounded, causing possibly unbounded solutions to $\tilde{LP}(z^0)$. In practical situations it is always possible to set upper and lower bounds on each variable, e.g.,

$$\underline{z} \leq z \leq \bar{z} \quad \text{componentwise}$$

Therefore, for practical purposes there is no loss of generality in assuming that Z is bounded by adding the bounds to the constraints.

7.4 Convex Cutting Plane Algorithms

The reader may have realized by now that the idea of Dembo's method is to generate a sequence of linear programs $\tilde{LP}^k \equiv \tilde{LP}(z^k)$ whose solutions will hopefully converge to a point z^* solving TGP or DCP. Generation of the sequence of problems is facilitated by introducing a "cut" at each iteration, which excludes part of the preceding feasible set from the new problem. The basic ideas of convex (or concave) cutting plane methods were formalized by Zangwill [55]. We review here the convex analog of his method.

The problem considered is

$$\begin{aligned} &\text{Minimize } c'x \\ &\text{Subject to } g_k(x) \leq 0, \quad k = 1, 2, \dots, K. \end{aligned}$$

The functions $g_k(x)$ for all k are convex and differentiable. There is no loss of generality in assuming that the objective function is linear since it can always be replaced by a single variable and added to the constraints. A solution is obtained when at some iteration k , $x^k \in G$, where G is the solution set.

Starting with a set U^1 approximating G externally (that is $G \subseteq U^1$), we solve the following problem LP^1 :

$$\begin{array}{ll} \text{Minimize} & cx \\ \text{Subject to} & x \in U^1 \end{array}$$

In general, at iteration k we solve

Problem LP^k

$$\begin{array}{ll} \text{Minimize} & cx \\ \text{Subject to} & x \in U^k. \end{array}$$

A solution x^k , to LP^k is then tested to see if $x^k \in G$.

If so, the algorithm terminates. Otherwise, a mapping Δ is used to generate a point v^k . Usually v^k is simply x^k , i.e., Δ is an identity mapping. Using v^k , a half space $H^k \equiv H(v^k)$ is generated, having the property that $G \subseteq H^k$ and $x^k \notin H^k$. The set U^{k+1} is then defined by

$$U^{k+1} = U^k \cap H^k$$

and we proceed to solve LP^{k+1} . The half space H^k is generated by a plane called a cutting plane.

Dembo's Algorithm (DA) replaces the objective function $g_0(x)$ by a single variable x_0 and the constraint $g_0(x) \leq x_0$. In the dual chemical problem DCP, since the objective is linear in z , this substitution is not necessary. This fact, and other changes

to his method discussed later, were incorporated in a modified version of the algorithm, henceforth referred to as MDA.

7.5 Generation of Cutting Planes

In this section we show how condensation techniques are applied in DA and MDA in view of the preceding section. Specifically we show how the half spaces H^k , defined by the cutting planes, are generated. We change our notation slightly, denoting by g_0 the linear objective function of TGP and DCP. The linearized program at iteration k will be denoted LP^k .

Following the notation of Chapter 5, we denote the minimizing set (solution set) of problem LP^k by $D(g_0|\tilde{Z}^k)$, where \tilde{Z}^k is the feasible set of LP^k . Note that the superscript k and the subscript k are not related. The former is iteration number; the latter, constraint index.

Lemma 7.7: Let $z \in D(g_0|\tilde{Z}(z^k))$. Then, either there exists a k such that $g_k(z) > 1$ or $z \in M(g_0|Z)$ where Z is the solution set.

The proof is trivial, by definition of Z .

Lemma 7.8: Let $\tilde{Z}^k \equiv \tilde{Z}(z^{k-1})$.

Let $z^k \in D(g_0|\tilde{Z}^k)$ and suppose $z^k \notin D(g_0|Z)$.

Define $L \equiv \{k | g_k(z) > 1\}$.

For $l \in L$ let

$$H_l(z^k) \equiv \{z \in \mathbb{R}^m \mid \log[\tilde{g}_l(z, z^k)] \leq 0\} \quad (1)$$

Define $\tilde{Z}^{k+1} \equiv \tilde{Z}^k \cap H_l(z^k)$.

Then the following hold:

- (i) $\tilde{Z}^{k+1} \subset \tilde{Z}^k$
- (ii) $z^k \notin \tilde{Z}^{k+1}$
- (iii) $Z \subseteq \tilde{Z}^{k+1}$

Proof:

(ii) If $z^k \in \tilde{Z}^{k+1}$ then $z^k \in H_l(z^k)$ and thus $\tilde{g}_l(z^k, z^k) \leq 1$ would imply $g_l(z^k) \leq 1$ by Lemma 7.3. This is a contradiction since $l \in L$. So $z^k \notin \tilde{Z}^{k+1}$.

(i) Follows from the definition of \tilde{Z}^{k+1} , with strict containment since (ii) holds.

(iii) Clearly $Z \subseteq \tilde{Z}^k$ by Proposition 7.5. From Lemma 7.3 (iii) it follows that $Z \subseteq H_l(z^k)$, hence $Z \subseteq \tilde{Z}^{k+1}$.

This completes the proof. \square

Lemma 7.7 describes the solution test and supplies a stopping criterion--when $g_k(z) \leq 1$ for all k . Lemma 7.8 shows that if a stopping criterion has not been reached, a cutting plane can be generated which meets the requirements of cutting plane methods.

Thus, a new problem LP^{k+1} is generated with a feasible set \tilde{Z}^{k+1} .

Dembo's algorithm (DA) performs all its computation in the untransformed variables t . In contrast, MDA works in the transformed variables z , for reasons discussed later. We show now that in the z -space the generation of cutting planes, and in fact, condensation in general, is actually a first order Taylor's expansion.

Recall that

$$g_k(z) = \sum_{j \in \langle k \rangle} \exp[-c_j + \sum_i a_{ij} z_i]$$

We expand $\log[g_k(z)]$ near z^0

$$\begin{aligned} \log[g_k(z)] &\approx \log[g_k(z^0)] + (z-z^0) \cdot \nabla g_k(z^0) \cdot \frac{1}{g_k(z^0)} \\ &= \log\left\{ \sum_{j \in \langle k \rangle} \exp[-c_j + \sum_i a_{ij} z_i^0] \right\} \\ &\quad + \sum_{i=1}^m \frac{(z_i - z_i^0)}{g_k(z^0)} \cdot \left\{ \sum_{j \in \langle k \rangle} a_{ij} \exp[-c_j + \sum_i a_{ij} z_i^0] \right\} \\ &= \log g_k(z^0) + \sum_{i=1}^m (z_i - z_i^0) \left(\sum_{j \in \langle k \rangle} a_{ij} w_j(z^0) \right) \\ &= \log g_k(z^0) + \sum_i z_i \cdot \alpha_{ik}(z^0) - \sum_i z_i^0 \cdot \alpha_{ik}(z^0) \end{aligned} \quad (2)$$

Here $w_j(z^0)$ and $\alpha_{ik}(z^0)$ are those defined by (2.3) and (3.7), respectively. From Lemma 7.3 and definitions (3.8) and (3.9) we obtain

$$\log[g_k(z^0)] = \log[g_k(z^0, z^0)] = \log \tilde{C}_k(z^0) + \sum_i z_i^0 \cdot \alpha_{ik}(z_i^0)$$

Combining with (2) we have

$$\log g_k(z) \approx \sum_i z_i \cdot \alpha_{ik}(z_i^0) + \log \tilde{C}_k(z^0) \quad (3)$$

The half space $H_\ell(z^k)$ is

$$H_\ell(z^k) = \{z \in \mathbb{R}^m \mid \tilde{A}_\ell(z^k) \cdot z + \log \tilde{C}_\ell(z^k) \leq 0\} \quad (4)$$

7.6 The MDA Algorithm

We assume that z is bounded, i.e., $\underline{z} \leq z \leq \bar{z}$ componentwise. The bounds need not be explicit as constraints since in the linear programs, we can use upper bounding techniques. The lower bounds can be eliminated by redefining z and requiring nonnegativity. The problem solved is DCP described in Section 7.3. We continue to use the term condensation although the description is in terms of the transformed variables z . Except for the z -notation and the fact that the original objective function $-bz$ is used, the description below is essentially Dembo's Algorithm [20].

Algorithm MDA

1. Select an arbitrary point z^0 and construct $\tilde{Z}^1 = \tilde{Z}(z^0)$ by total condensation. Set $k = 0$.

2. Increase k by 1.

Solve LP^k : minimize $-bz$

subject to $z \in \tilde{Z}^k$

Let $z^k \in \mathcal{D}(-bz | \tilde{Z}^k)$.

Compute $g_\ell(z^k)$, $\ell = 1, 2, \dots, K$.

Let $L = \{\ell | g_\ell(z^k) > 1, \ell = 1, 2, \dots, K\}$.

If L is empty terminate, since $z^k \in \mathcal{D}(-bz | Z)$.

3. Otherwise, select $\ell \in L$ (the most violated constraint is chosen).

Construct $H^k \equiv H_\ell(z^k)$ by equation (5.4).

4. Define $\tilde{Z}^{k+1} = \tilde{Z}^k \cap H^k$. Continue with Step 2.

The proof of convergence of this algorithm follows directly from Zangwill's convergence proof [55] and will not be repeated. However, his proof depends crucially on the nesting property, i.e., the fact that

$$Z \subset \dots \subset \tilde{Z}^{k+1} \subset \tilde{Z}^k \subset \dots \subset \tilde{Z}^1$$

This property can be insured, of course, by retaining all previous cuts throughout the computation. With the addition of cuts, the size of the problem may increase considerably, and much of the advantage of the method is lost. Topkis [51,52], and Eaves and Zangwill [27] investigated conditions under which the nesting requirement can be relaxed. Topkis showed that if the objective function is strictly quasiconvex (ours is not), one can relax nesting requirements and

retain only the active constraints of LP^{k-1} for problem LP^k . Retaining all previous cuts not only increases the dimension of the problem, it may also cause ill-conditioning in the LP matrices, especially when the generated cuts are "close" to each other.

The nesting requirement can still be relaxed if one is willing to assume, as we are, that z^k is the unique solution to LP^k . In this case, maintaining all the active constraints of LP^k in LP^{k+1} we have

$$-bz^{k+1} \geq -bz^k$$

since $z^k \notin \tilde{Z}^{k+1}$ and hence $\mathcal{D}(-bz|\tilde{Z}^{k+1}) \subset \mathcal{D}(-bz|\tilde{Z}^k)$. In the relaxed version of the algorithm only step 4 is changed where \tilde{Z}^{k+1} is defined by the active constraints of \tilde{Z}^k and the halfspace H^k . The assumption that each z^k be the unique solution to LP^k is equivalent to assuming nondegeneracy for each k .

Our experience showed that the relaxed version was more efficient. The reduced overhead of maintaining a large matrix more than compensated for the slight increase in the number of iterations.

In the rest of this chapter we discuss computational aspects of the algorithm and present a comparative study on several test problems, each solved by MDA and the RAND code [48].

7.7 Computational Aspects of the Algorithm

(a) Starting Point

The RAND, DA, and MDA methods require starting points x^0 , t^0 , z^0 respectively. For chemical problems t^0 and z^0 may be difficult to guess (although they are not required to be feasible). The overused phrase of nonlinear algorithms "Given a starting point which is close enough to the solution, the method converges..." is true also here. We found that the cutting plane method is, in general, more sensitive than the RAND code to this aspect. The RAND code uses projection to generate a feasible point x^0 if a nonfeasible guess is given. If no guess is given, it generates a positive feasible point by defining

$$x_j = y_j + \xi$$

and solving the linear program:

Program CLP

Maximize ξ

Subject to $Ay + A \cdot \xi \cdot 1_n = b$

$$y \geq 0$$

where 1_n is an n-vector of 1's. This method, due to Clasen [12],

has several advantages:

- (i) It generates a positive vector x if one exists.
- (ii) It detects degeneracies and dependent rows indicating infeasibility when no $x \geq 0$ exists.
- (iii) It does not require starting guesses.
- (iv) It can generate a dual point t or z by using the Kuhn-Tucker conditions and a least-squares approach (see Section 6.1).

Despite these advantages, the initial point generated was in many cases very far from optimal, due especially to incorrect distribution of species appearing in several phases. For example, the bulk of water is usually in the liquid phase, but the procedure sometimes assigns almost all the water to the vapor phase (see example below). To remedy the situation, the following modified linear program was solved:

Program LLP

$$\begin{aligned} & \text{Minimize } cy + c l_n \xi \\ & \text{Subject to } Ay + l_n \xi = b \\ & \xi \geq \epsilon \end{aligned}$$

where c is the vector of free energy coefficients and ϵ is a small positive number (say .001) to insure a positive solution if one exists.

The new procedure gives remarkably improved starting points, as shown by the example below. The composition vector x and the mole fractions \hat{x} are given for the two methods and compared to the equilibrium values. Notice that the CLP solution has practically all the water in the gas phase whereas LLP correctly assigns the bulk of the water to the liquid phase. Note also that the dimensionless free energy F/RT obtained by LLP is within 0.2% of the minimum value. Several problems were tested with LLP and the example is quite typical of the improvement achieved.

Example 7.8: Starting Points for Soda Pop Model (Appendix A.4)

Species	CLP		LLP $\epsilon = 0.002$		Equilibrium	
	x	\hat{x}	x	\hat{x}	x	\hat{x}
O ₂	5.256	.039	5.274	.056	5.276	.053
CO ₂	.020	1.5 (10 ⁻⁴)	6.042	.064	6.043	.060
N ₂	82.560	.613	82.578	.879	82.580	.826
H ₂ O	46.766	.347	.002		6.108	.061
Gas	134.602		93.896		100.0	
O ₂	.020	.045	.002		4.4 (10 ⁻⁵)	
CO ₂	.020	.045	.002		1.3 (10 ⁻³)	
N ₂	.020	.045	.002		3.8 (10 ⁻⁴)	
H ⁺	.051	.113	.002		3.0 (10 ⁻⁸)	
OH ⁻	.020	.045	.002		5.7 (10 ⁻⁷)	
Cl ⁻	.080	.179	.080	.0015	8.0 (10 ⁻²)	.0017
Na ⁺	.088	.197	.088	.0017	8.8 (10 ⁻²)	.0019
K	.048	.108	.048	.0009	4.8 (10 ⁻²)	.0010
H ₂ O	.020	.045	52.806	0.995	46.70	.995
HCO ₃ ⁻	.020	.045	.026		2.9 (10 ⁻²)	
H ₂ CO ₃	.020	.045	.002		1.8 (10 ⁻⁶)	
CO ₃ ⁼	.020	.045	.002		4.8 (10 ⁻⁵)	
Glucose	.020	.045	.002		2.0 (10 ⁻²)	
Liquid	.447		53.082		46.97	
F/RT	-2772.4		-3122.9		-3128.9	

The (primal) starting point generated by LLP is used directly to generate weights by $w_j = \hat{x}_j$. Unlike DA, which starts by condensing at z^0 , these weights are used in MDA to condense and linearize, so that Z^1 is generated without any point z^0 .

