Room 14-0551
77 Massachusetts Avenue
Cambridge, MA 02139
Ph: 617.253.5668 Fax: 617.253.1690
Email: docs@mit.edu
http://libraries.mit.edu/docs

MATHEMATICAL PROGRAMMING

and

ELECTRICAL NETWORKS

by

JACK BONNELL DENNIS

S.B. and S.M., Massachusetts Institute of Technology
(1954)

SUBMITTED IN PARTIAL FULFILLMENT OF THE

REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF SCIENCE

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
August, 1958

Signature of Author ____________________
Department of Electrical Engineering

Certified by ____________________
Thesis Supervisor

Accepted by ____________________
Chairman, Departmental Committee on Graduate Students

July 1, 1958

Mr. Jack B. Dennis
Room 6-220
M.I.T.

Dear Mr. Dennis:

This letter is to give you permission to print additional copies of your thesis by the Multilith process, and to submit to the Department of Electrical Engineering copies thus produced, in lieu of the typed copies normally required.

A copy of this letter is to be reproduced by the same process and is to be placed in each copy of the thesis immediately following its title page.

Sincerely yours,

S. H. Caldwell
for the
Department Graduate Committee

SHC:mp
cc: Professor Arden

MATHEMATICAL PROGRAMMING
and
ELECTRICAL NETWORKS

by

JACK BONNELL DENNIS

Submitted to the Department of Electrical Engineering on September 5, 1958, in partial fulfillment of the requirements for the degree of Doctor of Science.

ABSTRACT

A novel approach to mathematical programming is presented herewith which is based on an analogy with electrical networks. Mathematical programming concerns the minimization of a function of variables while satisfying inequality constraints on other functions of the variables. It is shown that any direct current electrical network made up of voltage sources diodes, resistors and ideal d-c transformers is equivalent to a pair of dual quadratic programs -- a set of voltages and currents forming a solution of the network provides automatically the optimal vectors of the pair of programming problems.

With this analogy as a foundation, several new developments in theory and method are presented: i) The duality principle of linear programming is extended to quadratic programming and concave programming (with linear constraints). ii) A simple algorithm for solving networks constructed of diodes and voltage and current sources is given. This versatile method is applicable to the well-known transportation problem as well as capacity-limited network flow problems. iii) The terminal characteristic (breakpoint curve) of a source-diode-resistor-transformer network is derived and an algebraic procedure for "tracing" the breakpoint curve is developed. iv) Electrical models for general linear and quadratic programs are given. The breakpoint tracing method is applied to these models yielding several algorithms for obtaining optimal vectors to programming problems. Some of these are new interpretations of familiar methods -- others are new. An adaptation of the method of steepest descent is proposed as an efficient means of solving general non-linear programming problems. Specifically it is demonstrated that the direction of steepest descent for a non-linear program may be determined by solving a quadratic program. A final chapter traces the history of the ideas developed in the thesis and points out possible extensions of the work.

Thesis Supervisor: Dean N. Arden

Title: Assistant Professor of Electrical Engineering

A C K N O W L E D G E M E N T

The author wishes to note with appreciation the contributions made to the creation of this thesis from many sources.

Special acknowledgement goes to Associate Professor Samuel Mason who was first at M.I.T. to observe the analogy between the network flow problem and diode-source electrical networks. Since it was this discovery which was the starting point and inspiration for the developments in the thesis, it may truly be said that Professor Mason is largely responsible for the success of the entire thesis project. To Assistant Professor Dean Arden, who supervised the work and has given gladly of his time for enthusiastic discussion and encouragement, I want to express my utmost appreciation. A special note of thanks goes to former Associate Professor William K. Linvill who introduced the author to the area of linear programming and also to the author's coworker, Dr. William S. Jewell, who gave helpful suggestions and consultation.

This thesis could not have been undertaken without the financial support provided by the Union Carbide Corporation and the U.S. Army Office of Ordnance Research through the M.I.T. Operations Research Project. The generous use of facilities and freedom of action granted by Project Supervisor Professor Philip M. Morse and Assistant Director Dr. Herbert P. Galliher have been most helpful in stimulating the research. Finally, a word of thanks to Mrs. Beth Sutherland and Miss Betty Cullen whose careful work shaped the thesis into its final form.

TABLE OF CONTENTS

		Page
CHAPTER ONE - Preface		
1 - 1	Introduction	1 - 1
1 - 2	Synopsis of Chapters	1 - 3
1 - 3	Notation	1 - 4
CHAPTER TWO — The Nature of Programming Problems		
2 - 1	The General Programming Problem	2 - 1
2 - 2	Linear and Quadratic Programs	2 - 4
2 - 3	Network Problems	2 - 5
2 - 4	Local and Global Minima — Concavity and Convexity	2 - 7
2 - 5	Lagrange Multipliers	2 - 10
2 - 6	The Generalized Lagrangian Problem	2 - 11
2 - 7	Equivalent Formulations — Duality	2 - 18
2 - 8	The Legendre Transformation	2 - 18
2 - 9	Duality in Constrained Minimization	2 - 23
2 - 10	Duality in Mathematical Programming	2 - 25
CHAPTER THREE - Properties and Equivalent Formulations for Networks Containing Diodes		
3 - 1	Graphs and Incidence Matrices	3 - 1
3 - 2	Electrical Networks - Laws and Devices	3 - 3
3 - 3	Equivalent Problems for Diode-Source-Resistor Networks	3 - 7
3 - 4	Reduced Networks	3 - 10
3 - 5	Electrical Networks Containing Transformers	3 - 11
3 - 6	Non-Linear Devices	3 - 14

3 - 7 Terminal Solutions and Breakpoint Curves 3 - 16

CHAPTER FOUR - An Algorithm for Solving Diode Networks

4 - 1 Existence Condition and Reduced Networks 4 - 1

4 - 2 First Algorithm for Diode Source Networks 4 - 8

4 - 3 Second Algorithm for Diode Source Networks 4 - 23

4 - 4 Altering Solution for New Parameter Values 4 - 27

4 - 5 An Application to Network Flow Problems 4 - 29

CHAPTER FIVE - A Breakpoint Tracing Procedure

5 - 1 The Electrical Model of the Terminal Pair System 5 - 2

5 - 2 Basic Solutions 5 - 5

5 - 3 Unit Solutions - Superposition 5 - 7

5 - 4 Computing the New Basis Inverse 5 - 12

5 - 5 Tracing the Breakpoint Curve 5 - 14

5 - 6 The Method in Case of Degeneracy 5 - 14

5 - 7 The Method When Q is Null 5 - 16

5 - 8 The Case When Q and b or c Are Null 5 - 19

CHAPTER SIX - Breakpoint Tracing Methods for General Linear and Quadratic Programs

6 - 1 Electrical Models for General Linear and Quadratic Programs 6 - 1

6 - 2 A Valve Algorithm for Solving the Current Reduced Network 6 - 5

6 - 3 A By-Pass Algorithm for Solving the Current Reduced Network 6 - 8

6 – 4 Solving the Voltage Reduced Network 6 – 11

6 – 5 A Valve Algorithm for Linear Programming 6 – 12

6 – 6 A By-Pass Algorithm for Linear and Quadratic Programming 6 – 14

CHAPTER SEVEN – An Application of Quadratic Programming to the General Programming Problem

7 – 1 Gradient Methods of Minimization 7 – 1

7 – 2 Minimization in the Presence of Equality Constraints 7 – 4

7 – 3 The Directions of Steepest Descent with Inequality Constraints 7 – 6

7 – 4 Equivalence with a Quadratic Program 7 – 8

CHAPTER EIGHT – Historical Notes and Discussion

8 – 1 On the Theoretical Developments 8 – 2

8 – 2 Concerning the Methodology 8 – 4

8 – 3 On the General Programming Problem 8 – 8

8 – 4 Conclusion 8 – 10

APPENDIX A – Geometrical Elements in Euclidian Space

APPENDIX B – A Fundamental Theorem on Linear Inequalities

APPENDIX C – The Theory of Non-Linear Programming

APPENDIX D – The Fundamental Theorems of Quadratic Programming

APPENDIX E – Some Properties of Concave Functions

APPENDIX F – Duality Relations in Concave Programming

APPENDIX G – Properties of Terminal Pair Solutions

APPENDIX H – The Generalized Breakpoint Stepping Method

LIST OF ILLUSTRATIONS

		Page
2 - 1	A network flow problem	2 - 6
2 - 2	Illustrating a concave function of a single variable	2 - 9
2 - 3	The convex set defined by a convex function	2 - 9
2 - 4	Illustrating the Lagrangian problem for unconstrained minimization	2 - 12
2 - 5	Illustrating the Lagrangian problem with constraint functions	2 - 15
2 - 6	Geometric interpretation of the Lagrangian problem	2 - 19
2 - 7	The Legendre transformation	2 - 20
3 - 1	A graph and its incidence matrix	3 - 2
3 - 2	Branch polarity convention	3 - 2
3 - 3	Relations imposed by electrical branch types	3 - 6
3 - 4	Non-linear resistive device	3 - 15
3 - 5	Network with observable terminal pair	3 - 15
3 - 6	Example of a breakpoint curve	3 - 19
3 - 7	Network with diodes separated	3 - 19
3 - 8	Demonstrating the continuity of the set of terminal solutions	3 - 21
4 - 1	Network violating the first non-redundancy assumption	4 - 3
4 - 2	Network violating the second non-redundancy assumption	4 - 3
4 - 3	Network violating the first existence condition	4 - 5
4 - 4	Network violating the second existence condition	4 - 5

4 - 5	Diode-source-resistor network and its associated reduced networks	4 - 7
4 - 6	Diode-source network with tree and potential distribution for starting phase I	4 - 10
4 - 7	The situation if node M is labelled	4 - 13
4 - 8	The situation if node M is not labelled	4 - 13
4 - 9	Solution of voltage reduced network and current distribution for starting phase II	4 - 17
4 - 10	Breakpoint curve of the solved subnetwork viewed from the selected diode	4 - 19
4 - 11	The situation if node M is labelled	4 - 21
4 - 12	The situation if node M is not labelled	4 - 21
4 - 13	Linear flow branch and its electrical analogue	4 - 30
4 - 14	Quadratic flow branch and its electrical analogue	4 - 30
5 - 1	Electrical model of the terminal pair system	5 - 3
5 - 2	Simplified drawing of terminal pair system model	5 - 4
5 - 3	Black box representation of the terminal pair system	5 - 9
5 - 4	Unit incremental solutions of a terminal pair system	5 - 9
5 - 5	A breakpoint and the superposition of unit solutions	5 - 13
5 - 6	Possible breakpoint curves for a primal reduced terminal pair system	5 - 21
5 - 7	Possible breakpoint curves for a completely degenerate terminal pair system	5 - 21
6 - 1	Electrical model of the dual pair of quadratic programs	6 - 2
6 - 2	A more general quadratic programming model	6 - 4

6 - 3	Second linear programming model	6 - 4
6 - 4	Valve method for solving the current reduced network	6 - 7
6 - 5	By-pass method for solving the current reduced network	6 - 9
6 - 6	Valve method for completing the solution of the linear programming model	6 - 13
6 - 7	By-pass method for linear and quadratic programming	6 - 15
7 - 1	The method of steepest descent	7 - 2
7 - 2	Flow diagram for the method of steepest descent	7 - 2
7 - 3	The method of steepest descent with equality constraints	7 - 7
7 - 4	Illustrating the direction of steepest descent for three variables	7 - 9
7 - 5	The method of steepest descent with inequality constraints	7 - 12

C H A P T E R O N E

Preface

1 - 1 Introduction

This thesis is based on the idea that by viewing a subject from a new direction one will gain insight into the subject and quite possibly discover new concepts and broader principles which are not readily apparent from the more heavily trodden paths of investigation. Here, the subject is mathematical programming--the new viewpoint is based on the analogy of these problems with the behavior of electrical networks.

The discipline of mathematical programming, although only a decade and a half old, has been subject to intensive study by many authors. Mathematical programming concerns the determination of values for a set of variables such that an objective function of the variables is minimized (or maximized) while a number of given constraint relations involving the variables are satisfied. Programming problems differ from minimization problems with constraint equations in that the constraint relations may also be inequalities.

Until the present, essentially all of the interest and research effort in mathematical programming has stemmed from applications to industrial operations and economic systems--the field of "operations research". Initially study was confined to linear programming in which the objective and constraint functions are assumed to be linear in the variables, and an extensive mathematical theory was developed. Also several algorithms for computing optimal solutions to linear programs were formulated. Later, useful results were presented relating to general programming problems. Currently the theory and methodology of quadratic programming are receiving much attention.

Quite recently it has been observed that certain simple linear programming problems could be solved by setting up an equivalent electrical network containing voltage sources, current sources, and diodes. If the voltage and current sources are set to values corresponding to the data of the programming problem, the optimal solution may be found by simply measuring currents and voltages at appropriate points in the network--the electrical network automatically establishes currents and voltages which form the optimal solution of the programming problem. This observation is the starting point of the new viewpoint taken in this presentation.

The two approaches have the following important distinction--the operations analyst nearly always views a programming problem as an <u>extremum</u> problem, that is, he thinks in terms of minimization or maximization. On the other hand, the electrical scientist in solving a network thinks in terms of finding a set of quantities which satisfy the individual devices of the network and satisfy the conditions imposed by their interconnection--he does not usually think in terms of minimization. This difference in approach is evident in the algorithms that have been developed for solving programming problems. The original and most widely known method for linear programming, the simplex method, is closely tied to the concept of minimization. In the new methods proposed by the author in this thesis the concept of minimization does not enter. The idea here is that of starting with a set of quantities which satisfy the interconnection requirements of a network and successively modifying this partial solution until all requirements of the devices of the network have been satisfied.

1 - 2 Synopsis of Chapters

In the body of this thesis an attempt has been made to present the pertinent concepts in as straightforward manner as possible. The proofs of well-known results and the more complicated original developments have been relegated to appendices.

In Chapter Two the theory of general mathematical programming is developed, leading to the formulation of a general Lagrangian problem. The concept of duality is introduced for ordinary minimization problems and is extended to yield duality properties for quadratic programs. From this the central theorems of linear and quadratic programming are derived. The extension of the duality principle to a class of concave programs is indicat

A sketch of the important concepts concerning electrical networks is given in Chapter Three. Then it is shown that the solution of any direct current network made up of voltage and current sources, ideal d-c transfor and ideal diodes is equivalent to a pair of dual linear programs. Also, i is shown that a network made up of these elements plus linear resistors is equivalent to a pair of dual quadratic programs. The concept of a breakpoint curve, the volt-ampere characteristic observed at a terminal pair of a network containing diodes, is introduced. In Chapter Four a simple new algorithm for solving d-c networks containing only sources and diodes is described in electrical terms. This algorithm has the advantage that modifications of the solution caused by changes in parameter values are easily computed from a previous solution. In view of the equivalence demonstrated in Chapter Three, this algorithm may be used to solve certain simple linear programs--the well-known transportation problem, for instance. In Chapter Five an algebraic method for "tracing" the breakpoint curve observed at the terminal pair of a network of sources, transformers, resistors and diodes is given. That electric network models may be constructed for any linear or quadratic program is shown in Chapter Six. The breakpoint tracing algorithm is then applied to these models yielding several algorithms for solving general linear and quadratic programs. Some of these algorithms are simply different ways of interpreting well-known methods. Others are new contributions. Chapter Seven treats a proposal for computing an optimal solution of a general programming problem. Specifically it is shown that determining the _direction_ of steepest descent for a general programming problem is equivalent to solving a quadratic program. Finally an attempt is made in Chapter Eight to trace the development of the ideas presented in the thesis and to give due credit to those persons whose work is most responsible for inspiring the contributions reported here.

1 - 3 Notation

The following conventions apply throughout the thesis except in Chapter Four where matrix notation is not employed.

Lower case letters x, y, ... will represent vector quantities and will be uniformly treated as column matrices. A vector inequality

$$x \geq y \quad \text{or} \quad x > y$$

always means that the indicated inequality relationship must hold between each pair of corresponding components of x and y. Rectangular matrices wil be represented by capital letters. The individual components of a vector c the columns or elements of a rectangular matrix will be distinguished by numerical subscripts

$$A = \begin{bmatrix} A_1 \; A_2 \; \cdots \; A_n \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & & \\ - & - & - & - \\ a_{m1} & \cdots & & a_{mn} \end{bmatrix} .$$

The transpose of a matrix will be indicated by a superscript T. Hence the dot product of two vectors

$$x \cdot y = \sum_i x_i \, y_i$$

will be written $x^T y$. Sometimes the notation x, y will also be used. Occasionally it will be necessary to identify the submatrices of a partitioned matrix or vector. This will be done with upper case subscripts, for instance

$$i = \begin{bmatrix} i_C \\ i_V \\ i_D \end{bmatrix} .$$

Scalar quantities will be denoted by lower case Greek letters.

The letter designating a function will indicate whether it is a scalar function or a vector function. Thus $\phi(\eta)$ is a scalar function of a scalar variable, $f(\xi)$ is a vector function of a scalar, $\phi(x)$ is a scalar function of a vector and $f(x)$ maps vectors into vectors.

The gradient of a scalar function of a vector $\phi(x)$ will be represented by

$$\phi(x) \equiv \begin{bmatrix} \frac{\partial \phi}{\partial x_1} \\ \frac{\partial \phi}{\partial x_2} \\ \cdot \\ \frac{\partial \phi}{\partial x_n} \end{bmatrix}$$

The differential of a map of vectors into vectors $f(x)$ will be denoted by

$$df \equiv \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdot\ \cdot & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & & \\ - & - & - & - \\ \frac{\partial f_m}{\partial x_1} & \cdot & \cdot\ \cdot & \frac{\partial f_m}{\partial x_n} \end{bmatrix} .$$

In Chapter Two and in the proofs given in the appendices geometrical interpretations will be made of the mathematical reasoning. As an introduction to the concepts of geometry in many dimensional space Appendix A defines some of the geometrical terms used here and in other literature on mathematical programming.

CHAPTER TWO

The Nature of Programming Problems

2 - 1 The General Programming Problem

The most general type of problem that will be considered here and in the rest of the thesis is the following: to select values of a number of variables so that an objective function is minimized among all choices of values that satisfy a group of inequality constraints on functions of the variables. Mathematically the problem may be written as

General Programming Problem:

Minimize the objective function

$$\phi(x) \qquad (2 - 1a)$$

subject to the constraints

$$g(x) \geq 0$$
$$x \geq 0 . \qquad (2 - 1b)$$

In this statement, x is a column vector (matrix) of n components and g(x) represents the column vector of m functions of x

$$g(x) = \begin{bmatrix} g_1(x) \\ g_2(x) \\ - \\ \vdots \\ - \\ g_m(x) \end{bmatrix} .$$

To introduce some terminology which will be used throughout the thesis, any vector x which satisfies the constraints (2 - 1b) is called a feasible vector of the programming problem. The set of all feasible vectors is termed the constraint set of the problem. A feasible vector which produces as low a value of the objective function as any other feasible vector is known as an optimal vector. The constraint relations (2 - b) of a programming problem may take one of the following alternate forms.

B) $g(x) = 0$ C) $g(x) \geq 0$

$x \geq 0$ x unrestricted

Simple considerations show that a problem expressed in any one of the three forms can be expressed in the other two forms. The principles are: 1) an equality constraint relation can be replaced by a pair of inequalities, and 2) an unrestricted variable can be replaced by the difference of two non-negative variables. The precise relations between the three forms are given in Table 2 - 1.

The usual analytic methods of solving minimization problems in the presence of constraint equations--solving the constraint equations, substituting in the objective function and differentiating, or the Lagrange multiplier technique--do not work with programming problems. These methods do not take account of the inequality constraints which characterize mathematical programming problems. It does not seem likely that the solution of a programming problem will never be expressed in closed analytic form. However, the Lagrange multipliers, suitably generalized, play an important role in the theoretical and computational aspects of programming. Moreover, the generalized Lagrange multipliers have great physical significance as will be evident in later chapters.

Old	New	Formulation of New Problem in Terms of Old		
		$\bar{g}(\bar{x})$	$\bar{\phi}(\bar{x})$	$\bar{x}$
A	B	$g(x) + Iw$	$\phi(x)$	$\begin{bmatrix} x \\ w \end{bmatrix}$
B	A	$\begin{bmatrix} g(x) \\ -g(x) \end{bmatrix}$	$\phi(x)$	x
A	C	$\begin{bmatrix} g(x) \\ Ix \end{bmatrix}$	$\phi(x)$	x
C	A	$g(x^+ - x^-)$	$\phi(x^+ - x^-)$	$\begin{bmatrix} x^+ \\ x^- \end{bmatrix}$ $x = x^+ + x^-$

I = identity matrix of appropriate order

Table 2 - 1—Relations between problem fornulations.

2 - 2 Linear and Quadratic Programs

A very important special case of the general programming problem is the linear programming problem in which $\phi(x)$ is simply a linear combination of the variables and $g(x)$ is a linear transformation.

Linear programming problem:

Minimize

$$\phi(x) = c^T x$$

with

$$A x \geq b$$

$$x \geq 0$$

Here the constant vector b of the linear transformation has been moved to the right-hand side of the constraint relations.

A somewhat more general case and undoubtably the simplest form of non-linear program is the quadratic program in which the objective function is a second degree form in the variables.

Quadratic programming problem:

Minimize

$$\frac{1}{2} x^T Q x + c^T x$$

with

$$A x \geq b$$

$$x \geq 0 .$$

These two special cases form the major subject matter of this thesis although a technique for solving general programming problems is proposed in the last chapter.

One reason for the importance of the two special cases is that there are computational schemes which produce a solution in a small, finite number of iterative steps. The last iteration produces an exact solution of the linear or quadratic program. On the other hand, methods for solving general programming problems approach closer to the solution with each iteration, but the exact solution is never attained in a finite number of steps although convergence may be very rapid.

2 - 3 Network Problems

One important and interesting class of programming problems may be stated in terms of flow in a network. An example of such a problem is the following. One is concerned with a network (which may be visualized as a transportation system, for instance) of the form of Figure 2 - 1. With a branch directed from node i to node j a flow variable X_{ij} is associated which is not allowed to be negative. Each branch has a given maximum capacity for flow M_{ij} and there is a cost C_{ij} for each unit of flow in the branch. At each node j of the network there is a stated demand D_j for flow which may be positive, negative or zero. It is desired to find a flow distribution which conserves flow at the nodes of the network, satisfies the restrictions imposed by each branch and has the smallest total cost. In symbols this may be written

Network flow problem:

Minimize

$$\sum_{i,j} C_{ij} X_{ij}$$

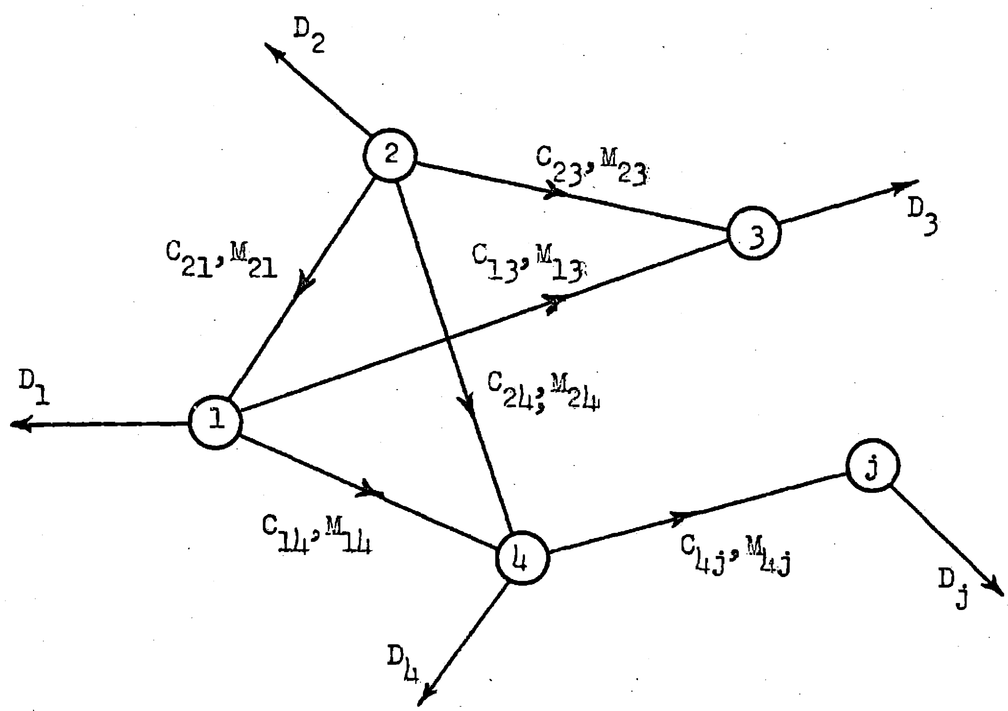


Figure 2 - 1--A network flow problem.

with

$$\sum_i X_{ij} - \sum_i X_{ji} = D_j$$

and

$$0 \leq X_{ij} \leq M_{ij} .$$

It is evident that this is a linear programming problem. If a quadratic component of cost were associated with each branch, this would become a problem in quadratic programming. Problems of this form have an elegant analogue in electrical networks as will be seen in Chapter 4. The network flow problems are important because the coefficients of the variables in the constraint relation are all either plus one, minus one or zero. This feature makes their solution much easier than is the case with general linear or quadratic programs.

2 - 4 Local and Global Minima--Concavity and Convexity

An important question in mathematical programming is when is a local minimum equivalent to a global minimum? To understand these terms, consider the programming problem as being phrased in an n-dimensional space with one coordinate axis for each variable. Then the constraint set of a programming problem consists of a certain group of points in this space. A point yields a *local* or *relative minimum* of the objective function if it is in the constraint set and no "nearby point" in the constraint set gives a lower value of the objective function. A point in the constraint set yields a *global minimum* if no other point in the constraint set gives a lower value of the objective.

Of course, someone looking for the solution of a particular programming problem wants to find a global minimum. Unfortunately the common methods for obtaining solutions are designed to find local minima. Hence it is desirable to know in which cases the two are equivalent. The ideas of concave functions and convex sets yields a simple answer to this question.

Let a function of n variables be represented by the $n + 1^{th}$ coordinate of points on a surface in an $n + 1$ - dimensional space in which the $n + 1^{th}$ axis is "vertical". The function is <u>concave</u> if the surface representing it is curved downward at no point. More precisely, the straight line segment joining any two points of this surface must be beneath the surface at no point. This is illustrated in Figure 2 - 2. A function will be termed strictly concave if the line segments mentioned above lie above the surface except for their end points. A function is <u>convex</u> if its negative is concave. Note that according to these conventions, the linear function $\phi(x) = c^T x$, which is represented by a hyperplane in $n + 1$ - dimensional space, is both concave and convex, but it is neither strictly concave nor strictly convex.

A <u>convex set</u> is defined as a set of points in n-dimensional space which contains all points on the line segment joining any two points in the set. The set of points in n-space for which $\Theta(x) \geq 0$ is a convex set if Θ is a convex function as is indicated in Figure 2 - 3. As an extension of this property, the constraint set of a programming problem is a convex set if the functions g_1 through g_m are convex since this set consists of all points common to the convex sets determined by each of the g_i (Lemma C - 1).

The important result made possible by these concepts in connection with mathematical programming is the following property which is proved as Theorem C - 1.

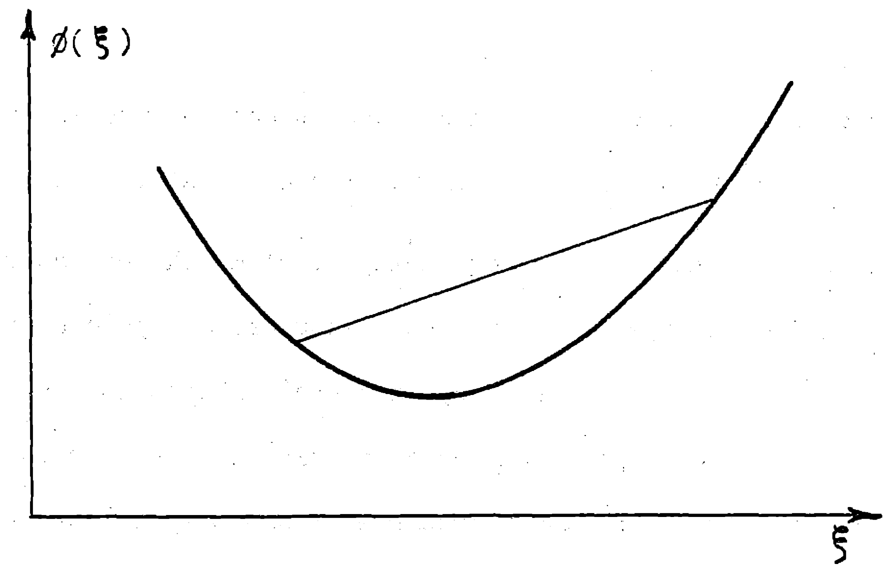


Figure 2 - 2--Illustrating a concave function of a single variable.

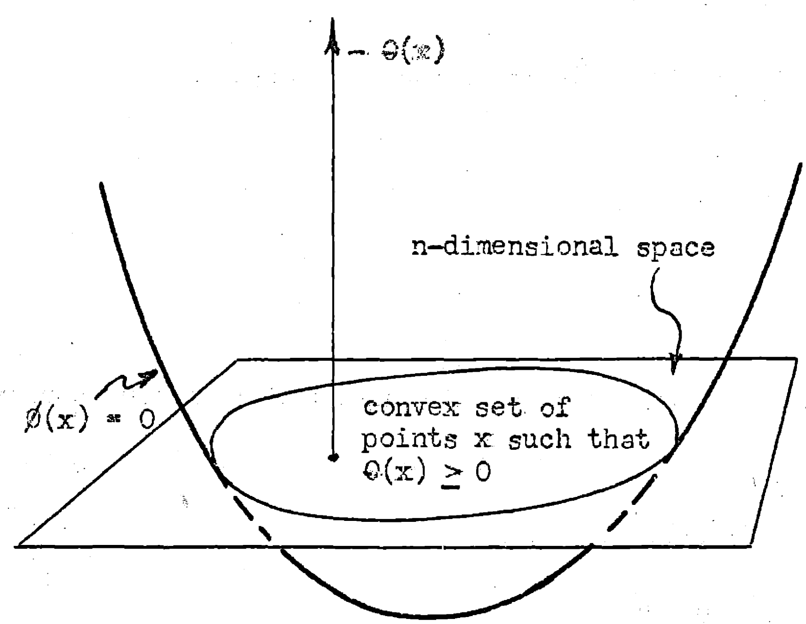


Figure 2 - 3--The convex set defined by a convex function.

For a programming problem with a concave objective function and a convex constraint set, a relative minimum is also a global minimum.

In the case of the linear programming problem, the objective function is always concave and the constraint functions are always convex. It follows that a relative minimum is always a global minimum in linear programming. With quadratic programming the constraint set is likewise convex. However, the objective function is concave if and only if the matrix Q is <u>positive semi-definite</u>, that is, if $x^T Q x$ is non-negative for all choices of x (Lemma D - 1). If this is so, then relative and global minima are equivalent for quadratic programming.

2 - 5 <u>Lagrange Multipliers</u>

Consider the constrained minimization problem

<u>Minimization Problem</u>:

Minimize $\phi(x)$

with

$$g(x) = 0 \qquad (2 - 2b)$$

Following the Lagrange multiplier rule one multiplies each constraint function g_i by a multiplier $- y_i$ and adds the results to the objective function to obtain the <u>Lagrangian function</u>

$$\psi(x, y) = \phi(x) - y^T g(x) .$$

Differentiating this expression and coupling the result with the constraints (2 - 2b) one obtains the

Lagrangian problem:

Find an x and y such that

$$g(x) = 0$$

$$\partial\phi(x) - y^T \, dg(x) = 0$$

In this statement $\partial\phi$ is the gradient of the objective function and dg is the differential of the transformation g. The Lagrange multiplier rule states:

i) If $\bar{\bar{x}}$ yields a local minimum of ϕ, then there is a vector $\bar{\bar{y}}$ such that $(\bar{\bar{x}}, \bar{\bar{y}})$ is a solution of the Lagrangian problem.

ii) If $(\bar{\bar{x}}, \bar{\bar{y}})$ is a solution of the Lagrangian problem, and $\phi(x)$ is concave and the constraint set convex in the vicinity of $\bar{\bar{x}}$, then $\bar{\bar{x}}$ yields a relative minimum of ϕ.

Thus the Lagrange multiplier rule gives a second equivalent formulation of any ordinary minimization problem.

2 - 6 The Generalized Lagrangian Problem

In this paragraph the manner in which the Lagrange multiplier rule generalizes to mathematical programming will be studied. A rigorous proof of the result involves application of the theory of inequalities and is given in Appendix C. However, the reasonableness of the result can be made clear through consideration of some simple cases.

Consider first the simple minimization problem

Minimize $\phi(\xi)$

with $\xi \geq 0$.

As is indicated in Figure 2 - 4 two cases are possible. Either the minimum

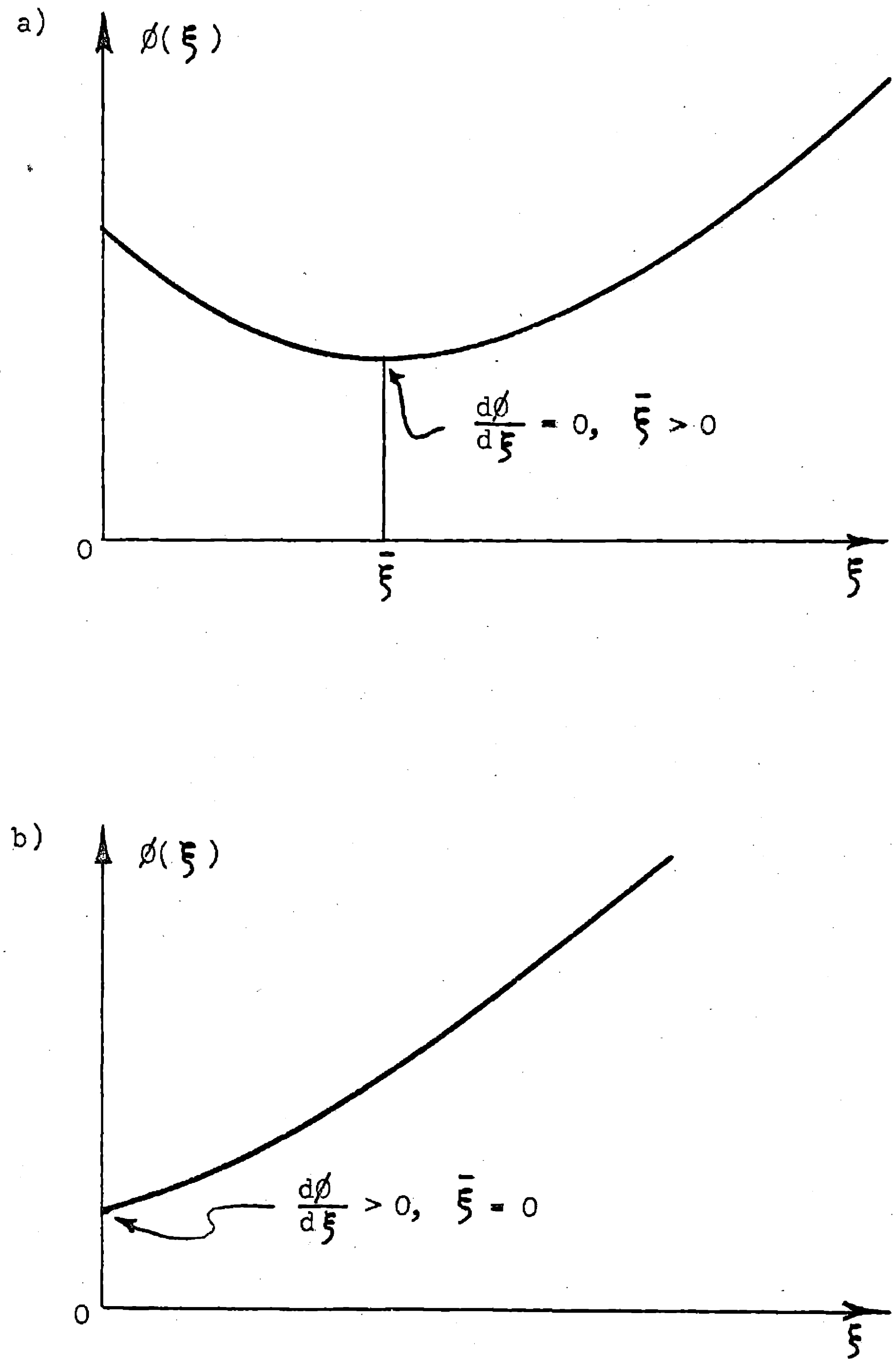


Figure 2 - 4—Illustrating the Lagrangian problem for unconstrained minimization.

occurs for some $\bar{\xi} \geq 0$ at which point $\frac{d\phi}{d\xi}(\bar{\xi}) = 0$, or the minimum is at the point $\bar{\xi} = 0$. In the latter case $\frac{d\phi}{d\xi}(0)$ must not be negative. Summing up, a necessary condition that $\bar{\xi}$ yield a minimum of $\phi(\xi)$ subject to $\xi \geq 0$ is that

$$\frac{d\phi}{d\xi}(\bar{\xi}) - \bar{\omega} = 0$$

$$\bar{\xi} \geq 0 \quad \bar{\omega} \geq 0 \qquad (2 - 3b)$$

$$\bar{\omega}\,\bar{\xi} = 0 . \qquad (2 - 3c)$$

The relation (2 - 3c) requires that either $\bar{\xi}$ be zero or $\bar{\omega}$ be zero and hence that only one of the inequalities (2 - 3b) can be satisfied in the strict inequality sense. If $\bar{\xi}$ is greater than zero, $\bar{\omega}$ must be zero and the first case holds; if $\bar{\omega}$ is greater than zero, $\bar{\xi}$ must be zero and the second case holds. Because of this relation between the original inequality restraint and the new one obtained as a necessary condition for a minimum, the relation (2 - 3c) is known as the <u>complementary slackness condition</u>. This condition is related to the behavior of the diode circuit element in electrical networks as will be seen in the next chapter.