(b) Bounds on Variables

To insure compactness of Z^k , both DA and MDA require bounds (upper and lower) on all variables. RAND requires lower bounds on all x_j to insure positivity. Bounds which are too strict may render the problem infeasible, whereas bounds which are too wide are likely to cause numerical problems, and slow convergence. Upper bounds on dual variables can be found using the techniques described in Section 6.2. In many cases these can be found by inspection.

Finding lower bounds is a somewhat more difficult problem, but since the dual variables represent (theoretical) concentrations of subspecies, a bound which is in the order of \bar{x}/A_0 where A_0 is the Avogadro number 6.02×10^{23} , is certainly low enough, since A_0 is the total number of molecules in one mole of any species.

(c) Generating Cutting Planes by Condensation

A major advantage of MDA over DA is that it works in the transformed space (the z -space) instead of the usual dual-chemical t -space. Consequently, exponentiation operations in DA are multiplications in MDA; similarly, multiplications are reduced to additions. Computations in the transformed space also eliminate overflow and underflow problems caused by large values of the free energy coefficients.

(d) Solving the Subproblems

The RAND code solves a linear system to determine a direction for improvement of the composition-vector. DA and MDA solve LP problems which are relatively small (in the unnested version) and easy to solve because the dual simplex method is used, so that an added cut is actually an added variable. A feasible basis is available from the preceding iteration. Since there are m dual variables, at most m constraints are active in each LP^k . The bounds, of course, are not treated as constraints. The lower bound is eliminated by a transformation of variables. The upper bound is handled by standard upper bounding methods. With the added slacks, the subproblems are of the order $m \times 2m$. The matrix \tilde{A}^k is usually dense, so that sparse matrix techniques are not applicable.

(e) Recondensation

As pointed out earlier, an inherent problem in cutting plane methods is ill-conditioning due to accumulation of "close" cutting planes. This was apparently the reason for the failure of the dual methods in the larger problems (see the computational results which follow). Some ways to accelerate convergence are discussed by Dembo [20].

A possible (untested) remedy is to "recondense" the problem after a fixed number of cuts, discarding most of the existing constraints and generating several new ones, by condensation of the original constraints. This approach is analogous in some ways to

reinversion of matrices in LP codes. In principle, recondensation could take place at each iteration, with a result equivalent to block pivoting. This approach was tested and did not prove fruitful.

(f) Selection of Cuts

The cutting plane H^k , generated in the k -th iteration is not necessarily unique, since any violated constraint g_ℓ , $\ell \in L$, can be condensed. The common practice of selecting the "most violated" constraint is not necessarily the best policy. The situation is somewhat analogous to the selection of entering column in the simplex algorithm, based on the most negative reduced cost.

Faster convergence (at the cost of a slight increase in computation at each iteration) can be obtained by condensing the constraint leading to the largest change in g_0 (in our case, an increase in g_0). In the simplex algorithm this can be easily computed, by multiplying the ratio b_i/a_{ij} by the reduced cost \bar{c}_j , and looking for the minimum over j such that $\bar{c}_j < 0$ and i such that $a_{ij} > 0$. For our purposes, the reduced cost for a constraint can be substituted by its lagrange multiplier. Since the latter is not known either, an approximate multiplier--that of the condensed constraint--can be used. This multiplier is available only if the constraint was linearized in a previous iteration and not discarded. Then the multiplier $\tilde{\lambda}_\ell$ of the most recent linearization of g_ℓ can be obtained from the solution of $(LP)^k$. We shall choose to condense constraint g_s in iteration k when

$$[g_s(z^k) - 1] \cdot \tilde{\lambda}_s = \max_{\ell \in L} [g_\ell(z^k) - 1] \cdot \tilde{\lambda}_\ell$$

We have not tested this scheme in our algorithm.

7.8 Comparative Test Results

The tables in this section present results of test runs obtained with MDA compared to results obtained with the RAND code [48]. Both codes were compiled in FORTRAN H, Optimization Level 2, and run on an IBM 360/67. All computations and results are in double precision. Execution times in seconds of CPU include input but not output time. The version of MDA used here did not use automatic starting points or bounds. Recondensation and acceleration methods were not applied.

The following are assumed, unless stated otherwise:

1. RAND's starting point was generated by program CLP.
2. MDA's starting point was based on a rough guess, as described in Section 6.2.
3. The unnested version was used.
4. Stopping criteria:

RAND: RMS error in mass balance $\leq 3 \times 10^{-5}$

RMS error in mass action $\leq 3 \times 10^{-5}$

MDA: For all k , $g_k(z) \leq 1 + \epsilon$, $\epsilon = 10^{-3}$.

5. Accuracy:

RAND's results are judged somewhat more accurate based on the smaller final errors and the more stringent termination criteria.

Two sets of problems were solved, one including chemical problems and the other including posynomial geometric programs. There is some indication that MDA did better with posynomial problems than with chemical ones, due perhaps to the former's "dual" structure.

All test problems and their solutions appear in Appendix A. In addition to the comparison of the codes, tests were run to:

- (i) Compare nested vs unnested versions.
- (ii) Test the effects of starting points on convergence.

Analysis of Test Results

Table 1 shows the CPU times for chemical problems. For small problems MDA is about 20% slower than RAND, but for the larger problems the RAND program is considerably faster than MDA, mostly due to MDA's slow convergence, attributable to the large number of cuts and the resulting ill-conditioning of LP^k . For comparison, Table 2 shows results obtained for problems which are geometric programs in nature (the last three are artificial). It seems that in this class MDA fares a little better. The last two problems, one with a dependent variable and the other with a loose constraint, demonstrate the advantages of dual methods. RAND was unable to handle the former, while the latter, although solved, caused several underflow warnings, since the values of x and \bar{x} corresponding to the loose constraint approached 0, causing numerical difficulties in the computation of $F(x)$.

TABLE 1
TEST RESULTS--CHEMICAL PROBLEMS

Problem	Dimensions			CPU(sec.)		cuts
	m	n	K	RAND	MDA	
Small Problem A.1	2	4	2	0.25	0.28	0
Soda Water A.2	5	13	2	0.57	0.89	23
Hydrazine A.3	3	10	1	0.43	0.50	12
Soda-Pop A.4	9	17	2	0.82	2.06	61
Respiratory1 A.5	12	30	3	2.07	5.80	150 ¹
Respiratory2 A.5	11	30	3	1.63	4.49	110
Plasma A.6	16	21	1	1.34	9.16	150 ¹
Fetus A.7	19	51	7	5.09	10.05	133 ²

- Notes: 0. The problem names are followed by their index number.
 1. Failed to satisfy the convergence criterion after 150 cuts, terminated within 0.1% of optimum.
 2. Failed to converge, terminated within 0.2% of optimum.

Dimensions:

- m = number of subspecies
 n = number of species
 K = number of phases.

TABLE 2

TEST RESULTS--POSYNOMIAL GEOMETRIC PROBLEMS

Problem	Dimensions			CPU(sec.)		cuts
	m	n	K	RAND	MDA	
Sea Power A.8	7	10	4	0.87	1.14	35
Reactor A.9	3	5	1	0.32	0.41	11
Condenser A.10	4	8	2	0.63	0.58	13
Stochastic Condenser A.11	9	13	6	1.28	1.15	22
Decomposition A.12	10	13	3	1.48	2.17	66
Dependent Variables A.13	3	4	0	* ¹	0.43	11
Loose Constraint A.14	2	5	1	0.56 ²	0.41	12

Notes: 1. RAND program cannot handle dependent variables.

2. Several underflow warnings were raised during computation.

Dimensions:

m = number of variables

n = number of posynomial terms

K = number of posynomial constraints, not including the objective function.

Effects of Nesting Relaxation

Table 3 compares the run times of MDA with and without nesting. As expected, relaxation of the nesting requirement increases the number of cuts, but the reduced dimensions of the subproblems offset this loss by reduced time per iteration. The results show that the overall effect is a reduction in computing time.

Effects of Starting Point

Four problems were tested with no starting point. In these tests the program automatically assumed that $t_i^* = 1$ for all i . The results are shown in Table 4--compared to the previous runs which used a rough starting point obtained by inspection of the matrix and the free energy coefficients. The results show a marked improvement with the better starting point. It is expected that incorporation of the technique discussed in 7.7(a) to generate a starting point will significantly improve the run times.

Convergence Rates

Due to the complexity of the mappings in the algorithm, no attempt was made to find a theoretical rate of convergence. Topkis analyzes convergence rates for cutting plane algorithms [51,52]. A simple test was made to check the "practical" rate of convergence. The objective function $F(x)$ was computed after each cut H^k . The value at iteration k being F^k . With the initial estimate F^0 and the final value F^* known, we computed the " α -cut" k_α , defined as

TABLE 3
NESTED VS UNNESTED TEST RUNS

Problem	Nested		Unnested	
	CPU(sec.)	cuts	CPU(sec.)	cuts
Soda-Water	0.72	12	0.66	14
Soda-Pop	2.62	53	2.06	61
Reactor	0.45	11	0.41	11
Stochastic Condenser	1.25	20	1.15	22
Decomposition	2.30	60	2.17	66

TABLE 4
EFFECTS OF STARTING POINT (S.P.)

Problem	No S.P.		Rough S.P.	
	CPU	cuts	CPU	cuts
Soda-Water	0.89	23	0.66	14
Hydrazine	0.62	16	0.50	12
Soda-Pop *	5.11	90	2.62	53
Respiratoryl *	11.76	116	8.36	88

* NOTE: In these two cases the nested version was used.

$$k_{\alpha} \equiv \min \left\{ k \left| \frac{F^k - F^0}{F^* - F^0} \geq \alpha \right. \right\}$$

The numbers can serve as indicators of convergence rate.

Table 5 lists the results for $\alpha = 0.5$, $\alpha = 0.95$ and $\alpha = 0.99$. In this table F^* was the final value reached, which was assumed to be the optimal value. In some cases where the starting point was particularly "bad," F^0 was taken as the first value after some stabilization was obtained.

It is hard to draw any definite conclusion from these results, but the need for acceleration techniques is evident as some of the problems with the larger number of cuts show slow convergence at the tail, with the last 1% improvement requiring as many cuts as the preceding 99%.

TABLE 5

CONVERGENCE INDICATORS

Problem	$-F^0$	$-F^*$	$k_{0.5}$	$k_{0.95}$	$k_{0.99}$	$k_{1.0}$
Soda-Water	2241.8	2253.2	3	4	4	14
Hydrazine	42.822	47.711	1	4	6	12
Soda Pop ¹	3012.4	3127.6	2	13	29	53
Soda Pop	3012.4	3127.6	2	13	29	61
Respiratory1 ²	1826.8	1835.1	14	45	75	150
Respiratory2	1809.0	1835.2	10	17	54	110
Plasma ²	830.19	830.46	46	133	145	150
Fetus ²	1866.7	1870.8	38	126	133	133
Sea Power	8.42	126.47	6	21	30	35
Reactor	259.49	334.26	1	4	9	11
Stochastic	383.89	883.48	1	4	8	22
Decomposition	1.0	18.25	13	32	46	66

$$k_{\alpha} = \min \left\{ k \left| \frac{F^k - F^0}{F^* - F^0} \geq \alpha \right. \right\}$$

- Notes: 1. Nested version
2. Algorithm terminated without satisfying convergence criteria.

CHAPTER 8

TRANSCENDENTAL GEOMETRIC PROGRAMMING

8.1 Introduction and Formulation

This chapter presents an extension to the theory of geometric programming to a wider class of functions, namely, forms including variables appearing as exponents (or in logarithms). Although the principal motivation is to extend the applicability of primal geometric programs, our results extend also to the dual, of interest in chemical equilibrium problems.

This chapter departs from chemically oriented terminology; the primal problem here will be the primal geometric program and its transcendental extension, called a transcendental program. The equivalent of the chemical equilibrium problem will be henceforth called the "dual problem." As will be evident from the following section, many of the useful properties of geometric programming, especially the unimodality of primal posynomial functions, no longer hold in transcendental programs. In this sense, transcendental programs have difficulties similar to those arising in signomial [39] or complementary geometric programs [5]--these are the extensions handling negative coefficients.

Since neither primal nor dual transcendental programs are convex or convex-transformable, Kuhn-Tucker conditions are no longer sufficient for optimality. The resulting duality theory will therefore be concerned only with stationary points of the Lagrangean, which may be minima, maxima, or neither. Figure 2 shows a simple, single variable transcendental function of the form of interest in this chapter. It shows that even in simple cases multiple local optima may exist.

We consider positive functions in positive variables $t \in \mathbb{R}_+^m$ and $\theta \in \mathbb{R}_+^p$. The distinction between t and θ is mostly for convenience. The functions are called posyentials (posynomial-exponential), and as with posynomial functions, we can speak of posyential terms. Thus the form of a posyential function is

$$g_k(t, \theta) = \sum_{j \in \langle k \rangle} P_j(t) Q_j(\theta) R_j(\theta) \quad (1)$$

where

$$P_j(t) = C_j \prod_{i=1}^m t_i^{a_{ij}} \quad (2)$$

$$Q_j(\theta) = \prod_{l=1}^p \theta_l^{b_{lj}} \quad (3)$$

$$R_j(\theta) = \prod_{l=1}^p \exp[d_{lj}\theta_l] \quad (4)$$

Here C_j are positive constants a_{ij} , b_{lj} , and d_{lj} are fixed real numbers for $i = 1, 2, \dots, m$, $j = 1, 2, \dots, n$, $l = 1, 2, \dots, p$.

$\langle k \rangle$ is a subset of consecutive integers of $N \equiv \{1, 2, \dots, n\}$. The

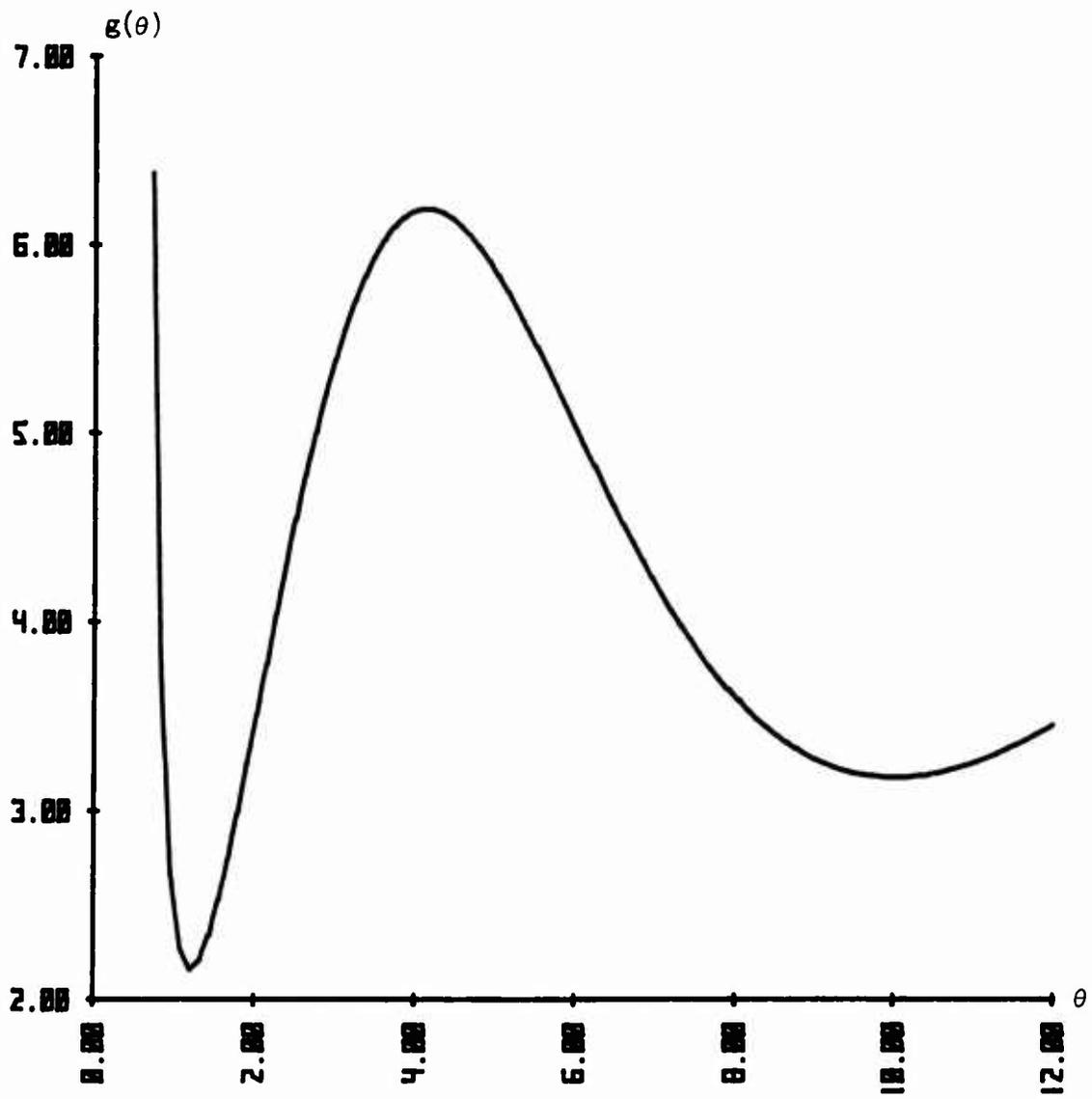


Figure 2

Posyniential Function

$$g(\theta) = \theta^{-5} + \theta^4 \cdot e^{-\theta} + e^{-0.1 \cdot \theta}$$

terms $P_j(t)$ and $Q_j(\theta)$ are posynomials, and so is their product. The variables t are called the ordinary variables, while the variables θ appearing also in the exponents are the transcendental variables.

We note that if either $d_{lj} = 0$ for all l and j or $b_{lj} = 0$ for all l and j , then the problem reduces to a standard geometric program. In the latter case this reduction is achieved simply by letting $t_{m+l} = \exp(\theta_l)$.

The Primal Transcendental Geometric Program is

Program PTP

$$\text{Minimize } g_0(t, \theta) \quad (5)$$

$$\text{Subject to } g_k(t, \theta) \leq 1, \quad k = 1, 2, \dots, k \quad (6)$$

$$t > 0, \quad \theta > 0 \quad (7)$$

This problem can be characterized by

- (i) a positive n -vector of coefficients C
- (ii) exponent matrices $A \in R^{m \times n}$, $B \in R^{p \times n}$, $D \in R^{p \times n}$
- (iii) a partition of the integer set $N = \{1, 2, \dots, n\}$ which defines the sets $\langle k \rangle$, $k = 0, 1, \dots, K$.

Before examining the necessary conditions for optimality, we give several examples of problems, which can be stated as transcendental programs. For convenience we define $u_j(t, \theta) = P_j(t) Q_j(\theta) R_j(\theta)$. Since $g_k(t, \theta) = \sum_{j \in \langle k \rangle} u_j(t, \theta)$ we shall be interested in the form of the terms $u_j(t, \theta)$.

1. Let $u_j(t) = C_j \left[\prod_{i=1}^m t_i^{a_{ij}} \right] \cdot \left[\prod_{i=1}^m (\log t_i)^{b_{ij}} \right]$ where $t_i > 1$ for all i .

Define $\theta_i \equiv \log t_i$, $i = 1, 2, \dots, m$, then

$$u_j(\theta) = C_j \left[\prod_i \theta_i^{b_{ij}} \right] \cdot \left[\prod_i e^{a_{ij} \theta_i} \right]$$

2. Let $v_j(t) = C_j \prod_i t_i^{a_{ij}}$ (a posynomial term) and suppose that we have a function of the form

$$g_0(t) = \exp\left(\sum_{j=1}^r v_j(t)\right) \cdot v_s(t)$$

Let $\theta_1 \equiv \sum_{j=1}^r v_j(t)$. The function can now be written

$$g_0(t, \theta) = e^{\theta} \cdot v_r(t)$$

subject to

$$\sum_j \theta_1^{-1} \cdot v_j(t) \leq 1$$

3. Let

$$g_0(t) = \left\{ \sum_{j=1}^r v_j(t) \right\} \cdot [\log(v_s(t))]$$

(assuming $v_s(t) > 1$). The function becomes

$$g_0(t, \theta) = \sum_j \theta v_j(t)$$

subject to

$$[v_s(t)]^{-1} \leq 1$$

$$e^{-\theta} \cdot v_s(t) \leq 1$$

In the three cases above, the program is reduced to a transcendental program.

8.2 Necessary Conditions for Minima

Proposition 8.1: Let problem PTP be superconsistent, i.e., having an interior feasible point.

If $(t, \theta) > 0$ is a (local) minimum for problem PTP, there exist nonnegative multipliers $(\delta_1, \delta_2, \dots, \delta_n)$ and $(\lambda_0, \lambda_1, \lambda_2, \dots, \lambda_K)$ satisfying the following conditions:

$$A\delta = 0 \quad (1)$$

$$(B + \Theta D)\delta = 0 \quad (2)$$

$$\sum_{j \in \langle k \rangle} \delta_j = \lambda_k \quad (3)$$

$$\lambda_0 = 1 \quad (4)$$

Here A , B , and D are the matrices in PTP, δ and λ are the multipliers (vectors); Θ is a $p \times p$ diagonal matrix with $\Theta_{\ell\ell} = \theta_\ell$. The conditions above are:

- (1) the ordinary orthogonality conditions;
 - (2) the transcendental orthogonality conditions;
- and
- (4) the normality condition.

Proof: We apply the necessary conditions to the Lagrangean

$$L(t, \theta, \mu) \equiv \mu_0 g_0(t, \theta) + \sum_{k=1}^K \mu_k [g_k(t, \theta) - 1] \quad (5)$$

Here μ_k are the Lagrange multipliers. $\mu_0 \equiv 1$. Let (t^*, θ^*) be a

minimizing point for PTP. By Kuhn-Tucker conditions there exists a $\mu^* \geq 0$ such that

$$\frac{\partial L(t^*, \theta^*, \mu^*)}{\partial t_i} = 0, \quad i = 1, 2, \dots, m \quad (7)$$

$$\frac{\partial L(t^*, \theta^*, \mu^*)}{\partial \theta_l} = 0, \quad l = 1, 2, \dots, p \quad (8)$$

$$\mu_k^* [g_k(t^*, \theta^*) - 1] = 0, \quad k = 1, 2, \dots, K \quad (9)$$

Since $t^* > 0$ and $\theta^* > 0$ we have by (9)

$$L(t^*, \theta^*, \mu^*) > 0$$

So we can replace conditions (7) and (8) by the logarithmic conditions

$$\frac{\partial \log[L(t^*, \theta^*, \mu^*)]}{\partial \log t_i} = 0$$

$$\frac{\partial \log[L(t^*, \theta^*, \mu^*)]}{\partial \log \theta_l} = 0$$

Let $u_j(t^*, \theta^*) \equiv P_j(t^*) Q_j(\theta^*) R_j(\theta^*)$

$$\frac{\partial \log L(t^*, \theta^*, \mu^*)}{\partial \log t_i} = \frac{t_i^*}{L(t^*, \theta^*, \mu^*)} \left[\frac{1}{t_i^*} \sum_{k=0}^K \mu_k \left(\sum_{j \in \langle k \rangle} a_{ij} u_j(t^*, \theta^*) \right) \right] = 0 \quad (10)$$

$$\frac{\partial [\log L(t^*, \theta^*, \mu^*)]}{\partial \log \theta_l} = \frac{\theta_l^*}{L(t^*, \theta^*, \mu^*)} \left[\frac{1}{\theta_l^*} \sum_{k=0}^K \mu_k \left(\sum_{j \in (k)} b_{lj} u_j(t^*, \theta^*) \right) + \sum_{k=0}^K \mu_k \sum_{j \in (k)} d_{lj} u_j(t^*, \theta^*) \right] = 0 \quad (11)$$

where $i = 1, 2, \dots, m$, $l = 1, 2, \dots, p$. Let

$$\delta_j^* = \mu_k^*(j) \frac{u_j(t^*, \theta^*)}{g_0(t^*, \theta^*)}, \quad j = 1, 2, \dots, n \quad (12)$$

Note that by (9), $L(t^*, \theta^*, \mu^*) = g_0(t^*, \theta^*)$. Interchanging the double sum and substituting in (10) we obtain the ordinary orthogonality conditions (1). Substitution of δ in (11) and rearrangement yields

$$\sum_{j=1}^n (b_{lj} + \theta_l^* d_{lj}) \delta_j^* = 0, \quad l = 1, 2, \dots, p$$

in matrix form these are the transcendental orthogonality conditions (2). Condition (3) is a definition for λ_k . Condition (4) is satisfied since $\mu_0^* = 1$. Note that the transcendental orthogonality conditions are linear in the primal variables θ . Observe also that each of these constraints has only one transcendental variable. \square

Lemma 8.2: For any feasible (t, θ) and any $\delta(\theta), \lambda(\theta)$ satisfying the conditions of proposition (8.1)

$$g_0(t, \theta) \geq v(\delta, \lambda) \equiv \left[\prod_j \left(\frac{c_j}{\delta_j} \right)^{\delta_j} \right] \left[\prod_k \lambda_k \right] \left[\prod_l \left(- \frac{B_{l.} \delta}{e \cdot D_{l.} \delta} \right)^{B_{l.} \delta} \right] \quad (13)$$

Here e is the base of natural logarithms, B_{ℓ} and D_{ℓ} are the ℓ -th rows of B and D , respectively. Since $\delta = \delta(\theta)$ we can write $v(\delta, \lambda) = v(\delta, \lambda, \theta)$.

Proof: The proof is based on the geometric inequality (Section 7.2) and is analogous to the proof of the main lemma of geometric programming [25, p. 114]. Clearly, for feasible (t, θ) and λ

$$\prod_{k=1}^K [g_k(t, \theta)]^{\lambda_k} \leq 1$$

Therefore

$$g_0(t, \theta) \geq \prod_{k=0}^K [g_k(t, \theta)]^{\lambda_k} \quad (14)$$

but

$$[g_k(t, \theta)]^{\lambda_k} \geq \left[\prod_{j \in \langle k \rangle} \left(\frac{u_j(t, \theta)}{\delta_j} \right)^{\delta_j} \right]^{\lambda_k} \quad (15)$$

where $x^x \equiv x^{-x} \equiv 1$ for $x = 0$, and $u_j(t, \theta) = P_j(t) Q_j(\theta) R_j(\theta)$.

Thus

$$g_0(t, \theta) \geq \prod_k \left\{ \prod_{j \in \langle k \rangle} \left[\frac{u_j(t, \theta)}{\delta_j} \right]^{\delta_j} \right\}^{\lambda_k} \quad (16)$$

Developing the right hand side we get

$$\begin{aligned}
& \left\{ \prod_{j=1}^n \left[\frac{P_j(t)}{\delta_j} \right]^{\delta_j} \right\} \left\{ \prod_k \lambda_k^{\lambda_k} \right\} \left\{ \prod_j [Q_j(\theta) R_j(\theta)]^{\delta_j} \right\} \\
&= \left\{ \prod_{j=1}^n \left(\frac{C_j}{\delta_j} \right)^{\delta_j} \right\} \left\{ \prod_k \lambda_k^{\lambda_k} \right\} \prod_j \left[\prod_{\ell} \theta_{\ell}^{b_{\ell j}} e^{d_{\ell j} \theta_{\ell}} \right]^{\delta_j} \\
&= \left\{ \prod_{j=1}^n \left(\frac{C_j}{\delta_j} \right)^{\delta_j} \right\} \left\{ \prod_{k=1}^K \lambda_k^{\lambda_k} \right\} \left\{ \prod_{\ell} [\theta_{\ell}^{B_{\ell} \cdot \delta} \exp(\theta_{\ell} D_{\ell} \cdot \delta)] \right\}
\end{aligned}$$

but by (2)

$$D_{\ell} \cdot \theta_{\ell} \delta = -B_{\ell} \cdot \delta$$

so that the last product is

$$\prod_{\ell} [\theta_{\ell}^{B_{\ell} \cdot \delta} \exp(-B_{\ell} \cdot \delta)] = \prod_{\ell} \left[\frac{\theta_{\ell}}{e} \right]^{B_{\ell} \cdot \delta}$$

A second application of (2) shows $\theta_{\ell} = -B_{\ell} \cdot \delta / D_{\ell} \cdot \delta$, which after substitution in the preceding equations leads to the result. \square

This lemma demonstrates the major difficulty in establishing a duality theory for transcendental programs. The dual is not a pure dual; its variables include the primal transcendental variables, and therefore it cannot be maximized independently of the primal. The appearance of e in the expression for the dual function suggests the following generalization.

Corollary 8.3: Let Problem PTP(α) be defined by replacing e by $\alpha > 0$ in PTP, that is to say, $PTP(e) \equiv PTP$, and PTP(α) has terms of the form

$$u_j(t, \theta) = P_j(t) Q_j(\theta) \cdot \prod_{\ell} \alpha^{d_{\ell j} \theta_{\ell}}$$

Then all the preceding results in this chapter hold for PTP(α), provided that $D = \{d_{\ell j}\}$ is replaced by $\tilde{D} = D \log \alpha$.