The exact same ideas generalize to the problem of minimizing a function of several variables which are constrained to be non-negative.

Minimize $\phi(x)$

with $x \geq 0$

The corresponding Lagrangian problem is to find vectors x and w such that

$$\phi(x) - w = 0$$

$$x \geq 0, \quad w \geq 0$$

$$x^T w = 0 .$$

This simply repeats the condition of the single variable problem for each component of x in the many variable case. Note that the complementary slackness relation $x^T w = 0$ requires that $x_i \, w_i = 0$ for all i when x and w are constrained to be non-negative.

Next a more complicated case will be considered, namely the problem of minimizing a function of two variables subject to a single equality constraint and the condition that the variables be non-negative.

Minimize $\phi(x) = \phi(x_1, x_2)$

with

$$\theta(x) = \theta(x_1, x_2) = 0$$

$$x_1 \geq 0, \quad x_2 \geq 0$$

If no inequality constraints were present one could find a necessary condition for a minimum by forming the Lagrangian expression

$$\psi\,(x, \eta) = \phi(x) - \eta\, g(x)$$

and differentiating to obtain

$$\partial\phi(x) - \eta\, \partial\theta(x) = 0 \qquad (2 - 4a)$$

$$\theta(x) = 0\,.$$

This condition is illustrated geometrically in Figure 2 - 5a. It requires that at the point $\bar{x}$, the gradient of the objective function must be some multiple η of the gradient of the constraint function.

In the presence of inequality constraints cases in which the minimizing x lies on the boundary of the constraint set must be taken into account. The generalized Lagrangian problem is to find vectors x and w and a scalar η (the multiplier) such that

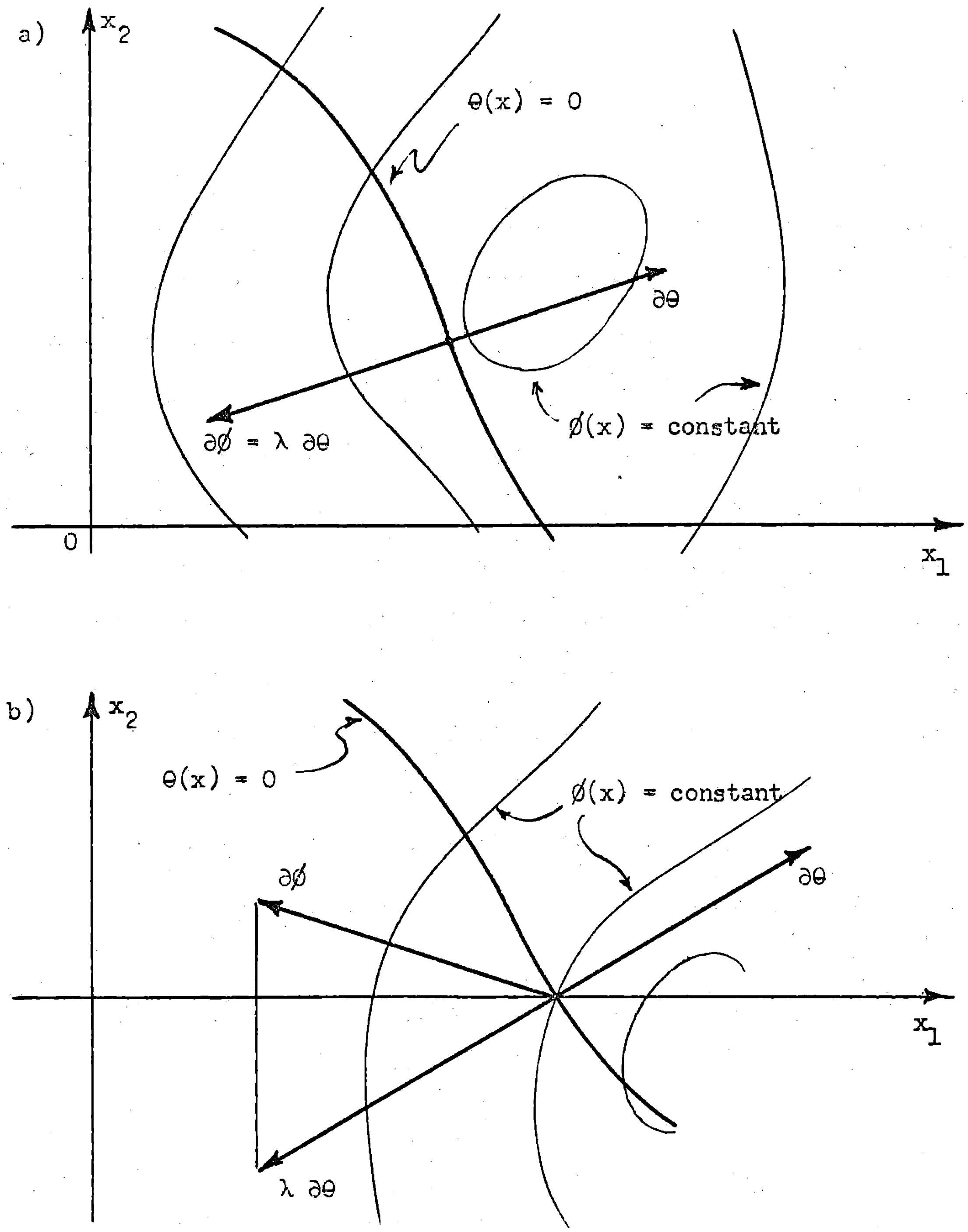


Figure 2 - 5—Illustrating the Lagrangian problem with constraint functions.

$$\eta \; \partial\theta(x) - \partial\phi(x) + v = 0 \qquad (2 - 5a)$$

$$\theta(x) = 0$$

$$x \geq 0, \quad v \geq 0$$

$$v^T x = 0. \qquad (2 - 5d)$$

If the minimizing x is not on the boundary of the constraint set then $x_1 > 0$, $x_2 > 0$ and the complementary slackness condition (2 - 5d) requires $v_1 = v_2 = 0$. Then this problem becomes exactly (2 - 4). On the other hand, consider the situation indicated in Figure 2 - 5b where the constrained minimum has $x_1 > 0$, $x_2 = 0$. According to the complementary slackness condition w_2 may be greater than zero, but w_1 must be equal to zero. Relation (2 - 5a) then states that the x_1 component of $\partial\phi$ must be equal to some multiple of the x_1 component of $\partial\theta$ while the x_2 component of $\partial\phi$ must not be less than the same multiple of the x_2 component of $\partial\theta$. This condition is evidently fulfilled for the case illustrated in the figure. If the condition were not satisfied the situation would be as in Figure 2 - 5c where it is evident that x does not minimize ϕ subject to the constraints. Similar considerations verify the correctness of the Lagrangian problem when the minimum lies on the x_2 axis or at the point $x_1 = x_2 = 0$.

For the case of minimizing a function of many variables subject to many equality constraints

Minimize $\phi(x)$

with $g(x) = 0$

$x \geq 0$

The generalized Lagrangian problem is to find vectors x, y and v such that

$$y^T \, dg(x) - \partial\phi(x) + v = 0 \qquad (2 - 6a)$$

$$g(x) = 0$$

$$x \geq 0, \quad v \geq 0$$

$$v^T x = 0.$$

These relations have the same form as in the single constraint case with one term for each constraint equation in (2 - 6a). The form of the Lagrangian problem associated with the general programming problem in the form (2 - 1) may be found by use of the relations given in Table 2 - 1. The result is

<u>Generalized Lagrangian problem</u>:

Find vectors x, v, u and y such that

$$y^T\, dg(x) - \partial\phi(x) + v = 0 \qquad (2 - 7a)$$

$$g(x) - u = 0$$

$$x \geq 0 \quad v \geq 0 \quad u \geq 0 \quad y \geq 0$$

$$v^T x = 0 \qquad y^T u = 0 \qquad (2 - 7d)$$

The generalized Lagrange multiplier (the fundamental theorem for general programming--Theorem D - 2) rule then states:

i) If x is a locally optimal solution of the general programming problem, then there exist vectors v, u and y such that (x, v, u, y) is a solution of the Lagrangian problem.

ii) If (x, v, u, y) is a solution of the Lagrangian problem and $\phi(x)$ is concave and the constraint set convex in the vicinity of x, then x yields a relative optimal solution of the programming problem.

A geometric interpretation of the Lagrangian problem can be made if (2 - 7a) is rewritten as

$$\phi(x) = dg(x)^T\, y + I\, v \, . \qquad (2 - 8)$$

The rows of the rectangular matrix dg(x) are vectors normal to the constraint surfaces $g_i(x) = 0$ at the point x, and the columns of the identity matrix are vectors normal to the constraint surfaces $x_i = 0$. The relation (2 - 8) requires that the gradient of the objective function be expressed as a positive linear combination of the normals of the constraint surface. The complementary slackness conditions (2 - 7d) allow only normals to equality satisfied constraints to participate in the linear combination. Figure 2 - 6 gives an illustration of this principle.

2 - 7 Equivalent Formulations - Duality

The relation developed above between extremum problems on the one hand and the systems of relations constituting the Lagrangian problems on the other is often observed in the physical systems. Moreover, it frequently happens that two distinctly different extremum formulations lead to the very same Lagrangian problem. One group of physical examples where this occurs is the electrical networks which are studied extensively in the next few chapters. In such cases the two extremum formulations are equivalent in the sense that they both give a complete description of the same physical situation--hence they are said to bear a dual relationship to each other. The Legendre transformation has a dominant role in this duality relation. Therefore a digression will be made to introduce it and its properties.

2 - 8 The Legendre Transformation

In the study of geometry there is a strong dual relationship between points and planes in three dimensional space or between points and hyperplanes in a multidimensional space. The specification of a surface in a multidimensional space is one example of this duality principle. In an n + 1 dimensional space the usual way of describing a surface is to state the value of the $n + 1^{th}$ coordinate of the surface as a function of the first n coordinates, i.e.,

$$\xi = \phi(x) .$$

This is the point description of the surface. There is an alternate means of describing the surface--in terms of its tangent planes. If the surface

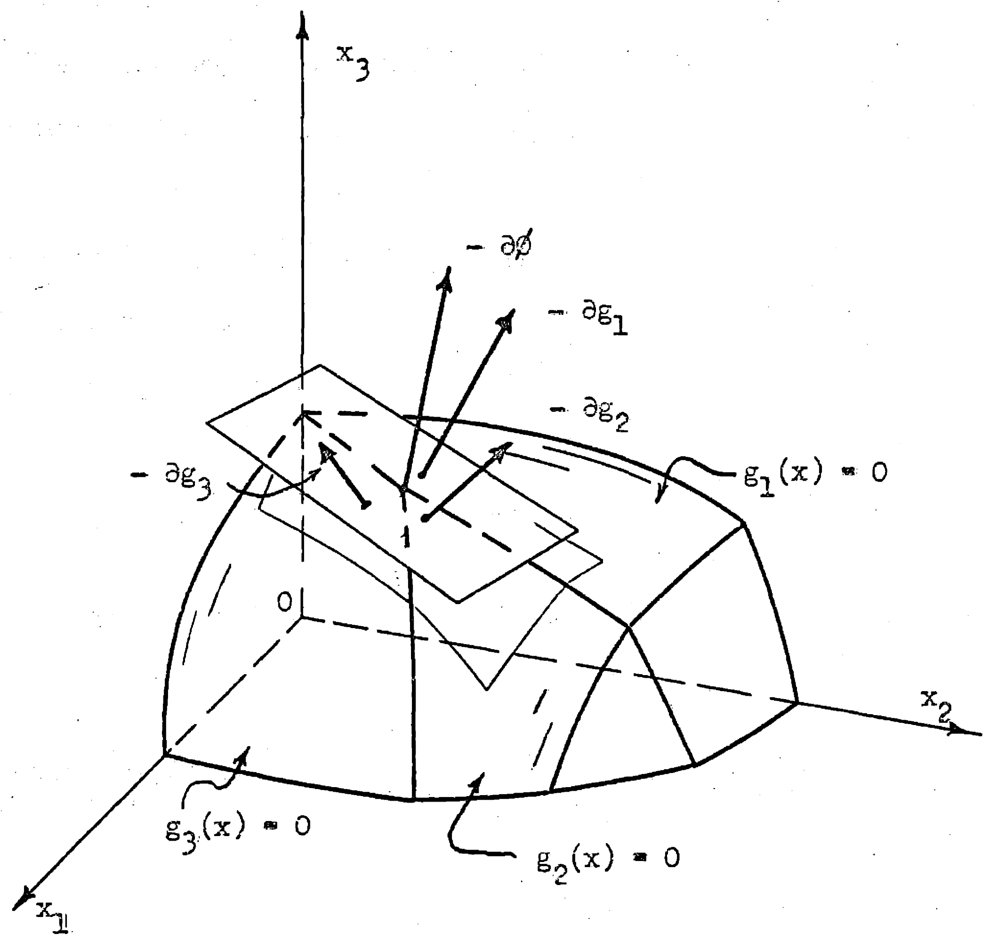


Figure 2 - 6—Geometric interpretation of the Lagrangian problem.

is strictly concave, then no two points on the surface can have tangent planes with identical normal vectors. Thus, the surface may be described by stating for each possible tangent vector the intercept of the corresponding tangent plane with the $n + 1^{th}$ axis. Take the $n + 1^{th}$ component of the tangent vector to be always minus one and represent its first n components by the vector y. Then if η is the value of the intercept along the $n + 1^{th}$ axis, the tangent plane description is given by stating η as a function of y.

$$\eta = \theta(y)$$

These two descriptions of the surface are illustrated by Figure 2 - 7.

Let us consider the relation between these two descriptions. First, recall that the normal vector of the tangent plane at the point $\bar{x}$ is

$$\begin{bmatrix} y \\ -1 \end{bmatrix} = \begin{bmatrix} \phi(\bar{x}) \\ -1 \end{bmatrix} .$$

Hence the equation of the tangent plane to the surface at $\bar{x}$ is

$$(x - \bar{x})^T \ \partial\phi(\bar{x}) - (\eta - \phi(\bar{x})) = 0 . \qquad (2 - 9)$$

The intercept $\bar{\eta}$ of this plane with the $n + 1^{th}$ coordinate axis may be found be setting x = 0 in (2 - 9).

$$\bar{\eta} = \phi(\bar{x}) - \bar{x}^T \ \partial\phi(\bar{x}) \qquad (2 - 10)$$

This expression for $\bar{\eta}$ can be put in terms of y provided that the relation

$$y = \partial\phi(x)$$

can be solved to give x in terms of y.

$$x = \partial\phi^{-1}(y)$$

That this is possible when ϕ is strictly concave is a consequence of the fact that there is a unique tangent plane for each point on the surface (Theorem E - 2). Substituting this relation in (2 - 10), the functional

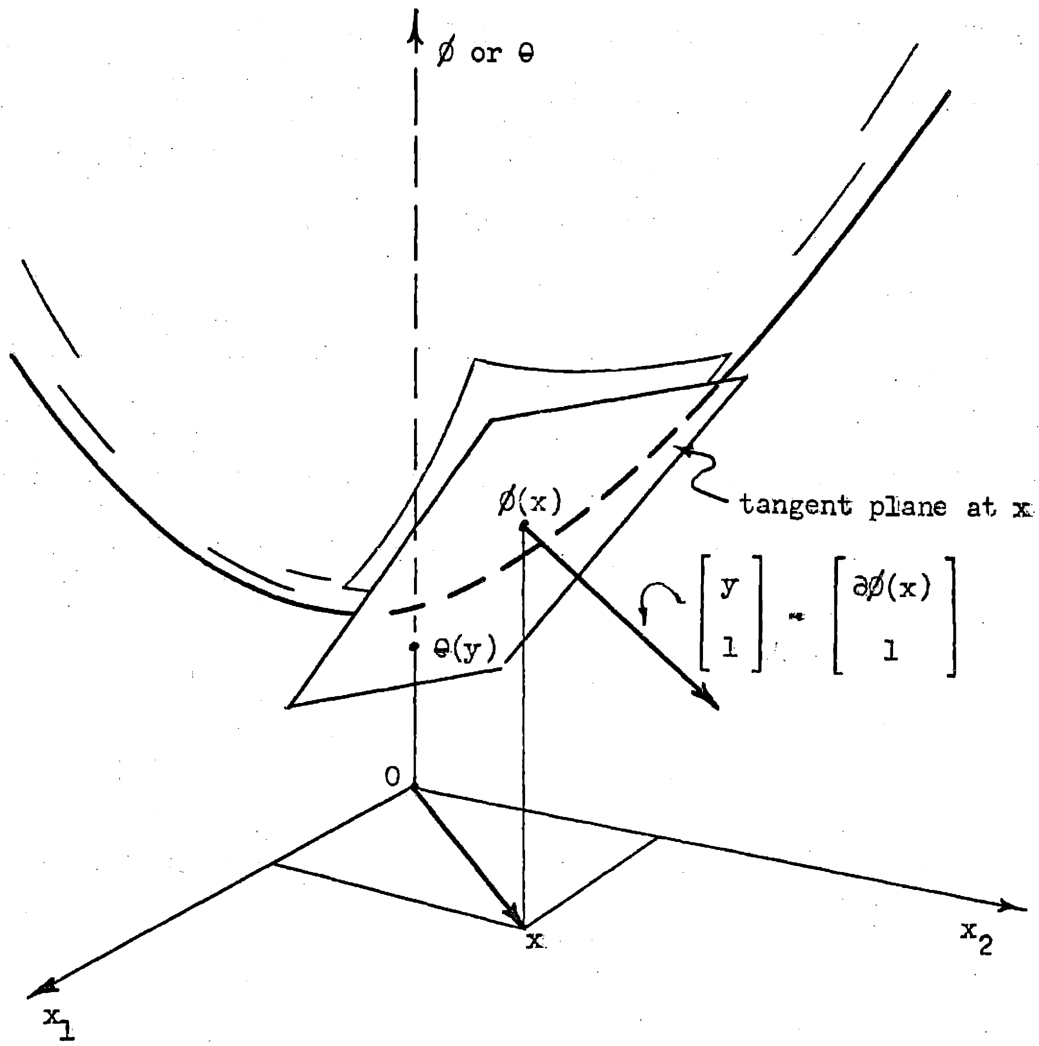


Figure 2 - 7—The Legendre transformation.

relation for the tangent plane description of the surface is obtained in terms of the functional relation for the point description

$$\eta = \Theta(y) = \phi\left[\partial\phi^{-1}(y)\right] - y^T\left[\partial\phi^{-1}(y)\right] \qquad (2 - 11)$$

This function is the <u>Legendre transformation</u> of ϕ and turns out to be strictly convex (Theorems E - 1 and E - 3) if ϕ is strictly concave.

An important property of the Legendre transform is that its gradient is the negative inverse of the gradient of ϕ. Specifically.

$$\partial\Theta(y) = -\partial\phi^{-1}(y)$$

This is easily shown by application of the chain rule for differentiating composite functions (see Appendix E).

As an example of the Legendre transformation consider the case where ϕ is a positive definite quadratic form (and hence a strictly concave function).

$$\phi(x) = \frac{1}{2}x^TP\,x$$

Then the gradient of ϕ is

$$\partial\phi(x) = P\,x.$$

Since P is positive definite, its inverse exists and is also positive definite. Therefore

$$\partial\phi^{-1}(y) = P^{-1}y.$$

The Legendre transform may now be evaluated using (2 - 11).

$$\Theta(y) = \frac{1}{2}(P^{-1}y)^TP\,P^{-1}y - \left[P^{-1}y, y\right]$$

$$= -\frac{1}{2}y^T\,P^{-1}y$$

The transform $\Theta(y)$ is obviously a strictly convex function. Its gradient is

$$\partial\Theta(y) = -P^{-1}y$$

in agreement with the statement made above.

2 - 9 Duality in Equality Constrained Minimization

In order to introduce the discussion of duality relationships in programming problems, a simple problem will be used to illustrate duality in an ordinary equality constrained minimization problem. Consider the problem of minimizing a positive definite quadratic form subject to a system of linear constraints.

Quadratic minimum problem:

Minimize

$$\frac{1}{2} x^T P x + c^T x$$

with

$$A x = b$$

Applying the Lagrange multiplier rule, the following equivalent problem is obtained.

Quadratic Lagrangian problem:

Find vectors x and z such that

$$A^T z - P x = c$$

$$A x = b$$

Now, as was shown above the Legendre transform of $\phi(x) = \frac{1}{2} x^T P x$ is $\theta(y) = -\frac{1}{2} y^T P^{-1} y$. Consider a second extremum problem concerning the maximization of the transform.

Quadratic maximum problem:

Maximize

$$-\frac{1}{2} y^T P^{-1} y - b^T z$$

with

$$A^T z - y = c$$

The corresponding result of use of the multiplier rule is the following.

Quadratic Lagrangian problem:

Find an x, y and z such that

$$A^T z - y = c$$

$$x = P^{-1} y$$

$$A x = b .$$

The two Lagrangian problems are identical except for the introduction of the extra variables y in the latter case. This indicates that the two extremum formulations are simply two different but equivalent ways of stating the same problem. The close connection between the two extremum problems is further brought out by the following property.

> If x is an optimal vector of the minimum problem and (y, z) is an optimal vector of the maximum problem, then the objective functions are equal.

To demonstrate this let (x, y, z) be a solution of the Lagrangian problem. It follows that x is an optimal solution of the minimization problem and (y, z) is an optimal solution of the maximization problem. The following relations are easily obtained from the Lagrangian problem.

$$c^T x = x^T A^T z - x^T y$$

$$b^T y = z^T A x = x^T A^T z .$$

It follows that

$$c^T x - b^T z = - x^T y . \qquad (2 - 12)$$

The difference between the objective functions of the two extremum problems is

$$\Delta = \frac{1}{2} x^T P^{-1} x + c^T x - \left[- \frac{1}{2} y^T P^{-1} y + b^T z \right]$$

or, using (2 - 12)

$$\Delta = \frac{1}{2} x^T P x - x^T y + \frac{1}{2} y^T P^{-1} y$$

and finally, using the relation $x = P^{-1}y$,

$$\Delta = \frac{1}{2} x^T y - x^T y + \frac{1}{2} y^T x = 0 .$$

Hence the objective functions take on the same value for the optimal solution. This principle of equality of objective functions carries over into the duality relations in linear and quadratic programming which are present in the next section.

2 - 10 Duality in Mathematical Programming

The duality relation established above for a simple quadratic minimum problem can be generalized to a class of quadratic programming problems. Consider the following pair of quadratic programs.

Primal quadratic program:

Minimize

$$\frac{1}{2} x_Q^T P x_Q + c_Q x_Q + c_L^T x_L$$

with

$$A_Q x_Q + A_L x_L \geq b$$

$$x_Q \geq 0, \quad x_L \geq 0$$

Dual quadratic program:

Maximize

$$-\frac{1}{2} y_Q^T P^{-1} y_Q + b^T y_L$$

with

$$A_Q^T y_L - y_Q \leq c_Q$$

$$A_L^T y_L \leq c_L$$

$$y_L \geq 0, \quad y_Q \text{ unrestricted}$$

Each of the obejctive functions consists of a positive definite quadratic form in one set of variables plus a linear combination of other variables and the quadratic form in the dual problem is the Legendre transform of the quadratic form in the primal. These two extremum problems have the same generalized Lagrangian problem, namely the following.

Generalized quadratic Lagrangian problem:

Find x, v, u and y such that

$$A_Q x_Q + A_L x_L - u = b \qquad \text{primal}$$

$$x_Q \geq 0 \qquad x_L \geq 0 \qquad \text{constraints}$$

$$A_Q^T y_L - y_Q + v_Q = c_Q \qquad \text{dual}$$

$$A_L^T y_L + v_L = c_L \qquad \text{constraints}$$

$$y_L \geq 0$$

$$x_Q = P^{-1} y_Q \qquad \text{primal-dual coupling}$$

$$x_Q^T v_Q = 0 \qquad \text{complementary}$$

$$x_L^T v_L = 0 \qquad \text{slackness conditions}$$

$$y_L^T u = 0$$

The important properties of the dual pair of quadratic programs are summarized in three fundamental theorems which are proved in Appendix D.

Duality Theorem: A feasible vector of the primal program is optimal if and only if there is a feasible vector of the dual program such that the primal and dual objective functions are equal. A feasible vector of the dual program is optimal if and only if there is a feasible vector of the primal program for which the objective functions are equal.

Existence Theorem: If both the primal and the dual program possess a feasible vector, then both have optimal vectors. If either of the two programs has no feasible vector, then neither has an optimal vector.

Complementary Slackness Principle: i) If (x_Q, x_L) is an optimal solution of the primal program and (y_Q, y_L) is an optimal solution of the dual program, then (x_Q, x_L, y_Q, y_L) is a solution of the Lagrangian problem. In particular the complementary slackness conditions are satisfied. ii) The values of the variables x_Q and y_Q are unique in optimal solutions of the primal and dual, respectively.

An important point in connection with quadratic programming is that the associated Lagrangian problem consists only of linear relations. This feature allows linear computations to be used in obtaining optimal solutions, and allows finite step algorithms to be designed. This would not be true for any more general class of programming problems.

The pair of quadratic programs stated above become a dual pair of linear programs if the matrices P, A_Q, c_Q, x_Q and y_Q are omitted.

Primal linear program:

Minimize $c^T x$

with $A x \geq b$

$x \geq 0$

Dual linear program:

Maximize $b^T y$

with $A^T y \leq c$

$y \geq 0$

On the other hand, the duality relationship can be extended by means of the Legendre transformation to yield a pair of dual concave programs. The primal concave program is obtained by replacing the positive definite form of the primal quadratic program by a concave function $\phi(x_Q)$. The dual convex program is obtained by substituting the Legendre transform of $\phi(x_Q)$ for the negative definite form of the dual quadratic program. The pertinent theory is given in Appendix F.

C H A P T E R   T H R E E

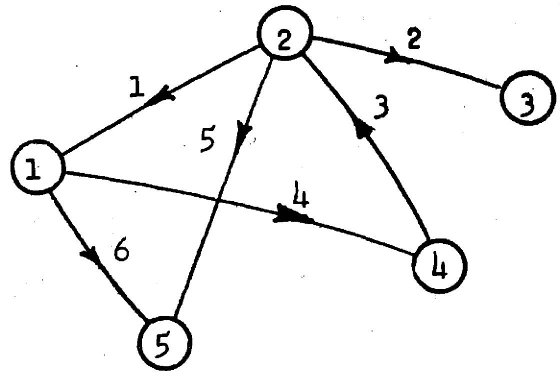
Properties and Equivalent Formulations for Electrical Networks Containing Diodes

The intimate connection between mathematical programming and the behavior of electrical networks containing diode is the central topic of this chapter. First the incidence matrix is introduced as a convenient way of expressing the topology of a network. Then the conditions for the solution of electrical networks are briefly reviewed and the characteristics of the idealized electrical devices which will concern us are defined. With the formal demonstration of the equivalence of mathematical programming and the solution of certain electrical networks, physical interpretation of some of the main theoretical ideas of programming theory are given. Finally, the concepts of terminal pair systems and breakpoint curves, which are the basis for later developments, are introduced.

3 - 1 Graphs and Incidence Matrices

A graph is a set of points which are called nodes connected by a set of line segments known as branches. Each branch of the graph connects two distinct nodes and has an associated direction indicated by an arrowhead. A graph is illustrated in Figure 3 - 1a. We will only be concerned with connected graphs in which any node is connected to each other node by some sequence of branches of the graph. Generally when one is using the concept

a)



b)

nodes \ branches	1	2	3	4	5	6
1	1			-1		1
2		-1	1		-1	
3		1				
4			-1	1		
5					1	-1

Figure 3 - 1—A graph and its incidence matrix.

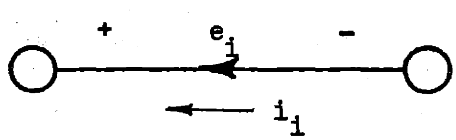


Figure 3 - 2—Branch polarity convention.

of a graph, it is only the topology of the graph--the description of which branches connect which nodes--that is important. This information may be conveniently given in a rectangular array having one row for each node and one column for each branch of the graph. In each column of the array a minus one is placed in the row corresponding to the node on which the branch originates--a plus one is placed in the row corresponding to the node on which the branch terminates. The remaining positions in the column are zero. The array so constructed is called the node incidence matrix of the graph. An example is given in Figure 3 - 1b for the graph in Figure 3 - 1a. The node incidence matrix description of a graph will be used in all mathematical formulations developed in this chapter. An alternate topological description may be obtained by listing the incidence of branches on loops and leads a network dual development. However this will not be used in the present work.

3 - 2 Electrical Networks--Laws and Devices

An electrical network is a graph in which various electrical devices are associated with the branches. With each branch i of the graph is associated a voltage e_1 and a current i_i. The sense chosen for these variables is indicated in Figure 3 - 2. It is such that the product e_i i_i is the power delivered to the network by the branch. With each node j is associated a node potential u_j. In order that a given set of values of e_i, i_i and u_j be a solution of the network it is necessary and sufficient that three conditions be satisfied. They are

1) (Kirchoff's node law) The branch currents must satisfy a continuity condition at the nodes.

$$\sum_i i_i - \sum_k i_k = 0$$

i covers all branches directed toward node j.

j covers all branches directed away from node j.

Using the node incidence matrix N this may be written as

$$N\, i = 0$$

where i is the column matrix of branch currents.

2) (Kirchoff's loop Law) The difference between potentials of the nodes on which a branch is incident must equal the branch voltage. This requires

$$u_j - u_i = e_k$$

if branch k is directed from node i to node j. This may be written, using the node incidence matrix, as

$$N^T\, u = e$$

in which u and e are the column matrices of node potentials and voltages, respectively. This condition is clearly equivalent to the requirement that the sum of the branch voltages around any loop equal zero.

3) Each branch current-branch voltage pair must satisfy any conditions imposed by the electrical device associated with that branch of the graph.

There are five different types of electrical devices that will be considered in this chapter. Their symbols and the conditions which they impose on the branch voltage-branch current pair are shown in Figure 3 - 3. The current source maintains a constant branch current regardless of what may happen to the branch voltage. Similarly the voltage source maintains a constant branch voltage. The diode acts like a switch. It maintains zero voltage as long as the current is greater than zero, and zero current if the voltage is greater than zero. The diode requirements may be stated concisely as

$$e_d \geq 0, \quad i_d \geq 0, \quad e_d\, i_d = 0.$$

The last of these relations states that a diode always delivers zero power to the network. A diode will be considered as having three mutually exclusive states:

open	$e_d > 0, \; i_d = 0$
breakpoint	$e_d = 0, \; i_d = 0$
closed	$e_d = 0, \; i_d > 0$

The resistor imposes a linear relation between branch current and voltage. The power delivered by the resistor is

$$e_r\, i_r = - \frac{e_r^2}{r} = - i_r^2\, r$$

which is a quadratic function of the branch voltage or current. The ideal d - c transformer is a device associated with two branches of the graph of a network. One branch is associated with the primary winding of the transformer and has current i_p and voltage e_p. The other branch is associated with the secondary winding with current i_s and voltage e_s. From the conditions imposed by the transformer,

$$e_s = t\, e_p, \quad i_p = - t\, i_s,$$

it follows that

$$e_p\, i_p = - e_s\, i_s,$$

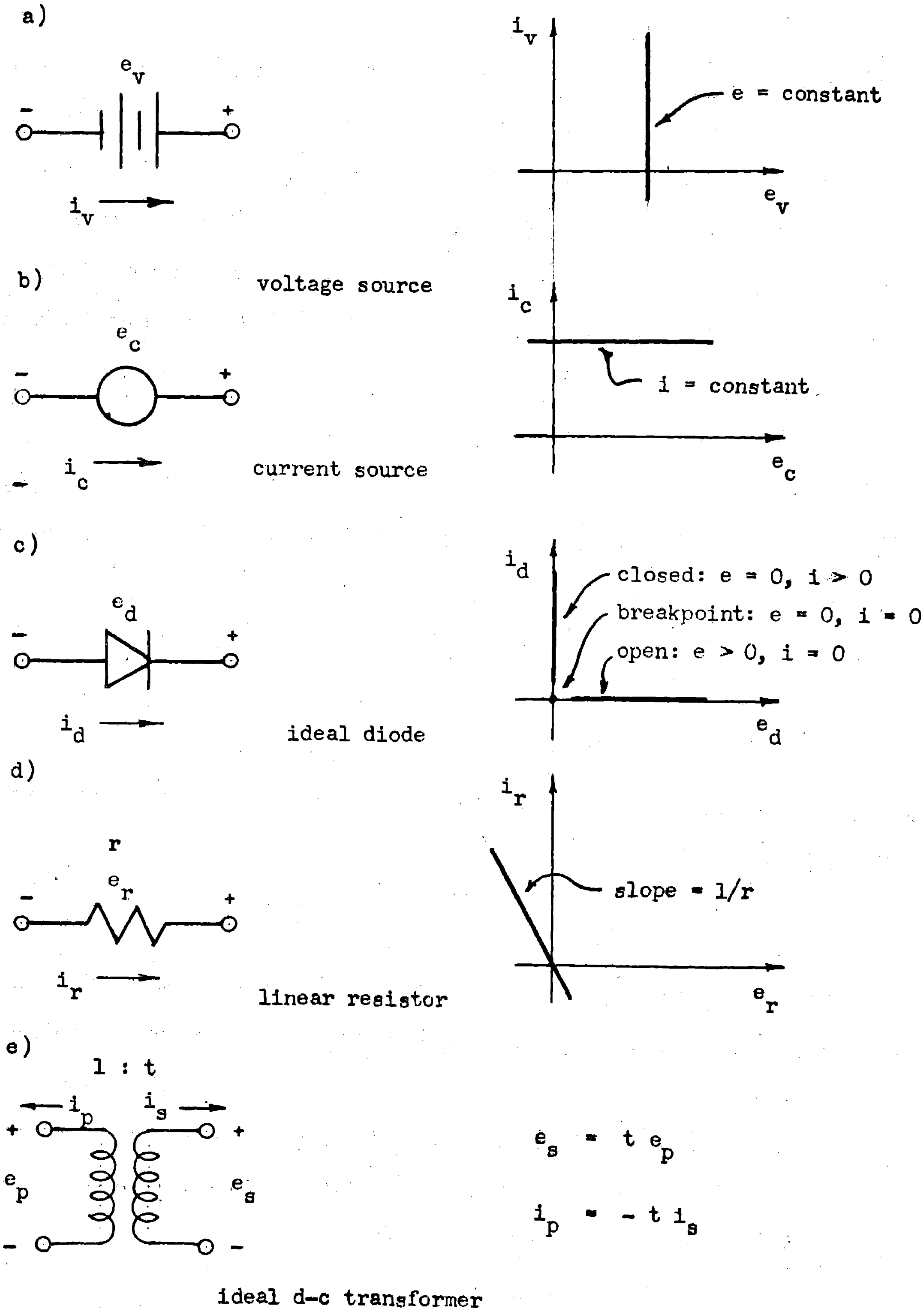


Figure 3 - 3—Relations imposed by electrical branch types.

that is, a transformer delivers zero power to a network, although it may transfer it from one branch to another.