Proof:

$$\alpha^{d_{\ell j} \theta_{\ell}} = \exp[\log \alpha \cdot d_{\ell j} \theta_{\ell}] = \exp[\tilde{d}_{\ell j} \theta_{\ell}] \quad \square$$

8.3 Duality in Transcendental Programs

Lemma 8.4: (Main lemma of transcendental programming). Let (t, θ) be feasible for PTP, and let $\delta(\theta), \lambda(\theta)$ satisfy (2.1)-(2.4).

Then

$$g_0(t, \theta) \geq v(\delta, \lambda) \quad (1)$$

Moreover, under these conditions

$$g_0(t, \theta) = v(\delta, \lambda)$$

if and only if

$$\delta_j = \begin{cases} \frac{u_j(t, \theta)}{g_0(t, \theta)} & j \in \langle 0 \rangle \\ \lambda_k(\theta) \cdot u_j(t, \theta) & j \in \langle k \rangle, k = 1, 2, \dots, K \end{cases} \quad (2)$$

Proof: Inequality (1) follows directly from Lemma 10.2.

Suppose $g_0(t, \theta) = v(\delta, \lambda)$. Then in Lemma 10.2 the inequalities (2.14), (2.15), (2.16) are all equalities. In particular, we have $g_k(t, \theta)^{\lambda_k} = 1$ which implies $\lambda_k = 0$ if $g_k(t, \theta) < 1$ for $k = 1, 2, \dots, K$. According to the geometric inequality (Section 7.2) there are constants $\eta_k, k = 0, 1, \dots, K$, such that

$$\delta_j = \eta_k u_j(t, \theta), \quad j \in \langle k \rangle \quad (3)$$

Summing over $j \in \langle k \rangle$ we arrive at

$$\lambda_k = \eta_k g_k(t, \theta), \quad k = 0, 1, \dots, K \quad (4)$$

so that $\eta_0 = 1/g_0(t, \theta)$ and hence (2) is satisfied for $j \in \langle 0 \rangle$.

If $\eta_k = 0$ we have by (3) $\delta_j = 0, j \in \langle k \rangle$ and hence $\lambda_k = 0$ so that (2) is satisfied for $j \in \langle k \rangle$ and $\eta_k = 0$.

If $\eta_k > 0$ then $\lambda_k > 0$ by (4) but then $g_k(t, \theta)^{\lambda_k} = 1$ implies $g_k(t, \theta) = 1$ so that $\lambda_k = \eta_k$ by (4), and again, using (3) we see that (2) holds for $\eta_k > 0, k = 0, 1, 2, \dots, K$.

Now suppose that (t, θ) is primal feasible, $(\delta(\theta), \lambda(\theta))$ satisfy (2.1-2.4) and that conditions (2) hold. It follows from the geometric inequality (Section 7.2), that for each $k > 1$ the inequality (2.15) is an equality, since if $\lambda_k(\theta) = 0$ then $\delta_j(\theta) = 0$ for all $j \in \langle k \rangle$ and then certainly

$$g_k(t, \theta)^{\lambda_k} = \prod_{j \in k} \left[\frac{u_j(t, \theta)}{\delta_j} \right]^{\delta_j} = 1$$

If $\lambda_k(\theta)$ is positive, then so is $\delta_j(\theta)$ for all $j \in \langle k \rangle$, $k > 1$.

But then (2) implies that

$$g_k(t, \theta) = \sum_{j \in \langle k \rangle} u_j(t, \theta) = \frac{1}{\lambda_k} \sum \delta_j = 1$$

and consequently

$$1 = \prod_{j \in \langle k \rangle} \left[\frac{u_j(t, \theta)}{\delta_j} \right]^{\delta_j} \cdot \lambda_k$$

Thus the right hand side of (2.15) is always 1 for $k > 1$. For

$g_0(t, \theta)$ the geometric inequality leads to

$$g_0(t, \theta) = \prod_{j \in \langle 0 \rangle} \left(\frac{c_j}{\delta_j} \right)^{\delta_j}$$

Combining the observations above, we get

$$g_0(t, \theta) = \left[\prod_{j=1}^n \left(\frac{u_j(t, \theta)}{\delta_j} \right)^{\delta_j} \right] \left[\prod_{k=0}^K \lambda_k \right] \quad (5)$$

and since $\delta(\theta)$, $\lambda(\theta)$ satisfy conditions (2.1-2.4), the right hand side equals $v(\delta, \lambda)$. \square

This lemma, which is analogous to the main lemma of geometric programming, does not provide an analogous dual program. Note that maximizing $v(\delta, \lambda)$ subject to (2.1)-(2.4), $\delta \geq 0$, $\lambda \geq 0$, $\theta > 0$, does not provide any solution to the primal since θ may be primal-infeasible. Moreover, even if θ is primal feasible, the solution may be a local minimum, a stationary point, or even a local maximum. Under more restrictive hypotheses we can achieve a more tangible result.

First we wish to verify that for any fixed θ for which both primal and dual problems are feasible, the duality relations reduce to the usual geometric programming duality. To see this, let θ be fixed so that for every $1 \leq j \leq n$, $Q_j(\theta)$, $R_j(\theta)$ is fixed. Define $\tilde{C}_j = C_j Q_j(\theta) R_j(\theta)$ and note that PTP is now a geometric program, PGP (Section 4.1), with a new coefficient vector \tilde{C} . Looking at the dual function $v(\delta, \lambda)$:

$$\begin{aligned}
 v(\delta, \lambda) &= \left[\prod_{j=1}^n \left(\frac{u_j(t, \theta)}{\delta_j} \right)^{\delta_j} \right] \left[\prod_{k=1}^K \lambda_k \right] \\
 &= \left[\prod_{j=1}^n \left(\frac{C_j}{\delta_j} \right)^{\delta_j} \right] \left[\prod_{j=1}^n (Q_j(\theta) R_j(\theta))^{\delta_j} \right] \left[\prod_{k=1}^K \lambda_k \right] \\
 &= \left[\prod_j \left(\frac{C_j Q_j(\theta) R_j(\theta)}{\delta_j} \right)^{\delta_j} \right] \left[\prod_k \lambda_k \right] \\
 &= \left[\prod_j \left(\frac{\tilde{C}_j}{\delta_j} \right)^{\delta_j} \right] \left[\prod_k \lambda_k \right]
 \end{aligned} \tag{6}$$

Since θ is fixed, $\partial L / \partial \theta \equiv 0$, and the optimality conditions on θ in the dual are now redundant. The remaining constraints are the usual dual geometric program constraints, so that the dual problem is the geometric programming dual, with the new coefficient vector \tilde{C} .

Our next theorem is the analog of the duality theorem of geometric programming. To state and prove it, we first state the dual transcendental program.

Program DTP

$$\min_{\theta} \max_{\delta, \lambda} v(\delta, \lambda, \theta) = \left[\begin{array}{c} n \\ \Pi \\ j=1 \end{array} \left(\frac{C_j}{\delta_j} \right)^{\delta_j} \right] \left[\begin{array}{c} K \\ \Pi \\ k=1 \end{array} \lambda_k \right] \left[\begin{array}{c} \Pi \\ j \end{array} \left\{ \Pi_{\ell} \theta_{\ell}^{b_{\ell j}} e^{d_{\ell j} \theta} \right\} \delta_j \right]$$

subject to: θ is primal feasible (7)

$$A\delta = 0 \quad (8)$$

$$(B + \Theta D)\delta = 0 \quad (9)$$

$$\sum_{j \in \langle k \rangle} \delta_j = \begin{cases} \lambda_k, & k = 1, 2, \dots, K \\ 1, & k = 0 \end{cases} \quad (10)$$

$$\delta \geq 0, \quad \theta > 0 \quad (11)$$

Theorem 8.5: Suppose the primal program PTP is superconsistent and attains its constrained minimum at a feasible point (t^*, θ^*) . Then

- (i) The corresponding dual program DTP is consistent and has a feasible solution $(\delta^*, \lambda^*, \theta^*)$.
- (ii) $g_0(t^*, \theta^*) = v(\delta^*, \lambda^*, \theta^*)$.
- (iii) the relations (3.2) hold at $(t^*, \theta^*, \delta^*, \lambda^*)$.

Proof:

(1) By Proposition 8.1 the dual is consistent. We now show that the vector δ^* defined by

$$\delta_j^* = \begin{cases} u_j(t^*, \theta^*)/g_0(t^*, \theta^*) & j \in \langle 0 \rangle \\ \mu_k^* \frac{u_j(t^*, \theta^*)}{g_0(t^*, \theta^*)} & j \in \langle k \rangle \end{cases}$$

is dual feasible where μ_k^* are the Lagrange multipliers whose existence is assured by Proposition 8.1. Clearly $\delta_j^* \geq 0$ and by hypothesis $\theta^* > 0$. Thus

$$\sum_j a_{ij} \delta_j^* = \frac{1}{g_0(t^*, \theta^*)} \sum_{k=0}^K \mu_k^* \sum_{j \in \langle k \rangle} a_{ij} u_j(t^*, \theta^*), \quad i = 1, 2, \dots, m$$

By Proposition 8.1 the right-hand side equals

$$\frac{\partial \log[L(t^*, \theta^*, \mu^*)]}{\partial \log t_i}$$

which must be 0 when (t^*, θ^*, μ^*) are the optimal values. Similarly

$$\begin{aligned} & \sum_j (b_{lj} + \theta^* d_{lj}) \delta_j^* \\ &= \frac{1}{g_0(t^0, \theta^0)} \left\{ \sum_{k=0}^K \mu_k^* \sum_{j \in \langle k \rangle} b_{lj} u_j(t^*, \theta^*) \right. \\ & \quad \left. + \theta^* \sum_{k=0}^K \mu_k^* \left[\sum_{j \in \langle k \rangle} d_{lj} u_j(t^*, \theta^*) \right] \right\} = 0 \quad \text{by (2.11)} \end{aligned}$$

It remains to define

$$\lambda_k^* = \sum_{j \in \langle k \rangle} \delta_j^* = \frac{\mu_k^*}{g_0(t^*, \theta^*)}$$

to complete this part of the proof.

By Lemma 8.4 the relations between δ_j^* and $u_j(t^*, \theta^*)$ imply

$$g_0(t^*, \theta^*) = v(\delta^*, \lambda^*, \theta^*)$$

It follows from Proposition 8.2 that

$$v(\delta^*, \lambda^*, \theta^*) = \max_{\delta, \lambda} v(\delta, \lambda, \theta^*) \quad (12)$$

subject to (8)-(11) with

$$\Theta = \Theta^*.$$

Now suppose that there is another pair $(t^\#, \theta^\#)$ satisfying the necessary conditions for a minimum but with $g_0(t^\#, \theta^\#) > g_0(t^*, \theta^*)$.

Using arguments similar to those above, we infer the existence of $\delta^\#, \lambda^\#$, satisfying the dual constraints and conditions (4.2) so that we must have

$$\begin{aligned} g_0(t^\#, \theta^\#) &= v(\delta^\#, \lambda^\#, \theta^\#) \\ &= \max_{\delta, \lambda} v(\delta, \lambda, \theta^\#) \end{aligned}$$

subject to (8)-(11) with

$$\Theta = \Theta^\#$$

Thus $\max v(\delta, \lambda, \theta^\#) > \max v(\delta, \lambda, \theta^*)$. Furthermore, given θ^* , we have shown earlier that $\min_t g_0(t, \theta^*)$ subject to $t > 0$ and $g_k(t, \theta^*) < 1$, is an ordinary geometric program having a corresponding dual problem

$$\max_{\delta, \lambda} v(\delta, \lambda, \theta^*)$$

subject only to (8), (10), (11) without the constraint

$$(B + \Theta D)\delta = 0$$

It follows from (12) that if θ^* is optimal then

$$\max_{\delta, \lambda} \{v(\delta, \lambda, \theta^*) \mid (8)-(11)\} = \max_{\delta, \lambda} \{v(\delta, \lambda, \theta^*) \mid (8), (10), (11)\}$$

Hence (9) can be eliminated, as long as θ is primal feasible.

Therefore, one can view the primal problem as

Program PTP

$$\min_{\theta} \{ \min_t g(t, \theta) \} \tag{13}$$

$$g_k(t, \theta) \leq 1 \tag{14}$$

$$t, \theta > 0 \tag{15}$$

In this representation the dual is

Problem DTP1

$$\min_{\theta} \max_{\delta, \lambda} v(\delta, \lambda, \theta) \tag{16}$$

$$\text{subject to} \quad A\delta = 0 \quad (17)$$

$$\sum_{j \in \langle 0 \rangle} \delta_j = 1 \quad (18)$$

$$\sum_{j \in \langle k \rangle} \delta_j = \lambda_k, \quad k = 1, 2, \dots, K \quad (19)$$

$$\delta, \lambda \geq 0 \quad (20)$$

$$\theta \text{ primal feasible} \quad (21)$$

Since we have shown redundancy of (9), DTP1 is equivalent to DTP when θ is primal feasible. This completes the proof. \square

The dual problem DTP is unsatisfactory in its current form. The reader may justly ask how the condition of primal feasibility could be incorporated in a dual problem, without attaching the whole primal constraint set to the problem. The restriction on θ serves to insure global minimum of $g(t^*, \theta^*)$, but the theorem sacrifices much in elegance to obtain it. In some situations, θ may not be severely restricted by the primal constraints, for instance, in primal unconstrained problems. In this case conditions (7) can be ignored.

We formulated the dual problem with the variables θ appearing in the objective function. It is easy to express θ explicitly in terms of B and D by (9) and then substitute in v to obtain an apparently "pure" dual. Specifically, we find $\theta_\ell = -B_\ell \delta / D_\ell \delta$.

as shown in the proof of Lemma 8.2, and we can use Equation (2.13).

The dual problem obtained this way is

Program DTP2

$$\max v(\delta, \lambda) = \left[\prod_j \left(\frac{c_j}{\delta_j} \right)^{\delta_j} \right] \left[\prod_k \lambda_k \right] \left[\prod_\ell \left(- \frac{B_{\ell} \cdot \delta}{e D_{\ell} \cdot \delta} \right)^{B_{\ell} \cdot \delta} \right] \quad (22)$$

$$\text{subject to} \quad A\delta = 0 \quad (23)$$

$$\sum_{j \in \langle k \rangle} \delta_j = \lambda_k, \quad k = 1, 2, \dots, K \quad (24)$$

$$\lambda_0 = 1 \quad (25)$$

$$\delta \geq 0 \quad (26)$$

This problem is not equivalent to DTP. In particular $v(\delta, \lambda)$ may be undefined even for positive vector δ due to the last product in (22). If a finite solution δ^*, λ^* exists to DTP2, one can define

$$\theta^* \equiv - \frac{B_{\ell} \cdot \delta^*}{D_{\ell} \cdot \delta^*}$$

All conditions for a stationary point in the Lagrangean $L(t, \theta)$ exist, except (perhaps) feasibility of θ . At any rate, these are necessary but not sufficient for a minimum of the primal.

8.4 Some Properties of Transcendental Programs

(a) Recovering primal variables from the dual solution

As noted earlier, θ need not appear explicitly in the dual, although its implicit value appears in the dual objective function and must be primal feasible. If we have a solution δ^*, λ^* to the dual, θ^* can be easily recovered by

$$\theta_l^* = - \frac{b_l \delta_l^*}{d_l \delta_l^*}$$

Once θ^* is known it can be incorporated into the constants C_j , and the problem reduces to the ordinary geometric programming primal-dual relations, discussed in Section 4.3.

(b) Degrees of difficulty in transcendental programs

Although the constraint set of DIP is more restrictive than that of the ordinary dual geometric program, the ordinary orthogonality constraints $A\delta = 0$, and the normality constraints which relate to the nontranscendental variables, still determine the dimension of the dual space. Indeed if $n = m+1$ and $\text{rank } A = m$, these equations will have a unique solution (which may or may not be feasible even if $\delta > 0$). We define, therefore, the degree of difficulty in the same way as for ordinary geometric programs, $dd \equiv n-m-1$. If $dd = 0$, and the dual is feasible, the solution is unique. This striking feature is demonstrated by the following example.

Example 8.6:

$$\begin{aligned} \text{PTP } \min g_0(t, \theta) \\ = 2 \cdot t_1 \cdot t_4 + t_1^{-1} \cdot t_2 \cdot \theta_1 \cdot \theta_2 \cdot \exp(\theta_1) + t_2^{-4} \cdot t_3 \cdot t_4^{-1} \cdot \theta_1 \cdot \exp(-3\theta_1 + \theta_2) \end{aligned}$$

subject to

$$g_1(t, \theta) = t_1 \cdot t_2 \cdot \theta_1^{-1} + t_2 \cdot t_3^{-1} \cdot \theta_2^{-1} \cdot \exp(-2\theta_2) \leq 1$$

$$t, \theta > 0$$

The dual constraints are:

1.	δ_1	+	δ_2	+	δ_3		$= 1$	normality
2. t_1 :	δ_1	-	δ_2		δ_3	+	δ_4	$= 0$
3. t_2 :			δ_2	-	$4\delta_3$	+	δ_4	$\delta_5 = 0$
4. t_3 :					δ_3	-	δ_5	$= 0$
5. t_4 :	δ_1			-	δ_3			$= 0$
6. θ_1 :			$(1 + \theta_1)\delta_2$	+	$(1 - 3\theta_1)\delta_3$	-	δ_4	$= 0$
7. θ_2 :			δ_2	+	$\theta_2\delta_3$		$-(1 + 2\theta_2)\delta_5$	$= 0$

$\left. \begin{array}{l} \text{ordinary} \\ \text{orthogonality} \end{array} \right\}$

$\left. \begin{array}{l} \text{transcendental} \\ \text{orthogonality} \end{array} \right\}$

The first five equations have the unique solution

$$\delta^* = \left(\frac{1}{4}, \frac{1}{2}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4} \right)$$

Substituting into equation (6), we have

$$\theta_1^* = \frac{\delta_4 - \delta_3 - \delta_2}{\delta_2 - 3\delta_3} = 2$$

From equation (7) we have

$$\theta_2^* = \frac{\delta_5 - \delta_2}{\delta_3 - 2\delta_5} = 1$$

The dual objective function is

$$\begin{aligned} v(\delta) &= \left[\prod_i \left(\frac{C_i}{\delta_i} \right)^{\delta_i} \right] \left[\prod_k \lambda_k \right] \prod_l \left[- \frac{B_{l, \delta}}{e D_{l, \delta}} \right]^{B_{l, \delta}} \\ &= \left(\frac{2}{\delta_1} \right)^{\delta_1} \left(\frac{1}{\delta_2} \right)^{\delta_2} \left(\frac{1}{\delta_3} \right)^{\delta_3} \left(\frac{1}{\delta_4} \right)^{\delta_4} \left(\frac{1}{\delta_5} \right)^{\delta_5} \cdot (\delta_4 + \delta_5)^{(\delta_4 + \delta_5)} \\ &\quad \times \left[\frac{\delta_4 - \delta_2 - \delta_3}{e(\delta_2 - 3\delta_3)} \right]^{-\delta_4 + \delta_2 + \delta_3} \left[\frac{\delta_5 - \delta_2}{e(\delta_3 - 2\delta_5)} \right]^{-\delta_5 + \delta_2} \\ &= (8 \cdot 4 \cdot 4 \cdot 4 \cdot 4)^{1/4} \cdot (1/4)^{1/4} \cdot (2/e)^{1/2} \cdot (1/e)^{1/4} \\ &= 4 \cdot 8^{1/4} \cdot e^{-3/4} = g_0(t^*, \theta^*) \end{aligned}$$

From the primal-dual relations at the optimum:

$$2t_1 t_4 = \frac{1}{4} g_0(t^*, \theta^*) = 8^{1/4} e^{-3/4}$$

$$t_1^{-1} t_2 \cdot 2e^2 = \frac{1}{2} g_0(t^*, \theta^*) = 2 \cdot 8^{1/4} e^{-3/4}$$

so

$$t_1^{-1} t_2 = 8^{1/4} e^{-11/4}$$

$$t_2^{-4} t_3 t_4^{-1} \cdot 2 \cdot e^{-5} = 8^{1/4} e^{-3/4}$$

$$t_1 t_2 \cdot \frac{1}{2} = \frac{1}{2}$$

∴

$$t_1 t_2 = 1$$

$$t_2/t_1 = 8^{1/4} e^{-11/4}$$

Hence

$$t_2 = 8^{1/8} e^{-11/8}$$

$$t_1 = 8^{-1/8} e^{11/8}$$

$$t_4 = \frac{1}{2} 8^{1/4} e^{-3/4} / (8^{-1/8} e^{11/8}) = \frac{1}{2} 8^{3/8} e^{-17/8}$$

$$t_2 t_3 e^{-2} = \frac{1}{2}$$

$$t_3 = 2e^{-2} t_2 = 2 \cdot 8^{1/8} \cdot e^{-27/8}$$

Thus

$$t^* = (8^{-1/8} e^{11/8}, 8^{1/8} e^{-11/8}, 2 \cdot 8^{1/8} e^{-27/8}, \frac{1}{2} 8^{3/8} e^{-17/8})$$

To verify that this point is not a maximum, let t_4 increase without limit.

8.5 Condensing Transcendental Programs

In much the same way that condensation was applied to geometric programs in Chapter 9, it can be applied also to transcendental programs. In fact, the proof of Lemma 8.2 already showed a condensation.

In general, let

$$g_k(t, \theta) = \sum_{j \in \langle k \rangle} u_j(t, \theta)$$

and let (t^0, θ^0) be some positive point. Define the weights

$$w_j^0 \equiv w_j(t^0, \theta^0) \equiv \frac{u_j(t^0, \theta^0)}{g_k(t^0, \theta^0)} \quad (1)$$

then

$$g_k(t, \theta) \geq \prod_{j \in \langle k \rangle} \left(\frac{u_j(t, \theta)}{w_j^0} \right)^{w_j^0} \quad (2)$$

The right-hand side is a single term posynomial having new exponents

$$\alpha_{ik} \equiv \alpha_{ik}(t^0, \theta^0) = \sum_{j \in \langle k \rangle} a_{ij} w_j^0 \quad (3)$$

$$\beta_{lk} \equiv \beta_{lk}(t^0, \theta^0) = \sum_{j \in \langle k \rangle} b_{lj} w_j^0 \quad (4)$$

$$\gamma_{lk} \equiv \gamma_{lk}(t^0, \theta^0) = \sum_{j \in \langle k \rangle} d_{lj} w_j^0 \quad (5)$$

and a new coefficient

$$\tilde{c}_k \equiv \tilde{c}_k(t^0, \theta^0) = \prod_{j \in \langle k \rangle} \left(\frac{c_j}{w_j^0} \right)^{w_j^0}$$

We write

$$\tilde{g}_k(t, \theta, t^0, \theta^0) = \tilde{c}_k \prod_i t_i^{\alpha_{ik}} \prod_l \theta_l^{\beta_{lk}} \prod_l \exp(\theta_l \gamma_{lk}) \quad (6)$$

Taking logarithms, we obtain

$$\log \tilde{g}_k = \log \tilde{c}_k + \sum_i \alpha_{ik} \log t_i + \sum_l \beta_{lk} \log \theta_l + \sum_l \gamma_{lk} \theta_l \quad (7)$$

This form is a linear-logarithmic function.

A condensed transcendental program $\tilde{\text{FTP}}$ can thus be viewed as a General Linear-Logarithmic problem. In this form, however, we do

not have the product $\theta \ln \theta$ which characterizes Clasen's Linear Logarithmic problems [12]. It can be easily verified that:

$$(i) \quad \tilde{g}_k(t^0, \theta^0, t^0, \theta^0) = g_k(t^0, \theta^0) \quad (8)$$

$$(ii) \quad \nabla \tilde{g}_k(t^0, \theta^0, t^0, \theta^0) = \nabla g_k(t^0, \theta^0) \quad (9)$$

$$(iii) \quad \tilde{g}_k(t, \theta, t^0, \theta^0) \leq g_k(t, \theta) \quad \text{for all } (t, \theta) > 0 \quad (10)$$

The proof is analogous to that of Lemma 7.3.

8.6 Solving Transcendental Programs by Condensation

In principle, we could apply the same ideas used in the cutting plane method of Chapter 9 to transcendental problems. Regrettably, our subproblems are not linear, but are linear-logarithmic (or linear-exponential). Trying to "linearize" (7) by defining

$$z_i \equiv \ln t_i$$

$$y_l \equiv \ln \theta_l$$

leads to the function

$$\tilde{g}_k(z, y) \equiv \sum_i \alpha_{ik} z_i + \sum_l \beta_{lk} y_l + \sum_l r_{lk} e^{y_l} + \log \tilde{C}_k$$

Noting that e^x is a convex function on \mathbb{R} , we can still linearize by observing that

$$e^y \geq e^{y_0} + (y - y_0)e^{y_0}$$

Letting $y_\ell^0 = \log \theta_\ell^0$ the condensed linearized form of the constraint $g_k(t, \theta) \leq 1$ becomes

$$\tilde{g}_k^L(z, y, t^0, \theta^0) \equiv \sum_i \alpha_{ik} z_i + \sum_\ell (\beta_{\ell k} + \bar{\gamma}_{\ell k}) y_\ell + \bar{c}_k \leq 0$$

where

$$\begin{aligned} \bar{\gamma}_{\ell k} &\equiv \theta_\ell^0 \gamma_{\ell k}, \\ \bar{c}_k &\equiv \ln \tilde{c}_k + \sum_\ell \gamma_{\ell k} \theta_\ell^0 (1 - \log \theta_\ell^0) \end{aligned}$$

If $\bar{\gamma}_{\ell k} \geq 0$ for all ℓ then

$$g(z, y) \geq \tilde{g}(z, y, t^0, \theta^0) \geq \tilde{g}^L(z, y, t^0, \theta^0)$$

and all the cutting plane arguments remain valid, so this could serve as a solution method. If $\gamma_{\ell k}$ are not nonnegative (and recall that these are functions of the condensation point and not constants!), then the right inequality above need not be **valid**, in fact, we may have $\tilde{g}^L > g$. The algorithm in this case no longer guarantees the nesting property and the "cuts" need not exclude the previous point, so that the method is no longer applicable. If $d_{ij} \geq 0$ for all i and j then, by using the transformation $z_i = \log t_i$ and $y_\ell = \log \theta_\ell$ we obtain a convex transformed transcendental program with transformed posynomial terms of the form:

$$\begin{aligned}
 u_j(\mathbf{z}, \mathbf{y}) &= C_j \exp\left(\sum_i a_{ij} z_j\right) \exp\left(\sum_\ell b_{\ell j} y_\ell\right) \exp\left(\sum_\ell d_{\ell j} e^{y_\ell}\right) \\
 &= C_j \exp\left(\sum_i a_{ij} z_j + \sum_\ell [b_{\ell j} y_\ell + d_{\ell j} e^{y_\ell}]\right)
 \end{aligned}$$

In this case condensation at any $(t, \theta) > 0$ will yield nonnegative $\gamma_{\ell k}$, so that linearization is possible.

It seems that in general, the primal transcendental program is easier to solve, as the transformed problem is well defined for all y, z . The dual problem may be ill-defined even for positive δ . The resulting computational difficulties may offset any advantages offered by linearity of the dual constraints.

8.7 Nonideal Systems and Transcendental Programs

Most of the computational and theoretical treatment of chemical equilibrium presented so far, was confined to ideal system. Reviewing some notions of Chapter 2, recall that the (dimensionless) free energy for a general system is

$$F(T, P, \mathbf{x}) = \sum_j x_j \mu_j(T, P, \mathbf{x}) \quad (1)$$

where $\mu_j(T, P, \mathbf{x})$ is the chemical potential or the partial molar free energy. For fixed temperature T and pressure P we obtain from (2.4.5) after dividing by RT

$$\mu_j = \mu_j^0 + \log a_j \quad (2)$$

In principle, all of the nonideality in the system is expressed in a_j , the activity of species S_j . In the ideal case $a_j = \hat{x}_j$. In the nonideal case, without loss of generality, the nonideality can be lumped in the activity coefficient $\gamma_j(x)$ by defining

$$a_j = \gamma_j(x) \cdot \hat{x}_j \quad (3)$$

By this definition of $\gamma_j(x)$, the free energy of the system becomes

$$F(x) = \sum_j x_j (c_j + \log \gamma_j(x) + \log \hat{x}_j) \quad (4)$$

The term $\log \gamma_j(x)$ represents the deviation from ideality. It would be nice, theoretically, if $F(x)$ retained its cherished properties--homogeneity of degree one and convexity. In practical applications this need not be the case, as $\gamma_j(x)$ is sometimes expressed by empirical relations. Still, in the majority of cases at least the homogeneity is preserved. Recall that $\gamma_j(x)$ must be homogeneous of degree 0 to maintain homogeneity of F , since it is multiplied by x_j . This suggests, as is indeed true by chemical arguments, that $\gamma_j(x)$ is a function of concentration, independent of the total amount of one species or another. That is to say,

$$\gamma_j(x) = \gamma_j(\hat{x})$$

Several relations of this type appear in the chemical literature; some of the useful relations are summarized by Prausnitz et al. [42].