3 - 3 Equivalent Problems for Diode-Source-Resistor Networks

In this section electrical networks containing voltage and current sources, diodes and resistors will be considered. The vectors e and i representing the branch voltages and currents will be partitioned according to type of branch (by possibly renumbering the branches).

$$e = \begin{bmatrix} e_V \\ e_C \\ e_D \\ e_R \end{bmatrix} \qquad i = \begin{bmatrix} i_V \\ i_C \\ i_D \\ i_R \end{bmatrix}$$

The rows of the node incidence matrix will be similarly partitioned so that

$$N = \begin{bmatrix} N_V & N_C & N_D & N_R \end{bmatrix} .$$

The network laws as applied to a diode-source-resistor network may now be written down.

Diode-source-resistor Lagrangian problem

Find e, i and u such that

$$N_V\, i_V + N_D\, i_D + N_R\, i_R = -N_C\, i_C \qquad \text{Kirchoff node law}$$

$$\left.\begin{aligned} N_V^T u &= e_V \\ N_D^T u - e_D &= 0 \\ N_R^T u - e_R &= 0 \\ N_C^T u - e_C &= 0 \end{aligned}\right\} \qquad \text{Kirchoff loop law}$$

$e_R - R\ i_R$ Resistor requirements

$e_D \geq 0, \quad i_D \geq 0, \quad e_D^T\ i_D = 0$ Diode requirements

Here R is the diagonal matrix of resistance values. The terms $N_C\ i_C$ and e_V have been put on the right as they are given constants. The important point is that this problem is exactly the Lagrangian problem associated with the following pair of dual quadratic programs.

Diode-source-resistor primal problem:

Find an i that minimizes

$$\frac{1}{2}\, i_R^T\ R\ i_R - e_V^T\ i_V$$

subject to

$$N_V\ i_V + N_D\ i_D + N_R\ i_R = -\ N_C\ i_C$$

$$i_D \geq 0$$

Diode-source-resistor dual problem:

Find an e and u that maximizes

$$-\frac{1}{2}\ e_R^T\ G\ e_R + i_C^T\ e_C$$

subject to

$$N_V^T\ \ u = e_V$$

$$N_D^T\ \ u - e_D = 0$$

$$N_R^T\ \ u = e_R = 0$$

$$N_C^T\ \ u - e_C = 0$$

$$e_D \geq 0$$

In the dual problem, $G = R^{-1}$ is the diagonal matrix of conductance values of the resistance branches. In the primal formulation the variables are branch currents and the node potentials are the Lagrange multipliers. Conversely, the dual problem concerns voltage quantities and the branch currents are the Lagrange multipliers. The coupling relation of the Lagrangian problem is a consequence of the resistor requirements. If the network contains no resistors, the coupling relation disappears and the problem statements reduce to a pair of dual linear programs. The above two problems have the following corresponding statements in words:

> <u>Primal</u>: Find a feasible current distribution in the network which minimizes the sum of the power absorbed by the voltage sources and one half the power absorbed by the resistors.
>
> <u>Dual</u>: Find a feasible voltage distribution in the network which maximizes the power delivered to the network by the current sources minus one half the power absorbed by the resistors.

If there are no diodes or voltage sources in the network the above primal problem becomes the well-known rule that the solution of a source-resistor network is that current distribution which minimizes the power dissipated in the resistors among all current distributions satisfying Kirchoff's node law. With voltage sources but no diodes, the primal statement is a more general rule given by Maxwell.

This correspondence between programming problems and electrical network problems yields some very nice physical interpretations of the mathematical principles of programming.

1) The inequality constraints in the programming problem are embodied in the diode requirements in the electrical network.

2) The complementary slackness principle corresponds to the requirement that diodes deliver zero power to an electrical network.

3) The duality theorem states that the objective functions in the primal and dual formulations are equal at the solution point. In the electrical network problem this requires

$$e_V^T i_V - \frac{1}{2} i_R^T R i_R - \frac{1}{2} e_R^T G e_R + e_C^T i_C = 0$$

which is clearly a statement of the principle of conservation of energy.

3 - 4 Reduced Networks

An interpretation of the existence theorem of mathematical programming can be given in terms of the reduced networks associated with a given source-diode-resistor network.

The voltage reduced network associated with a given source-diode resistor network is obtained by replacing each resistor or current source of the original network with an open circuit.

The current reduced network associated with a given source-diode resistor network is obtained by replacing each resistor or voltage source of the original network with a short circuit.

The equations for the voltage reduced network are obtained by setting i_C equal to zero and replacing the resistor conditions with $i_R = 0$ in the Lagrangian problem. It is evident by inspection that if (i, e, u) is a solution of the voltage reduced network, then i satisfies the constraints of the primal programming formulation. Conversely, if i satisfies the primal constraints, then (i, e = 0, u = 0) is evidently a solution of the voltage reduced network. Hence, the primal constraints are feasible if and only if the voltage reduced network has a solution. The equations for the current reduced network are obtained by setting e_V equal to zero and replacing the resistor conditions with $e_R = 0$ in the Lagrangian problem. As above one concludes that the constraints of the dual programming formulation are feasible if and only if the current reduced network possesses a solution. The existence theorem of quadratic (or linear) programming (Theorem D - 3) then has the following physical statement in terms of electrical networks.

> An electrical network made up of sources, diodes and resistors possesses a solution if and only if the associated voltage reduced and current reduced networks both possess solutions.

The fact that the voltage reduced network has a solution apparently guarantees that none of the voltage sources are "shorted out" in the original network. The fact that the current reduced network has a solution guarantees that all the current produced by the current sources has a "place to go".

The correspondence between network problems and programming problems also yields nice physical interpretations of algorithms for solving linear and quadratic programs as will be seen in the following chapters.

3 - 5 Electrical Networks Containing Transformers

The general case of a network made up of voltage and current sources, diodes, resistors and ideal d-c transformers will be considered next. In this case the current vector i will be partitioned to include also i_P and i_S, the primary and secondary transformer currents. The branch voltage vector e will also contain e_P and e_S, the primary and secondary transformer voltages. The incidence matrix is expanded to include N_P and N_S giving the incidence of transformer primary and secondary windings on nodes. The diagonal matrix

T gives the turns ratios of the transformers. Expressing the network laws in these terms gives the Lagrangian problem.

<u>Transformer network Lagrangian problem</u>:

Find e, i and u such that

$$N_V\, i_V + N_D\, i_D + N_R\, i_R + N_P\, i_P + N_S\, i_S = -\, N_C\, i_C \qquad \text{Kirchoff's node law}$$

$$N_V^T\, u = e_V$$

$$\left.\begin{aligned} N_D^T\, u - e_D &= 0 \\ N_R^T\, u - e_R &= 0 \\ N_P^T\, u - e_P &= 0 \\ N_S^T\, u - e_S &= 0 \end{aligned}\right\} \qquad \text{Kirchoff's loop law}$$

$$e_R + R\, i_R = 0 \qquad \text{resistor conditions}$$

$$e_S - T\, e_P = 0,\ i_P + T\, i_S = 0 \qquad \text{transformer conditions}$$

$$e_D \geq 0,\ \ i_D \geq 0,\ \ e_D^T\, i_D = 0 \qquad \text{diode conditions}$$

This is the Lagrangian problem corresponding to the following pair of dual quadratic programs.

<u>Transformer network primal problem</u>:

Find an i that minimizes

$$\frac{1}{2} i_R^T R i_R - e_V^T i_V$$

subject to

$$N_V i_V + N_D i_D + N_R i_R + N_P i_P + N_S i_S = - N_C i_C$$

$$i_P + T i_S = 0$$

$$i_D \geq 0$$

<u>Transformer network dual problem</u>:

Find an e and u that maximizes

$$- \frac{1}{2} e_R^T G e_R + i_C^T e_C$$

subject to

$$N_V^T u = e_V$$

$$N_D^T u - e_D$$

$$N_R^T u - e_R = 0$$

$$N_P^T u - e_P = 0$$

$$N_S^T u - e_S = 0$$

$$e_S - T e_P = 0$$

$$e_D \geq 0$$

Thus every electrical network containing the five types of devices treated here is equivalent to a pair of dual quadratic programs. It will be shown later (in Chapter six) that the reverse is also true--every quadratic (or linear) program can be represented by an electrical network containing only these five types of elements. All of the remarks made above concerning the physical interpretation of the principles of mathematical programming apply without modification when transformers are included.

The important point of difference between networks without transformers and networks with transformers is the following: In the former, the coefficients in the constraint relations of the primal and dual formulations are all either plus one, minus one or zero. Indeed, they are the elements of an incidence matrix. In linear programming this fact allows algorithms to be devised in which the arithmetic involves only additions and subtractions. The importance of this will be seen in the next chapter where such an algorithm is developed. With transformers included, the turns ratio matrix T appears in the constraint relations and the coefficients therefore need not even be integers.

3 - 6 Non-Linear Devices:

The correspondence between programming problems and network problems can be carried even one step further. In place of the linear resistors included in the networks discussed above, a general class of non-linear resistors will be allowed. The voltage-current relation of the device will be required to be a "strictly decreasing" curve. Precisely, if (i_r, e_r) and $(\bar{i}_r, \bar{e}_r)$ are any two points on the curve, then

$$i_r < \bar{i}_r \text{ if an only if } e_r > \bar{e}_r.$$

Otherwise, the curve is arbitrary. The characteristics of such a device are illustrated in Figure 3 - 4. Obviously the linear resistor is a

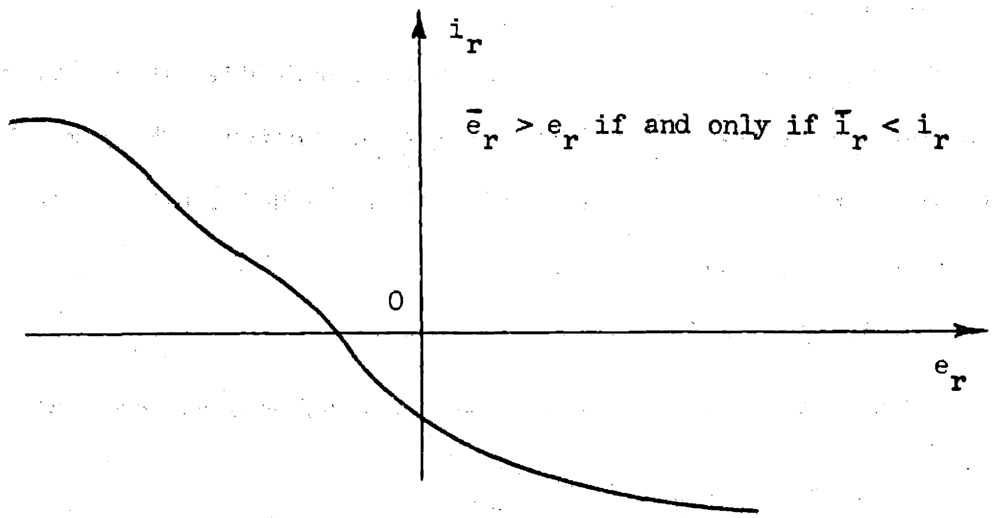


Figure 3 - 4--Non-linear resistive device.

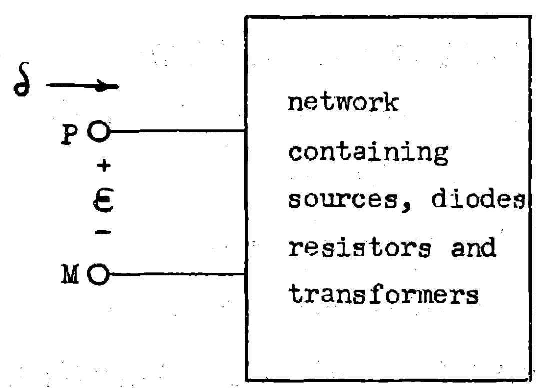


Figure 3 - 5-- Network with observable terminal pair.

special case of this device. For this class of devices the relation between current and voltage is one-to-one.

For a network containing voltage and current sources, diodes, transformers and the general resistive devices, the Lagrangian problem is the same as for the transformer network above except that the non-linear resistors are described by a relation

$$e_R = f(i_R)$$

where f is a one-to-one transformation. As a consequence we may also write

$$i_R = f^{-1}(e_R) = - g(e_R).$$

From the properties of convex functions (Appendix E) one concludes that f is the differential of some concave function $\emptyset$ and g is the differential of some convex function Θ. The net result is that the solution of a network of this type is equivalent to a pair of dual concave programming problems which are identical to the quadratic programs stated above except that the quadratic forms are replaced by $\emptyset(i_R)$ in the primal formulation and by $\Theta(e_R)$ in the dual formulation.

3 - 7 Terminal Solutions and Breakpoint Curves

The concepts of terminal solutions and breakpoint curve are of utmost importance in the algorithms which will be described in succeeding chapters. Hence, these ideas will be introduced here.

Consider an arbitrary network containing sources, diodes, resistors and transformers in which one pair of nodes (the terminal pair) is available for external observation (See Figure 3 - 5). The current and voltage at the terminal pair will be represented by δ and ε, respectively. The nodes which comprise the terminal pair will be indicated by the one column incidence matrix n_T which contains a minus one in the position corresponding to node M and a plus one in the position corresponding to node P. By a <u>terminal solution</u> is meant a pair of values of δ and ε which corresponds to some solution of the network. Terminal pair solutions thus corresponds to solutions of the <u>terminal pair system</u> of relations given below.

<u>Terminal Pair System</u>

$N\ i + n_T\ \delta = 0$ Kirchoff's node law

$N^T\ u = e$
$N_T^T\ u = -\ \varepsilon$ } Kirchoff's loop law

$e_R = -\ R\ i_R$ resistor conditions

$e_S = T\ e_P,\ i_P = -\ T\ i_S$ transformer conditions

$i_D \geq 0,\ e_D \geq 0,\ e_D^T\ i_D = 0$ diode conditions

This system of relations may also be set down in tabular form as follows:

ε	i_R	i_V	i_D	i_P	i_S	δ	u	e_P	e_D	
	N_R	N_V	N_D	N_P	N_S	n_T				$-N_C\ i_C$
				$-I$	$-T$					0
	$-R$						N_R^T			0
							N_V^T			e_V
							N_D^T		I	0
							N_P^T	$-I$		0
							N_S^T	$-T$		0
1							n_T^T			0

The line joining the vector variables i_D abd e_D indicates that these variables are required to satisfy the "diode conditions". The system above is a special case of the <u>general terminal pair system</u>

η	x	ξ	y	v	
	A	e			b
	$-Q$		A^T	I	c
1			e^T		0

An electrical model for this mathematical system will be given in Chapter Five. Its properties are studied in Appendix G.

A <u>breakpoint</u> curve is a continuous curve made up of straight line segments none of which has negative slope. An example is given in Figure 3 - 6. In Appendix G it is demonstrated that the set of values of (ξ, η) which correspond to solutions of the general terminal pair system form a breakpoint curve in the $\xi - \eta$ plane. In the following an attempt will be made to make this plausible by means of physical arguments relating to electrical networks. For this purpose it is helpful to view the network as in Figure 3 - 7, separating the diodes from the network by placing each one at its own terminal pair. For any solution of the terminal pair system each diode is either "open" and it has zero current and behaves as an open circuit, or it is "closed", has zero voltage and behaves as a short circuit. (disregarding the breakpoint state for the present). For a given state of each diode (open or closed) it is clear physically that since the network contains only linear devices, there will be a unique incremental resistance observed at the terminal pair. This incremental resistance may be zero or it might be an open circuit.

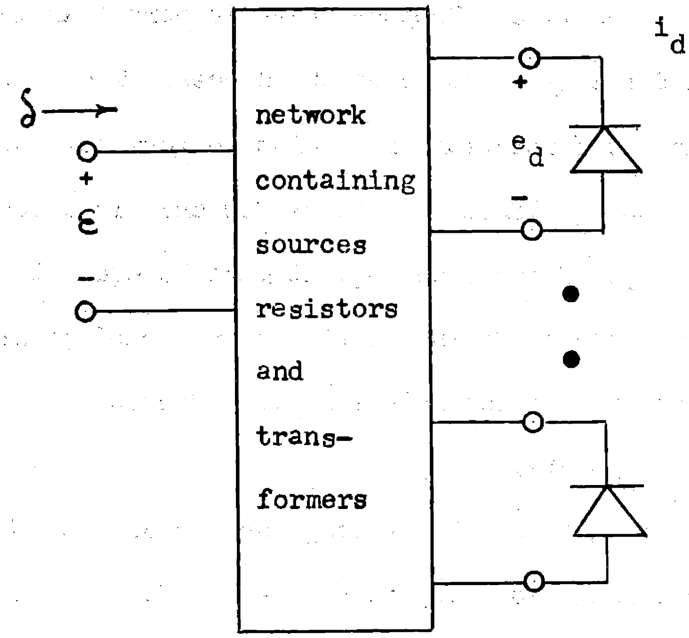


Figure 3 - 7—Network with diodes separated.

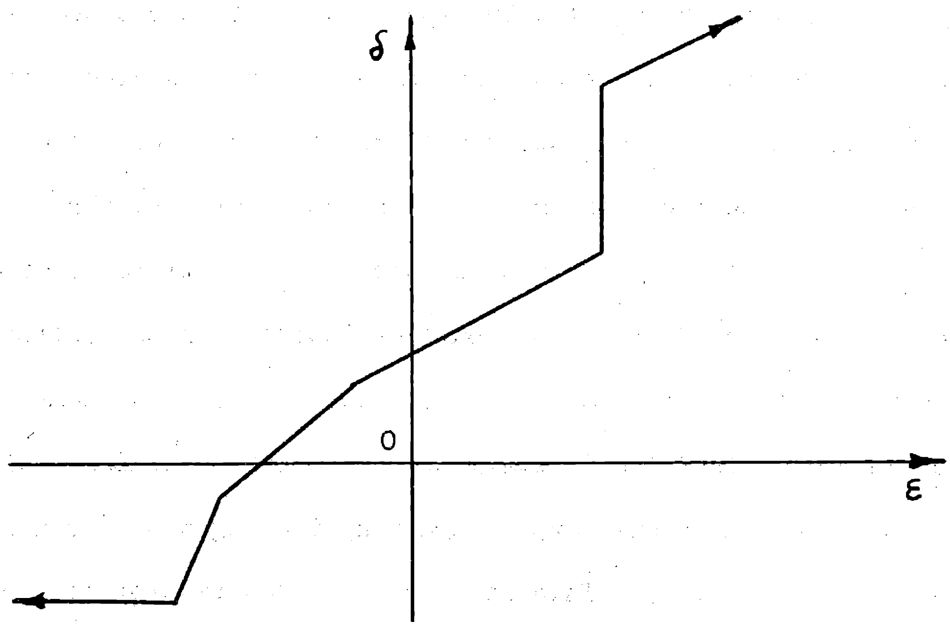


Figure 3 - 6—An example of a breakpoint curve.

However, it can never by negative. Suppose there is more than one terminal solution of the system for a given state of the diodes. Then all points on the line joining these two solutions must also be terminal solutions because the system is a set of linear equations for any given configuration of diode states. The slope of the line segment must be equal to the incremental resistance. Since there are only a finite number of possible choices of diode states, the set of terminal solutions must consist of a finite number of straight line segments.

Now consider connecting a variable voltage source and one-ohm resistor to the terminal pair, as shown in Figure 3 - 8, and suppose that the voltage has been adjusted so that the network has a solution. As the variable voltage α is raised the terminal solution will move along the line segment corresponding to the present state of the diodes. At some point a diode current or voltage will reach zero and the diode will shift from the open state, through the breakpoint state to the closed state, and the terminal solution will be on a new line segment. As the variable voltage is further increased, the states of the diodes will change generating the successive line segments of the breakpoint curve. Eventually, since there are only a finite number of combinations of diode states, a state will be reached where no amount of increase in the voltage α will cause any diode to shift to a new state. Thus the last segment of the breakpoint curve must extend indefinitely toward positive terminal voltage and current. The same arguments apply regarding the negative portion of the set of terminal solutions. This is a physical description of the mathematical technique used to prove the corresponding properties of the general terminal pair system in a rigorous manner in Appendix G.

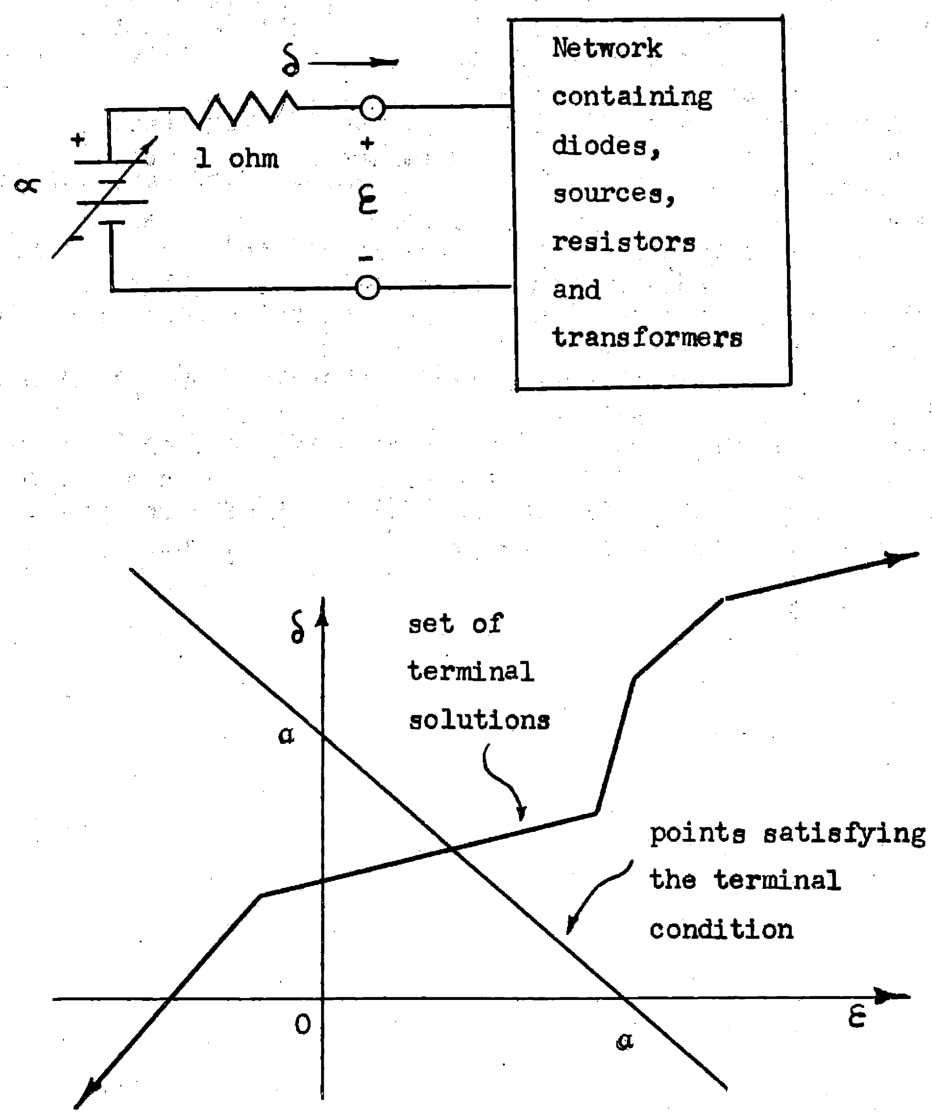


Figure 3 - 8--Demonstrating the continuity of the set of terminal solutions.

A couple of special cases of the terminal pair system are sufficiently important to be worth special comment. First, if the network contains no resistors, the incremental resistance observed at the terminal pair must always be either zero or an open circuit. Hence the breakpoint curve would contain only horizontal and vertical line segments. Moreover, if the voltage sources of the network are set to zero or removed, all horizontal line segments must be coincident with the current axis. This is true because the vertical positions of such line segments are determined by the voltage sources. Similarly, all vertical line segments of the breakpoint curve collapse to the voltage axis if all current sources are set to zero. These results are demonstrated for the general terminal pair system as Theorems G - 2 and G - 3.

C H A P T E R   F O U R

An Algorithm for Solving Diode-Source Networks

In this chapter an algorithm for finding a solution of any electrical network containing ideal diodes, current sources and voltage sources is developed. A necessary and sufficient condition for the existence of solutions is given and the algorithm is shown to yield a solution for any network satisfying the condition. The reasoning used in this development is based in its entirety on the ideas and concepts of electrical engineeri

In the presentation the terms path, loop and tree will be used to describe certain groups of branches in the graph of a network. Their definitions are:

<u>path between node i and node j</u>--a sequence of branches and nodes in the graph starting with node i and terminating with node j.

<u>loop</u>--a path which starts and terminates on the same node.

<u>tree</u>-- a set of branches and nodes of the graph which contains a path between each pair of nodes in the set, but contains no loops

<u>full tree</u>--a tree which contains every node of the graph.

4 - 1 Existence Conditions and Reduced Networks

For the purpose of the present discussion it will be supposed that certain redundant branches have been removed from any network considered. Specifically, the following assumptions will be made.

First non-redundancy assumption:

The nodes of the network cannot be divided into two groups such that the only branches connecting members of the two groups are current sources.

In particular, this means that no single node of the network may have current sources only incident on it. If this assumption were not valid, then either the values of these current sources do not add up to zero and the network possesses no solution, or, the values of the current sources cancel, and the two sections of the network may be solved separately. (See Figure 4 - 1).

Second non-redundancy assumption

The network has no loop on which only voltage sources are incident.

If this assumption did not hold, one of the voltage sources could be removed without affecting solutions except for an arbitrary current in the loop. (See Figure 4 - 2).

The following statement gives a necessary and sufficient condition for the existence of solutions for an electrical network constructed of voltage and current sources, ideal diodes and resistors.

An electrical network whose branches are voltage and current sources, diodes and resistors possesses a solution if and only if the following two conditions are both satisfied.

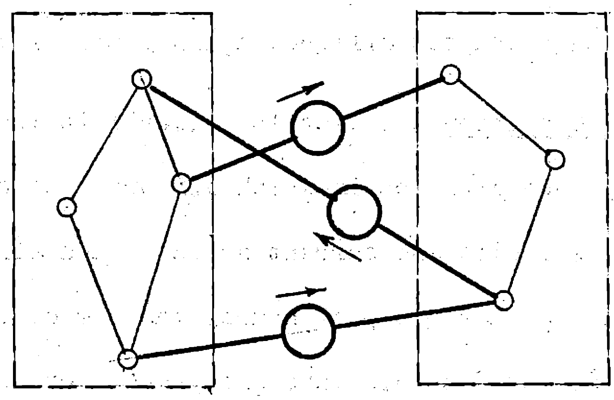


Figure 4 - 1--Network violating the first non-redundancy assumption.

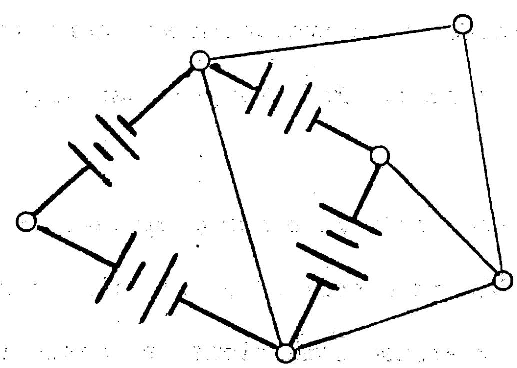


Figure 4 - 2—Network violating the second non-redundancy assumption.

First Existence Condition: For each loop of the network on which only diodes and voltage sources are incident, and all diodes are alligned in the same way with the loop, the sum of values of the voltage sources alligned with the diodes must not be greater than the sum of values of voltage sources alligned against the diodes.

Second Existence Condition: Divide the nodes into two groups in any manner. For any such division in which all branches connecting nodes of both groups are diodes and current sources, and all diodes are directed towards the same group of nodes, the sum of values of the current sources directed toward this group must not exceed the sum of values of the current sources directed away.

These conditions are simply a restatement of the interpretation of the existence theorem of quadratic programming given in section 3 - 4. The necessity of these conditions is readily apparent. If the first condition were violated for some loop, then no assignment of potentials to the nodes of the loop could satisfy the branches of the loop. If the second condition were violated for some division of nodes, then no permissible assignment of currents to the diodes could produce a net current between the two groups of nodes equal to zero. For examples see Figures 4 - 3 and 4 - 4.

The sufficiency of these conditions for a diode-source network will be evident when the algorithm is described as it will be shown to yield a solution whenever the two existence conditions are satisfied. Of course, the sufficiency is also followed from the existence theorem.

Suppose it is required that a certain network N of voltage and current sources, diodes and resistors be solved. From this network a new network V may be obtained by removing all current sources and resistors. This network

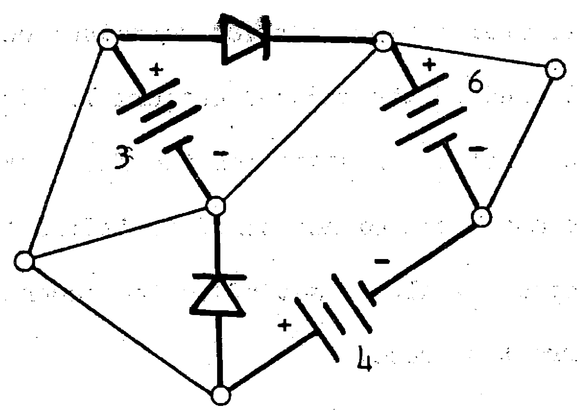


Figure 4 - 3—Network violating the first existence condition.

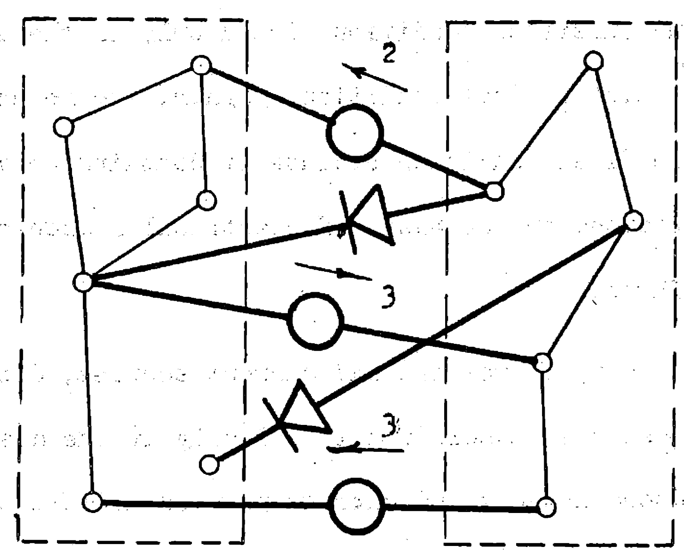


Figure 4 - 4—Network violating the second existence condition.

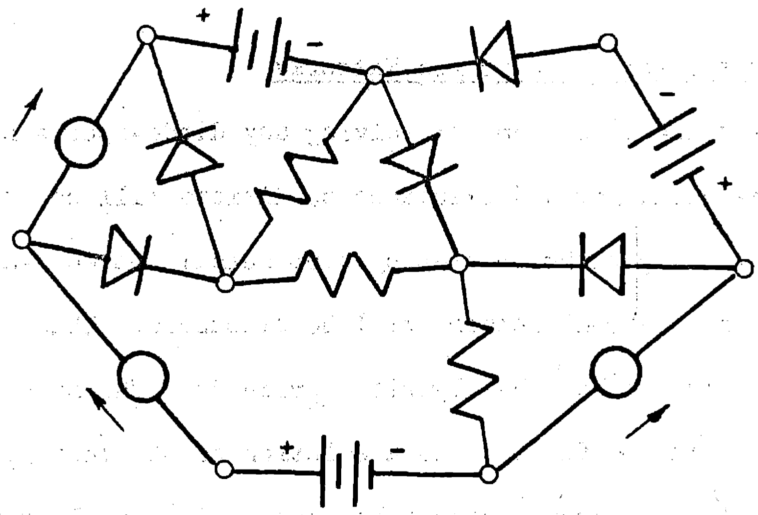
will be called the voltage reduced network associated with N. The voltage reduced network for the network in Figure 4 - 5a is shown in Figure 4 - 5b. Now every loop in the original network N which contained only voltage sources and diodes, appears unaltered in the reduced network V. In fact, V contains only those loops. Hence a network N satisfies the first existence condition if and only if the associated reduced network V satisfies that condition. Note that the reduced network V always satisfies the second condition because it contains no current sources.

Similarly, we may form the current reduced network C from N by replacing all voltage sources and resistors in N by short circuits. The current reduced network for the network in Figure 5a is shown in Figure 5c. Then for any division of the nodes of N in which only diodes and current sources connect the two groups of nodes, the corresponding division of the nodes of C will yield the same situation. In fact, every division of the nodes of C corresponds to some such division of the nodes of N. Thus a network N satisfies the second existence condition if and only if the associated reduced network C satisfies that condition. Again, the reduced network C always satisfies the first condition because it contains no voltage sources.

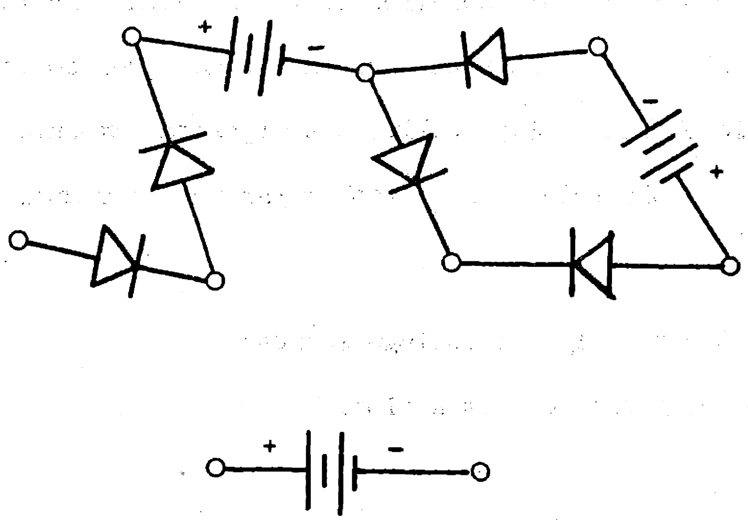
Thus we have reached on the basis of electrical reasoning the conclusion stated in Chapter Three:

> A network made up of voltage and current sources, diodes and resistors possesses a solution if and only if the associated voltage reduced network and current reduced network both possess solutions.

a)



b)



c)

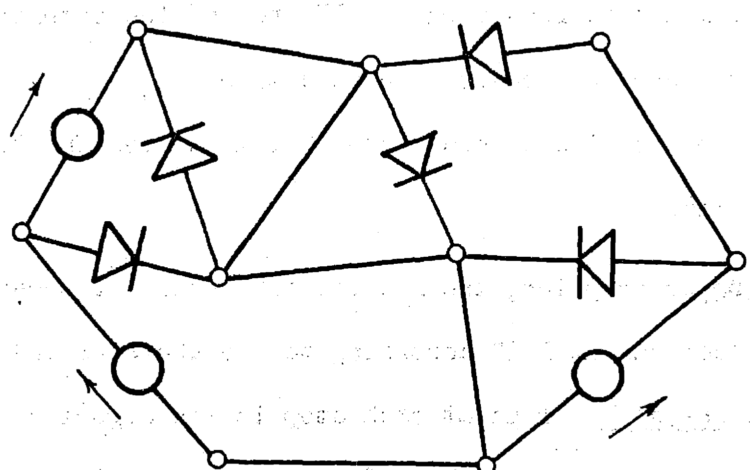


Figure 4 - 5—Diode-source-resistor network and its associated

4 - 2 First Algorithm for Diode-Source Networks

In this section the algorithm for solving any diode-source network satisfying the non-redundancy and existence conditions will be presented. Before describing the algorithm in step by step detail the equivalent operations performed on a real network will be discussed. This should give a good physical picture of how the algorithm proceeds. The method involves two distinct phases: In the first phase a solution of the voltage reduced network is found. In the second phase, the result of phase I is used as a starting point to form a complete solution of the network. In terms of linear programming, a feasible vector is found for the primal problem and is then modified until it is optimal. Physically, the algorithm amounts to assembling the components of the network in a rather particular manner as follows.

Phase I

1) Start with only the voltage sources.

2) Add the diodes one at a time.

Phase II

3) Short circuit sufficient diodes so that the short circuits and the voltage sources form a full tree of the complete network.

4) Add the current sources one at a time

5) Restore the diodes which were short-circuited in step three one at a time.

In the following description, the network consisting of those components which have already been "assembled" according to the above procedure will be called the solved subnetwork. Thus at each step in the algorithm one has a set of branch currents and voltages which form a solution of the pertinent solved subnetwork. By following the order of assembly given above it is always possible, as will be demonstrated below, to find a solution for the combination of the solved subnetwork and the next component to be added. This statement is true, of course, only if the network to be solved satisfies the the non-redundancy and existence conditions given above.