For a vapor phase, perhaps the most commonly used correction is by the virial equation. With it, the correction takes the form

$$\log \gamma_j(x) = \frac{1}{v} \sum_i \hat{x}_i \beta_{ji} + \frac{3}{2v^2} \sum_{i,l} \hat{x}_i \hat{x}_l \gamma_{jil} + \dots \quad (5)$$

Here v is the molar volume of the phase, β_{ji} and γ_{jik} are empirical constants, usually determined by complex relations to other parameters.

For liquid phase, the Wilson equation is

$$\log \gamma_j(x) = 1 - \log \left[\sum_i \hat{x}_i \Lambda_{ji} \right] - \sum_l \left[\frac{\hat{x}_l \Lambda_{lj}}{\sum_l x_l \Lambda_{il}} \right] \quad (6)$$

where again, the constants Λ_{ji} are determined by complex relations.

We shall see that the dual transcendental program leads to corrections of the form $\gamma_j(x) = \gamma_j(\hat{x})$. The exact form of $\gamma_j(x)$ and its meaning depend on appropriate choice of coefficients in the matrices B and D of the transcendental program. In this sense, transcendental programs are a generalization of the ideal chemical equilibrium problem. We consider here program DTP2 of Section 8.3. Taking the logarithm of the dual function, replacing δ by x and λ by \bar{x} to conform with chemical terminology, we find

$$F(\mathbf{x}) = -\log v(\mathbf{x}, \bar{\mathbf{x}})$$

$$\begin{aligned}
&= \sum_j x_j [-\log C_j + \log x_j] - \sum_k \bar{x}_k \log \bar{x}_k \\
&\quad - \sum_l \left\{ \left(\sum_j b_{lj} x_j \right) \cdot \log \left(- \frac{\sum_j d_{lj} x_j}{e \cdot \sum_j b_{lj} x_j} \right) \right\} \\
&= \sum_j x_j [-\log C_j + \log \hat{x}_j] - \sum_j \left[\sum_l (B_{l \cdot} \cdot \mathbf{x}) \log \left(- \frac{B_{l \cdot} \cdot \mathbf{x}}{e \cdot D_{l \cdot} \cdot \mathbf{x}} \right) \right] \\
&= \sum_j x_j \left\{ -\log C_j + \log \hat{x}_j - \sum_l b_{lj} \left[\log \left(- \frac{B_{l \cdot} \cdot \mathbf{x}}{D_{l \cdot} \cdot \mathbf{x}} \right) - 1 \right] \right\} \\
&= \sum_j x_j \left\{ c_j + \log \hat{x}_j + \sum_l b_{lj} \log \left(- \frac{D_{l \cdot} \cdot \mathbf{x}}{B_{l \cdot} \cdot \mathbf{x}} \right) \right\}
\end{aligned}$$

Here $c_j = -\log C_j + \sum_l b_{lj} \cdot B_{l \cdot}$, $D_{l \cdot}$ are the l -th rows of B and D , respectively. With c_j taken as the new free energy coefficient, the sum

$$\sum_l b_{lj} \log \left(- \frac{\sum_j d_{lj} x_j}{\sum_j b_{lj} x_j} \right) \quad (7)$$

is the logarithm of the "correction" term $\gamma_j(\mathbf{x})$, accounting for non-ideality. Thus

$$\gamma_j(\mathbf{x}) = \prod_l \left[- \frac{\sum_j d_{lj} x_j}{\sum_j b_{lj} x_j} \right]^{b_{lj}} \quad (8)$$

The bracketed term must be positive for $\gamma_j(\mathbf{x})$ to be well defined.

Based on the structure of B and D we can now examine several cases. Note that B and D have n columns and p rows where p can be arbitrarily chosen (equivalently, the primal program has p transcendental variables). We shall choose p to yield meaningful expressions for $\gamma_j(x)$.

1. $p = K$ (the number of phases)

Let

$$b_{lj} = \begin{cases} 1 & j \in \langle l \rangle \\ 0 & \text{otherwise} \end{cases}$$

For example

$$B = \begin{bmatrix} 1 & 1 & 0 & \dots & \dots & \dots & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & \dots & 0 \\ 0 & \dots & \dots & 0 & 1 & 0 & \dots & 0 \\ 0 & \dots & \dots & \dots & 0 & 1 & 1 & 1 & 1 \end{bmatrix}$$

Each row represents a phase. There are 1's in the row in the locations corresponding to species in the phase. Then for species S_j in phase Φ_k we obtain by expanding (8)

$$\gamma_j(x) = \frac{-\sum_{i=1}^n d_{ki} x_i}{\sum_{i \in \langle k \rangle} x_i} = - \sum_{i \in \langle k \rangle} d_{ki} \hat{x}_i \equiv \gamma_k$$

This introduces for each phase, a single correction factor γ_k for all species in the phase. The factor depends on the concentrations

Let $S_j \in \Phi_k$ and expand (8) with the given B and D to get

$$r_j(x) = \left[\frac{1}{d_j} \left(\frac{-b_{jj}x_j}{\bar{x}_k} \right) \right]^{-b_{jj}} \equiv \beta_j (\hat{x}_j)^{\eta_j}$$

where

$$\beta_j = -b_{jj}/d_j$$

$$\eta_j = -b_{jj}$$

By appropriate selection of d_j , β and η are independent. $r_j(x)$ is then a two parameter correction. However, taking the logarithm

$$\log r_j(x) = \log \beta_j + \eta_j \log \hat{x}_j$$

we see that $\log \beta_j$ can be lumped with c_j and the remaining part is a linear term in $\log \hat{x}_j$. This is perhaps the simplest and most useful correction as it corrects the "ideal" $\log \hat{x}_j$ to $(\eta_j+1) \cdot \log \hat{x}_j$ i.e.,

$$F(x) = \sum_j x_j \cdot c_j + \bar{\eta}_j \log \hat{x}_j$$

where c_j includes $\log \beta_j$ and $\bar{\eta}_j = \eta_j + 1$. It is relatively easy to construct empirical corrections of this type. In some sense, this is a "first order" correction for nonideality.

In a similar fashion, by selecting more complex structures for B and D, more elaborate corrections can be obtained. Notice that in our corrections, $r_j(x)$ is homogeneous of degree zero, as

required by theoretical considerations. We do not propose here a solution technique, nor do we imply that transcendental duality promises a new theory of chemical duality for nonideal system, analogous to the theory for ideal systems. Still, the examples above, and potential applications of transcendental programs in engineering design, suggest that the first steps developed here merit further study.

APPENDIX A

TEST PROBLEMS AND SOLUTIONS

Seven chemical problems and seven geometric programming problems are presented here for reference. Most of the problems were found in the literature but in some cases these were slightly modified. All the solutions, except as noted, are those obtained by RAND's Chemical Composition Code [48]. MDA solutions vary slightly and are somewhat less accurate.

Data contained for chemical problems:

1. Name and source of problem.
2. Dimensions: m = number of subspecies,
 n = number of species,
 k = number of phases.
3. F^*/RT = minimal dimensionless free energy.
4. Subspecies table:
 i = serial number,
 B_i = name (formula) of subspecies,
 b_i = right hand side (moles of B_i),
 z_i^* = optimal multiplier (dual variable).

5. Species table:

- k = phase number,
- j = species number (sequential),
- S_j = species name (formula),
- c_j = free energy coefficient,
- formula = composition of S_j by subspecies,
- x_j^* = equilibrium composition (moles).

After each phase its name and its total number of moles \bar{x}_k are listed.

Problems A.8-A.14 are geometric programming problems. Data for these problems includes:

1. Name and source of problem.
2. Dimensions: m = number of variables,
 n = number of terms,
 k = number of constraints (without objective function).

3. Terms table:

- k = constraint number (0 = objective function),
- j = term number (sequential)
- C_j = coefficient
- Exponents = variables and their exponents in term j ,
- a_j^* = optimal value of the term.

Each constraint is followed by the value of its multiplier λ_k at the optimal solution.

4. Variables table:

i = serial number,

symbol = symbol of variable i ,

Low bd. = lower bound used in MDA computations,

Upper bd. = upper bound used in MDA computations,

t_i^* = optimal value.

Note: The notation $2.5(10^{-5})$ means 2.5×10^{-5} .

A.1 SMALL PROBLEM (Clasen [14])

Dimensions: $m = 2, n = 4, k = 2$

$$F^*/RT = -5.6755$$

i	B_i	b_i	z_i^*
1	R1	3.5	-1.1032
2	R2	4.5	-0.4032

k	j	S_j	c_j	Formula	x_j^*
1	1	C1	0.7	R2	0.8340
	2	C2	-0.7	R1	1.6795
				PHASE1 $\bar{x}_1 =$	2.5136
2	3	C3	0.0	R2	3.6060
	4	C4	0.0	R1	1.8205
				PHASE2 $\bar{x}_2 =$	5.4865

A.2 SODA-WATER (De Haven [19])

Dimensions: $m = 5$, $n = 13$, $k = 2$.

$$F^*/RT = -2253.21$$

i	B_i	b_i	z_i^*
1	O ₂	0.6469	-12.8783
2	CO ₂	0.2602	-10.6341
3	N ₂	3.7058	-11.8027
4	OH ⁻	55.8103	-24.8080
5	OH ⁺	55.8103	-14.5820

k	j	S_j	c_j	formula	x_j^*
1	1	O ₂	-10.85	O ₂	0.6468
	2	CO ₂	-7.69	CO ₂	0.2588
	3	N ₂	-11.52	N ₂	3.7054
	4	H ₂ O	-36.61	H ⁺ + OH ⁻	0.3050
				GAS PHASE $\bar{x}_1 =$	4.9159
2	5	O ₂	0.0	O ₂	1.42 (10 ⁻⁴)
	6	CO ₂	0.0	CO ₂	1.34 (10 ⁻³)
	7	N ₂	0.0	N ₂	4.15 (10 ⁻⁴)
	8	H ⁺	0.0	H ⁺	2.58 (10 ⁻⁵)
	9	OH ⁻	0.0	OH ⁻	9.34 (10 ⁻¹⁰)
	10	H ₂ O	-39.39	H ⁺ + OH ⁻	55.5053
	11	HCO ₃ ⁻	-20.86	CO ₂ + OH ⁻	2.58 (10 ⁻⁵)
	12	H ₂ CO ₃	-33.61	CO ₂ + H ⁺ + OH ⁻	4.13 (10 ⁻⁶)
	13	CO ₃ ⁼	6.73	CO ₂ - H ⁺ + OH ⁻	5.78 (10 ⁻¹¹)
				LIQUID PHASE $\bar{x}_2 =$	55.5073

A.3 HYDRAZINE (White, Johnson and Dantzig [54])

Dimensions: $m = 3, n = 10, k = 1$

$$F^*/RT = -47.7611$$

i	B_i	b_i	z_j^*
1	H	2.0	- 9.7851
2	O	1.0	-15.2221
3	N	1.0	-12.9689

k	j	S_j	c_j	formula	x_j^*
1	1	H	- 6.089	H	$4.06 (10^{-2})$
	2	H ₂	-17.164	2H	$1.48 (10^{-1})$
	3	H ₂ O	-34.054	2H + O	$7.83 (10^{-1})$
	4	N	- 5.914	N	$1.41 (10^{-3})$
	5	N ₂	-24.721	2N	$4.85 (10^{-1})$
	6	NH	-14.936	N + H	$6.93 (10^{-4})$
	7	NO	-24.100	N + O	$2.72 (10^{-2})$
	8	O	-10.708	O	$1.79 (10^{-2})$
	9	O ₂	-26.662	2O	$3.73 (10^{-2})$
	10	OH	-22.179	O + H	$9.69 (10^{-2})$
				GAS PHASE $\bar{x} =$	1.638

A.4 SODA-POP (Shapley and Cutler [48])

Dimensions: $m = 9$, $n = 17$, $k = 2$

$$F^*/RT = -3128.86$$

i	B_i	b_i	z_i^*
1	O ₂	5.27583	- 2.9421
2	CO ₂	6.07349	-10.4964
3	N ₂	82.58040	-11.7115
4	H ⁺	52.81000	-21.1660
5	OH ⁻	52.83950	-18.2997
6	Cl ⁻	0.08005	- 6.3746
7	Na ⁺	0.08813	- 6.2785
8	K ⁺	0.04829	- 4.2834
9	GLUCOSE	0.02000	- 7.7625

A.4 SCDA-POP Cont.

k	j	S _j	c _j	formula	x _j *	
1	1	O ₂	0.0	O ₂	5.2758	
	2	CO ₂	- 7.69	CO ₂	6.0427	
	3	N ₂	-11.52	N ₂	82.5800	
	4	H ₂ O	-36.60	H ⁺ + OH ⁻	6.1075	
			GAS PHASE	$\bar{x}_1 =$	100.000	
2	5	O ₂	10.94	O ₂	4.39 (10 ⁻⁵)	
	6	CO ₂	0.0	CO ₂	1.30 (10 ⁻³)	
	7	N ₂	0.0	N ₂	3.85 (10 ⁻⁴)	
	8	H ₂ O	-39.39	H ⁺ + OH ⁻	46.70	
	9	H ⁺	0.0	H ⁺	3.02 (10 ⁻⁸)	
	10	OH ⁻	0.0	OH ⁻	5.69 (10 ⁻⁷)	
	11	Cl ⁻	0.0	Cl ⁻	8.00 (10 ⁻²)	
	12	Na ⁺	0.0	Na ⁺	8.81 (10 ⁻²)	
	13	K ⁺	2.60	K ⁺	4.83 (10 ⁻²)	
	14	HCO ₃ ⁻	-21.35	CO ₂ + OH ⁻	2.94 (10 ⁻²)	
	15	H ₂ CO ₃	-32.84	CO ₂ + H ⁺ + OH ⁻	1.85 (10 ⁻⁶)	
	16	CO ₃ ⁼	6.26	CO ₂ - H ⁺ + OH ⁻	4.86 (10 ⁻⁵)	
	17	GLUCOSE	0.0	GLUCOSE	2.00 (10 ⁻²)	
				LIQUID PHASE	$\bar{x}_2 =$	46.97

A.5 RESPIRATORY SYSTEM (Dantzig, DeHaven and Sams [17])

Dimensions: $m = 12$, $n = 30$, $k = 3$

$$F^*/RT = 1835.24$$

i	B_i	b_i	z_i^*
1	O ₂	0.013317	-12.8354
2	CO ₂	0.022709	-15.0412
3	N ₂	0.024854	-11.7188
4	H ⁺	46.700000	-25.9376
5	OH ⁻	46.719700	-13.2999
6	Cl ⁻	0.081400	- 5.6638
7	Na ⁺	0.080920	-23.5015
8	k ⁺	0.050000	- 6.1512
9	Hb1 ⁻	0.009090	-11.2836
10	HPp ⁻	0.088000	-12.2357
11	HPr-	0.011900	- 7.5867
12	Z	0.0	17.8266

Note: The problem called RESPIRATORY2 was generated from this problem by eliminating the last row Z.

A.5 RESPIRATORY SYSTEM Cont.

k	j	s_j	c_j	formula	x_j^*
1	1	O ₂	-10.89	O ₂	4.397 (10 ⁻⁵)
	2	CO ₂	- 7.69	CO ₂	1.974 (10 ⁻⁵)
	3	N ₂	-11.49	N ₂	2.447 (10 ⁻²)
	4	H ₂ O	-36.44	H ⁺ + OH ⁻	1.875 (10 ⁻⁵)
			AIR OUT	$\bar{x}_1 =$	3.076 (10 ⁻²)
2	5	O ₂	0.0	O ₂	6.284 (10 ⁻⁵)
	6	CO ₂	0.0	CO ₂	6.923 (10 ⁻⁶)
	7	N ₂	0.0	N ₂	1.919 (10 ⁻⁴)
	8	H ⁺	0.0	H ⁺ + Z	7.080 (10 ⁻⁵)
	9	OH ⁻	0.0	OH ⁻ - Z	7.154 (10 ⁻¹³)
	10	Cl ⁻	0.0	Cl ⁻ - Z	1.482 (10 ⁻⁷)
	11	Na ⁺	0.0	Na ⁺ + Z	8.092 (10 ⁻²)
	12	H ₂ O	-39.23	H ⁺ + OH ⁻	23.41
	13	HCO ₃ ⁻	-21.20	CO ₂ + OH ⁻ - Z	3.383 (10 ⁻¹⁰)
	14	H ₂ CO ₃	0.0	CO ₂ + OH ⁻ + H ⁺	6.305 (10 ⁻²⁵)
	15	CO ₃ ⁼	6.25	CO ₂ + OH ⁻ - H ⁺ - 2Z	1.350 (10 ⁻¹⁸)
	16	HPp ⁻	0.0	HPp ⁻ - Z	8.800 (10 ⁻⁴)
			PLASMA	$\bar{x}_2 =$	23.583

A.5 RESPIRATORY SYSTEM Cont.

k	j	S_j	c_j	formula	x_j^*
3	17	O_2	0.0	O_2	$6.252 (10^{-5})$
	18	CO_2	0.0	CO_2	$6.888 (10^{-6})$
	19	N_2	0.0	N_2	$1.910 (10^{-4})$
	20	H^+	0.0	H^+	$1.276 (10^{-10})$
	21	OH^-	0.0	OH^-	$3.929 (10^{-5})$
	22	Cl^-	0.0	Cl^-	$8.140 (10^{-2})$
	23	K^+	0.0	K^+	$5.000 (10^{-2})$
	24	H_2O	-39.23	$H^+ + OH^-$	23.29
	25	HCO_3^-	-21.20	$CO_2 + OH^-$	$1.858 (10^{-2})$
	26	H_2CO_3	0.0	$CO_2 + H^+ + OH^-$	$6.273 (10^{-23})$
	27	$CO_3^{=}$	6.25	$CO_2 - H^+ + OH^-$	$4.095 (10^{-3})$
	28	HbI^-	0.0	HbI^-	$2.951 (10^{-4})$
	29	$HbI O_2$	-16.23	$O_2 + HbI^-$	$8.795 (10^{-3})$
	30	HPr^-	0.0	HPr^-	$1.190 (10^{-2})$
			RED CELLS	$\bar{x}_3 =$	23.463

A.6 PLASMA MODEL (Clasen [14])

Dimensions: $m = 16$, $n = 21$, $k = 1$

$F^*/RT = -832.57$

i	B_i	b_i	z_i^*
1	R1	$1.9073 (10^{-5})$	-13.9225
2	R2	$1.1423 (10^{-2})$	-10.5336
3	R3	$1.7458 (10^{-4})$	-11.7084
4	R4	21.108	-21.9996
5	R5	21.119	-18.5960
6	R6	$4.2369 (10^{-2})$	- 6.2166
7	R7	$5.8757 (10^{-2})$	- 5.8896
8	R8	$1.8593 (10^{-3})$	- 9.3429
9	R9	$7.3176 (10^{-4})$	-10.2754
10	R10	$3.9100 (10^{-4})$	-10.9021
11	R11	$1.1382 (10^{-4})$	-12.1362
12	R12	$2.9362 (10^{-4})$	-11.6898
13	R13	$7.1448 (10^{-4})$	-10.2993
14	R14	$1.1643 (10^{-3})$	- 9.8109
15	R15	$1.0177 (10^{-3})$	- 9.9455
16	R16	$3.8600 (10^{-4})$	-10.9150

A.6 PLASMA MODEL Cont.

k	j	S _j	c _j	formula	x _j *
1	1	C1	0.0	R1	1.91 (10 ⁻⁵)
	2	C2	0.0	R2	5.65 (10 ⁻⁴)
	3	C3	0.0	R3	1.75 (10 ⁻⁴)
	4	C4	0.0	R4	1.61 (10 ⁻⁸)
	5	C5	0.0	R5	2.18 (10 ⁻⁷)
	6	C6	0.0	R6	4.24 (10 ⁻²)
	7	C7	0.0	R7	5.88 (10 ⁻²)
	8	C8	0.0	R8	1.86 (10 ⁻³)
	9	C9	0.0	R9	7.32 (10 ⁻⁴)
	10	C10	0.0	R10	3.91 (10 ⁻⁴)
	11	C11	0.0	R11	1.14 (10 ⁻⁴)
	12	C12	0.0	R12	1.78 (10 ⁻⁴)
	13	C13	0.0	R13	7.14 (10 ⁻⁴)
	14	C14	0.0	R14	1.16 (10 ⁻³)
	15	C15	0.0	R15	1.02 (10 ⁻³)
	16	C16	-21.35	R2 + R5	1.08 (10 ⁻³)
	17	C17	-32.84	R2 + R4 + R5	8.04 (10 ⁻⁷)
	18	C18	6.26	R2 - R4 + R5	1.46 (10 ⁻⁵)
	19	C19	-39.39	R4 + R5	21.1082
	20	C20	0.0	R16	3.86 (10 ⁻⁴)
	21	C21	-20.57	R4 + R12	1.16 (10 ⁻⁴)
				PLASMA $\bar{x}_1 =$	21.23

A.7 FETUS MODEL (Clasen [14])

Dimensions: $m = 19$, $n = 51$, $k = 7$

$F^*/RT = -1869.55$

1	B_1	b_1	z_1^*
1	R1	0.61550	- 2.092
2	R2	0.28000	-10.610
3	R3	37.370	-11.792
4	R4	46.113	-22.003
5	R5	0.00581	- 6.832
6	R6	46.137	-17.391
7	R7	0.07925	- 4.219
8	R8	0.07487	-39.995
9	R9	0.08292	- 6.297
10	R10	0.01022	- 6.242
11	R11	0.05720	- 0.875
12	R12	0.00250	-18.561
13	R13	0.0	2.171
14	R12	0.0	- 2.119
15	R15	0.0	- 3.846
16	R16	0.0	- 0.449
17	R17	0.0	- 0.200
18	R18	0.00295	-11.827
19	R19	0.01431	- 8.681

A.7 FETUS MODEL Cont.

k	j	S _j	c _j	formula	x _j *
1	1	C1	0.0	R1	.6057
	2	C2	- 7.690	R2	.2646
	3	C3	-11.520	R3	3.7367
	4	C4	-36.600	R4 + R6	.3001
			PHASE1	$\bar{x}_1 =$	4.9070
2	5	C5	10.940	R1	1.01 (10 ⁻⁴)
	6	C6	0.0	R2	1.14 (10 ⁻³)
	7	C7	0.0	R3	3.48 (10 ⁻⁴)
	8	C8	0.0	R5 - R13	5.66 (10 ⁻³)
	9	C9	0.0	R4 + R13	1.12 (10 ⁻⁷)
	10	C10	0.0	R6 - R13	1.47 (10 ⁻⁷)
	11	C11	0.0	R7 - R13	7.72 (10 ⁻²)
	12	C12	0.0	R8 + R13	1.72 (10 ⁻¹⁵)
	13	C13	2.196	R9 + R13	8.27 (10 ⁻²)
	14	C14	-39.390	R4 + R6	45.83
	15	C15	-21.350	R2 + R6 - R13	6.78 (10 ⁻³)
	16	C16	-32.840	R2 + R4 + R6	1.62 (10 ⁻⁶)
	17	C17	6.260	R2 - R4 + R6 - 2R13	2.84 (10 ⁻⁶)
	18	C18	0.0	R10 - R13	1.02 (10 ⁻²)
	19	C19	0.0	R13 + R18	2.95 (10 ⁻³)
	20	C20	1.568	R13 + R19	1.43 (10 ⁻²)
			PHASE2	$\bar{x}_2 =$	46.03

A.7 FETUS MODEL Cont.

k	j	S _j	c _j	formula	x _j *	
3	21	C21	0.0	R5	1.48 (10 ⁻⁴)	
	22	C22	10.450	R1	4.90 (10 ⁻⁷)	
	23	C23	0.0	R2	3.38 (10 ⁻⁶)	
	24	C24	- 0.500	R3	1.71 (10 ⁻⁶)	
	25	C25	0.0	R4	3.82 (10 ⁻¹¹)	
	26	C26	0.0	R6	3.84 (10 ⁻⁹)	
	27	C27	0.0	R7	2.02 (10 ⁻³)	
	28	C28	-39.389	R8	7.49 (10 ⁻²)	
	29	C29	0.0	R9	2.53 (10 ⁻⁴)	
	30	C30	3.373	R4 + R6	3.66 (10 ⁻²⁰)	
	31	C31	-21.490	R2 + R6	2.04 (10 ⁻⁴)	
	32	C32	-32.840	R2 + R4 + R6	4.82 (10 ⁻⁹)	
	33	C33	6.120	R2 - R4 + R6	7.49 (10 ⁻⁷)	
	34	C34	0.0	R11	5.72 (10 ⁻²)	
	35	C35	0.0	R12 - 4R4	5.72 (10 ⁻⁶)	
	36	C36	- 1.903	R1 + R12 - 3R14 - R15	2.66 (10 ⁻⁵)	
	37	C37	- 2.900	2R1 + R12 - 2R14 - 2R15	5.01 (10 ⁻⁵)	
	38	C38	- 3.362	3R1 + R12 - R14 - 3R15	5.51 (10 ⁻⁵)	
	39	C39	- 7.485	4R1 + R12 - 4R15	2.36 (10 ⁻³)	
	40	C40	2.606	R18	7.40 (10 ⁻⁸)	
	41	C41	0.0	R19	2.33 (10 ⁻⁵)	
				PHASE3	$\bar{x}_3 =$	0.137

A.7 FETUS MODEL Cont.

k	j	S _j	c _j	formula	* x _j
4	42	C42	-15.640	R4 + R14	5.34 (10 ⁻⁸)
	43	C43	0.0	R14	3.10 (10 ⁻⁵)
	44	C44	21.810	-R4 + R14 - 4R16	2.27 (10 ⁻⁴)
				PHASE4 $\bar{x}_4 =$	2.58 (10 ⁻⁴)
5	45	C45	-16.790	R4 + R15	1.13 (10 ⁻⁶)
	46	C46	0.0	R15	2.08 (10 ⁻⁴)
	47	C47	18.980	-R4 + R15 - 4R17	9.53 (10 ⁻³)
				PHASE5 $\bar{x}_5 =$	9.74 (10 ⁻³)
6	48	C48	0.0	R16	5.79 (10 ⁻⁴)
	49	C49	11.960	R2 - R4 + R16	3.29 (10 ⁻⁴)
				PHASE6 $\bar{x}_6 =$	9.08 (10 ⁻⁴)
7	50	C50	0.0	R17	3.12 (10 ⁻²)
	51	C51	12.900	R2 - R4 + R17	6.92 (10 ⁻³)
				PHASE7 $\bar{x}_7 =$	3.81 (10 ⁻²)

A.8 SEA POWER (Duffin, Peterson and Zener [25, p. 127])

Dimensions: $m = 7, n = 10, k = 4$

k	j	C_j	Exponents	u_j^*
0	1	1.0	A	126.7476
			OBJECTIVE =	126.7476
1	2	40.000	$A^{-1} * Q * U^{-1} * \beta^{-1}$	1.0000
			$\lambda_1 =$	0.8265
2	3	0.1800	$A^{-1} * Q * (\beta')^{-1}$	1.000
			$\lambda_2 =$	0.4205
3	4	44.5000	$Q^{-1} * \alpha^{-1}$	0.7753
	5	$6.00 (10^{-8})$	$A * U^3 * Q^{-1} * \alpha^{-1}$	0.1915
	6	$2.15 (10^{-8})$	$U^2 * \alpha^{-1} * \gamma^{-1}$	0.0332
			$\lambda_3 =$	1.2898
4	7	1.0	α	0.5000
	8	1.0	β	0.3204
	9	1.0	β'	0.1630
	10	1.0	γ	0.0166
			$\lambda_4 =$	2.5795

A.8 SEA POWER Cont.

i	Symbol	Low bd.	Upper bd.	t_i^*
1	A	1.000	500.0	126.7476
2	Q	1.000	300.0	114.7878
3	U	10.000	1000.0	113.0615
4	α	0.001	1.0	0.5000
5	β	0.001	1.0	0.3204
6	β'	0.001	1.0	0.1630
7	γ	0.001	1.0	0.0166

A.9 REACTOR DESIGN (Avriel [1])

Dimensions: $m = 3, n = 5, k = 1$

k	j	C_j	Exponents	u_j^*
0	1	400.0	$t_1^{0.4}$	222.3149
	2	4.183	$t_1^{-1.2} * t_2^{1.2}$	47.0450
	3	10.00	$t_1^{-0.5} * t_2^{0.5} * t_3^{-1}$	64.9443
			OBJECTIVE =	334.3042
1	4	1	t_2^{-1}	0.5779
	5	1	t_3	0.4221
			$\lambda_1 =$	0.4603

i	Symbol	Low bd.	Upper bd.	t_i^*
1	t_1	0.1	1.0	0.2303
2	t_2	1.0	10.0	1.7303
3	t_3	0.1	1.0	0.4221

A.10 CONDENSER DESIGN (Avriél and Wilde [2])

Dimensions: $m = 4, n = 8, k = 2$

k	j	C_j	Exponents	u_j^*
0	1	172,400	$DO^{-1} * L^{-4/3} * N^{-8/7}$	98.6717
	2	97,790	$DI^{4/5} * L^{-1} * N^{-1/5}$	171.0548
	3	1.570	$DO * L * N$	410.6575
	4	0.0382	$DI^{-4.8} * L * N^{-1.8}$	54.7279
	5	38,380	$DI^{-1} * L^{-1} * N^{-1}$	162.6811
			OBJECTIVE =	897.8829
1	6	$8.17 (10^{-3})$	DO^{-1}	0.0980
	7	1.0	$DI * DO^{-1}$	0.9020
			$\lambda_1 =$	0.3563
2	8	12.0	DO	1.0000
				$\lambda_2 =$

i	Symbol	Low bd.	Upper bd.	t_i^*
1	DI	0.02	0.080	0.0752
2	DO	0.03	0.084	0.0833
3	L	1.00	200.0	28.0715
4	N	1.00	1000.0	111.8138

NOTE: DI and DO are each considered a single symbol.