In the figures the letters C, B and O will be used to indicate that a diode is in the closed, breakpoint or open state, respectively.

Phase I - Solving the voltage reduced network

Step 1) Select as the initial solved network a full tree of the network being solved by

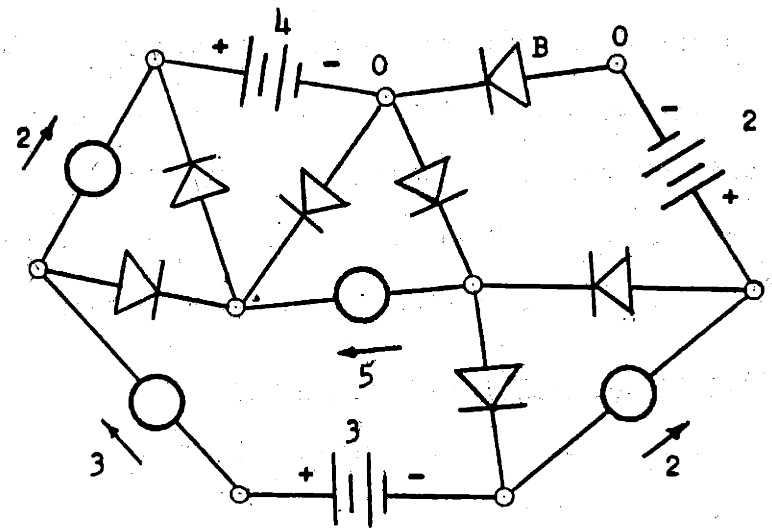
a) including all voltage source branches

b) adding sufficient diode branches so that every node is covered while producing no loops.

It is readily apparent that such a tree exists. First, according to the first non-redundancy assumption, there is no possible division of the nodes of the network into two groups such that only current sources connect members of the two groups. This assures that there is a tree of the network containing only voltage sources and diodes. The second non-redundancy assumption state that there are no loops containing only voltage sources. This assures the existence of a full tree containing diodes and all voltage sources. A diode-source network and a full tree for it with a potential distribution are shown in Figure 4 - 6.

Step 2) Consider each diode of the initial solved network as being in the breakpoint state (zero voltage, zero current). Compute a consistent set of node potentials for this network (tree). Consider all diodes whose voltage is zero or greater than zero as part of the solved subnetwork and in the breakpoint or open state, respectively.

a)



b)

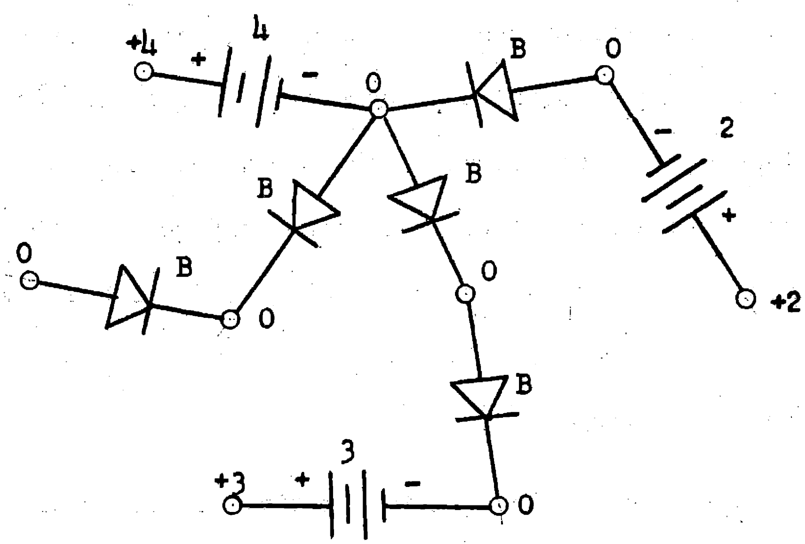


Figure 4 - 6—Diode-source network with tree and potential distribution for starting phase I.

This may be done by setting the potential of an arbitrarily chosen initial node equal to zero. Then compute the potentials of those nodes connected to the initial node by a branch of the tree. This will allow the potentials of other nodes to be evaluated from values previously computed. Continuance of the process will lead to the assignment of a potential value to each node of the network. Since by definition a tree contains no loops, no inconsistency can ever arise--that is, it is always possible to compute the required set of node potentials.

<u>Step 3</u>) Select a diode not in the solved subnetwork. If there is none Phase I is complete and the present set of node potentials form a solution of the voltage reduced network. Proceed with Phase II. If there is such a diode, designate the nodes on which its plus and minus terminals are incident by P and M, respectively. Let E be the magnitude of the (negative) voltage of the selected diode.

In Steps 4, 5, 6 and 7 a solution will be found for a new subnetwork consisting of the solved subnetwork plus the diode chosen in Step three. The method involves "labelling" those nodes at which current is "available" over some path from node P. At the termination of the labelling, a new solution is indicated which is nearer to satisfying the diode selected in Step 3.

<u>Step 4</u>) Label node P.

Step 5) Label node j if node i is labelled and there is a branch of the subnetwork connecting nodes i and j which is

i) a voltage source

ii) a diode in the breakpoint state directed from node i to node j.

Repeat until no more nodes can be labelled. If node M is not labelled proceed with Step 6. If node M is labelled, the network possesses no solution.

In the latter case there is a path between the plus and minus terminals of the chosen diode containing only voltage sources and breakpoint diodes (See Figure 4 - 7). Together with the selected diode the branches of this path form a loop containing only voltage sources and similarly alligned diodes with the total loop voltage equal to E. The first existence condition is thereby violated and the network has no solution.

If node M is not labelled, then the nodes of the network can be divided into two groups--labelled and unlabelled. Any branch running from a labelled node i to an unlabelled node j must be

i) a diode in the open state

ii) a breakpoint diode directed from node i to node j.

The situation is illustrated in Figure 4 - 8. If this were not true the labelling operation, Step 5, was not completed. Remember that in Phase I there are no diodes in the closed state because all currents are zero.

Now visualize similtaneously raising the values of the potentials on all of the labelled nodes by the amount ΔU. The new set of potentials will constitute a solution for the solved subnetwork as long as it is consistent with the demands of the individual branches. Those branches which join a pair of labelled nodes or join a pair of unlabelled nodes will be unaffected by the change. There are two classes of branches which connect labelled nodes with unlabelled nodes--breakpoint diodes and open diodes. In the case of the breakpoint diodes, they will change to the open state for any value of Δu greater than zero, and will be satisfied no matter how large ΔU is.

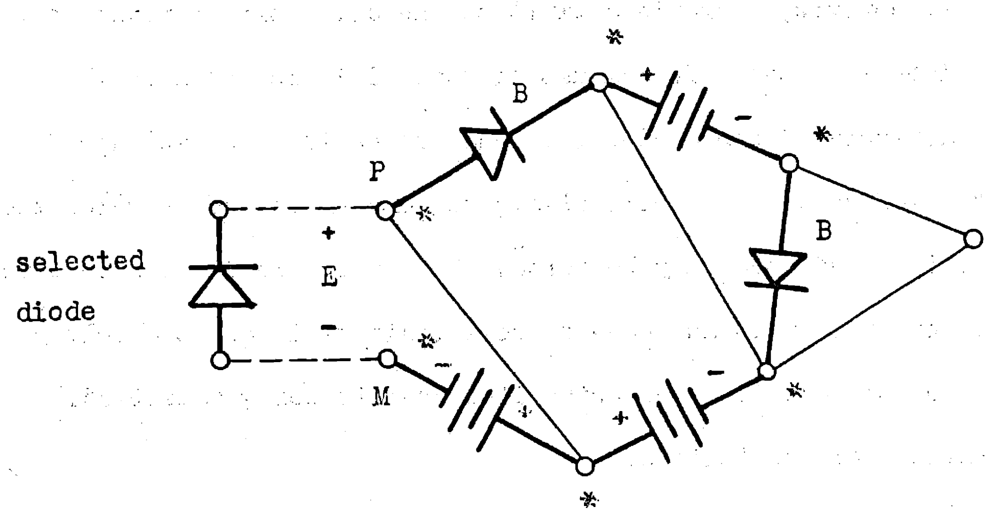


Figure 4 - 7—The situation if node M is labelled.

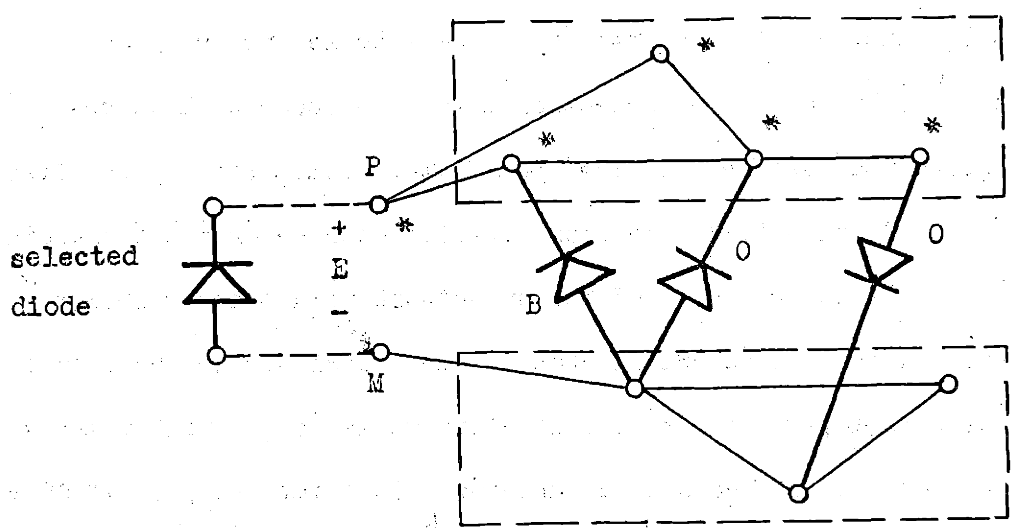


Figure 4 - 8—The situation if node M is not labelled.

Open diodes directed from unlabelled nodes to labelled nodes will remain satisfied for any positive value of ΔU. For each open diode directed toward an unlabelled node, however, there is a limit on the allowable increase in the potential of the labelled node. The amount of this limit is equal to the voltage for the particular diode. Since, by our convention, the voltage for an open diode is never zero, there will always be some amount ΔU by which the potentials of the labelled nodes may be raised which will yield a new solution of the solved subnetwork. This increase of potentials will make the voltage of the chosen diode more positive and bring it nearer to being satisfied. The precise operations are as follows:

Step 6) Consider those diodes in the subnetwork which are in the open state and directed from a labelled node to an unlabelled node. Let ΔV be the smallest voltage for any of these diodes. If ΔV is greater than E or there are no open diodes directed toward unlabelled nodes, set ΔU = E, otherwise set ΔU = ΔV.

Step 7) Increase the potentials of the labelled nodes by ΔU. Every breakpoint diode connecting a labelled node with an unlabelled node switches to the open state. Some open diodes directed toward unlabelled node from labelled ones enter the breakpoint state. All other diodes of the subnetwork retain their former state. After the potential change some diodes not in the solved subnetwork may no longer have negative voltages. Consider these diodes as part of the solved subnetwork in future steps. If ΔU = E a solution has been obtained for a new subnetwork consisting of the old subnetwork plus the selected diode. Remove all labels and go back to Step 3. Otherwise set E = E - ΔU and return to Step 5.

Each time a potential change is executed in Steps 6 and 7, either a solution is found satisfying the selected diode, or one of the open diodes joining a labelled node to an unlabelled node moves to the breakpoint state. If the latter is the case another node can always be labelled in Step 5. Therefore, for each diode chosen in Step 3, the number of labelling operations before the diode can be satisfied can never exceed the number of nodes in the network. Thus Phase I must terminate in a finite number of steps.

Phase II - The completion of the solution

Step 1) Consider any division of the nodes of the voltage reduced network into two groups such that the only branches joining members of the two groups are open diodes. Raise or lower the potentials of one of groups of nodes until at least one of the open diodes changes to the breakpoint state. Repeat until no division of the nodes as above is possible.

The object of Step 1 is to insure that it is possible to find a tree with the properties requested in Step 2. It is clear that each change of potentials in Step 1 gives a new set of potential values which is a solution of the voltage reduced network. For each change of potentials at least one diode changes from the open to the breakpoint state. While all breakpoint diodes remain at breakpoint. Hence Step one must terminate in a number of repetitions not greater than the initial number of open diodes.

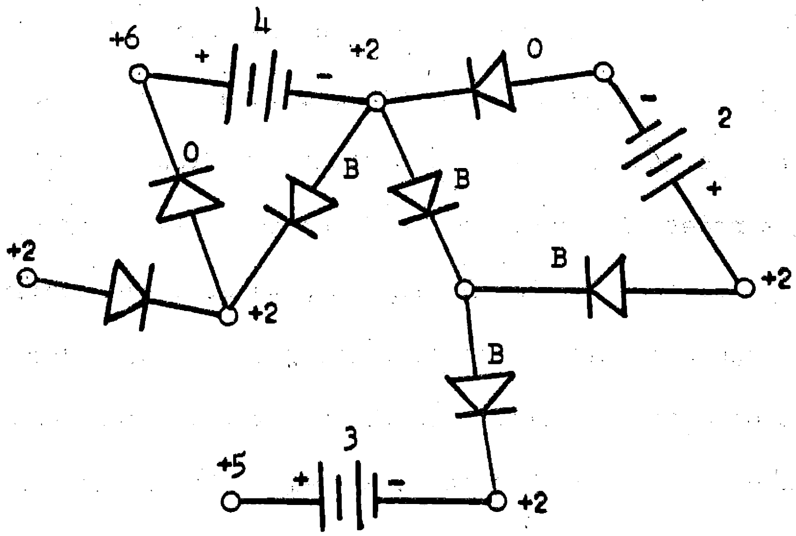
Step 2) Select from the voltage reduced network a full tree consisting only of voltage sources and breakpoint diodes. Assign a current distribution to the branches of this tree which satisfies the continuity condition and meets the requirements of the current sources. Identify those diodes of the tree which have currents greater than zero as being in the closed state. Short circuit those diodes in the tree which have current less than zero. The resulting network (with the current sources) is the solved subnetwork used as the starting point for the iterations of Phase II.

The process of assigning the current distribution to the tree might be carried out as follows. Select a current source and a path from the tree connecting the pair of nodes on which the current source is incident. Add the current demanded by the current source to the current in each branche alligned with the path. Subtract this current from the other branches in the path. Repeat for each current source. This method also shows that such an assignment is always possible. Figure 4 - 9a shows a solution of the voltage reduced network obtained from the network of Figure 4 - 6a. Figure 4 - 9b shows a tree satisfying the requirements of Step 2 with a current distribution satisfying the current source requirements.

Step 3) Select one of the diodes which is short circuited. If all of the short circuits have been removed, the present set of branch currents and voltages constitutes a solution of the network and Phase II of the algorithm has been completed. Otherwise, designate the nodes on which the plus and minus terminals of the diode are incident by P and M respectively. Denote the magnitude of the (reverse) current through this diode by I.

In Steps 4 through 9 a solution will be produced for a new network consisting of the solved subnetwork with the short circuit removed from the diode selected in Step three. Again nodes are labelled for which current is

a)



b)

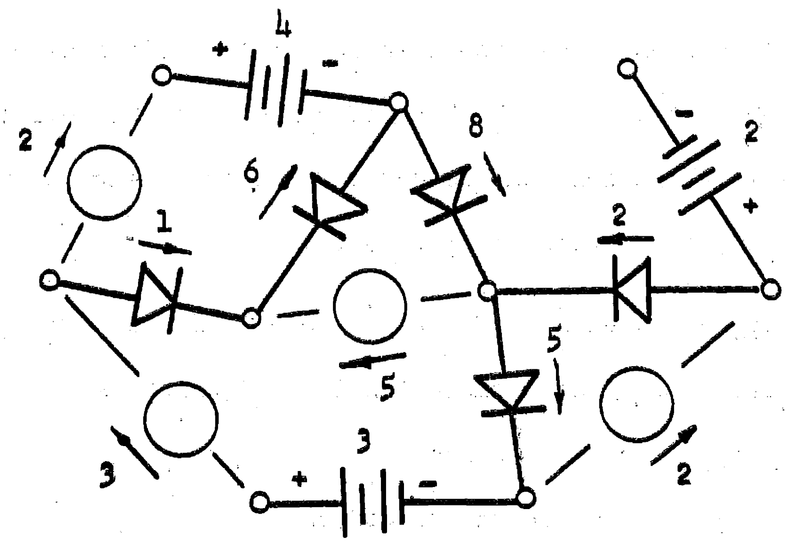


Figure 4 - 9—Solution of voltage reduced network and current distribution for starting phase II.

available over paths from node P. As a consequence of the labelling one of the following two results occur. Either a path is produced connecting nodes P and M over which an increment of current can flow bringing the current between these nodes closer to zero. Otherwise a means of changing the node potentials is indicated such that such a path will eventually be found. The meaning of these statements will be clearer if one considers nodes P and M as terminals of a diode-source network and examines the possible combinations of current and voltage at this terminal pair. As was indicated in Chapter Three, the set of terminal solutions of the solved subnetwork constitutes a breakpoint curve. Also, since the network contains no resistors, the breakpoint curve contains only horizontal and vertical line segments as is illustrated in Figure 4 - 10. Initially terminals P and M are short circuited and the network stands at point a in the figure. Point b is the intersection of the curves describing the relations imposed by the diode and by the network. Steps 4 through 9 move the solution of the solved network from point a to point b by "tracing" the breakpoint curve. The potential changes correspond to vertical motions on the curve--each increment of current along a path corresponds to a horizontal movement.

<u>Step 4</u>) Label node P. Label node j if node i is labelled and there is a branch connecting i and j which is

i) a voltage source

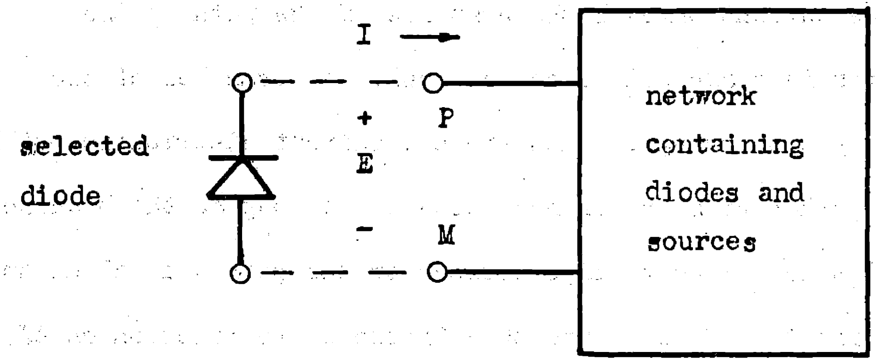
ii) a diode in the closed state

iii) a breakpoint diode directed toward node j

iv) a short circuit.

Repeat until no more nodes can be labelled, If node M is labelled, proceed with Step 5. If not, go to Step 8.

If node M is labelled the situation is as shown in Figure 11.



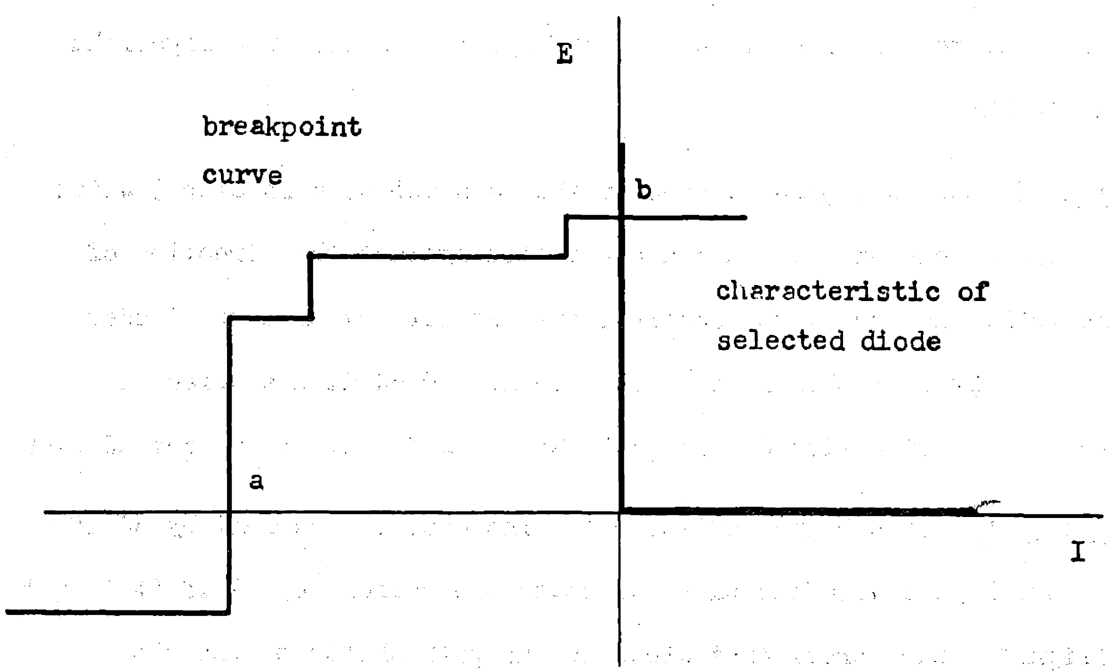


Figure 4 - 10—Breakpoint curve of the solved subnetwork viewed from the selected diode.

Step 5) Select a path running from node P to node M which contains only voltage sources, closed diodes, and breakpoint diodes alligned with the direction of the path.

Such a path must exist or else node M could never have been labelled. Now consider increasing the current through the branches of the path by the amount ΔJ. If the potentials are held constant, only the branches of the path will be affected by this change. Thus the new current distribution will still be a solution together with the branch voltages so long as the branches of the path remain satisfied. The voltage sources in the path will allow an arbitrary amount of current to flow, hence they impose no restriction on ΔJ. The breakpoint diodes and the closed diodes alligned with the direction of the path also allow an arbitrary increase in current. On the other hand, each closed diode oriented against the path places a limit on the amount of increase in current along the path. Each of these diodes will allow an amount of increase equal to the present current through the diode. The algorithm proceeds as follows:

Step 6) Consider those diodes of the path selected in Step 5 which are in the closed state and are oriented against the direction of the path. Let ΔH be the smallest current for any of these diodes. If ΔH is greater than I, or there are no closed diodes oriented against the direction of the path, set ΔJ = I. Otherwise set ΔJ = ΔH.

Step 7) For each branch of the path, increase the current by ΔJ if it is alligned with the path--decrease the current by ΔJ if it is not alligned. Each breakpoint diode in the path shifts to the closed state. Some of the closed diodes may move to the breakpoint state. If ΔJ = I a solution of the subnetwork has been found which satisfies the selected diode. Consider a new subnetwork in which this diode

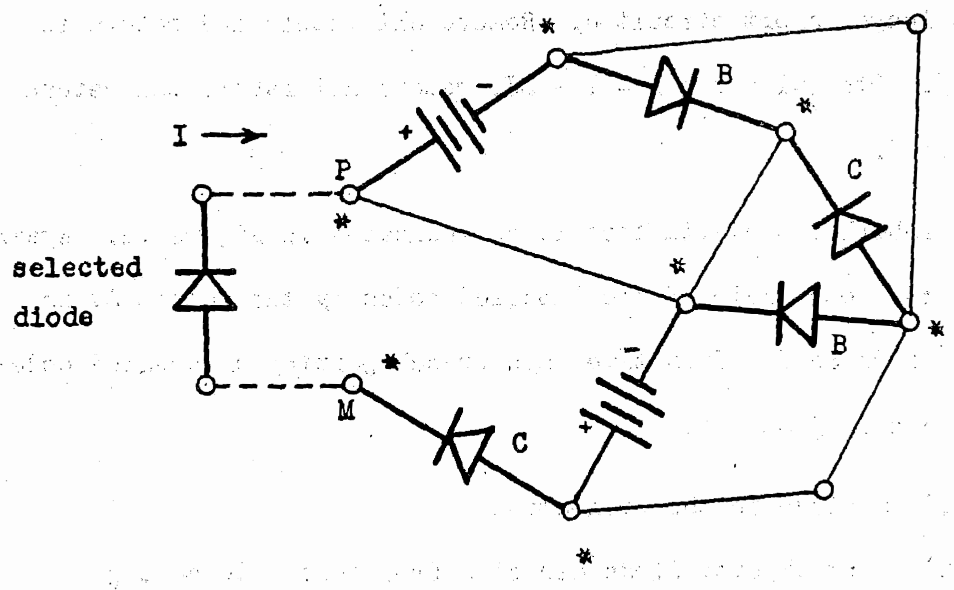


Figure 4 - 11--The situation if node M is labelled.

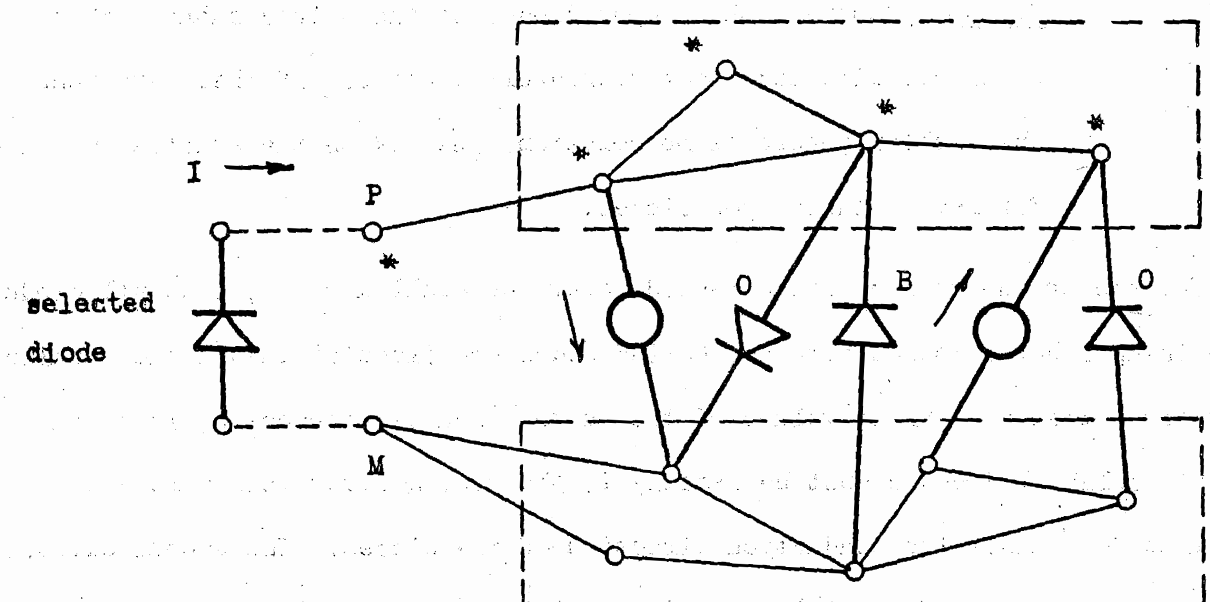


Figure 4 - 12--The situation if node M is not labelled.

is no longer short circuited. Remove all labels and return to Step 3. Otherwise set $I = I - \Delta J$, remove all labels and return to Step 4.

If node M is not labelled the situation is as indicated in Figure 12. Again consider raising the potentials of the labelled nodes by the amount ΔV as in Steps 6 and 7 of Phase I. This time each branch joining a labelled node i to an unlabelled node j must be

i) a diode in the open state

ii) a breakpoint diode directed from node i to node j

iii) a current source.

Otherwise node j should have been labelled. As before the open diodes directed toward unlabelled nodes determine the amount by which the potentials may be raised and still constitute, with the branch currents, a solution of the solved subnetwork.

<u>Step 8</u>) Consider those open diodes of the solved subnetwork which are directed towards unlabelled nodes. If there are none the problem possesses no solution. Let ΔV be the smallest voltage for any of these open diodes.

If there are no open diodes directed toward unlabelled nodes, then all diodes joining labelled nodes with unlabelled nodes are directed toward the labelled nodes. Also the sum of values of the current sources connecting labelled with unlabelled nodes must be exactly I, that is, greater than zero and directed against the conduction direction of the diodes. The second existence condition is therefore violated and no solution of the network exists.

Step 9) Increase the potential of each labelled node by ΔV. Each breakpoint diode connecting a labelled node to an unlabelled node will switch to the open state. At least one open diode will move to its breakpoint. Remove all labels and return to Step 4.

For each change in either the potential distribution or the branch current distribution during Steps 4 through 9 of Phase II, at least one diode of the solved subnetwork changes state. Also each change in potentials or currents moves the solution point uniformly upward or to the right along the breakpoint curve. Therefore no given combination of diode states can ever be repeated. This assures that the iterations terminate in a finite number of steps--with either a solution or a demonstration that no solutions exist.

4 - 3 Second Algorithm for Diode-Source Networks

The algorithm described in the previous section first produced a solution of the voltage reduced network corresponding to a given diode-source network. In this section a second algorithm will be described which first obtains a solution of the associated current reduced network. In mathematical programming terminology, a feasible vector is found for the dual linear program and is modified until it becomes optimal.

The second algorithm is equivalent to performing the following operations on the physical diode-source network.

Phase I - Solving the current reduced network

1) Replace all diodes and voltage sources of the network by short circuits. Thus the starting network contains only current sources.

2) Exchange each diode with the corresponding short circuit, one at a time until all diodes have been repositioned in the

Phase II - Completing the solution

3) Remove all diodes from the network which have zero current through them.

4) Exchange the voltage sources for their corresponding short circuits one at a time.

5) Restore the diodes that were removed in (3), one at a time.

The precise procedure is given below. The description has been shortened as the principles are essentially the same as in the first algorithm.

Phase I - Solving the current reduced network

Steps 1 and 2) Select a tree from the current reduced network which contains none of the current sources. Assign a current distribution to this tree which satisfies the current source requirements. Short circuit those diodes of the tree which have reverse (negative) current. The resulting network is the initial solved subnetwork for Phase I.

Step 3) Select one of the short circuited diodes. If none, Phase I is complete and the present current distribution is a solution of the current reduced network. Proceed with Phase II. If there is, let P and M represent the plus and minus nodes for this diode. Let ΔI be the current through this diode.

Step 4) Label node P.

Step 5) Label node j if node i is labelled and there is a branch joining nodes i and j which is

i) a diode in the closed state

ii) a breakpoint diode directed toward node j

iii) a short circuit.

Repeat until no more nodes can be labelled. If node M is labelled, proceed with Step 6. If not the network possesses no solution.

If M is not labelled, then the second existence condition is violated.

Step 6) Select a path from node P to node M made up of closed diodes and breakpoint diodes alligned with the path. Let ΔH be the smallest current of the closed diodes alligned against the path. If ΔH ΔI or there are no such closed diodes, set $\Delta J = \Delta I$. Otherwise set $\Delta J = \Delta H$.

Step 7) For each branch of the path, increase the branch current by ΔJ if it is alligned with the path, decrease by ΔJ if it is not. Each breakpoint diode in the path will become closed. Some of the closed diodes alligned against the path may move to the breakpoint state. After the change in the current distribution some of the short circuited diodes may no longer have negative currents. Remove the short circuits from these diodes and consider them part of the solved subnetwork. Remove all labels. If $\Delta J = \Delta I$, go back to Step 3, otherwise return to Step 4.

Phase II - Completing the solution

Step 1) Consider any loop of the current reduced network which, with the current distribution given by Phase I, contains only diodes in the closed state. If there are none, proceed with Step 2. Modify the currents in the branches of the loop by adding an increment ΔI to each branch alligned with the loop and subtracting ΔI

from each branch alligned against the loop, so that one of the diodes moves to the breakpoint state. Repeat Step 1 until all such loops are eliminated.

This assures that a tree as required in Step two does exist.

Step 2) Select a full tree for the current reduced network by

a) including all closed diodes and short circuits

b) adding sufficient breakpoint diodes to form a full tree.

Reinsert the voltage sources which were replaced by short circuits in forming the current reduced network. Compute a set of node potentials which satisfies the voltage sources and the other branches of the tree. This tree plus the current sources forms the initial solved subnetwork for Phase II. Add to and consider as part of the solved subnetwork any diodes not in the tree which do not have negative voltage.

Step 3) Select a diode not in the solved subnetwork. If there are none the present branch currents and voltages form a solution of the entire network and Phase II has been completed. If there is such a diode, let P and M designate the node on which its plus and minus terminals are incident. Denote by ΔE the magnitude of the (negative) voltage of this diode.

In Steps 4 through 9 of the second algorithm a solution is produced for a new subnetwork consisting of the solved subnetwork plus the diode selected in Step three. The method, as in the first algorithm is by tracing the breakpoint curve observed between nodes P and M. The steps are identical to those in Phase II of the first algorithm except for the following alterations.

a) In figuring the change in current in a path from node P to node M in Step 6, ΔJ is always set equal to ΔH. If there are no closed diodes directed against the path, the network has no solutions by violation of the first existence condition.

b) In figuring the amount of potential change in Step 8, set ΔV equal to ΔE if ΔV is greater than ΔE or there are no open diodes directed from labelled nodes to unlabelled nodes. After performing Step 9, go back to Step 3 if ΔV is equal to ΔE.

Similar arguments to those given in the previous section demonstrate that this procedure also will terminate in a finite number of steps.

4 - 4 Altering Solution for New Parameter Values

One appropriate question concerning any algorithm for solving diode-source networks is this: Given a solution for some diode-source network, how easy is it to find a solution for the same network but with different values of the current and voltage sources? Does one have to start over from the beginning of the algorithm or can one make use of the previous solution?

In the case of the algorithms described above the previous solution is readily used as a starting point for forming a new solution. Three cases are evident depending on whether the current sources, the voltage sources, or both take on new values.

Case I) Suppose that only current source values differ between the new and previous networks. Select a full tree of the network which, for the current and voltage values for the old solution, contains all voltage sources and all closed diodes, plus sufficient breakpoint diodes to form a complete tree. Assign a current distribution to the tree such that the new current source requirements are met. Keep the old set of node potentials. Some diodes of the tree may now have negative currents. Use the procedure of Phase II of the first algorithm to satisfy these diodes and thus obtain the new solution of the network.

Case II) Suppose that only voltage source values differ between the new and previous networks. Select a full tree of the network which, for the current and voltage values of the old solution, contains all voltage sources and all closed diodes, plus sufficient breakpoint diodes to complete the tree. Keep the previous distribution of branch currents but compute a set of node potentials that satisfies the branches of the tree for the new voltage source values. Some diodes not in the tree may not have negative voltage. Apply the method of Phase II of the second algorithm to satisfy these diodes and produce the solution of the new network.

That a full tree with the right properties to start these procedures can always be found is assured by the non-redundancy assumptions. Suppose that the nodes of the network could be divided into two groups such that any branches connecting nodes in both groups is either an open diode or a current ssource. This situation could prohibit the existence of the required tree. The first non-redundancy assumption assures that not all of these branches are current sources. Therefore the potentials of one of the groups of nodes could be altered so that one of the open diodes moves to the breakpoint state. On the other hand, suppose that the set of branches containing all of the voltage sources and all closed diodes contained a loop, This also could prohibit the construction of the required tree. However the second non-redundancy assumption assures that such a loop does not contain voltage sources only. Therefore the currents in the branches of the loop can be altered so that a closed diode moves to the breakpoint state. These considerations insures that given a diode-source network and a solution, either a full tree containing all of the voltage sources and all closed diodes, plus sufficient breakpoint diodes can be selected, or the solution can be easily modified so that such a tree can be selected.

Case III) Suppose both voltage source and current source values are different in the new network. Consider an intermediate network in which only the current source values have been changed from those in the previous network. Solve this network by the method indicated under Case I. Now change the voltage source values and proceed as under Case II. The intermediate network is always solvable if the new network has a solution. This is true because the voltage reduced network for the intermediate problem is identical to that of the previous problem while the current reduced network for the intermediate problem is identical to the one for the new problem. The former must possess solution if the previous network was solvable--the latter must possess solutions if the new network is solvable. Hence the intermediate network must also be solvable.

An alternate way of handling Case III problems is to solve an intermediat network in which only the voltage source values are changed and subsequently to modify this solution for the new current source values. As above the intermediate network will always be solvable.

4 - 5 An Application to Network Flow Problems

The network flow problem described in paragraph 2 - 3 concerns finding the least cost pattern of flow through a network of branches of the type shown in Figure 4 - 13a. The problem statement is (from Chapter Two)

Network flow problem:

Minimize $\sum_{ij} C_{ij} X_{ij}$

with

$$\sum_{i} X_{ij} - \sum_{i} X_{ji} = D_j \qquad \text{continuity at nodes}$$

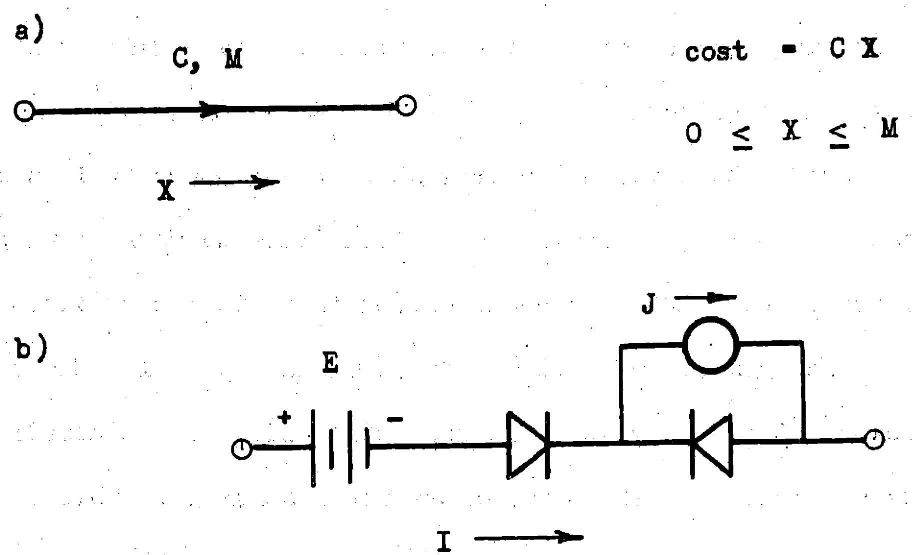


Figure 4 - 13—Linear flow branch and its electrical analogue.

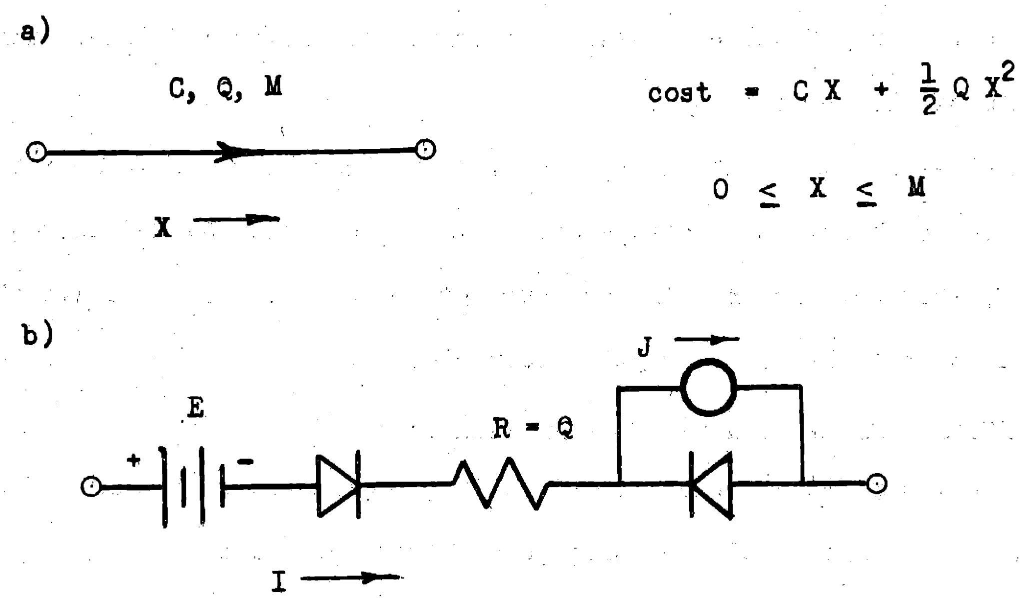


Figure 4 - 14—Quadratic flow branch and its electrical analogue.

$$0 \leq X_{ij} \leq M_{ij} \qquad \text{branch conditions}$$

Now consider the electrical branch shown in Figure 4 - 13b. Suppose an electrical network is made up of branches of this type and a current source of value H_j is associated with each node. The primal programming problem for this network as given by the principle stated in section 3 - 3 is

$$\text{Minimize} \quad \sum_{i,j} E_{ij}\, I_{ij}$$

$$\text{with} \quad 0 \leq I_{ij} \leq J_{ij}$$

$$\sum_i I_{ij} - \sum_i I_{ji} = H_j$$

Thus a complete analogy holds between the two problems and the variables are related in the following manner.

Flow variable	X	C	M	D
Electrical variable	I	E	J	H

Therefore the algorithms presented in this chapter can be applied directly to the flow network problem.

This analogy can be extended to flow networks in which a quadratic cost is associated with each branch. The corresponding electrical branch then includes a resistor as indicated in Figure 4 - 14.

C H A P T E R   F I V E

A Breakpoint Tracing Procedure

At the end of Chapter Three the terminal pair system of equations

η	x	ξ	y	v	
	A	e			b
	$-Q$		A^T	I	c
1			e^T		0

(5 - 1a)

in which the variables x and v are required to satisfy the complementary slackness condition

$$x \geq 0\ , \quad v \geq 0\ , \quad v^T x = 0 \qquad (5 - 1b)$$

was discussed. It was pointed out (and proved in Appendix G) that the solutions of such a system formed a breakpoint curve in the $\xi - \eta$ plane. In this chapter an algebraic method will be developed for "tracing" the breakpoint curve corresponding to a terminal pair system. In Chapter Six this method will be applied to general linear and quadratic programs.

A second form of terminal pair system

ξ	y	η	x	v	
	A^T	e		I	c
	$-Q$		A		b
1			e^T		0

(5 - 2)

and several other variations of (5 - 1) will also be employed in connection with the methods presented in Chapter Six. The sets of terminal solutions for all these variations have the same properties as are demonstrated in Appendix G for (5 - 1). The breakpoint tracing algorithm is formulated here with reference to the system (5 - 1). However, the method applies with little modification to any of the variations. The reader may easily work out the details.

5 - 1 The Electrical Model of the Terminal Pair System

Again, the electrical analogy will be employed to give a physical interpretation of the method. The terminal pair system 5 - 1 has the electrical model shown in Figure 5 - 1. The primal variables ξ and x are represented by currents in the model while the dual variables η, y and v are represented by voltages. The m by n rectangular matrix A becomes a rectangular array of ideal d-c transformers and Q becomes an array of resistors in the electrical analog. The column matrix e appears as a column of transformers with their secondaries connected in series. Voltage and current sources form the analog of the right hand side of the terminal pair system. In Figure 5 - 2 the diagram has been repeated, simplified by the use of symbols to represent arrays of transformers, resistors, diodes and sources. Heavy lines in the figure mean that many electrical circuits are represented — a light line means only one circuit is represented. These same conventions will be used in Chapter Six where electrical

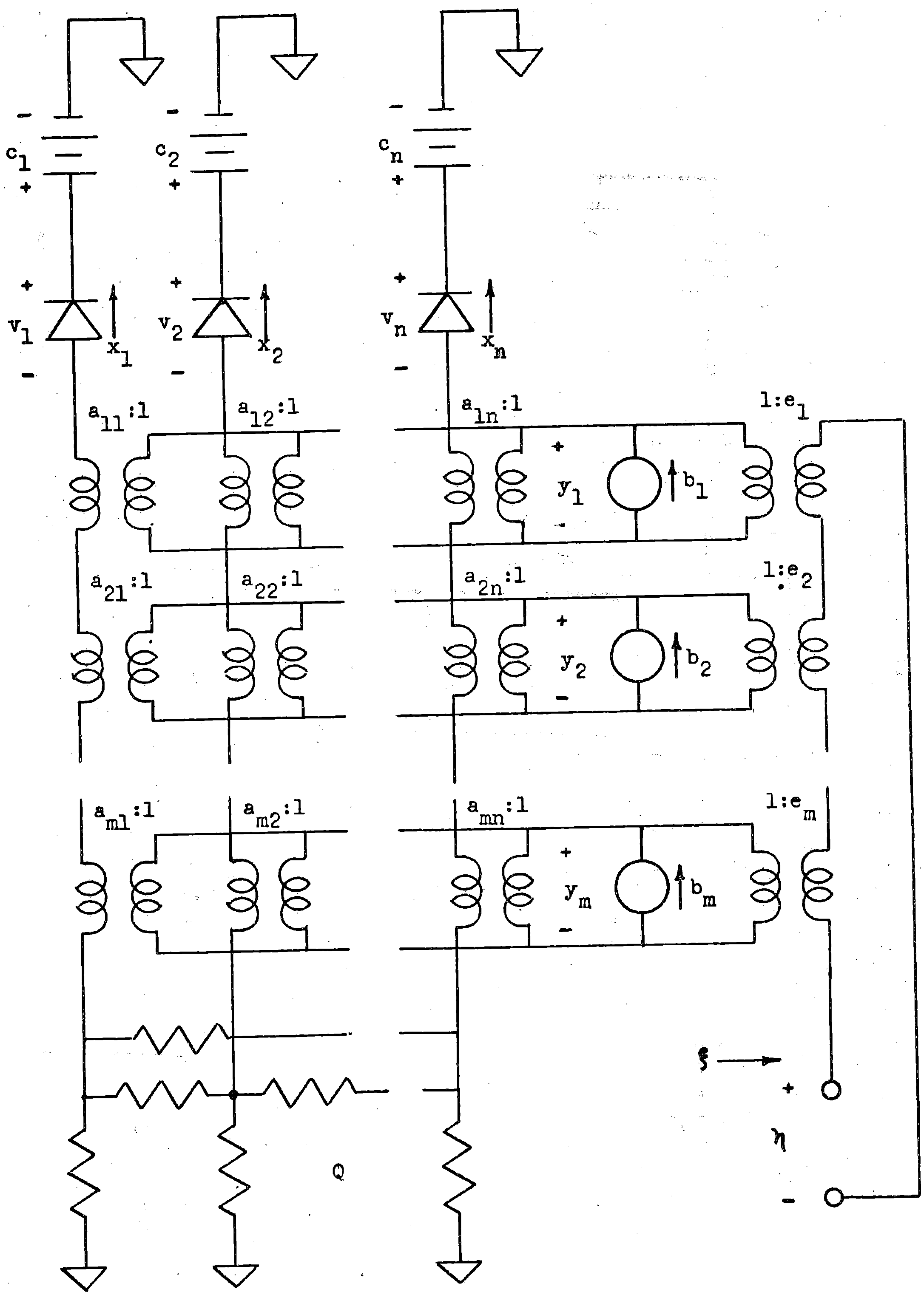


Figure 5 - 1--Electrical model of the terminal pair system.

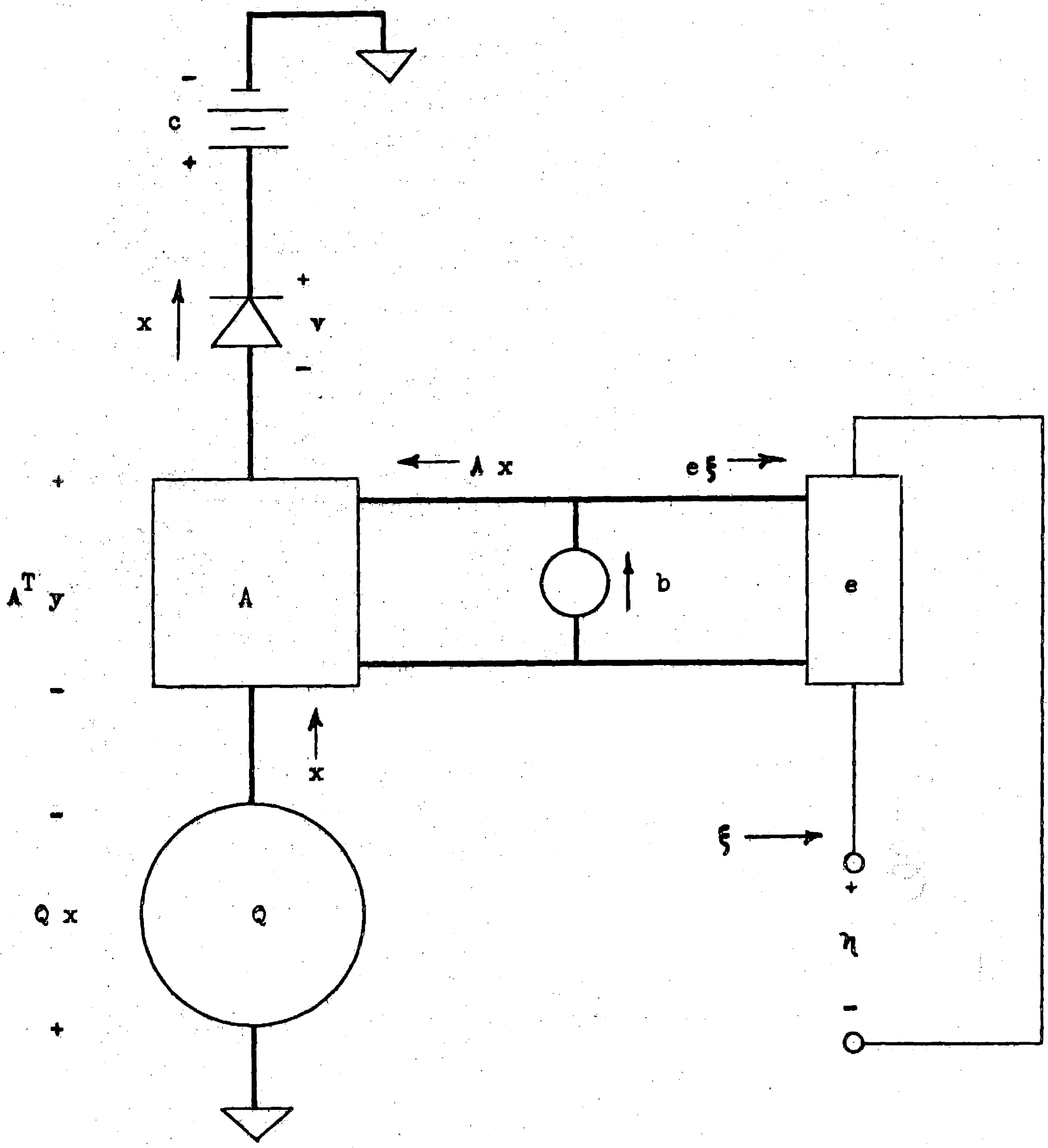


Figure 5 - 2—Simplified drawing of terminal pair system model.

models are constructed for general linear and quadratic programs.

The "black box" representation of the electrical model shown in Figure 5 - 3 will be employed in developing the breakpoint tracing scheme. Here all circuit elements with the exception of the n diodes are represented by a "black box" with n + 1 terminal pairs. One of the terminal pairs is the terminal pair of the original network with current ξ and voltage η. The diodes appear at the remaining n terminal pairs and have currents $x_1, \ldots, x_n$ and voltages $v_1, \ldots, v_n$. The variables y must now be regarded as unobservable voltages inside the black box.

5 - 2 Basic Solutions

For convenience the letter P will be used in this chapter to designate the matrix of a terminal pair system

$$P = \begin{bmatrix} P_\eta & P_X & P_\xi & P_Y & P_V \end{bmatrix} = \begin{bmatrix} 0 & A & e & 0 & 0 \\ 0 & -Q & 0 & A^T & I \\ 1 & 0 & 0 & e^T & 0 \end{bmatrix}$$

and the column vectors of variables and constants of the system will be denoted by

$$z = \begin{bmatrix} \eta \\ x \\ \xi \\ y \\ v \end{bmatrix}, \quad \text{and} \quad d = \begin{bmatrix} b \\ c \\ 0 \end{bmatrix}.$$

The equations of the terminal pair system become, with this notation,

$$P\,z = d\,. \tag{5 - 3}$$

The pair of variables (x_i, v_i) associated with the i^{th} diode of the electrical model will be called a <u>conjugate variable pair</u>. Certain solutions of a terminal pair system called <u>basic solutions</u> are particularly important in the tracing procedure. These solutions correspond to the changes in slope or <u>breakpoints</u> of the breakpoint curve. Specifically:

> A <u>basis</u> of a terminal pair system is a set B of m + n + 1 linearly independent columns of P which includes P_η, P_ξ and all the P_{y_i}, but does not contain both P_{x_i} and P_{v_i} for any i.
>
> A <u>basic solution</u> of a terminal pair system is a solution $\bar{z}$ in which the only non-zero components of $\bar{x}$ and $\bar{v}$ correspond to vectors in some basis of the system.

Since the vectors P_η, P_ξ and P_{y_i} are in every basis, exactly n - 1 of the vectors P_{x_i} and P_{v_i} must be members of each basis. In terms of the electrical model, this means that in any basic solution at least one of the diodes must have both $x_i = 0$ and $v_i = 0$, that is, must be in the breakpoint state, and we will say that this conjugate variable pair is <u>at breakpoint</u>. For the present it will be assumed that any basic solution has only one diode (conjugate variable pair) at breakpoint. The contrary is a possible situation. It is an "accidental" <u>degenerate</u> condition and will be discussed later.

The columns of a basis form an m + n + 1^{th} order square matrix. Since it was stipulated that the basis vectors were linearly independent, this matrix possesses an inverse. Let the <u>row</u> of the inverse matrix corresponding to the column P_i of the basis be represented by the <u>column</u> vector S_i. Then

$$P_i^T S_j = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}, \quad i, j \text{ in B} \qquad (5 - 4)$$

because the product of a matrix with its inverse is the identity matrix.

By means of the vectors of the basis inverse, any column of P not in the basis may be expressed as a linear combination of basis vectors, namely

$$P_j = I\,P_j = [\sum_{i \text{ in } B} P_i\, S_i^T]\, P_j$$

$$= \sum_{i \text{ in } B} (S_i^T\, P_j)\, P_i\,. \qquad (5-5)$$

5 - 3 Unit Solutions - Superposition

Let b indicate the diode which is at breakpoint for the basic solution $\bar{z}$ with basis $\bar{B}$. Consider the two variable vectors

$$\overset{o}{z} = \begin{bmatrix} \overset{o}{\eta} \\ \overset{o}{x} \\ \overset{o}{\xi} \\ \overset{o}{y} \\ \overset{o}{v} \end{bmatrix} \quad \text{and} \quad \overset{*}{z} = \begin{bmatrix} \overset{*}{\eta} \\ \overset{*}{x} \\ \overset{*}{\xi} \\ \overset{*}{y} \\ \overset{*}{v} \end{bmatrix}$$

associated with the non-basic variables x_b and v_b respectively. These vectors are defined by

$$\overset{o}{z}_i = \begin{cases} -S_i^T\, P_k\,, & i \text{ in } \bar{B} \\ 1\,, & i = k \\ 0\,, & \text{other } i\,, \end{cases} \qquad (5-6)$$

and

$$\overset{*}{z} = \begin{cases} -S_i^T\, P_h\,, & i \text{ in } \bar{B} \\ 1\,, & i = h \\ 0\,, & \text{other } i\,. \end{cases} \qquad (5-7)$$

where P_k and P_h are the columns of P associated with x_b and v_b, respectively. It follows that

$$P \overset{o}{z} = \sum_i \overset{o}{z}_i \, P_i$$

$$= P_k - \sum_{i \text{ in } \bar{B}} (S_i^T \, P_k) \, P_i \qquad (5 - 8)$$

$$= 0$$

by relation (5 - 5), and similarly

$$P \overset{*}{z} = 0 \, .$$

Thus z and z are solutions ofthe terminal pair system equations with the constant terms on the right set to zero. In terms of the electrical model $\overset{o}{z}$ and $\overset{*}{z}$ are incremental solutions of the network in the black box. That is, they are solutions of the black box when the current and voltage sources are "dead". Moreover, if the diodes which are in the open state for the basic solution $\bar{z}$ are replaced by open circuits and the closed diodes replaced by a short circuit, then $\overset{o}{z}$ is an incremental solution of this network in which a current of one ampere is forced in the terminal pair of the breakpoint diode while the voltage is maintained equal to zero by proper choice of $\overset{o}{\xi}$ and $\overset{o}{\eta}$. This is illustrated in Figure 5 - 4a. Similarly, $\overset{*}{z}$ is an incremental solution in which a unit voltage is applied to the breakpoint terminal pair while $\overset{*}{\xi}$ and $\overset{*}{\eta}$ are selected so that the current x_b is zero. See Figure 5 - 4b. The incremental solutions $\overset{o}{z}$ and $\overset{*}{z}$ are therefore called the unit solutions associated with the basic solution $\bar{z}$ and the variables x_b and v_b.

Now consider the variable vector formed by superimposing a positive multiple α of the unit solution $\overset{o}{z}$ on the basic solution $\bar{z}$

$$z(\alpha) = \bar{z} + \alpha \overset{o}{z} \, . \qquad (5 - 9)$$

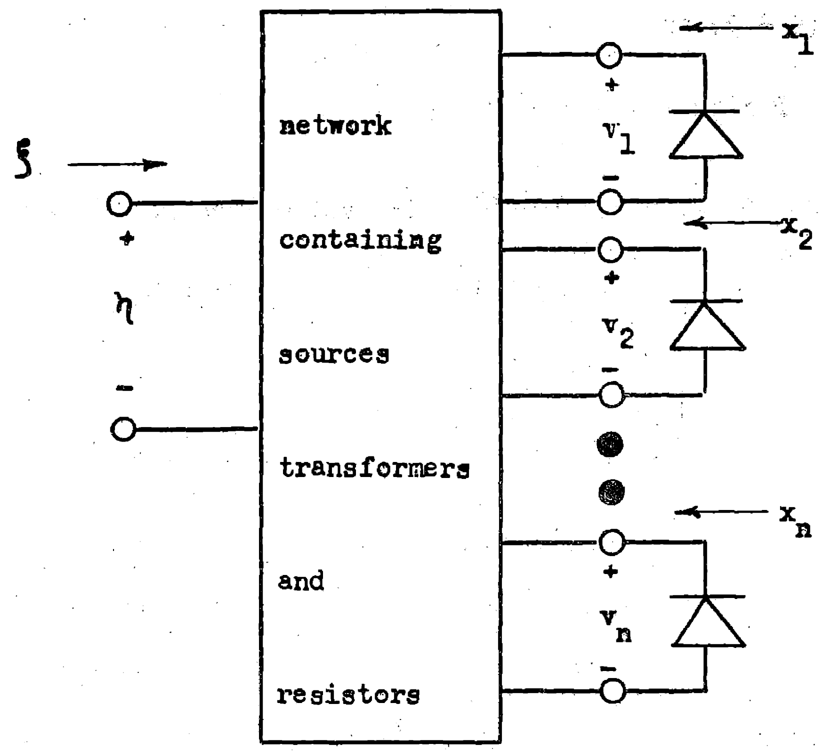


Figure 5 - 3—Black box representation of the terminal pair system.

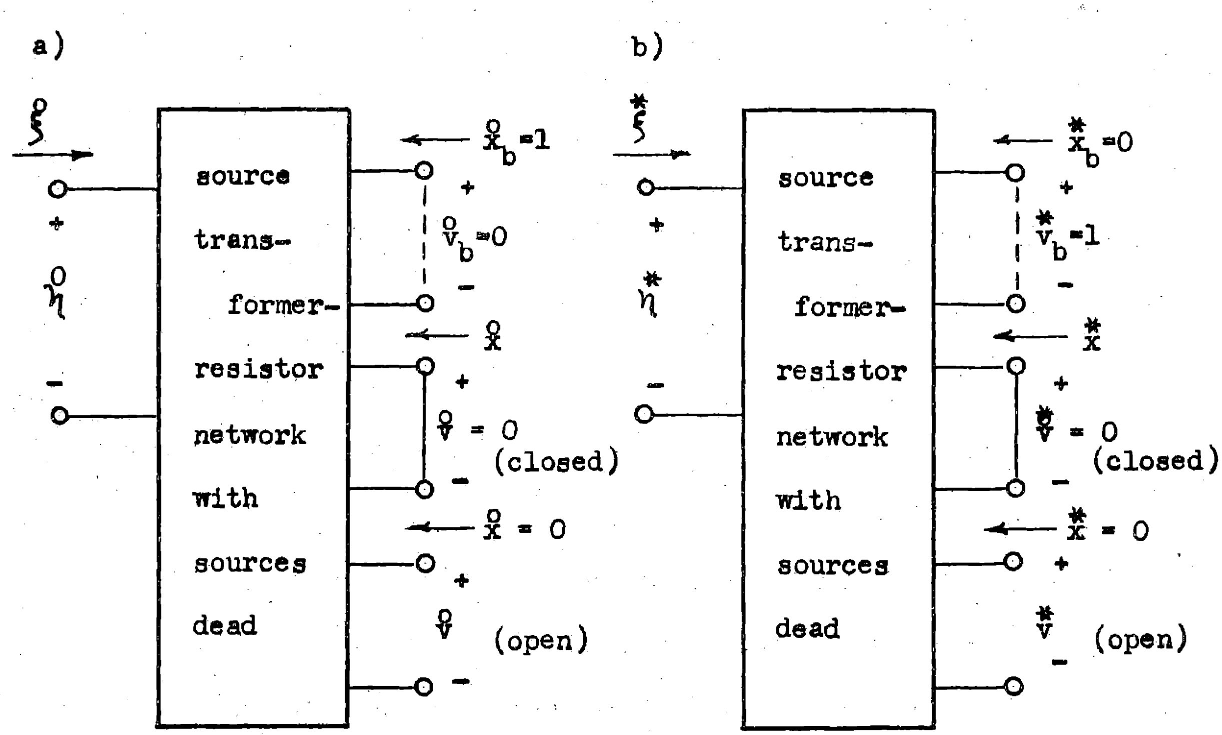


Figure 5 - 4—Unit incremental solutions of a terminal pair systema

Then, by using (5 - 3) and (5 - 8)

$$P\, z(\alpha) = P\,\bar{z} + \alpha P \overset{\circ}{z} = d \; .$$

Hence z is a solution of the terminal pair system equations. But also

$$z_i(\alpha) = \begin{cases} \bar{z}_i - \alpha\, S_i^{\,T} P_k \;, & i \text{ in } \bar{B} \\ \alpha \;, & i = k \\ 0 \;, & \text{other } i \;. \end{cases} \qquad (5 - 10)$$

Since z is a basic solution, it satisfies the complementary slackness conditions (5 - 1b). The only non-zero variable in $z(\alpha)$ which was not zero in $\bar{z}$ is $z_k(\alpha) = x_b = \alpha$. But this also satisfies the complementary condition because the variables pair $(x_b,\ v_b)$ was at breakpoint in $\bar{z}$, P_{v_b} is not in the basis and therefore $v_b(\alpha) = 0$. Under the assumption that only one diode is ever at its breakpoint, each basic x_i or v_i must be greater than zero. As a consequence each x_i of v_i or the sum $z(\alpha)$ is greater that or equal to zero for some value of α greater than zero as long as it is not too large. Thus

$$z(\alpha) = \bar{z} + \alpha \overset{\circ}{z} \qquad (5 - 11)$$

in a solution of the terminal pair system as long as α is less than some limiting value α which is greater than zero. This limiting value is given by

$$\bar{\alpha} = \underset{i}{\text{Min}} \left[\frac{\bar{z}_i}{S_i^{\,T} P_k} \right] , \qquad \begin{array}{l} i \text{ corresponding to} \\ \text{columns of } P_X \text{ or } P_V \\ \text{in } \bar{B} \text{ with } S_i^{\,T} P_k > 0. \end{array} \qquad (5 - 12)$$

If there is no index i satisfying the condition given above, it follows that (5 - 10) gives a solution of the terminal pair system for all positive values of α. This would mean that the last breakpoint of the system had been passed. Otherwise, for the value of α given by (5 - 12) one of the basic variables x_i

or v_i will become zero. Denote the column of P associated with this variable by P_r. Then for the solution $z(\overset{o}{\alpha})$ diode b has moved away from its breakpoint, while some other diode has reached its breakpoint state. There $z(\overset{o}{\alpha})$ is a new basic solution in which the vector P_k has been added to the old basis and P_r dropped. The values of the basic variables in the new basic solution are given by the formulae

$$z_i = \begin{cases} \bar{z}_i - \dfrac{S_i^T P_k}{S_r^T P_k} \bar{z}_r, & i \neq k \\ \\ \dfrac{1}{S_r^T P_k} \bar{z}_r, & i = k \end{cases}, \quad i \text{ in } B \qquad (5 - 13)$$

where B contains the vectors in the new basis.

In the same way

$$z(\alpha) = \bar{z} + \alpha \overset{*}{z} \qquad (5 - 14)$$

is found to be a solution of the terminal pair system for $0 \leq \alpha \leq \overset{*}{\alpha}$, where $\overset{*}{\alpha}$ is given by (5 – 12) with P_k replaced by P_h. Again $z(\overset{*}{\alpha})$ is a new basic solution.

On the breakpoint curve of the terminal pair system the two classes of solutions formed by (5 – 6) and (5 – 7) correspond to motion along the breakpoint curve in the directions away from the breakpoint associated with the basic solution z. It is shown in Lemma G – 7 that the unit incremental solutions $\overset{o}{z}$ and $\bar{z}$ must satisfy the relation

$$\overset{o}{\xi}\,\overset{*}{\eta} > \overset{*}{\xi}\,\overset{o}{\eta} \qquad (5 - 15)$$

Two points are made by this relation. Both $\overset{o}{\xi}$ and $\overset{o}{\eta}$ cannot be zero, nor can both $\overset{*}{\xi}$ and $\overset{*}{\eta}$. This means that the superposition (5 – 11) or (5 – 14) always moves the solution away from the breakpoint in the $\xi - \eta$ plane. Also, the slope of the line segment produced by superimposing the unit solution $\overset{*}{z}$

in which $v_b > 0$ must be greater than the slope of the segment formed with the unit solution $\overset{o}{z}$ in which $x_b > 0$. This is illustrated in Figure 5 - 5. In terms of the electrical model, diode b which is at breakpoint for the basic solution $\bar{z}$ shifts to the closed state when $\overset{o}{z}$ is added to $\bar{z}$ — or shifts to the open state when $\overset{*}{z}$ is added. Naturally the second case produces the greater incremental resistance as seen at the $\xi - \eta$ terminal pair, and hence the greater slope in the breakpoint curve.

5 – 4 Computing the New Basis Inverse

The basis for the new basic solution (5 - 13) differs from the previous one in that the vector P_k appears in the new basis while P_r does not. The rows of the inverse of the new basis S_i may be calculated from the rows of the previous inverse $\bar{S}_i$ by means of the formulae

$$S_i = \begin{cases} \bar{S}_i - \dfrac{\bar{S}_i^T P_k}{\bar{S}_r^T P_k} \bar{S}_r , & i \neq k \\ \\ \dfrac{1}{\bar{S}_r^T P_k} \bar{S}_r , & i = k \end{cases} \qquad (5-16)$$

This calculation can always be performed because $S_r^T P_k$ is always greater than zero. That the $\bar{S}_i$ formed this way actually are the inverse of the new basis may be verified by multiplying by the vectors of the new basis B.

If $i \neq k$,

$$S_i^T P_j = \bar{S}_i P_j - \frac{\bar{S}_i^T P_k}{\bar{S}_r^T P_k} \bar{S}_r^T P_j = \begin{cases} 1 , & i = j \\ 0 , & i \neq j \end{cases}$$

If $i = k$

$$S_k^T P_j = \frac{\bar{S}_r^T P_j}{\bar{S}_r^T P_k} = \begin{cases} 1 , & j = k \\ 0 , & j \neq k \end{cases} \qquad i, j \text{ in B.}$$

(Note that $\bar{S}_r^T P_j = 0$ if $j \neq k$.)

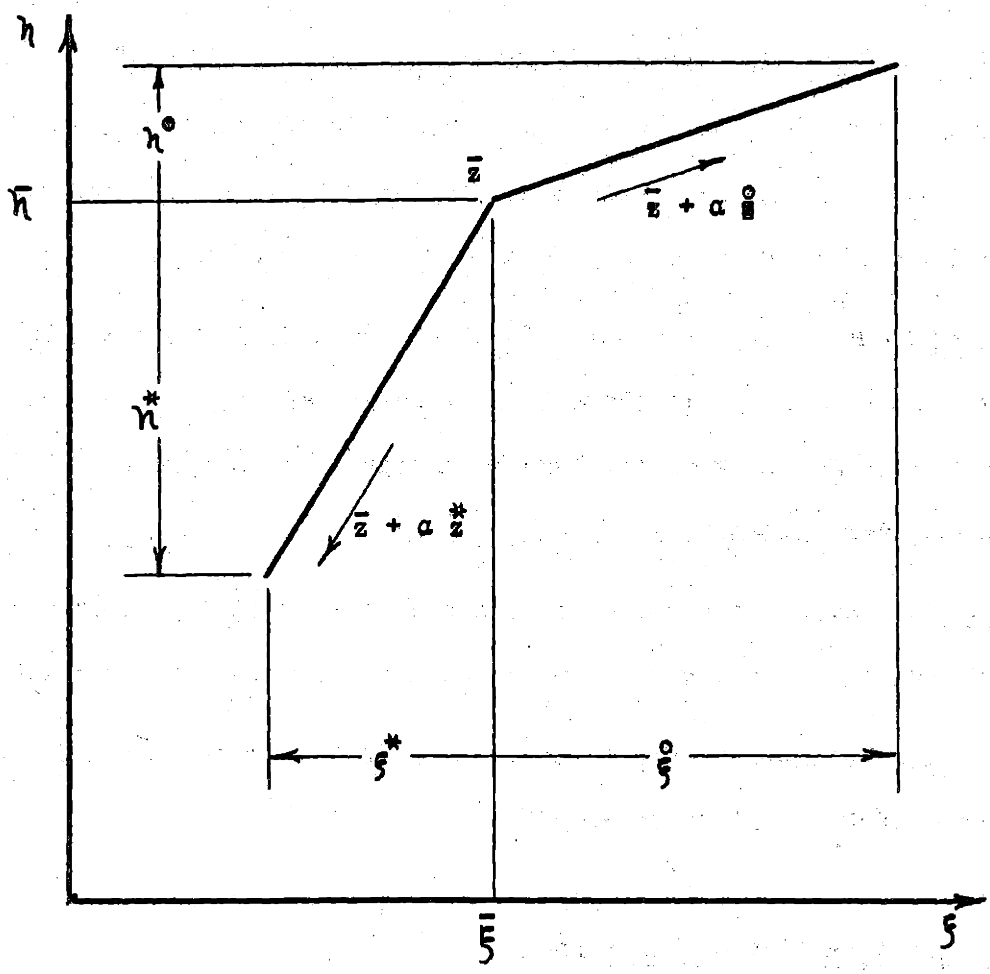


Figure 5 - 5—A breakpoint and the superposition of unit solutions.

5 - 5 Tracing the Breakpoint Curve

Suppose that one is interested in tracing the breakpoint curve of a terminal pair system P and has at hand a basic solution of the system z and the corresponding inverse vectors S_i. Then the steps involved are the following:

Step 1) Two unit incremental solutions are available at the given basic solution. Select as the first vector P_k to substitute in the basis the column associated with the incremental solution which proceeds in the desired direction along the breakpoint curve.

Step 2) Determine the vector P_r which drops from the basis by applying the rule (5 - 12). If there is no i corresponding to a vector from P_X or P_V with $S_i^T P_k > 0$, the last breakpoint has been reached and the algorithm terminates.

Step 3) Compute the values of the new basic variables using (5 - 13) and construct the new vectors of the basis inverse by means of (5 - 16).

Step 4) The next vector P_k to be substituted in the basis is the conjugate of P_r. Return to Step 2.

5 - 6 The Method in Case of Degeneracy

So far only one diode in the breakpoint state has been allowed for any basic solution. By means of this condition it is insured that there will be a unique vector P_r selected to drop from the basis in each iteration. In Appendix H the mathematical development is generalized by embedding the terminal pair system in a larger system in which, in a sense, each basic solution has only one pair of conjugate variables at breakpoint. It is shown there that the quantity corresponding to $\xi + \eta$ increases strictly with each step of the method. The only place where the procedure of the previous paragraph breaks down is in Step 2 where the vector to be dropped from the basis may not be uniquely determined. The development in Appendix H demonstrates that this procedure still works providing the rule (5 - 12) for determining the vector to be dropped in Step 2 is modified as follows:

<u>Step 2</u>) i) Compute the quantities

$$\alpha_i = \frac{\bar{z}_i}{S_i^T P_k},$$

i corresponding to vectors from P_X and P_V which are in the basis and have $S_i^T P_k > 0$.