A.11 STOCHASTIC CONDENSER (Avriel and Wilde [3])

Dimensions: $m = 9, n = 13, k = 7$

k	j	C_j	Exponents	u_j^*
0	1	39.805	t_2	425.1091
	2	1.570	$t_3 * t_4 * t_5$	404.7844
	3	0.309	t_6	53.9416
			OBJECTIVE =	883.8351
1	4	1.0	$t_2^{-1} * t_7$	0.2372
	5	1.0	$t_2^{-1} * t_8$	0.3947
	6	1.0	$t_2^{-1} * t_9$	0.3681
			$\lambda_1 =$	0.4810
2	7	2373.200	$t_3^{-0.2} * t_5^{-1} * t_8^{-1} * t_{10}^{0.8}$	1.0000
				$\lambda_2 =$
3	8	534.320	$t_3^{7/8} * t_4^{-3/4} * t_5^{-1} * t_7^{-3/4}$	1.0000
				$\lambda_3 =$
4	9	914.150	$t_5^{-1} * t_3^{-1} * t_9^{-1} * t_{10}^{-1}$	1.0000
				$\lambda_4 =$
5	10	0.124	$t_3^{-1.8} * t_5 * t_6^{-1} * t_{10}^{-4.8}$	1.0000
				$\lambda_5 =$

A.11 STOCHASTIC CONDENSER Cont.

k	j	C_j	Exponents	u_j^*
6	11	$8.17 (10^{-3})$	t_4^{-1}	0.0980
	12	1.0	$t_4^{-1} * t_{10}$	0.9020
				$\lambda_6 =$ 0.3527
7	13	12.000	t_4	1.0000
				$\lambda_7 =$ 0.0088

i	Symbol	Low bd.	Upper bd.	t_i^*
1	t_2	5.0	120.0	10.6798
2	t_3	5.0	1000.0	111.9377
3	t_4	0.8	1.0	0.0833
4	t_5	2.0	100.0	27.6394
5	t_6	50.0	1000.0	174.5680
6	t_7	0.5	40.0	2.5336
7	t_8	0.1	40.0	4.2152
8	t_9	0.1	40.0	3.9310
9	t_{10}	0.5	1.0	0.0752

A.12 DECOMPOSITION PROBLEM (Heyman and Avriel [37])

Dimensions: $m = 10, n = 13, k = 3$

k	j	C_j	Exponents	u_j^*
0	1	1.0	$t_1^2 * t_3^2$	5.1486
	2	0.1	$t_2^{-1} * t_9$	0.0628
	3	3.0	$t_1 * t_2^{-1} * t_4 * t_5$	3.2817
	4	2.5	$t_1^2 * t_4 * t_5^{-1}$	3.2817
	5	7.0	$t_1^{-6} * t_2^{-1} * t_4^{-3}$	2.1878
	6	0.9	$t_1^{1/3} * t_6 * t_8$	4.3001
			OBJECTIVE =	18.2627
1	7	0.4	$t_1^{-1} * t_2$	0.4011
	8	1.0	$t_1 * t_3^{-1}$	0.5989
			$\lambda_1 =$	0.8616
2	9	2.0	$t_1^{-3} * t_3^{1/3} * t_6 * t_7^{-3}$	0.2308
	10	1.0	$t_1^{-1} * t_6^{-1} * t_7 * t_8^{-1}$	0.6923
	11	1.0	$t_1^2 * t_2^{-1} * t_8^3$	0.0769
			$\lambda_2 =$	0.5102
3	12	0.5	$t_1^{-1} * t_3^{-1} * t_{10}$	0.6000
	13	3.0	$t_2^{-1} * t_3^{-1} * t_9^{-1} * t_{10}^{-3/2}$	0.4000
			$\lambda_3 =$	0.0086

A.12 DECOMPOSITION Cont.

i	Symbol	Low bd.	Upper bd.	t_i^*
1	t_1	0.01	100	1.1657
2	t_2	0.01	100	1.1690
3	t_3	0.01	100	1.9465
4	t_4	0.01	100	1.0294
5	t_5	0.01	100	1.0656
6	t_6	0.01	100	11.2241
7	t_7	0.01	100	3.6638
8	t_8	0.01	100	0.4045
9	t_9	0.01	100	0.7336
10	t_{10}	0.01	100	2.7229

A.13 DEPENDENT VARIABLE PROBLEM (Artificial)

Dimensions: $m = 3^*$, $n = 4$, $k = 0$ (*2 independent variables)

k	j	C_j	Exponents	u_j^*
0	1	2.0	t_1	2.8738
	2	1.0	$t_1 * t_2 * t_3$	1.0202
	3	4.0	$t_2 * t_3$	2.8375
	4	4.0	$t_1^{-1} * t_2^{-1} * t_3^{-1}$	3.9140
			OBJECTIVE =	10.6455

i	Symbol	Low bd.	Upper bd.	t_i^*
1	t_1	0.2	5.0	1.4381
2	t_2	0.2	5.0	0.2000
3	t_3	0.2	5.0	3.5500

NOTE: This problem was solved by MDA only.

A.14 LOOSE CONSTRAINT PROBLEM (Artificial)

Dimensions: $m = 2, n = 5, k = 1$

k	j	C_j	Exponents	u_j^*
0	1	2.0	t_1	2.8667
	2	1.0	$t_1 * t_2$	1.0272
	3	4.0	t_2	2.8667
	4	4.0	$t_1^{-1} * t_2^{-1}$	3.8939
			OBJECTIVE =	10.6546
1	5	0.3	$t_1 * t_2^{1.5}$	0.2609
			$\lambda_1 =$	0.0

i	Symbol	Low bd.	Upper bd.	t_i^*
1	t_1	0.2	5.0	1.4333
2	t_2	0.2	5.0	0.7166

REFERENCES

- [1] Avriel, M., "Fundamentals of Geometric Programming," in Applications of Mathematical Programming Techniques, E. M. L. Beale (Ed.), American Elsevier, New York (1970), 295-316.
- [2] _____ and Wilde, D. J., "Optimal Condenser Design by Geometric Programming," I & EC, Proc. Des. & Dev. 6 (1967), 256-263.
- [3] _____ and Wilde, D. J., "Engineering Design Under Uncertainty," I & EC, Proc. Des. & Dev. 8 (1969), 124-131.
- [4] _____, "Topics in Optimization: Block Search, Applied and Stochastic Geometric Programming." Doctoral Thesis, Stanford University (1966).
- [5] _____ and Williams, A. C., "Complementary Geometric Programming," SIAM J. Appl. Math. 19 (1970), 125-141.
- [6] _____ and Williams, A. C., "On the Primal and Dual Constraint Sets in Geometric Programming," J. Math. Anal. Appl. 32 (1970), 684-688.
- [7] Bigelow, J. H., "The Chemical Equilibrium Problem: A Special Case of Nonlinear Programming," Doctoral Thesis, Stanford University (1969).
- [8] _____, De Haven, J. C., and Shapiro, N. Z., "Chemical Equilibrium Problems with Unbounded Constraint Sets," SIAM J. Appl. Math. 18 (1970), 768-775.
- [9] _____ and Shapiro, N. Z., "Sensitivity Analysis in Chemical Thermodynamics," The RAND Corp., P-4628 (1971).
- [10] Brinkley, S. R., "Calculation of the Equilibrium Composition of Systems of Many Constituents," J. Chem. Phys. 15 (1947), 107-110.
- [11] Callen, H. B., Thermodynamics, John Wiley, New York (1960).

- [12] Clasen, R. J., "The Linear Logarithmic Programming Problem," The RAND Corp., RM-3707-PR. (1963).
- [13] _____, "The Numerical Solution of the Chemical Equilibrium Problem," The RAND Corp., RM-4345-PR (1965).
- [14] _____, "Some Experiments on a Proposed Method for Solving the Chemical Equilibrium Problem," The RAND Corp., 14918-PR. (1966).
- [15] Dantzig, G.B., Linear Programming and Extensions, Princeton University Press (1963).
- [16] _____, "New Mathematical Methods in Life Sciences," in Computers in Biomedical Research, Vol. I, R. W. Stacy (Ed.), Academic Press (1965), 33-46.
- [17] _____, De Haven, J. C. and Sams, C.F., "A Mathematical Model of the External Respiratory System," Proc. 4th Berkeley Symp. on Math. Stat. and Prob., University of California Press (1960), 181-196.
- [18] _____, Johnson, S. and White, W., "A Linear Programming Approach to the Chemical Equilibrium Problem," Management Sci. 5 (1958).
- [19] De Haven, J. C., "Prerequisites for Chemical Thermodynamic Models of Living Systems," The RAND Corp., RM-5691-PR. (1968).
- [20] Dembo, R. S., "Solution of Complementary Geometric Programs," M.Sc. Thesis, Technion, Israel (1972).
- [21] Denbigh, K., The Principles of Chemical Thermodynamics, Cambridge University Press (1961).
- [22] Dluzniewski, J. H., Adler, S. B., Ozkardesh, H. and Barner, H. E., "Aid to Correlation of Complex Equilibria," AIChE manuscript (1973).
- [23] Dorn, W. S., "Variational Principles for Chemical Equilibrium," J. Chem. Phys. 32 (1960), 1490-1492.

- [24] Duffin, R. J., "Linearizing Geometric Programs." SIAM Review 12 (1970), 227.
- [25] _____, and Zener, C., Geometric Programming, John Wiley, New York (1967).
- [26] _____, and Zener, C., "Geometric Programming, Chemical Equilibrium and the Anti-Entropy Function," Proc. Nat. Acad. Sci. 63 (1969).
- [27] Eaves, B. C. and Zangwill, W. I., "Generalized Cutting Plane Algorithms," SIAM J. Control 9 (1971), 529-541.
- [28] Eisenberg, E., "Duality in Homogeneous Programming," Proc. Am. Math. Soc. 12 (1961), 783-787.
- [29] Fox, B. L., "Stability of the Dual Cutting Plane Algorithm for Concave Programming," The RAND Corp., RM-6147-PR (1970).
- [30] Gibbs, J. W., The Scientific Papers of J. Willard Gibbs, Vol. 1, Dover, New York (1961).
- [31] Guggenheim, E. A., Thermodynamics, 5th Ed., North Holland (1967).
- [32] Hamala, M., "Geometric Programming in Terms of Conjugate Functions," CORE Discussion Paper No. 6811 (1968).
- [33] Hardy, G. H., Littlewood, J. E. and Polya, G., Inequalities, Cambridge Univ. Press (1964).
- [34] Heyman, M. and Avriel, M., "On a Decomposition for a Special Class of Geometric Programming Problems," J. Opt. Th. Appl. 3 (1969), 392 ff.
- [35] Kortanek, K. O., Rom, W. O. and Soyster, A. L., "Solution sets of Convex Programs Related to Chemical Equilibrium Problems," Opn. Res. 21 (1973), 240-246.
- [36] Kuhn, H. W. and Tucker, A. W., "Nonlinear Programming," in Proc. 2nd. Berkeley Symp. on Math. Stat. and Prob., J. Neyman (Ed.), University of California Press, Berkeley, Ca. (1950), 481-492.

- [37] Loomis, L. H. and Sternberg, S., Advanced Calculus, Addison Wesley (1968).
- [38] National Bureau of Standards, Selected Values of Chemical Thermodynamic Properties, Circular 500 (1952).
- [39] Passy, V. and Wilde, D. J., "Generalized Polynomial Optimization," SIAM J. Appl. Math. 15 (1967), 1344-1356.
- [40] _____ and Wilde, D. J., "Mass Action and Polynomial Optimization," J. Eng. Math. 3 (1969), 325-335.
- [41] _____ and Wilde, D. J., "A Geometric Programming Algorithm for Solving the Chemical Equilibrium Problem," SIAM J. Appl. Math. 16 (1968), 363-373.
- [42] Prausnitz, J. M., Eckert, C. A., Drye, R. V. and O'Connell, J. P., Computer Calculations for Multicomponent Vapor-Liquid Equilibria, Prentice Hall Inc., N. J. (1967).
- [43] Rockafellar, R. T., Convex Analysis, Princeton University Press (1970).
- [44] Shapiro, N. Z., "Conditions for a Homogeneous Mixture to be Ideal," The RAND Corp., RM-3677-PR (1963).
- [45] _____, "On the Behavior of a Chemical Equilibrium System When its Free Energy Parameters Are Changed," The RAND Corp., RM-4128-PR (1964).
- [46] _____, "A Generalized Technique for Eliminating Species in Complex Chemical Equilibrium Calculations," The RAND Corp., RM-4205-PR (1964).
- [47] _____ and Shapley, L. S., "Mass Action Laws and the Gibbs Free-Energy Function," J. SIAM 13 (1965), 353-375.

- [48] Shapley, M. and Cutler, L., "RAND's Chemical Composition Program-A Manual," The RAND Corp., R-495-PR (1970).
- [49] Theil, H., "Substitution Effects in Geometric Programming," Man. Sci. 19 (1972), 25-33.
- [50] Tisza, L., Generalized Thermodynamics, The MIT Press (1966).
- [51] Topkis, D. M., "Cutting Plane Methods Without Nested Constraint Sets," O. R. Center, University of California, Berkeley, ORC 69-14 (1969).
- [52] _____, "A Note on Cutting Plane Methods Without Nested Constraint Sets," O. R. Center, University of California, Berkeley, ORC 69-36 (1969).
- [53] Van Zeggeren, F. and Storey, S. H., The Computation of Chemical Equilibria, Cambridge University Press (1970).
- [54] White, W. B., Johnson, S. M. and Dantzig, G. B., "Chemical Equilibrium in Complex Mixtures," J. Chem. Phys. 28 (1958), 751 ff.
- [55] Zangwill, W. I., Nonlinear Programming: A Unified Approach, Prentice Hall, N. J., (1969).