If there are no such vectors, the last breakpoint has been reached. The vector

$$z = \bar{z} + \alpha \overset{o}{z}$$

is then a solution of the terminal pair system for all positive values of α.

ii) Let $\bar{\alpha}$ have the value of the smallest α_i computed above, and let R be the set of all i such that $\alpha_i = \bar{\alpha}$. If R contains only one member, then P_r is the corresponding vector. Otherwise take j = 1 and proceed.

iii) Evaluate the quantities

$$\alpha_{ij} = \frac{S_{ij}}{S_i^T P_k}$$

for each i in R. Let $\bar{\alpha}_j$ be the smallest of these.

iv) Remove from R all i except those for which $\alpha_{ij} = \bar{\alpha}_j$. If only one member remains in R, the corresponding vector is P_r. Otherwise repeat (iii) with j incremented by one unit. Any ties are always resolved by the time j = m + n + 1.

5 - 7 The Method When Q Is Null

Although the method developed in the preceding sections will work for any terminal pair system, the method simplifies sufficiently in certain important special cases to warrant special consideration. First, the case where the matrix Q is null will be considered. As was pointed out in section 3 - 7 the breakpoint curve will then consist of only horizontal and vertical line segments. The terminal pair system becomes the uncoupled system

η	x	ξ	y	v	
	A	e			b
			A^T	I	c
1			e^T		0

Consider the properties of a basis for this system. A basis must contain precisely m + n + 1 columns of the system matrix P (where A has m rows and n columns). However, a basis can contain no more than m linearly independent vectors from $[P_X \quad P_\xi]$ because each of these is zero except for the first m components. Similarly, a basis can contain no more than n + 1 vectors from $[P_\eta \quad P_Y \quad P_V]$. It follows that:

> Any basis for a terminal pair system in which Q is null must contain exactly m vectors from $[P_X \quad P_\xi]$ and exactly n + 1 vectors from $[P_\eta \quad P_Y \quad P_V]$.

Thus such a basis has the form

$$B = \begin{bmatrix} B_P & 0 \\ 0 & B_D \end{bmatrix} .$$

It is made up of a primal basis B_P containing m linearly independent columns from

$$[A \quad e]$$

and a <u>dual basis</u> B_D containing n + 1 linearly independent columns from

$$\begin{bmatrix} 0 & A^T & I \\ 1 & e^T & 0 \end{bmatrix}.$$

The inverse of this basis is obviously

$$\begin{bmatrix} B_P^{-1} & 0 \\ 0 & B_D^{-1} \end{bmatrix}$$

An immediate result of the special structure of the basis inverse concerns the unit solution associated with a non-basic vector P_k:

$$\overset{o}{\varrho}_i = \begin{cases} -S_i^T P_k, & i \text{ in } B \\ 1, & i = k \\ 0, & \text{other } i \end{cases}$$

If P_k is a column of P_X, then $\overset{o}{\eta} = \overset{o}{y} = \overset{o}{v} = 0$; if P_k is a column of P_V, then $\overset{o}{\xi} = \overset{o}{x} = 0$. Therefore on a step of the algorithm in which a column of P_X is substituted in the basis, only ξ and x may change while η, y and v must remain fixed. Hence the vector which drops from the basis must also be a column of P_X. Likewise, on a step in which a column of P_V is introduced, η, y and v may change but ξ and x are fixed. The vector dropped must be from P_V. Since the vector introduced on one step is the conjugate of the vector dropped on the previous step, it follows that the algorithm alternately selects the vector to be introduced from P_X and P_V. Furthermore the primal basis B_P stays fixed on a step in which the dual variables η, y and v are changed while the dual basis remains the same when the primal variables change.

Next it will be shown that it is only necessary to employ one of the inverses B_P^{-1} or B_D^{-1} in the breakpoint tracing computation. Suppose there is at hand a basic solution $\bar{z} = (\bar{\eta}, \bar{x}, \bar{\xi}, \bar{y}, \bar{v})$ with corresponding basis

$$\bar{B} = \begin{bmatrix} \bar{B}_P & 0 \\ 0 & \bar{B}_D \end{bmatrix},$$

but only the inverse vectors $(\bar{S}_P)_i$ of $\bar{B}_P$. Let $(P_P)_i$, i in $\bar{B}_P$, be the columns of $P_P = [A \quad e]$ which are in the primal basis $\bar{B}_P$. Then

$$(S_P)_i^T \, (P_P)_j = \begin{cases} 1\,, & i = j \\ 0\,, & i \neq j \end{cases}, \quad i, j \text{ in } \bar{B}_P$$

Suppose also that the basic solution $\bar{z}$ has been reached by a primal step in which the primal vector $(P_P)_r$ dropped from the basis. Now note that the vector $(P_P)_\xi = e$ is always a member of B_P, and consider the quantities

$$\overset{*}{y} = -(\bar{S}_P)_\xi$$

$$\overset{*}{v} = A^T(S_P)_\xi = A^T \overset{*}{y}$$

$$\overset{*}{} = e^T(\bar{S}_P)_\xi = -e^T \overset{*}{y} \;.$$

One has

$$\overset{*}{v}_i = -(P_P)_{x_i}^T \, (\bar{S}_P)_\xi = 0 \qquad \text{if } x_i \text{ is in } \bar{B}_P \,,$$

and

$$\overset{*}{\eta} = (P_P)_\xi^T \, (\bar{S}_P)_\xi = 1 \,.$$

It is apparent that $\overset{*}{z} = (\overset{*}{\eta}, 0, 0, \overset{*}{y}, \overset{*}{v})$ is a unit incremental solution of the terminal pair system.

Since a primal step has just been completed, the unit solution $\overset{*}{z}$ in which all primal variables are zero must be the unit solution to be superimposed on $\bar{z}$ to accomplish the dual step. Since $\overset{*}{\eta} > 0$, $\overset{*}{z}$ should be added to $\bar{z}$ if the breakpoint curve is being traced toward positive ξ and/or η and should be subtracted if the opposite course is being followed. In the former case the dual step of the tracing method is given by the formulae

$$y(\bar{\alpha}) = \bar{y} - \alpha(\bar{S}_P)_\xi$$
$$v_i(\bar{\alpha}) = \bar{v}_i , \qquad x_i \text{ in } \bar{B}_P$$
$$v_i(\bar{\alpha}) = \bar{v}_i - (P_P)_{v_i}^T (S_P)_\xi , \quad x_i \text{ not in } \bar{B}_P$$
$$\eta(\bar{\alpha}) = \bar{\eta} + \bar{\alpha} ,$$

where

$$\bar{\alpha} = \text{Min} \left[\frac{\bar{v}_i}{(P_P)_{v_i}^T \ (\bar{S}_P)_\xi} \right] ,$$

i such that x_i is not in $\bar{B}_P$.

As presented above, the specialization of the breakpoint tracing algorithm does not have provision to handle the problem of degeneracy. Unfortunately, the author has been unable to specialize the rule given in section 5 - 6 so that the information in the primal basis alone is sufficient to make all of the required decisions. For more discussion on this point, see Chapter Eight.

The method described in this section can also be formulated so as to make use of information contained in the dual basis B_D only.

5 - 8 The Case When Q and b or c Are Null

The second special case of interest is when one of the sets of constants b or c is null as well as Q. In the electrical model c being null is equivalent to replacing the voltage sources with short circuits. The resulting system will be called a primal reduced terminal pair system. It has a breakpoint curve in which (Theorem G - 3) all horizontal line segments are coincident with the horizontal axis, and therefore must have one of the forms shown in Figure 5 - 6. As a consequence all breakpoints must lie on the vertical axis. Since this case is a further specialization of the uncoupled system discussed in the previous section, the results obtained there also apply to the present case. In particular, for any basis B of the system, the dual portion B_D must contain exactly n + 1 linearly independent columns from

$$\begin{bmatrix} 0 & A^T & I \\ 1 & e^T & 0 \end{bmatrix}$$

The corresponding basic solution z must be such that

$$\sum_{i \text{ in } B_D} z_i \, (P_D)_i = 0 \, .$$

This can only be true if all of the dual variables η, y and v are zero. This means that the dual step of the tracing method does not change the values of the dual variables at all, but merely indicates which primal vector is to be entered into the basis next.

Similar remarks hold for the <u>dual reduced terminal pair system</u> in which both Q and b are null. In that case, the primal variables x and ξ are zero for any basic solution. If Q, b and c are all zero, then the terminal pair system is <u>completely degenerate</u> and all breakpoints are at the origin of the $\xi - \eta$ plane. Then the breakpoint curve must have one of the simple forms shown in Figure 5 - 7. Note that this is an extremely degenerate case for all pairs of conjugate variables are at breakpoint for each basic solution.

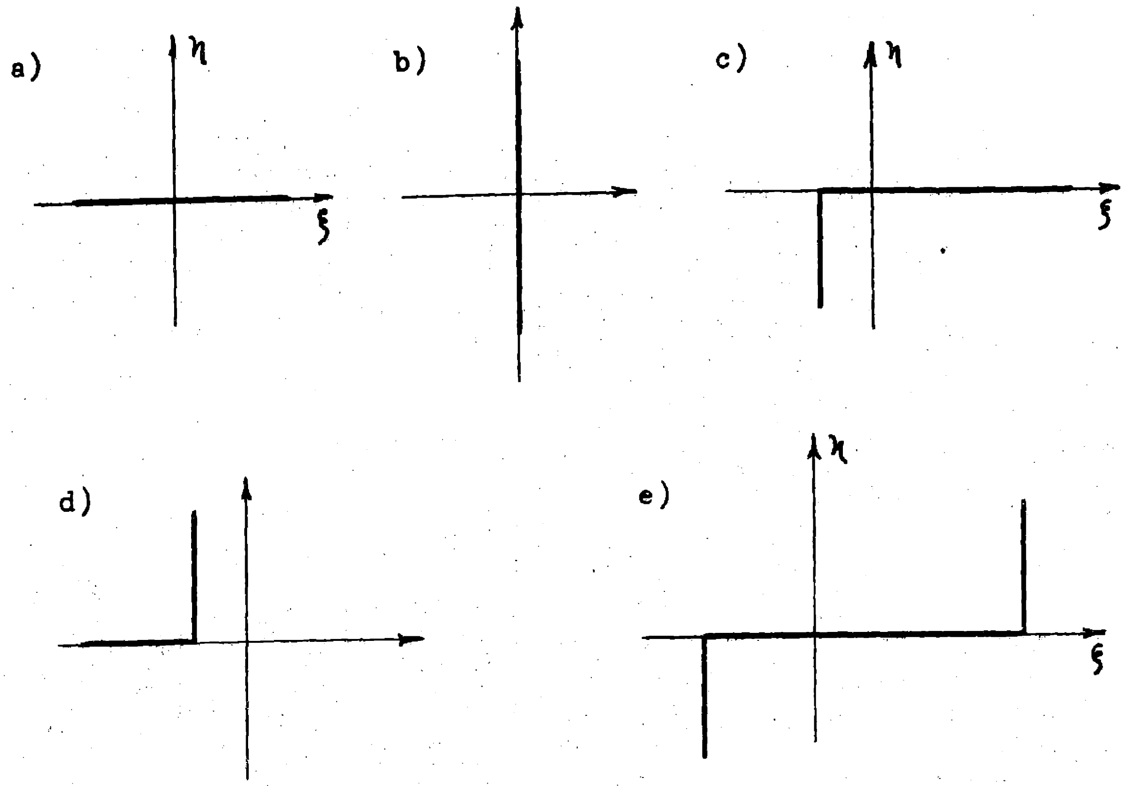


Figure 5 - 6—Possible breakpoint curves for a primal reduced terminal pair system.

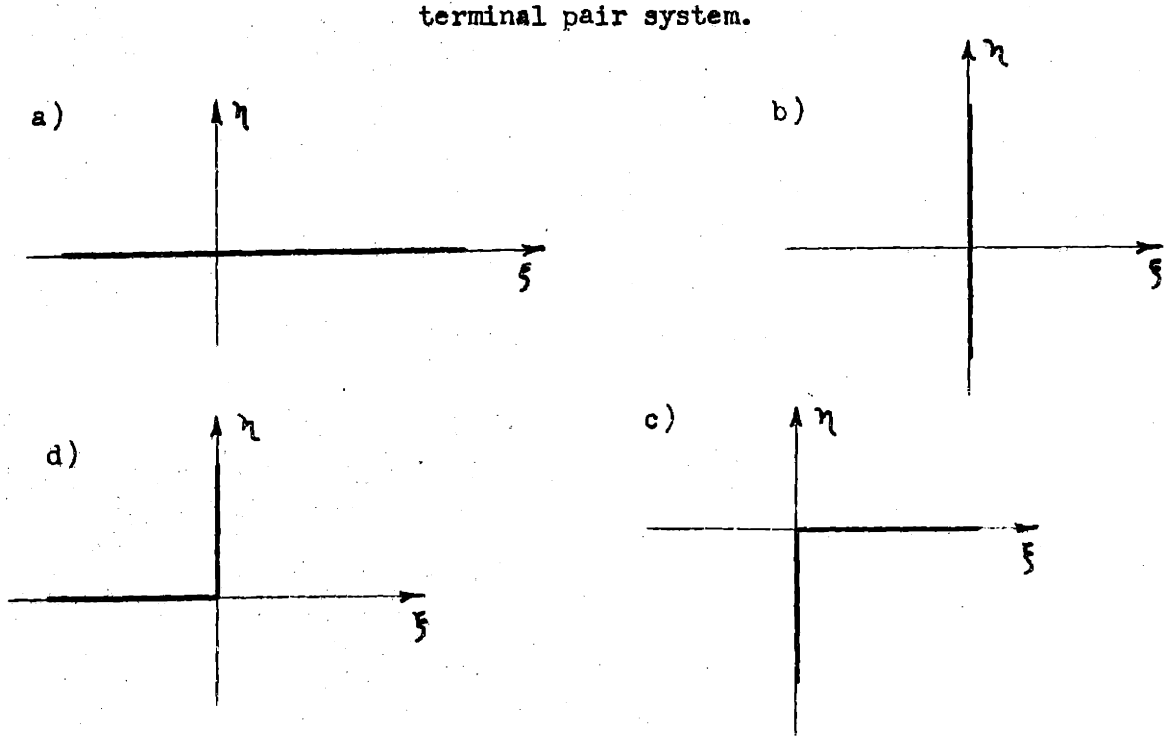


Figure 5 - 7—Possible breakpoint curves for a completely degenerate terminal pair system.

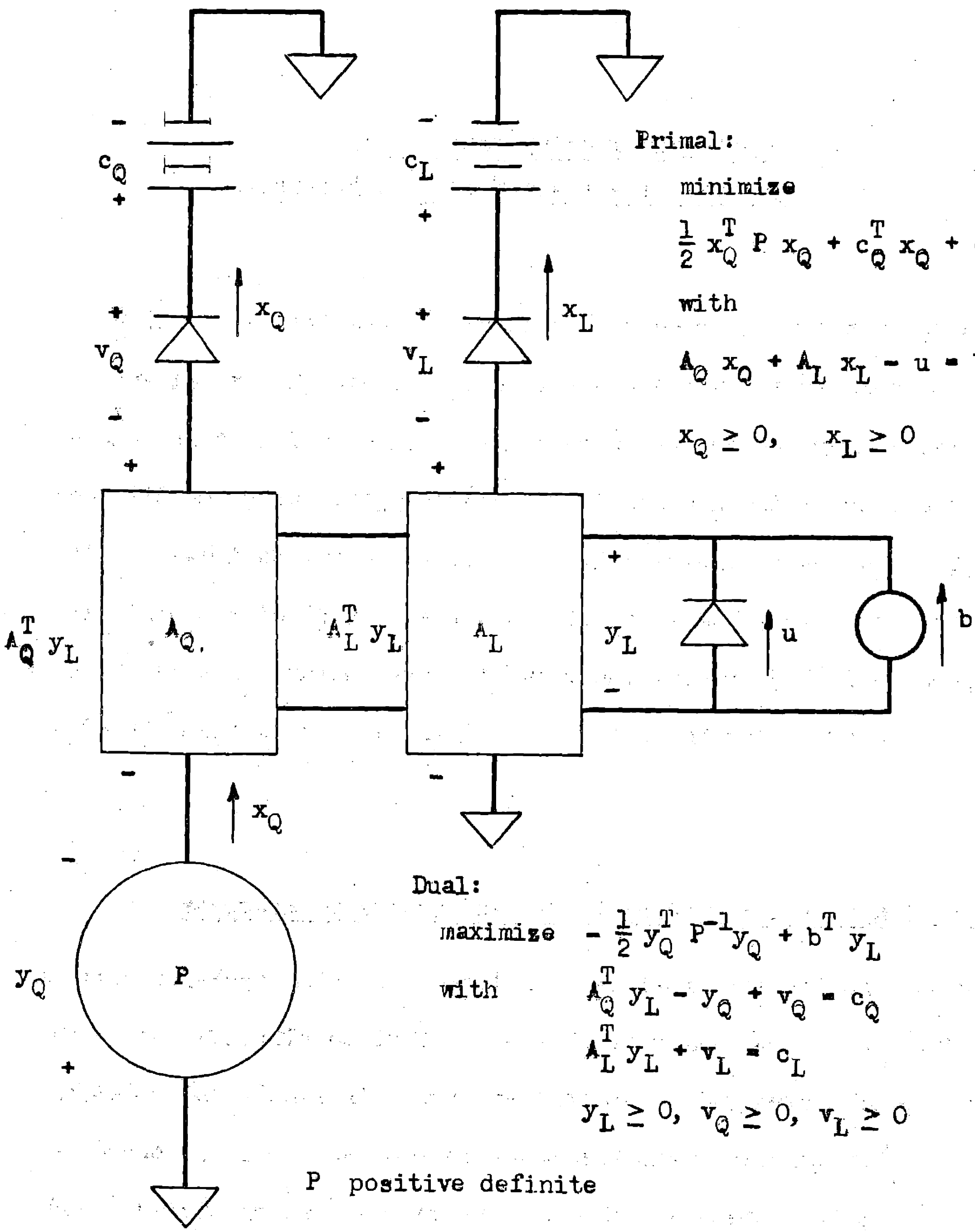
CHAPTER SIX

Breakpoint Tracing Methods for General Linear and Quadratic Programs

The breakpoint stepping procedure developed in Chapter Five will be employed here to construct several algorithms for obtaining optimal solutions to general linear and quadratic programs. To accomplish this, electrical network models of the general linear and quadratic programs will be formulated. Terminal pair systems will be constructed around these models in such a manner that i) a basic solution and the vectors of the basis inverse are obvious, and ii) tracing the resulting breakpoint curve leads to a solution of the network and hence to the optimal vectors of the dual programs.

6 – 1 Electrical Models for General Linear and Quadratic Programs

In order to simplify the drawings of networks in which banks of similar elements appear, the symbols employed in Chapter Five to simplify the model of the terminal pair system will also be used here. An electrical model of the dual pair of quadratic programs discussed in section 2 – 10 is shown in Figure 6 – 1. The primal variables x and u are represented by currents and the dual variables y and v by voltages. By inspecting the figure it is evident that any solution of the network model is also a solution of the Lagran-



Primal:

minimize

$$\frac{1}{2} x_Q^T P x_Q + c_Q^T x_Q + c_L^T x_L$$

with

$$A_Q x_Q + A_L x_L - u = b$$

$$x_Q \geq 0, \quad x_L \geq 0$$

Dual:

maximize $$-\frac{1}{2} y_Q^T P^{-1} y_Q + b^T y_L$$

with $$A_Q^T y_L - y_Q + v_Q = c_Q$$

$$A_L^T y_L + v_L = c_L$$

$$y_L \geq 0, \quad v_Q \geq 0, \quad v_L \geq 0$$

P positive definite

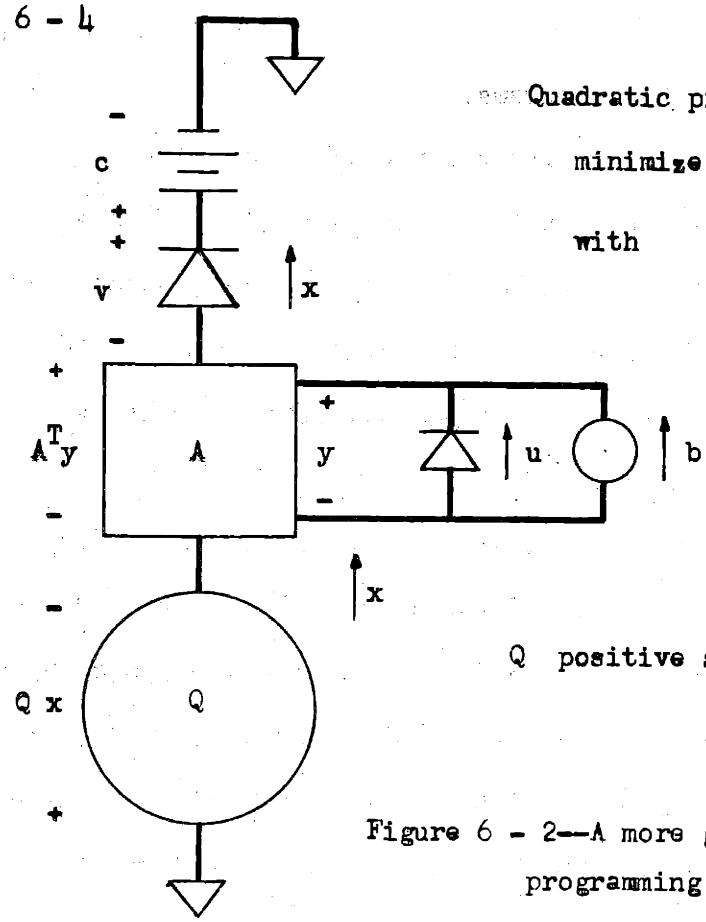
Figure 6 – 1—Electrical model of the dual pair of quadratic programs.

gian problem associated with the dual programs.

In correspondence with the physical interpretations given in Chapter Three, the electrical sources represent the constant vectors of the constraint inequalities, the diodes embody the principle of complementary slackness and the primal-dual coupling relation appears as the resistance array P. The conservation of electrical power is equivalent to the equality of objective functions required by the duality theorem.

The voltage reduced network associated with the model is obtained by "open circuiting" the resistance array P (setting $x_Q = 0$) and setting the the current source values to zero. Its significance, as was mentioned in Chapter Three, is that ($x_Q = 0$, $x_L = 0$, y_Q, y_L) is a solution of the voltage reduced network if and only if (y_Q, y_L) is a feasible vector of the dual programming problem. Similarly, the current reduced network is obtained by " short circuiting" the resistance array and setting the voltage source values to zero, and it has a solution (x_Q, x_L, $y_Q = 0$, $y_L = 0$) if and only if (x_Q, x_L) is a feasible vector of the primal program.

Note that for any quadratic program stated in the form given in Figure 6 - 1, the duality principle shows that any method for solving the primal formulation is in reality two methods, for it could also be applied to the dual formulation. This is not the case for the more general quadratic programming formulation stated and modelled in Figure 6 - 2. There the primal program yields a Lagrangian problem which gives the correct conditions for a solution of the network model. However, it is not clear how to construct a dual formulation for this more general problem. (It is possible by means of a change of variables to put any quadratic program with a positive semidefinite objective function into the formulation of Figure 6 - 1. But a neater resolution of the question is to be desired.)



Quadratic program:

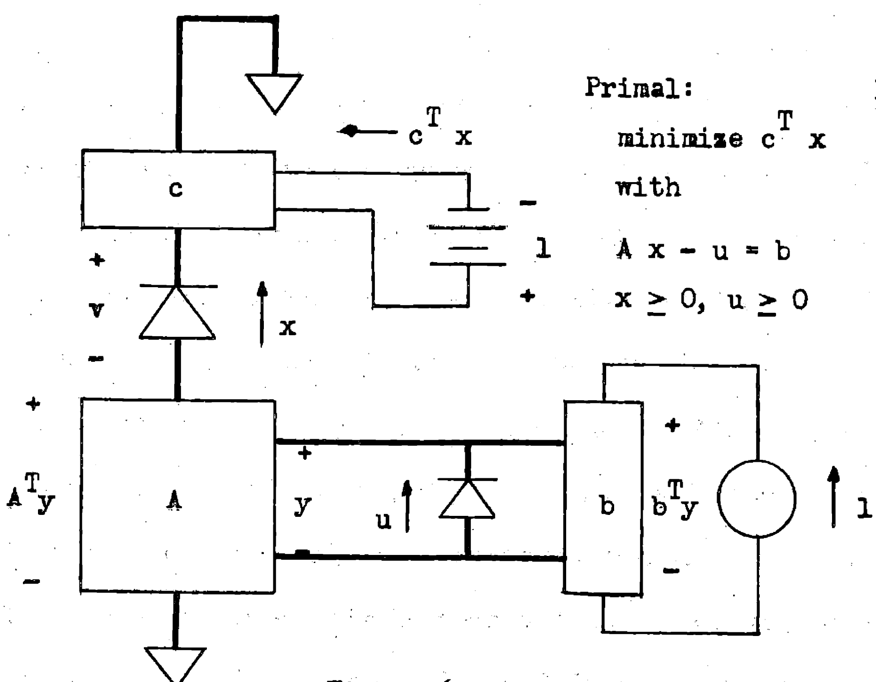
minimize $\frac{1}{2} x^T Q x + c^T x$

with $A x - u = b$

$x \geq 0, \quad u \geq 0$

Q positive semidefinite

Figure 6 - 2—A more general quadratic programming model.



Primal:	Dual:
minimize $c^T x$	maximize $b^T y$
with	with
$A x - u = b$	$A^T y + v = c$
$x \geq 0, \; u \geq 0$	$y \geq 0, \; v \geq 0$

Figure 6 - 3—Second linear programming model.

Figure 6 - 1 becomes a model of the dual pair of linear programs when the left side is discarded. A second model of the dual linear programs is formed by replacing the banks of voltage and current sources with transformer banks (Figure 6 - 3). This produces a model with just one voltage source and one current source. This device will enable us to design one algorithm for linear programming which apparently has no counterpart in quadratic programming.

In the following sections two types of algorithms for solving the network models will be discussed. In the first type the sources of the network are initially set to zero where the network has the obvious solution of zero current and voltage. The methods proceed by gradually increasing the source value to their assigned levels while keeping the network laws satisfied. It is as if one gradually "turned on the juice" with a potentiometer. They will therefore be called valve algorithms. In the second type the sources are maintained at the assigned values, but the effect of the sources is kept from the network by being shunted through a by-pass. The algorithm then gradually transfers the effect of the sources from the by-pass to the network. These will be called by-pass algorithms. The valve algorithms are based on the second electrical model for the general linear program -- the by-pass algorithms are based on the model of Figure 6 - 1.

6 - 2 A Valve Algorithm for Solving the Current Reduced Network

First the breakpoint tracing method will be applied to determining a solution of the current reduced network associated with the network model of a quadratic or linear program. In programming terminology this is equivalent to forming a feasible vector for the primal constraints

$$\begin{aligned} Ax - u &= b \\ x \geq 0 , \quad u &\geq 0 . \end{aligned} \qquad (6 - 1)$$

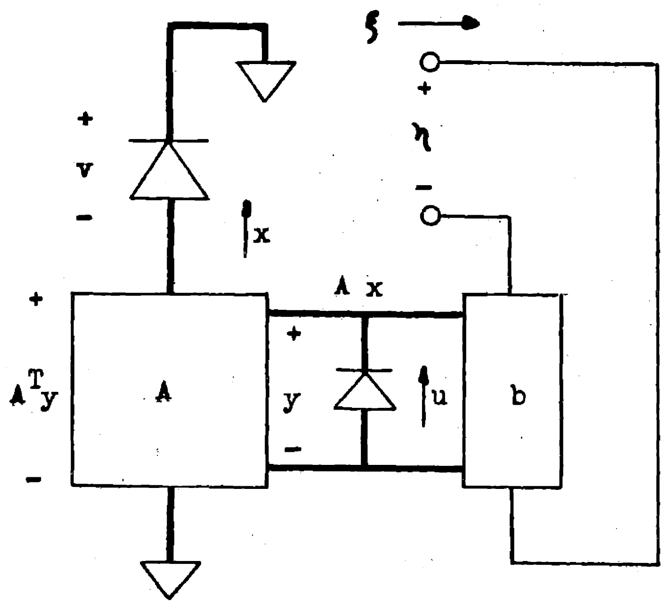
Consider the terminal pair system shown in Figure 6 – 4. Since the model contains no voltage or current sources, this system is completely degenerate and all its breakpoints are at the origin of the $\xi - \eta$ plane. Now suppose that the constraint relations (6 – 1) are feasible, that is, there is some $(\bar{x}, \bar{u})$ satisfying (6 – 1). Then $(\eta = 0, \bar{x}, \xi = 1, y = 0, \bar{u}, v = 0)$ is a solution of the terminal pair system with the terminal solution $(\xi = 1, \eta = 0)$. Conversely, if the system has a terminal solution (1, 0), then the corresponding x and u form a feasible vector of (6 – 1). It follows that

> The constraint relations (6 – 1) are feasible if and only if the breakpoint curve of the terminal pair system in Figure 6 – 4 passes through the point (1, 0).

A convenient basic solution of the terminal pair system is available by inspection, namely with the basic variables $\eta = 0$, $u = 0$, $v = 0$. With this choice the basis vectors B are the unit vectors of the $m + n + 1^{th}$ order identity matrix and the basis inverse S likewise consists of unit vectors. Note that the vector P_ξ is not in the basis as was required by the development in Chapter Five. In all other respects this is a proper basic solution. The one objection will be met by choosing P_ξ as the vector to be substituted into the basis to initiate the breakpoint tracing process. Note that the unit solution

$$\overset{o}{s}_i = \begin{cases} -S_i^T P_\xi, & i \text{ in } B \\ 1, & i = \xi \\ 0, & \text{other } i \end{cases}$$

corresponding to this substitution has $\overset{o}{\xi} = 1$ and therefore automatically starts the tracing in the correct direction along the breakpoint curve. Of course, because of the degenerate nature of the terminal pair system the



	η	x	ξ	y	u	v	
P		A	−b		−I		0
				A^T		I	0
	1			$-b^T$			0

Figure 6 - 4—Valve method for solving the current reduced network.

generalized procedure given in section 5 - 6 must be used.

The tracing process is carried on until Step 2 fails to indicate a vector to be dropped from the basis and a basic solution $\bar{z} = 0$ and unit solution $\overset{*}{z}$ are available such that

$$z = \bar{z} + \alpha \overset{*}{z} \qquad (6 - 2)$$

is a solution of the system for all $\alpha \geq 0$. There are two possibilities — since the terminal pair system is completely degenerate, the ray (6 - 2) is either the positive ξ-axis or the positive η-axis. If $\overset{*}{\eta} > 0$ the latter case holds, the breakpoint curve does not pass through the point (1, 0) and there is no feasible vector for (6 - 1). If $\overset{*}{\xi} > 0$, then $(\overset{*}{x}/\overset{*}{\xi}, \overset{*}{u}/\overset{*}{\xi})$ is the desired feasible vector.

Recall that a basis for a system in which Q is null must have the form

$$B = \begin{bmatrix} B_P & 0 \\ 0 & B_D \end{bmatrix} .$$

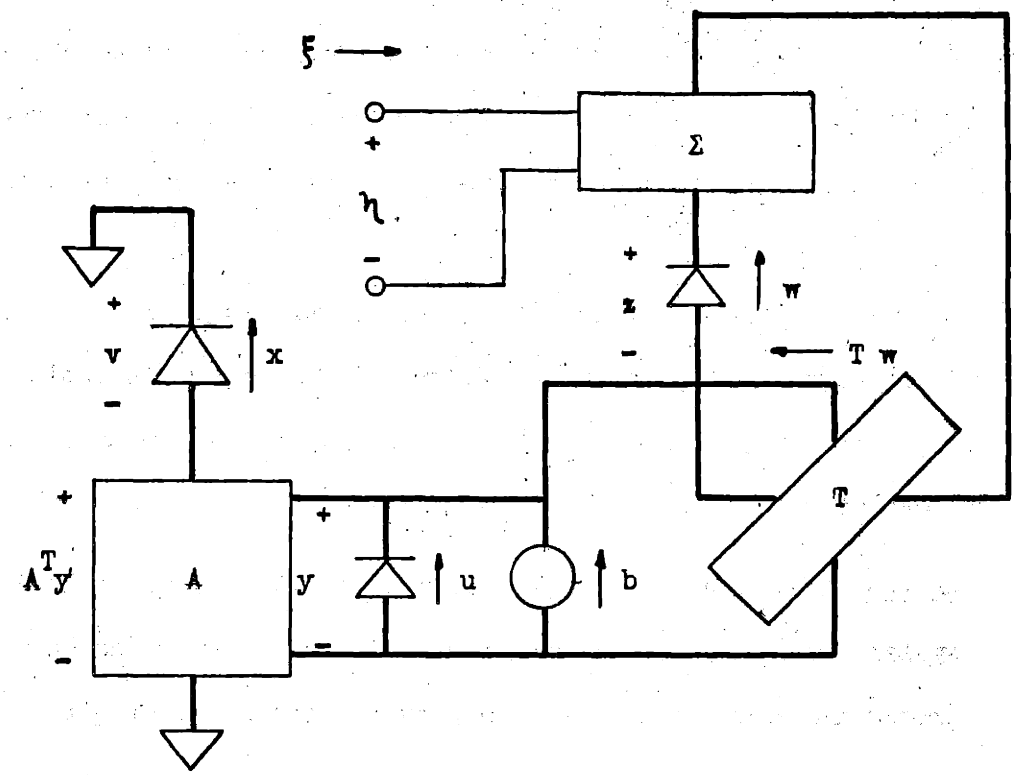
As a consequence the primal basis B_P associated with the basic solution $\bar{z}$ must consist of m linearly independent vectors from the primal vector set

$$[A \quad I]$$

The fact that this primal basis and its inverse S_P are available at the conclusion of the tracing process will be utilized in initiating the algorithms described in sections 5 - 5 and 5 - 6.

6 - 3 A By-Pass Algorithm for Solving the Current Reduced Network

A second way of employing the breakpoint tracing technique in solving the current reduced network is illustrated in Figure 6 - 5. In this terminal pair system T is a diagonal matrix consisting of plus and minus ones such



	ξ	y	η	x	w	v	z	u	
P		A^T	0			I			0
		T	Σ				I		0
				A	T			−I	b
	1			0	Σ^T				0

Figure 6 - 5—By-pass method for solving the current reduced network.

that

$$T_{ii} = \begin{cases} +1 & \text{if} \quad b_i \geq 0 \\ -1 & \text{if} \quad b_i < 0 \end{cases}$$

In the electrical model T is represented by a diagonal bank of one-to-one d-c transformers. Each component of the special column vector Σ is one so that $\Sigma^T w$ is the sum of the components of w. Here the transformer bank representing Σ has its secondaries connected in parallel since currents are being summed.

The terminal pair system is of the primal reduced form as it contains current sources but no voltage sources. Therefore its terminal solution set has all its breakpoints along the ξ-axis. By inspection of the terminal pair system it is evident that:

> The constraint relations (6 - 1) are feasible if and only if the system has a solution with w = 0, which is true if and only if the breakpoint curve of the system passes through the origin.

A convenient basic solution with which to start the breakpoint tracing method has as basic variables

$$v = 0, \quad z = 0, \quad w = Tb, \quad \xi = \Sigma^T w$$

The basis associated with this choice of basic variables is

$$B = \begin{bmatrix} I & 0 \\ 0 & B_D \end{bmatrix} \qquad B_D = \begin{bmatrix} T & 0 \\ \Sigma^T & 1 \end{bmatrix}$$

Although the inverse of B_D is not readily apparent, it is easily verified that

$$S = B_D^{-1} \begin{bmatrix} T & 0 \\ \Sigma^T T & 1 \end{bmatrix}$$

and is trivially calculable from B_D.

With the basic solution and inverse given above, the breakpoint tracing is set in motion by introducing the vector P into the basis. Since ξ cannot be greater than zero, the breakpoint curve must terminate with a basic solution $\bar{z}$ and a ray of solutions

$$z = \bar{z} + \alpha \overset{*}{z}$$

where $\bar{\xi} \leq 0$, $\bar{\eta} = 0$ and $\overset{*}{\xi} = 0$, $\overset{*}{\eta} > 0$. Therefore tracing is continued until a basic solution z is attained with $\xi = 0$, or until a ray of solutions is obtained parallel to the η-axis. In the first case, x and u constitute the desired feasible vector. Otherwise the constraints are infeasible.

As in the valve algorithm the final basis contains as its dual portion m linearly independent vectors from

$$[A \quad -I]$$

(If it contained a vector associated with some w_i this could be replaced with the vector associated with u_i.)

6 - 4 Solving the Voltage Reduced Network

Exactly the same techniques as were used in the previous two sections can be employed in solving the voltage reduced network of the linear or quadratic programming model. The mathematical form of the terminal pair systems for a valve algorithm and a by-pass algorithm are the same as before except for a switch of letters. They are

ξ	y	η	x	v	u	
	A^T	$-c$		I		0
			A		$-I$	0
1			$-c^T$			0

for the valve algorithm and

η	x	ξ	y	z	u	w	v	
	A	0			$-I$			0
	T	Σ				$-I$		0
			A^T	T			I	0
1			0	Σ^T				0

for the by-pass algorithm. Either method will yield a feasible solution (y, v) for the dual constraints

$$A^T y + v = c$$

$$y \geq 0, \quad v \geq 0$$

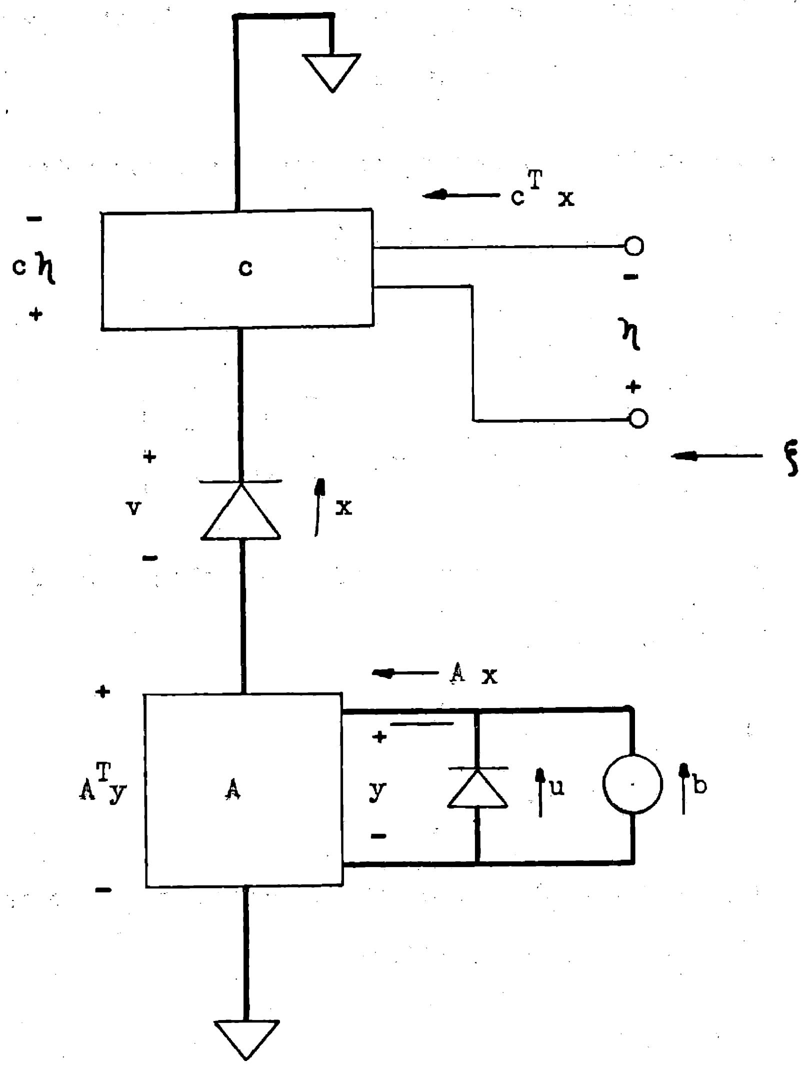
together with a corresponding basis of n linearly independent vectors from

$$[A^T \quad I]$$

and its inverse.

6 - 5 A Valve Algorithm for Linear Programming

Once a solution (x, u) of the corresponding current reduced network has been found, a complete solution of the linear programming model may be constructed by tracing the breakpoint curve of the terminal pair system shown in Figure 6 - 6. It will be necessary to utilize the partial basis M of m linearly independent vectors from $[A \quad -I]$ and its inverse which are available from the solution of the current reduced network. The terminal pair system in the figure contains no voltage sources. It is therefore a dual reduced system and all its breakpoints lie on the ξ-axis. By considering the figure it is evident that:



	ξ	y	η	x	v	u	
P		A^T	$-c$		I		0
				A		$-I$	b
	-1			$-c^T$			0

Figure 6 - 6--Valve method for completing the solution of the linear programming model.

A solution exists for the network model if an only if the breakpoint curve contains a point with $\eta = 1$.

As an initial basic solution we may take for the basic variables

$$v = 0 , \quad u = \bar{u} , \quad x = \bar{x} , \quad \xi = c^T \bar{x} .$$

The corresponding basis is

$$B = \begin{bmatrix} -I & 0 \\ 0 & B_D \end{bmatrix}, \qquad B_D = \begin{bmatrix} M & 0 \\ d^T & 1 \end{bmatrix}$$

where d contains the elements of $\begin{bmatrix} -c \\ 0 \end{bmatrix}$ corresponding to columns of $[A \quad -I]$ in M. The inverse of B_D is found to be

$$B_D^{-1} = \begin{bmatrix} M^{-1} & 0 \\ -d^T M^{-1} & 1 \end{bmatrix}$$

The tracing of the breakpoint curve is started by introducing P_η into the basis. The tracing is terminated when a ray of solutions

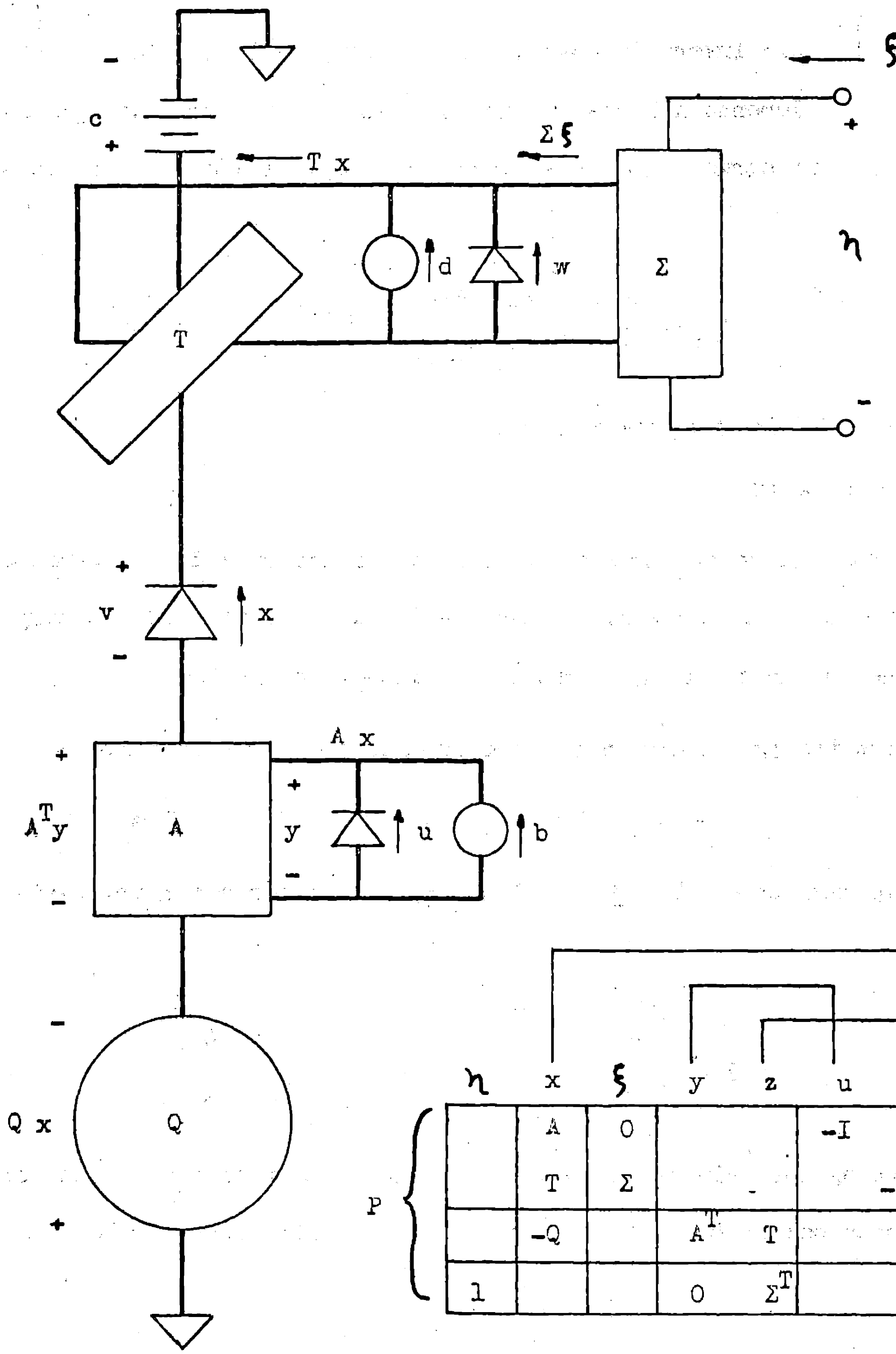
$$z = \bar{z} + \alpha \overset{*}{z} \qquad (6 - 3)$$

of the terminal pair system has been found. If $\overset{*}{\eta} > 0$ (6 - 3) gives the solution of the network with $\alpha = 1/\overset{*}{\eta}$. If $\overset{*}{\xi} > 0$ then the network has no solution.

The algorithm can equally well be applied to the dual formulation of the general linear program.

6 - 6 A By-Pass Algorithm for Linear and Quadratic Programming

The terminal pair system for a by-pass method of solving the linear programming network model is shown in Figure 6 - 7. To initiate the algorithm it is necessary to have a solution $(\bar{x}, \bar{u})$ of the current reduced network and



P

η	x	ξ	y	z	u	w	v	
	A	0			−I			b
	T	Σ				−I		0
	−Q		A^T	T			I	c
1			0	Σ^T				0

Figure 6 - 7—By-pass method for linear and quadratic programming.

the associated basis inverse M^{-1} as in the previous section. In the terminal pair system all elements are present and it therefore has the most general form of breakpoint curve. The elements of the diagonal matrix T are chosen so that

$$T_{ii} = \begin{cases} +1 & \text{if } (Qx + c)_i \geq 0 \\ -1 & \text{if } (Qx + c)_i < 0 . \end{cases}$$

and the value of d is given by $d = Tx$.

It is evident that:

The model network possesses a solution if and only if the terminal pair system has a solution with $z = 0$. This is true if and only if there is a point on the breakpoint curve with $\eta = 0$.

The breakpoint tracing procedure is started with the basic solution

$$x = \bar{x}, \quad u = \bar{u}, \quad z = T(Q\bar{x} + c), \quad \eta = -\Sigma^T z .$$

From the construction of T, $z \geq 0$. Also $\eta \leq 0$. The corresponding basis matrix is

$$B = \begin{bmatrix} -I & 0 \\ 0 & B_D \end{bmatrix}, \qquad B_D = \begin{bmatrix} M & 0 \\ R & I \end{bmatrix}$$

where R contains the columns of $[-Q \quad 0]$ associated with the columns of $[A \quad -I]$ contained in M. The inverse of B_D is readily found to be

$$B_D^{-1} = \begin{bmatrix} M^{-1} & 0 \\ -RM^{-1} & I \end{bmatrix},$$

which is easily calculated as M^{-1} is available from the solution of the current reduced network.

The first vector to be substituted into this basis is P_ξ which starts the tracing in the positive direction along the breakpoint curve. The algorithm is terminated when either a basic solution is found for which z = 0 and the network has been solved, or when a ray of solutions

$$z = \bar{z} + \alpha \overset{*}{z}$$

has been found with $\bar{\xi} < 0$, $\overset{*}{\eta} > 0$, $\overset{*}{\xi} = 0$, indicating that no solution exists.

With the most general formulation of the quadratic programming problem, this method is only applicable to the primal formulation as there is no simple dual. However, if the matrix Q has the form

$$Q = \begin{bmatrix} P & 0 \\ 0 & 0 \end{bmatrix}$$

then the problem has a dual formulation and the above method could be used either on the primal as described above, or on the dual starting with a solution of the voltage reduced network. Of course, in the special case of linear programming the method can always be used either way.

C H A P T E R   S E V E N

An Application of Quadratic Programming to the General Programming Problem

In this chapter an iterative method is presented for solving the general programming problem formulated in Chapter One. It is an extension of the gradient methods frequently employed to compute the unconstrained minimum of a function of many variables. Specifically it will be shown that the problem of finding the direction of steepest descent for a general programming problem is itself a quadratic programming problem. First, however, the gradient methods are applied to unconstrained and equation constrained minimization problems will be reviewed.

7 - 1 Gradient Methods of Minimization

For illustration consider the two variable minimization problem

$$\text{Minimize} \quad \phi\,(x_1, x_2).$$

It is desired to find a point in the $x_1 - x_2$ plane for which $\phi(x_1, x_2)$ takes on its smallest value. Let $\bar{x}$ in Figure 7 - 1 be the point at which ϕ takes on its minimum. It follows that near $\bar{x}$ the lines along which ϕ is constant must be closed curves encircling this point. Suppose that the point x^o is chosen as an initial guess as to the location of the minimum. A better approximation to the location of the minimum can be found by moving from the

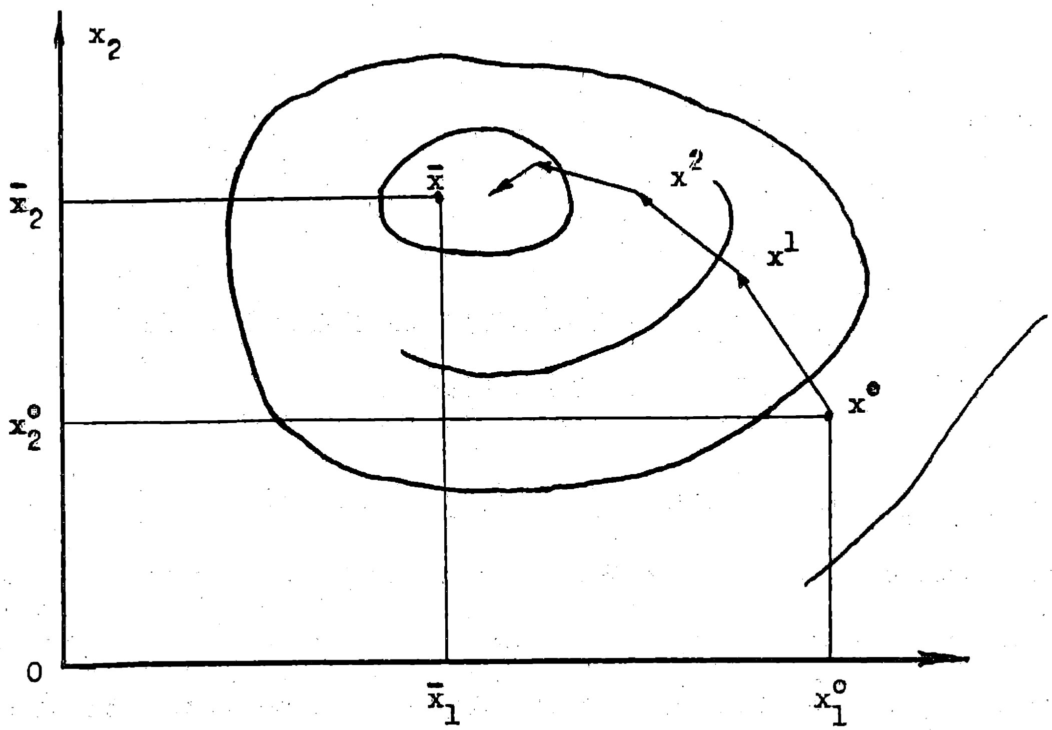


Figure 7 - 1—The method of steepest descent.

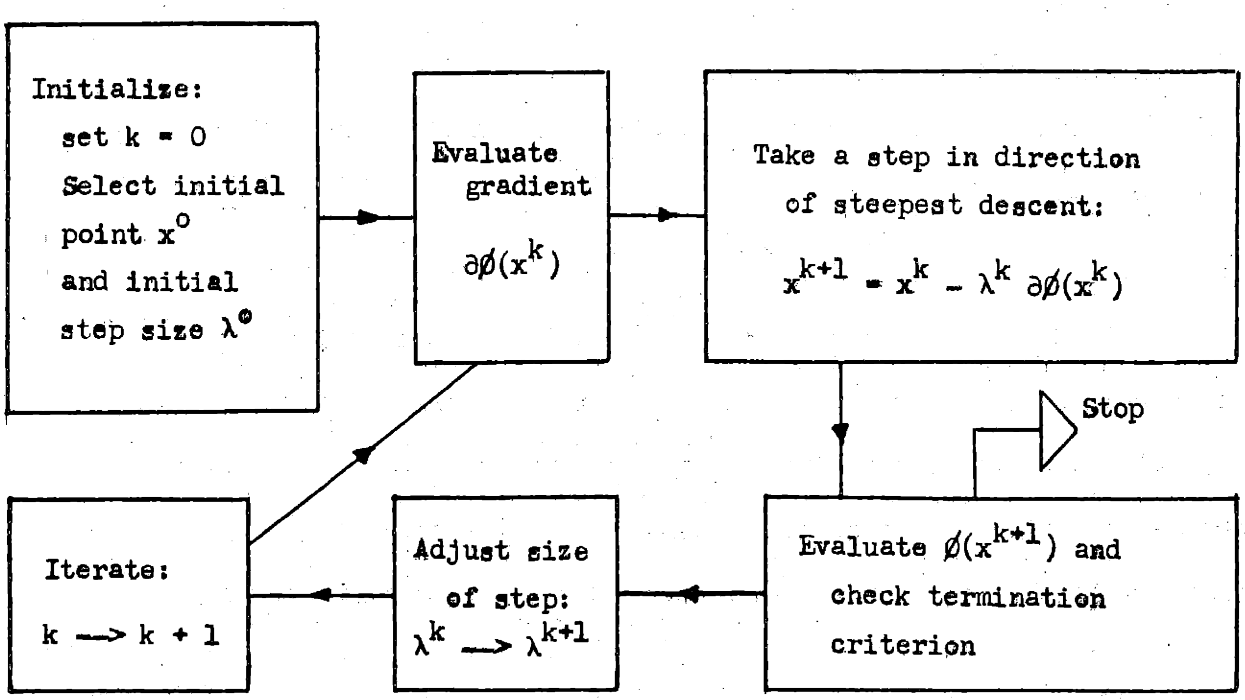


Figure 7 - 2—Flow diagram for method of steepest descent.

the initial point in the direction which yields the greatest rate of decrease in $\emptyset$. This direction, the direction of steepest descent, is directly opposite to the gradient of $\emptyset$ and is perpendicular to the line of constant $\emptyset$ passing through the initial point. Thus, the new point is given by

$$x^1 = x^o - \lambda \ \partial\emptyset(x^o) \tag{7 - 1}$$

where λ is a positive constant. One step in this manner, of course, will not in general take one to the minimum at $\bar{\bar{x}}$. Hence the gradient step defined by (7 - 1) should be repeated until some criterion is satisfied indicating that the minimum has been reached. A suitable criterion is that the change in $\emptyset$ on the last step be less than some chosen value. The value of the constant λ must be controlled carefully to avoid so large a value that the process does not converge or so small a value that an excessive number of iterations are required. One possible method of controlling λ is by observing the angle Θ between successive evaluations of the gradient vector. If this angle becomes too large λ is decreased; if it is too small λ is increased. In addition λ is never allowed to be so large that $\emptyset$ is increased. The method will work with functions of any number of variables with only the loss of the simple geometric interpretation given above. A flow diagram of the computation is given in Figure 7 - 2.

There are several drawbacks to gradient methods of minimization. One is that the method only finds a local minimum of the objective function $\emptyset$. If this function is concave the local minimum will also be a global minimum as pointed out in Chapter Two. Otherwise the initial point x^o must be chosen so that the "nearest" relative minimum is the desired solution. A second difficulty is that the method is not invariant with respect to a change in scale of the variables. If the scale or metric is poorly chosen the method may converge to the minimum at a very slow rate.

7 - 2 Minimization in the Presence of Equality Constraints

The gradient technique may be extended to problems in which the variables must satisfy equality or inequality constraints by modifying the notion of the direction of steepest descent.

> A vector u will be called an allowable direction at a point x in the constraint set of a general programming problem, if it is possible to move a small distance away from x in the direction u without leaving the constraint set. For a general programming problem the direction of steepest descent at a point x in the constraint set is the direction, among all allowable directions, for which the rate of decease of ϕ is greatest.

First the case of equality constraints will be considered.

Equality constrained minimization problem:

Minimize $\phi(x)$

with $g(x) = 0$

Suppose $\bar{x}$ is a point in the constraint set obtained at some point in the iterative procedure. Let the negative gradient of the objective function be represented by

$$t = - \phi(\bar{x})$$

and let A_i be the gradient of the i^{th} constraint function

$$A_i = g_i(\bar{x})$$

This is a vector normal to the tangent plane of the constraint surface at $\bar{x}$. Then a vector u will be allowable if it lies in the tangent plane to each of the constraint surfaces at $\bar{x}$. This will be true if u is orthogonal to the normal vector of each constraint surface at $\bar{x}$.

$$A_i^T u = 0 \qquad (7 - 2)$$

The direction of steepest descent is then given by the allowable vector of unit magnitude u which has the greatest component along the negative gradient of ϕ. Thus the problem of finding the direction of steepest descent is to

$$\text{Maximize} \quad t^T u$$

$$\text{with} \quad u^T u = 1$$

$$A^T u = 0 \, .$$

Employing multipliers v and $- \delta/2$, the Lagrangian expression for this problem is

$$\psi (u, v \lambda) = t^T u + v^T A^T u - \delta/2 \, u^T u \, .$$

The corresponding Lagrangian problem is found by differentiation to be

Find u, v and δ such that

$$u^T u = 1 \qquad (7 - 3a)$$

$$A^T u = 0 \qquad (7 - 3b)$$

$$t + A v - \lambda u = 0 \qquad (7 - 3c)$$

This system of relations can be solved by substituting (7 - 3c) in (7 - 3b) to obtain

$$- \frac{1}{\lambda} A^T A v = \frac{1}{\lambda} A^T t$$

or

$$v = - M^{-1} A^T t$$

where $M = A^T A$. Finally

$$u = \frac{1}{\lambda} \left[I - A M^{-1} A^T \right] t \qquad (7 - 4)$$

The matrix M is positive semi-definite because

$$x^T M x = x^T A^T A x$$

$$= \left[A x, A x \right] \quad 0, \text{ all } x.$$

If the columns of A are linearly independent, M is positive definite and the inverse M^{-1} can be computed and a unique λ u obtained. If the columns of A are linearly dependent, one of the normal vectors is a linear combination of the others and the corresponding constraint is redundant at $\bar{x}$. The multiplier δ in (7 - 4) must be chosen so that the magnitude of u is unity and (7 - 3a) is satisfied.

A flow diagram for minimizing a function subject to equality constraints is shown in Figure 7 - 3. Besides the difficulties mentioned in connection with the unconstrained case, there is the problem of keeping the constraints satisfied. The fact that the direction of movement satisfies (7 - 2) only guarantees that the constraints are not violated for an infinitesmal motion on the cirection of steepest descent, whereas finite size steps will actually be taken. Hence it is necessary to provide a block in the flow diagram which moves the point back into the constraint surfaces. Also there is the problem of determining the initial point x^0 in the constraint set. These problems will not be further discussed here.

7 - 3 The Direction of Steepest Descent with Inequality Constraints

In the case of inequality constraints a feasible point $\bar{x}$ does not necessarily lie in any particular constraint surface. If $\bar{x}$ does not lie in a constraint surface $g_i(x) = 0$, then this constraint has no effect on the direction of steepest descent at $\bar{x}$. On the other hand if $g_i(\bar{x}) = 0$, the direction of steepest descent may lie in the constraint surface or it may point into the interior of the constraint set. In order to be an allowable direction of motion away from $\bar{x}$, u must not have a negative component along the normal to any constraint surface. The direction of steepest descent is again the unit vector in an allowable direction which has the greatest component along the negative gradient of ϕ. If I is the set of constraint relations which are equality satisfied at the point $\bar{x}$, the direction

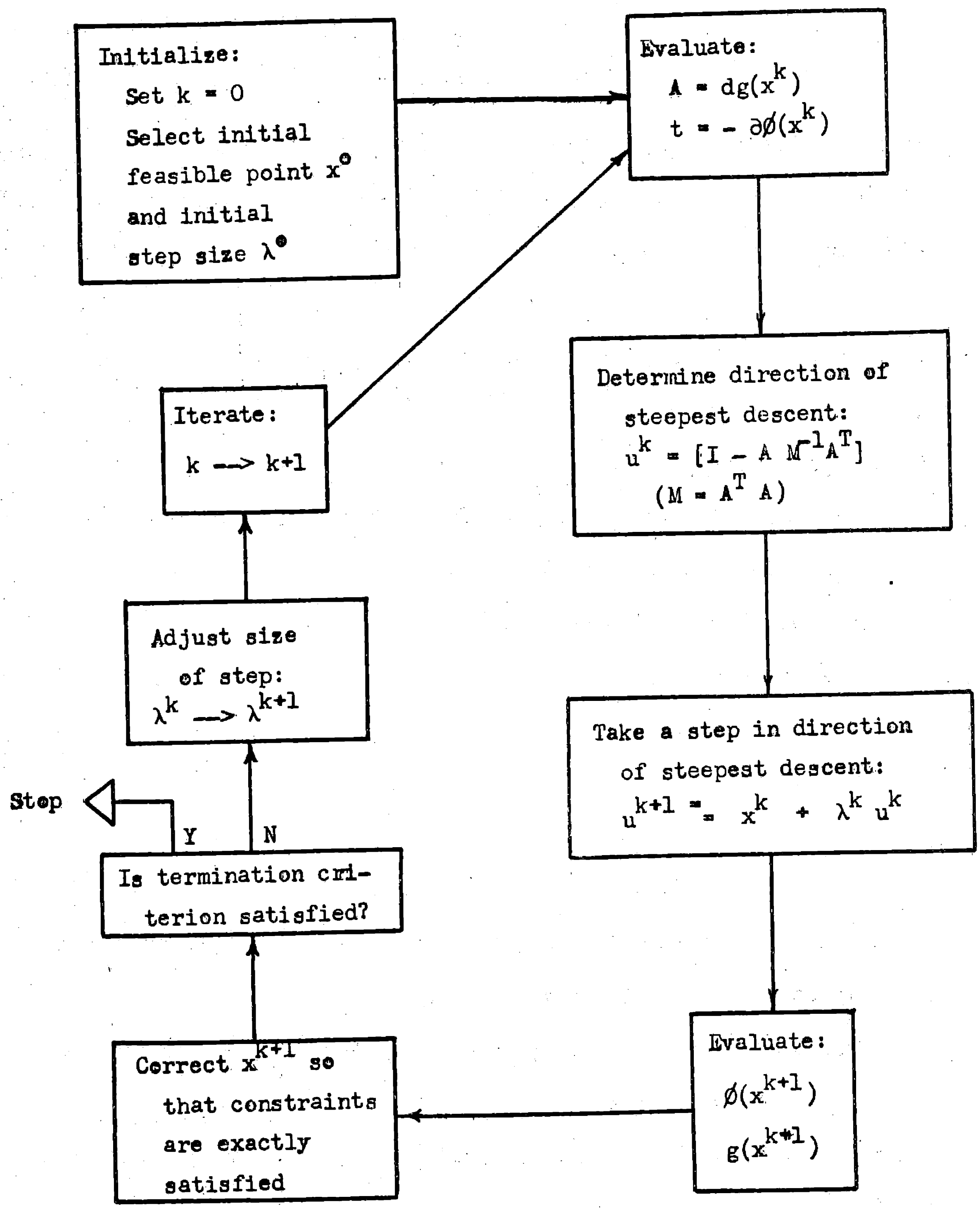


Figure 7 - 3—The method of steepest descent with equality constraints.

of steepest descent is given by the programming problem

$$\text{Maximize} \quad t^T u \tag{7 - 4a}$$

$$\text{with} \quad u^T u = 1 \tag{7 - 4b}$$

$$A_i^T u \geq 0, \ i \text{ in } I. \tag{7 - 4c}$$

This is illustrated in Figure 7 - 4 for the three dimension case. Because of the non-linear equality constraint (7 - 4b), this program does not have a convex constraint set and the development in Chapter Two does not guarantee that a relative minimum is a global minimum. However, in the next section this problem will be shown equivalent to a quadratic program which does have a unique solution.

7 - 4 Equivalence with a Quadratic Program

The Lagrangian problem associated with the program (7 - 4) is, using the result given in Chapter One,

Find v, u, z and δ such that

$$\bar{A}\, v - \delta\, u + t = 0 \tag{7 - 5a}$$

$$u^T u = 1 \tag{7 - 5b}$$

$$A^T u \geq 0, \tag{7 - 5c}$$

$$v \geq 0$$

$$v^T \bar{A}^T u = 0 \tag{7 - 5d}$$

where $\bar{A}$ is the matrix consisting of those columns of A corresponding to equality satisfied constraint surfaces.

This statement may be simplified if (7 - 5a) is solved for δ u

$$\delta\, u = t + A\, v$$

and the result substituted in (7 - 5c). Then the Lagrangian problem becomes, for $\delta > 0$,

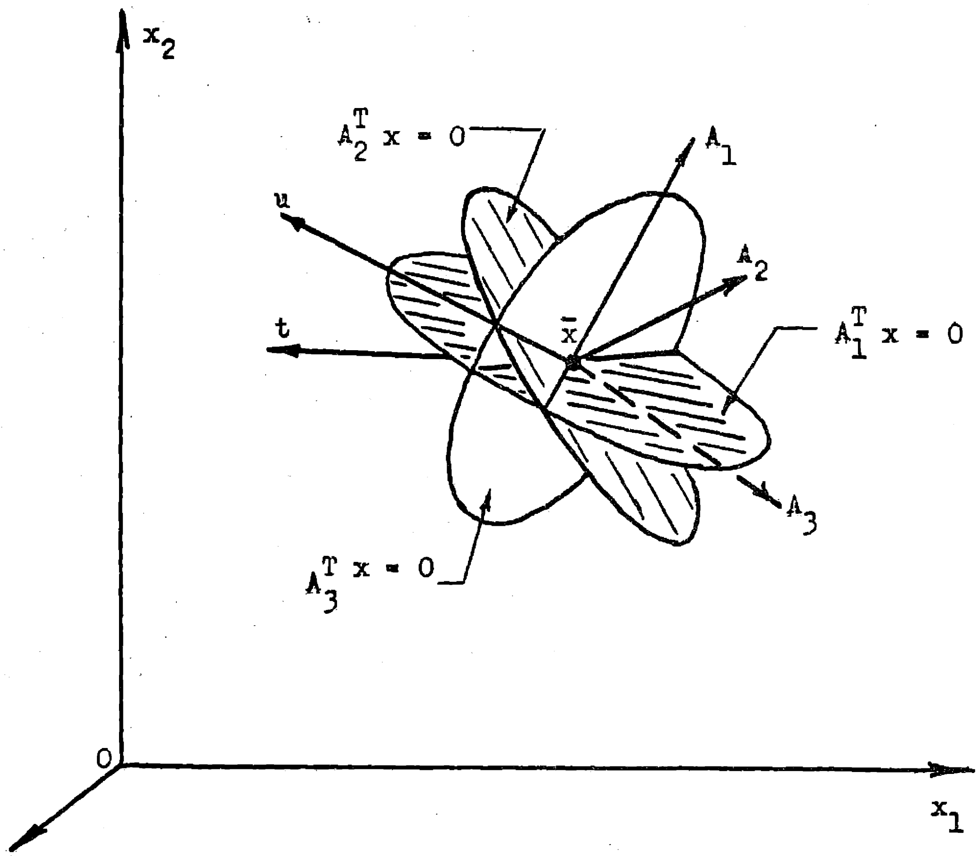


Figure 7 - 4—Illustrating the direction of steepest descent for three variables.

Find v, z such that

$$- M\, v + z = A^T t \qquad (7 - 6a)$$

$$z \geq 0, \quad v \geq 0 \qquad (7 - 6b)$$

$$z^T v = 0$$

where $M = A^T A$ as before.

Now (7 - 6) is exactly the Lagrangian problem associated with the following quadratic program.

$$\text{Maximize } \frac{1}{2} v^T M v + t^T A v$$

$$\text{subject to } \quad v \geq 0 \qquad (7 - 7)$$

As was shown previously, M is a positive definite matrix if the columns of A are linearly independent. Hence the objective function is always bounded from below. Since the constraint set of the quadratic program is obviously feasible, the theory of quadratic programming (Theorem D - 4) shows that (7 - 7) always has a unique optimal solution.

Next we will show that if $(\bar{v}, \bar{z})$ is a solution of (7 - 5) with $t^T \bar{u} > 0$, then no other allowable unit vector u has as great a component along t, i.e.,

$$t^T u < t^T \bar{u}$$

for all u not equal to $\bar{u}$ such that

$$u^T u = 1 \quad \text{and} \quad A^T u + z = 0, \quad z \geq 0. \qquad (7 - 8)$$

From relations (7 - 5) we have

$$\delta\, \bar{u}^T \bar{u} - t^T \bar{u} = \bar{u}^T A \bar{v}$$

$$\bar{v}^T A^T \bar{u} = \bar{v}^T \bar{z} \geq 0$$

and therefore, since $\bar{u}^T \bar{u} = 1$,

$$\delta = t^T \bar{u} > 0.$$

Let u be any vector not equal to $\bar{u}$ and satisfying (7 - 8). Then from (7 - 5a)

$$\delta\, u^T \bar{u} = t^T u + u^T A \bar{v}$$

$$\bar{v}\, A^T u = \bar{v}^T z \geq 0,$$

and therefore

$$t^T u = (t^T \bar{u}) (u^T u) - \bar{v}^T z.$$

Since u and u are not identical, $u^T \bar{u}$ must be less than 1, and it follows that $t^T u < t^T \bar{u}$.

From this discussion it is seen that if(7 - 5) has a solution $(\bar{\delta}, \bar{u}, \bar{v})$ with $\bar{\delta} = t^T \bar{u} > 0$, then u is the direction of steepest descent. But then $\bar{v}$ and $\bar{z} = A\ t + M\ \bar{v}$ form a solution of (7 - 6). It follows that $\bar{v}$ is the unique optimal solution of the quadratic program (7 - 7). Conversely, if $\bar{v}$ is the optimal solution of (7 - 7) and

$$\bar{\delta} = t^T t + t^T A\ \bar{v} > 0,$$

then $\bar{v}$, $\bar{u} = \dfrac{t = A\ \bar{v}}{\bar{\delta}}$ and $\bar{\delta}$ form a solution of (7 - 5) with $t^T u > 0$. Therefore we have the following principle.

> If the optimal vector $(\bar{v}, \bar{z})$ of the quadratic program has $t^T t + t^T A\ \bar{v} > 0$, then the vector $u = t + A\ \bar{v}$ is in the direction of steepest descent for the general programming problem at $\bar{x}$. If the optimal vector of the quadratic program has $t^T t + t^T A\ \bar{v} \leq 0$, then $\bar{x}$ is a local minimum of the general program.

One of the algorithms for solving quadratic programs presented in Chapter Six may be used to compute the direction of steepest descent for the general programming problem. A flow diagram of the steps in solving a general program is given in Figure 7 - 5. Since the quadratic program which must be solved at one step of the method of steepest descent is nearly the same as the one solved for the previous step, a very good starting solution is always available for the new solution. If an algorithm is employed which takes advantage of this fact, the amount of computation required to determine the new direction of steepest descent should be quite small.

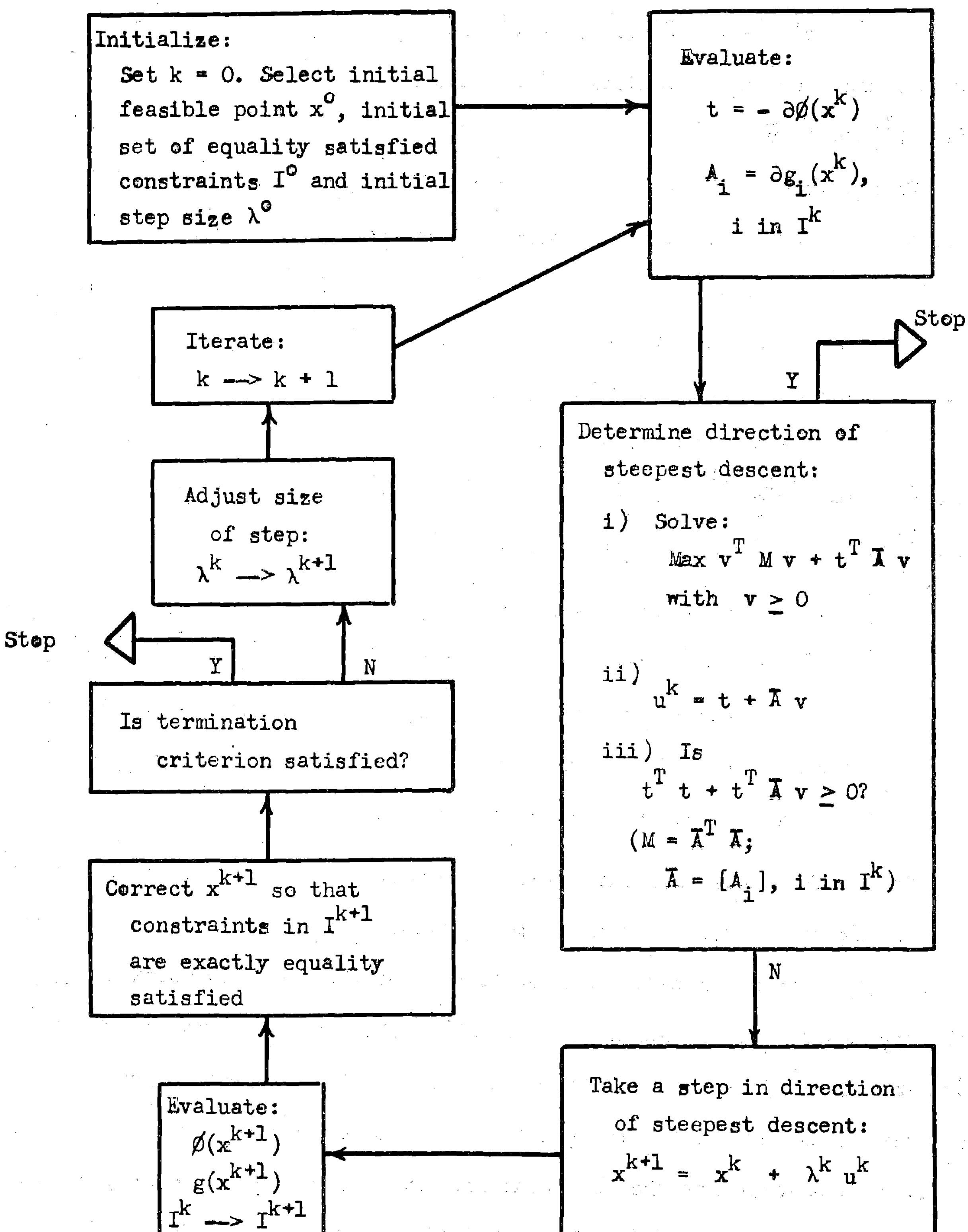


Figure 7 - 5—The method of steepest descent with inequality constraints.

C H A P T E R E I G H T

Historical Notes and Discussion

It was pointed out in Chapter One that essentially all of the effort in the area of mathematical programming has been concerned with industrial operations and economic systems. It is therefore no surprising that a different approach to the subject should yield significant results. A new viewpoint on the general area of mathematical programming was introduced to the author when Samuel J. Mason, Associate Professor of Electrical Engineering at M.I.T., pointed out the simple and elegant relation between network flow problems and electrical networks containing diodes and sources given in section 4 – 5. It was this observation that inspired the contributions to the theory and methodology of mathematical programming contained in this thesis.

In the following sections, the history of these contributions is briefly sketched. Those papers and other references from programming literature which were most responsible for guiding the author's work are acknowledged and several directions in which the work could be carried further are indicated.

First the author would like to acknowledge the sources for the theoretical background presented in the first three appendices. The theorem on linear inequalities (Theorem B – 1) which forms the foundation of the entire theory of mathematical programming was first given by Farkas (12) half a century ago.

The proof given in Appendix B is a relatively new one due to David Gale (24). The development of the theory of non-linear programming in Appendix C is based on the work of Kuhn and Tucker (20).

7 - 1 On the Theoretical Developments

It has long been known that there are several equivalent formulations for the problem of solving a network of resistors and sources. Maxwell noted in 1873 (23) that the current distribution in a network of resistors and current sources which produced the least "heating" was the one which satisfied the Kirchoff loop condition. However, the present author knows of no place in the literature where extremum formulations for electrical networks containing diodes are discussed. Noting the observation of Professor Mason, the author was able to write down a pair of dual linear programs for an arbitrary network of diodes and sources. Being familiar with the extremum principle of Maxwell, this led to the formulation of the diode-source-resistor network problem as the dual pair of quadratic programs given in section 3 - 3. This result raised the querie: Is there a more general duality formulation for quadratic programming? Further investigation produced the dual pair of quadratic programs presented in Chapter One. This duality relation bridges the gap between the principle stated by Courant and Hilbert (5) for equality constrained minimization problems and the familiar duality relation of linear programming. The equivalent problems for quadratic programming presented by Frank and Wolfe (15) and especially Hildreth (18) gave valuable clues to the formulation of this duality relation.

Unfortunately, the duality principle as stated here is not applicable to the most general class of quadratic programming problems because of the condition on the form of the matrix Q. It seems that it should be possible to formulate a dual problem for any quadratic program:

minimize

$$x^T Qx + b^T x$$

with

$$Ax \geq b$$

$$x \geq 0 .$$

The dual problem would have a number of constraint inequalities equal to n - r where n is the number of primal variables and r is the rank of Q. The manner of determining what these constraints should be, and what form the dual objective function would take is not clear, however.

The further extension of the duality principle to concave programming was inspired by acquaintance with the Legendre transformation in its applications to thermodynamics. The development in Appendix E is a slight generalization of the transformation presented by Courant and Hilbert (6). While their formulation requires that the initial function be twice differentiable, the author's development, by utilizing the concavity property, requires only first differentiability. The application of the Legendre transformation to the concave programming problem yields the duality principle derived in Appendix F.

The fact that the existence theorem of linear programming also applies in the quadratic case was observed by the author while experimenting with methods for solving diode-source-resistor networks as will be noted later. The method of proof used in Appendix D was based partially on the existence condition stated by Barankin and Dorfman (2). The reader will notice that no corresponding theorem has been included for the concave programming case. It appears that more stringent conditions than the simple concavity requirement must be met by the objective function in order to guarantee the applicability of the corresponding existence theorem. It seems that a suitable condition is that the primal objective function must not be "assymptotically linear" on any ray contained in the primal constraint set and similarly for the dual objective function.

Unknown to the author at the time of his own developments, Fenchel (14) has demonstrated the following very general duality principle: Given a concave function $\emptyset$ defined on an arbitrary convex set C, there is (subject to very slight qualification) a convex set D and a convex function Θ defined on D such that $\emptyset$ takes on a minimum over C if and only if Θ takes on a maximum over D, and these two extrema are the same. The rules given for the construction of D and Θ when C and $\emptyset$ are given make this the ultimate generalization of the Legendre transformation. In theory at least this is a duality principle applicable to the general programming problem. However, it is expressed only in terms of a general notion of convex sets. In order to apply the principle to practical problems it would be desirable to know, given the objective and constraint functions of a general programming problem, how to construct the objective function of the dual problem and the convex functions defining the dual constraint set. This the present author has at least partially done for the case of quadratic programming.

8 – 2 Concerning the Methodology

The fundamental idea on which the methodological contributions of this thesis are based also arose from the study of electrical networks. By considering the problem of finding a solution of a network containing diodes, sources, and resistors, the author almost immediately conceived the idea of

i) assuming an arbitrary state (open or closed) for each diode of the network.

ii) replacing the diodes by open circuits and short circuits.

iii) solving the resulting source-resistance network (a set of simultaneous equations)

iv) successively replacing each diode in the network by tracing the breakpoint curve observed at its terminal pair.

It was noted that the tracing could be accomplished by the superposition of unit incremental solutions as is described in Chapter Five. Also it was suspected and confirmed that the recursion formulae of the simplex method (10) could be used to obtain the new unit solutions as the tracing process progressed.

It was soon evident, however, that this approach could not be guaranteed to work for all situations -- the observed breakpoint curves do not necessarily intersect with the diode characteristics. Eventually, the author realized that the way to avoid the difficulty was to find first a feasible solution for one of the corresponding dual programs. This led to the development of the reduced network concept and the electrical interpretation of the existence theorem of linear programming. It also indicated that the existence theorem was valid for quadratic programming.

The network flow type of problem formulated in section 2 - 3 is important because it appears to be nearly the most general type of linear program which can be solved with the arithmetic operations of addition and subtraction alone. As a result it is possible to construct methods for solving them which do not require carrying an inverse matrix along with the computation. The original example of such a method is the "stepping stone" specialization of the simplex method (8) used to solve the familiar transportation problem. Recently Ford and Fulkerson (16) have presented a new algorithm for solving this problem. The basis for the extension of the method to apply to the network flow problem has been laid down by the same writers (17). They have not given the details of the complete algorithm, but these are stated in the doctoral thesis of Jewell (19). It was observed by Professor Mason and this writer that the latter method is equivalent to tracing the breakpoint curve of a certain diode-source network. This observation led to the method for solving an arbitrary diode-source network given in Chapter Four.

The next objective was to apply the same idea of satisfying one diode at a time to the general formulation of linear and quadratic programming. Elec-tircal models of these problems were constructed using another circuit element, the ideal d-c transformer. It was easily demonstrated that simplex change of basis formulae would apply equally well to tracing breakpoint curves for networks containing transformers. The question remained concerning the problem of degeneracy: What does one do if more than one diode moves to the breakpoint state simultaneously? In the case of the simplex method it is known that degeneracy can cause perpetual cycling without ever reaching an optimal vector. In the breakpoint tracing algorithm, the danger of degeneracy is that the procedure might ''turn around'' and proceed in the reverse direction along the breakpoint curve. The degeneracy question was resolved in the case of the simplex method by Charnes (4) who gave a rule for choosing the vector to leave the basis which would insure that cycling did not occur. A more elegant mathematical treatment of the matter has since been formulated by Wolfe (10). The same mathematical principles as are employed by Wolfe are used in Appendix H to resolve the degeneracy question for the breakpoint stepping algorithm.

The author's valve algorithm for linear programming (Figure 6 - 5) is similar to the simplex method of Dantzig. In the simplex method only the dual basis inverse B_D^{-1} from the author's terminal pair system is used in the computation. The vector to enter this basis is selected so that $\xi = c^T x$ will increase, while it is determined by the dual step in the author's method. In Dantzig's procedure there is no dual step; for this terminal pair system the dual step of the breakpoint tracing method does not change the values of any variables, but merely indicates which vector should be entered in the primal basis next. In a somewhat similar manner Lemke's dual simplex method (21) goes through the same steps, but uses only the primal basis inverse B_P^{-1} from the author's algorithm. The relation between these three methods is not completely clear and further study of the matter would seem warranted.

The author's by-pass algorithm as applied to linear programming is nearly equivalent to the general primal-dual method of Ford and Fulkerson (9) as applied to the dual problem. The primal-dual method differs in that dual steps are taken until η attains its highest possible value for the current values of the primal variables. At this point, a primal step is indicated which moves the solution point in the ξ-direction along the breakpoint curve. The author's method always alternates between primal and dual steps. In the non-degenerate case η is always maximized in one step of the primal-dual method and the two algorithms are identical. Thus they differ only in the way in which degeneracy is handled. In the Ford-Fulkerson method only information contained in the dual basis inverse B_D^{-1} is needed to avoid the cycles possible in degenerate cases, whereas the by-pass algorithm appears to require knowledge of both B_P and B_D to avoid degeneracy problems. It is evident that the primal-dual algorithm could be phrased in a form utilizing B_P^{-1} only— primal steps would be taken until ξ is maximized whereupon a dual step is indicated which would increase η.

The author's by-pass algorithm as applied to quadratic programming is equivalent to a generalization of the method recently developed by Wolfe (25). Wolfe's method traces the breakpoint curve of the terminal pair system of Figure 6 - 7 by successively superimposing unit incremental solutions which increase η until no further such unit solution can be found. Thus the method will terminate when it reaches any horizontal segment of the breakpoint curve. As a consequence it is only applicable in the form stated by Wolfe to problems where the breakpoint curve has no horizontal segment except at the solution point $\eta = 0$, i.e., when Q is positive definite. The method could be generalized by employing a dual step as in the primal-dual method whenever η has been maximized but is not yet zero. An obvious variant to this method would be to superimpose unit solutions which increase ξ until it is maximized and follow with a pure primal step to traverse vertical segments of the breakpoint curve. On the other hand the author's by-pass method avoids all mention of maximization and yields an algorithm in which all steps are executed in the same manner. All three of the methods would take precisely the same steps -- at least in the non-degenerate case -- while using quite dissimilar decision rules. This discussion shows, moreover, the close relation between Wolfe's method for quadratic programming and the linear programming method of Ford and Fulkerson.

The quadratic programming procedure proposed by Markowitz (22) can also be interpreted as a scheme for tracing a breakpoint curve. However, the algorithm developed by Frank and Wolfe (15) and a method given by Hildreth (18) do not fit in this category.

A third manner of employing the breakpoint stepping procedure would be to satisfy the diodes of the network model one at a time by successively tracing the breakpoint curves observed at their terminal pairs. This would be the generalization of the diode-source algorithm presented in Chapter Four. Time and space limitations prohibited its inclusion in the thesis.

8 - 3 On the General Programming Problem

The author's interest in non-linear programming was sparked by problems in the optimization of the design of electrical machinery. As a means of computing optimum designs, the application of the method of steepest descent (7) to the minimization of a function of n variables subject to m constraints was investigated. In the original method proposed by the author (11) m of the variables were considered as dependent and the gradient with respect to the remaining n - m variables was computed giving the direction of steepest descent. The employment of Lagrange multipliers in deriving the direction of steepest descent as given in section 7 - 2 avoids the necessity of choosing a set of dependent variables. This idea was also given by Feder in connection with the solution of non-linear equations (13).

The formulation for the direction of steepest descent given in Chapter Seven for the general programming problem was originally obtained through geometric considerations: The surface of the unit sphere centered at the feasible point x is projected onto a hyperplane perpendicular to the gradient of the objective function. The problem is to find the point on the surface of the sphere with the greatest component along the gradient and on the correct side of the tangent planes to the constraint surfaces passing through x. In the hyperplane this is equivalent to finding the point nearest the intersection of the gradient and the hyperplane which is on the correct side of the intersections of the hyperplane with the tangent planes.

Other proposals for computing solutions to general programs include the method of Arrow and Hurwicz (1) in which the problem is first converted into a saddle-value problem and a gradient procedure is used to search for the saddle point. Also Brown has investigated (3) the conversion of the inequality constraints to equality constraints by the addition of squared slack variables. Both of these methods have the disadvantage of introducing extra variables which would complicate and lengthen the computational procedure. In the former method it also appears troublesome to guarantee the convergence of the steps to a solution point. On the other hand Brown's approach has the disadvantage of introducing extraneous saddle-points near which a gradient procedure is likely to proceed at a very slow rate. The direct steepest descent approach appears to offer an efficient and straightforward approach to the problem. However, many details need to be worked out if it is to be applied with success.

8 - 4 Conclusion

In the author's view this thesis makes four contributions to the discipline of mathematical programming:

1) The extension of the duality principle to quadratic and concave programming.
2) A versatile algorithm for solving diode-source networks which is applicable to the transportation problem and capacity limited network flow problems.
3) The generalized breakpoint tracing method and particularly its application to quadratic programming.
4) The demonstration that the direction of steepest descent for a general programming problem can be determined as the solution of a quadratic program.

In all but the last of these the electrical analogy has been largely responsible. Hence it might be said that the most valuable contribution is the demonstration of how the concepts and scientific knowledge in one field (electrical networks) can aid in understanding the principles of another discipline (mathematical programming).

A P P E N D I X   A

Geometrical Elements in Euclidian Space

In this section, geometric concepts are presented which will be used in interpretation of algebraic developments, especially in Appendix B.

Definition: The Euclidian space of n dimensions is the set R^n of all n-tuples of real numbers. An element of R^n may be thought of as a point whose coordinates are the n real numbers or as a vector emanating from the origin whose components along the coordinate axes are the n real numbers. A point or vector x in R^n will be represented by the column matrix

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix} .$$

We will say that a vector y in R^n is greater than a vector x in R^n and write $y > x$ if each component of y is greater than the corresponding components of x, that is,

$$y > x \text{ implies } y_i > x_i , \qquad i = 1, 2, \ldots, n.$$

Similarly we will write $y \geq x$ if the relation holds for each component;

$$y \geq x \quad \text{implies} \quad y_i \geq x_i , \qquad i = 1, 2, \ldots, n.$$

Two vectors x and y are orthogonal if their dot product

$$x^T y = \sum_{i=1}^{n} x_i \, y_i$$

is equal to zero.

We will have occasion to consider the product of an m by n matrix A^T by a vector y. Geometrically this product is to be interpreted as the column vector whose components are the dot products of y with the corresponding columns of A.

$$A^T y = \begin{bmatrix} A_1^T y \\ A_2^T y \\ - \quad - \\ A_n^T y \end{bmatrix}$$

On the other hand, if we are concerned with the product of an n by m matrix A and a vector y having m components, we can visualize this in R^n as the vector sum of the columns of A each weighted by the corresponding component of .

$$Ax = x_1 A_1 = x_2 A_2 + \ldots + x_m A_m$$

Both of these interpretations will be useful in subsequent proofs.

<u>Definition</u>: A <u>line</u> in R^n is a set of all points x that satisfy a relation

$$\frac{x_1 - a_1}{b_1} = \frac{x_2 - a_2}{b_2} = \dots = \frac{x_n - a_n}{b_n}$$

where a and b are members of R^n. The line includes the point a and is parallel to the <u>direction</u> <u>vector</u> b.

<u>Definition</u>: A <u>hyperplane</u> in R^n is a set of all points x in R^n that satisfy a relation

$$a^T x = \beta$$

where a, a member of R^n, is the normal vector of the hyperplane and β is scalar.

Note that a hyperplane $a^T x = 0$ contains the origin and is the union (the set of all points in any) of all lines containing the origin whose direction vectors are orthogonal to a.

<u>Definition</u>: A <u>half-space</u> in R^n is the set of all points x that satisfy a relation

$$a^T x \geq \beta$$

where a is a member of R^n and β is a scalar. The hyperplane $a^T x = \beta$ is called the <u>bounding</u> <u>hyperplane</u> of the half-space.

It is reasonable to call such a set a half-space because for each point x strictly on one side of the bounding hyperplane, $a^T x < \beta$, we can produce a point y strictly on the other side

$$y = x - 2 \frac{a^T x - \beta}{|a|} \frac{a}{|a|}$$

$$a^T y = a^T x - 2a^T x + 2\beta$$

$$= \beta - (a^T x - \beta) \qquad \text{so that} \qquad a^T y > \beta ,$$

and this is a one-to-one correspondence between x and y.

Definition: A ray or half-line in R^n is a set containing all points x such that

$$x = a + \lambda b$$

for some $\lambda \geq 0$. Here a and b are fixed members of R^n.

Definition: A cone is a set of points in R^n such that if x is in the set then

$$y = \lambda x$$

is also in the set for any non-negative number λ.

Note that by this definition a cone always has its apex at the origin.

Definition: A convex set in R^n is one in which, given any two points, x and y in the set, all points on the line segment joining x and y are also in the set, that is

$$z = (1 - \alpha)x + \alpha y$$

is in the set for any number α between zero and one. The convex hull of a given set is the "smallest" convex set containing the given set. More precisely, it is the intersection of (all points common to) all convex sets which contain the given set. An extreme point of a convex set is a point in the convex set which does not lie on a line segment joining two distinct points of the set.

A polyhedron is a convex set. It is the convex hull of its extreme points which are its vertices.

Definition: A convex polyhedral set in R^n is the intersection of a finite number of half-spaces. It is thus the set of points x which satisfy the relations

$$A_1^T x \geq b_1$$
$$A_2^T x \geq b_2$$
$$- \quad - \quad -$$
$$A_m^T x \geq b_m$$

or

$$A^T x \geq b$$

where A is the n by m matrix whose columns are $A_1, \ldots, A_m$ and b is the m-component column vector

$$\begin{matrix} b_1 \\ \cdot \\ \cdot \\ \cdot \\ b_m \end{matrix} \quad .$$

In particular, the set of points x which satisfy the constraints of a linear program is a convex polyhedral set. A convex polyhedral set may contain points at an arbitrarily large distance from the origin, that is, it may not be bounded. This is why the term polyhedron is not used for such a set.

In order to justify the above definition we should show that any such set is actually convex.

Proposition: A convex polyhedral set is convex.

Proof: Suppose x and y are points in a convex polyhedral set. Then we have

$$A^T x \geq b$$

$$A^T y \geq b \, .$$

Consider the point on the line segment joining x and y.

$$z = (1 - \alpha)x + \alpha y$$

We have

$$A^T z = (1 - \alpha)A^T x + \alpha A^T y$$

$$\geq (1 - \alpha)b + \alpha b = b$$

$$A^T z \geq b \, .$$

Thus z is also in the convex polyhedral set, and the set is therefore convex.

Definition: A <u>convex polyhedral cone</u> is the intersection of a finite number of half-spaces whose bounding hyperplanes contain the origin. It is the set of points x which satisfy a relation of the form

$$A^T x \geq 0 \, .$$

Any convex polyhedral cone is a convex polyhedral set and, hence, is convex. We must show however that such a set is really a cone.

Proposition: A convex polyhedral cone is a cone.

Proof: Suppose x is a point of a convex polyhedral cone. Then

$$A^T x \geq 0 \, .$$

But then

$$A^T(\lambda x) = \lambda A^T x \geq 0$$

for all $\lambda \geq 0$. Thus it is a cone.

An important theorem of convex geometry demonstrates the equivalence of the above definition and a second definition of a convex polyhedral cone which we give now.

Definition: A convex polyhedral cone is the convex hull of a finite number of rays which emanate from the origin. If $A_1, A_2, \ldots, A_m$ are vectors having the direction of the rays, then the cone is precisely all positive linear combinations of these vectors, that is all points

$$z = A x$$

where $x \geq 0$.

That any convex polyhedral cone in a three dimensional space is consistent with both of these definitions is fairly obvious. In higher dimensional spaces, though the statement is correct, it is moderately difficult to prove. Since these geometric ideas will be used for illustration only, they are not essential to the rigor of the proofs and we will not give the proof of equivalence here.

APPENDIX B

A Fundamental Theorem on Linear Inequalities

In the following a famous theorem on homogeneous linear inequalities is proved which is the basis for the fundamental theorem of the general programming problem. The theorem asserts that there exist vectors x and y which satisfy the system of relations

$$A^T y \geq 0 \qquad (B-1a)$$

$$A x = 0 \qquad (B-1b)$$

$$x \geq 0 \qquad (B-1c)$$

$$A_1^T y + x_1 > 0 , \qquad (B-1d)$$

where A is an arbitrary m by n matrix.

It is easy to show that there are two mutually exclusive cases for this system:

Property B - 1: Any solution (x, y) of the system (B - 1) is such that either

1) $A_1^T y = 0 , \quad x_1 > 0$

or

2) $A_1^T y > 0 , \quad x_1 = 0$

<u>Proof</u>: From (B – 1a) and (B – 1b)

$$A_i^T y \geq 0 , \quad \text{all } i$$

$$x_i \geq 0 , \quad \text{all } i .$$

Therefore each term of the sum

$$\sum_i x_i A_i^T y \qquad (B - 2)$$

is non-negative. But

$$\sum_i x_i A_i^T y = x^T A^T y = y^T(Ax) = 0 ,$$

where the fact has been used that the transpose of a matrix product is the product of the transposes in reverse order. Hence each term of (B – 2) must be zero. In particular

$$x_1 A_1^T y = 0 .$$

This requires that one of x_1 and $A_1^T y$ be equal to zero. Relation (B – 1d) prohibits both being zero simultaneously. Thus the only possibilities are those stated above.

With this property, a nice geometrical interpretation of the theorem may be given. First write the system (B – 1) in the expanded form

$$\left.\begin{array}{l} A_1^T y \geq 0 \\ A_2^T y \geq 0 \\ - \; - \; - \\ A_m^T y \geq 0 \end{array}\right\} \; (B - 1a) \qquad \begin{array}{lr} 0 = x_1 A_1 + x_2 A_2 + \ldots + x_m A_m & (B - 1b) \\ x_i \geq 0 & (B - 1c) \\ & \\ A_1^T y + x_1 > 0 & (B - 1d) \end{array}$$

in line with the two interpretations for the product of a matrix and a vector given in Appendix A. Then (b - 1a) requires a hyperplane with normal vector y such that each A_i is either in the hyperplane or has a positive component along the normal vector y. Hence no point of the cone spanned by the A_i's can have a component along y of less than zero. The relations (B - 1b) and (B - 1c) require that the origin be expressed as a positive linear combination of the A_i's. Clearly these demands may be met for any matrix A by choosing $x = 0$, $y = 0$. Thus the relation (B - 1d) is what makes the theorem interesting and significant.

The geometric statement of the theorem is

Given a set of vectors $A_2, A_3, \ldots, A_m$ and a single vector $-A_1$, either

1) There exists a hyperplane which separates $-A_1$ from the cone determined by $A_2, A_3 \ldots, A_m$; specifically, there exists a hyperplane such that $-A_1$ is strictly on one side and all points of the cone are in or to the other side of the hyperplane.

or

2) The vector $-A_1$ is an element of the cone determined by $A_2, A_3, \ldots, A_m$.

This interpretation of the theorem is depicted in Figure B - 1.

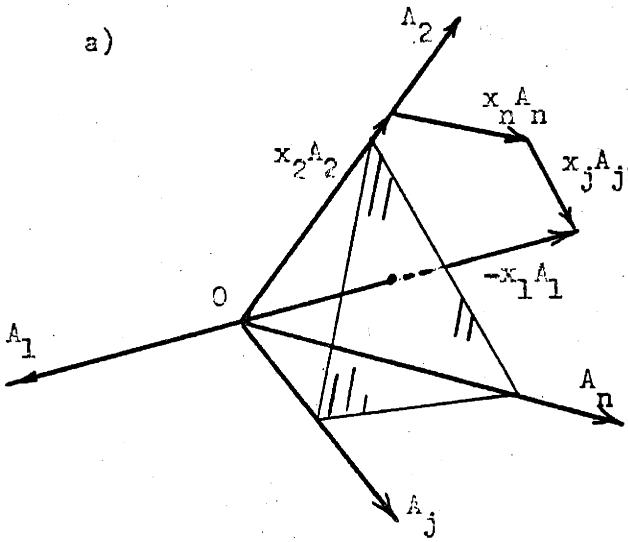
Theorem B - 1: The system of relations

$$A^T y \geq 0 \, , \quad Ax = 0 \, , \quad x \geq 0 \, , \quad A_1^T y + x_1 > 0$$

possesses a solution.

Proof: According to the property proved above any solution to the system must have either 1) $A_1^T y = 0$, $x_1 > 0$ or 2) $A_1^T y > 0$, $x_1 = 0$. Note that in the first case, if $(x, y) = (\bar{x}, \bar{y})$ is a solution with the required property, then $(x, y) = (\bar{x}, 0)$ also is. Similarly for the second case, if $(\bar{x}, \bar{y})$ is a solution, so is $(0, \bar{y})$. Our proof

a)



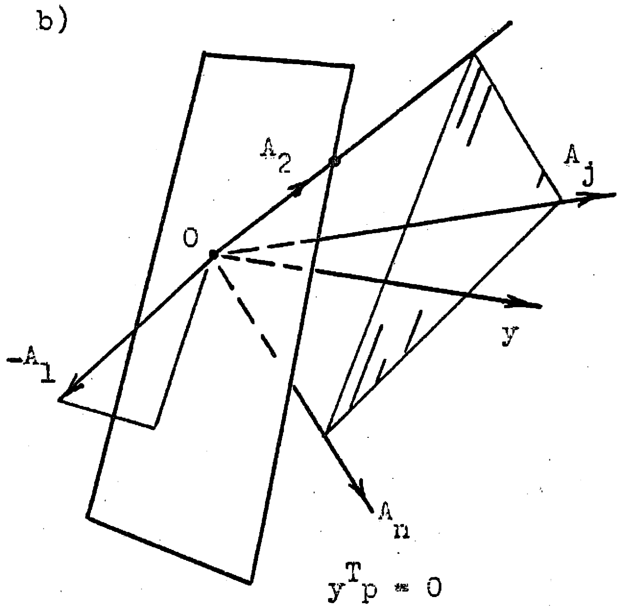
Either $- A_1$ is a positive linear combination of $A_2, \ldots, A_n$

$$- A_1 = \frac{x_2 A_2 + \ldots + x_n A_n}{x_1}$$

$$x_1 > 0,$$

that is, A_1 is in the cone spanned by $A_2, \ldots, A_n$,

b)



or, there is a hyperplane $p^T y = 0$ such that $A_2, \ldots, A_n$ are all in it or to one side of it, and $- A_1$ is strictly on the other side.

Figure B - 1--Illustrating the statement of Theorem B - 1.

will demonstrate the existence of a solution having one of these two forms.

The proof is by induction on the number m of columns in the matrix A.

A) We first consider the case where A contains one column. Then either

1) $A_1 = 0$ and $(\bar{y} = 0, \bar{x} = 1)$

is a suitable solution, or

2) $A_1 \neq 0$ and $(\bar{y} = A_1, \bar{x} = 0)$

is a suitable solution.

B) Next we assume that the theorem is true for a matrix A of m columns and demonstrate that it is true for a matrix $\bar{A} = [A, A_{m+1}]$ of m + 1 columns.

1) Suppose we have a solution (0, x) for the matrix A. Then the same solution with $x_{m+1} = 0$ is a solution for $\bar{A}$.

Geometrically the solution (0, x) for A means that $-A_1$ is contained in the convex cone determined by $A_2, A_3, \ldots, A_m$. Clearly if the cone is enlarged by the addition of an extra member to the set of generating rays, $-A_1$ will still be a member of the new cone.

2) Suppose we have a solution $(\bar{y}, 0)$ for the matrix A. Here there are two cases to consider.

a) Suppose $A_{m+1}^T \bar{y} \geq 0$. Then $(\bar{y}, 0)$ is still a suitable solution.

In this case we are given a hyperplane with normal vector $\bar{y}$ which separates $-A_1$ from the cone determined by the other columns of A. The new vector

A_{m+1} is on the same side of the hyperplane as the cone, hence the same hyperplane will serve to separate $-A_1$ from the cone defined by the enlarged set of rays.

b) Suppose $A_{m+1}^T y < 0$. In this case a solution for $\bar{A}$ cannot be constructed from the solution (y, 0) for A. We proceed by forming the matrix

$$B = [A_1 + \lambda_1 A_{m+1}, A_2 + \lambda_2 A_{m+1}, \ldots, A_m + \lambda_m A_{m+1}]$$

$$= [B_1, B_2, \ldots, B_m] ,$$

where

$$\lambda_j = - \frac{A_j^T y}{A_{m+1}^T y} \geq 0 .$$

This matrix has the property that

$$B_j^T y = A_j^T y + \lambda_1 A_{m+1}^T y$$

$$= A_j^T y - A_j^T y = 0 .$$

Thus the vectors B_j lie in the hyperplane with normal vector y. This situation is illustrated in Figure B – 2a. Each B_j is obtained from the corresponding A_j by adding sufficient amount of the vector A_{m+1} to place the result in the hyperplane.

Since B is a matrix of m columns and for the induction we have assumed the theorem valid for such a matrix, we are assured of the existence of two vectors v and u such that

$$B^T v \geq 0 , \quad Bu = 0 \qquad u \geq 0 \qquad B_1^T v + u_1 > 0$$

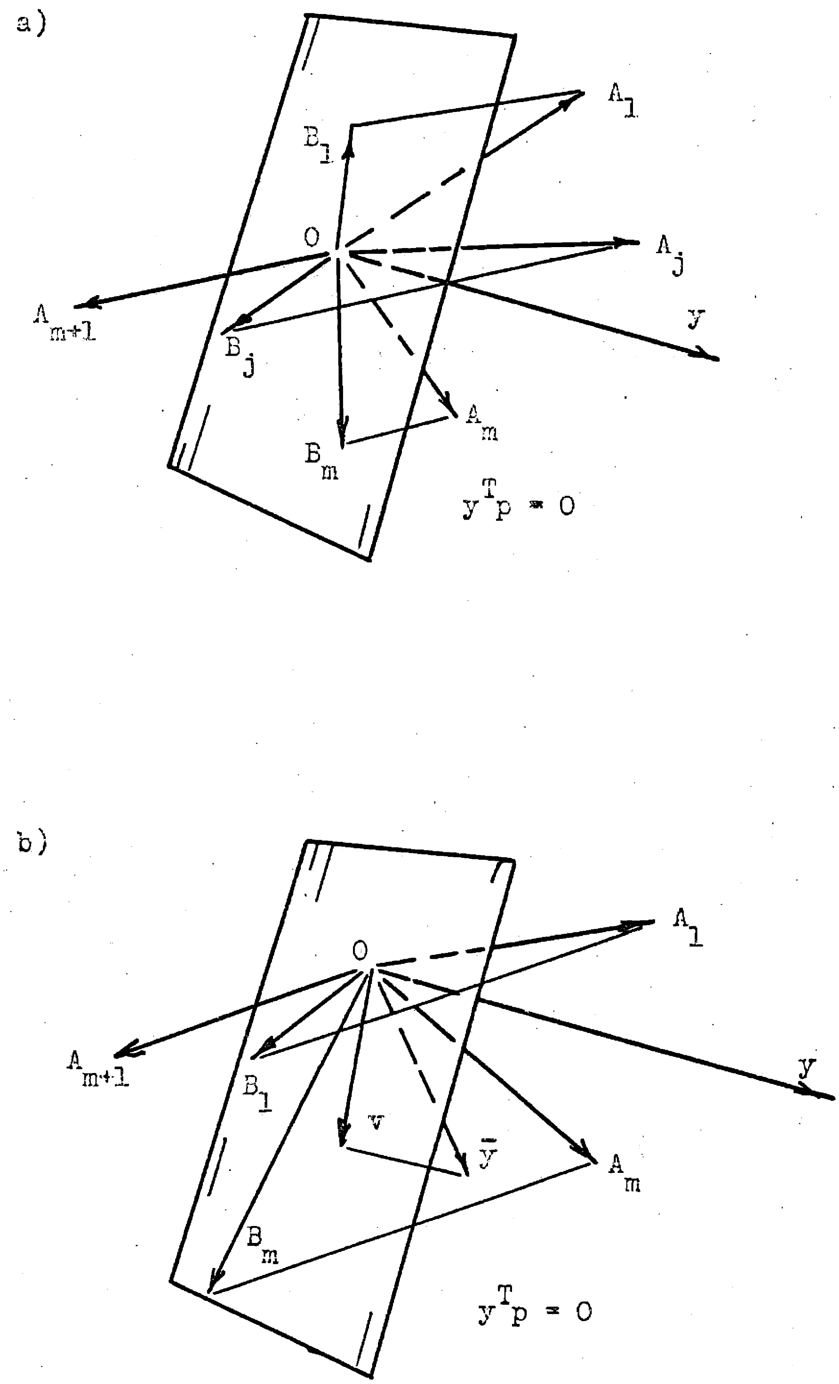


Figure B - 2—Illustrations for proof of Theorem B - 1.

We may take the solution for B to have one of the two forms

1) $B_1^T v = 0$, $u_1 > 0$ and $(0, u)$ is also a solution

2) $B_1^T v > 0$, $u = 0$ and $(v, 0)$ is also a solution

i) Suppose $(0, u)$ is a solution for B with $u_1 < 0$. We show that

$$(\bar{y} = 0, \bar{x} = (u, \sum_j \lambda_j u_j))$$

is a solution for A. Certainly

$$\bar{A}^T y \geq 0 ,$$

$x \geq 0$ because $u \geq 0$ and $\lambda_j \geq 0$,

$A_1^T y = 0$, $\bar{x}_1 = u_1 > 0$.

Now

$$\bar{A} \bar{x} = x_1 A_1 + x_2 A_2 + \ldots + x_m A_m + x_{m+1} A_{m+1}$$

$$= u_1 A_1 + u_2 A_2 + \ldots + u_m A_m + \sum_j \lambda_j u_j A_{m+1}$$

$$= u_1(A_1 + \lambda_1 A_{m+1}) + u_2(A_2 + \lambda_2 A_{m+1}) + \ldots + u_m(A_m + \lambda_m A_{m+1})$$

$$= Bu = 0$$

so that all requirements are satisfied.

The geometry of this case is shown in Figure B - 2a. Here $-B_1$ lies in the cone defined by $B_2, B_3, \ldots, B_m$. Hence $-B_1$ may be expressed as a positive linear combination of $B_2, B_3, \ldots, B_m$. By summing $A_2, A_3, \ldots, A_m$ with the same weight and adding a sufficient multiple of A_{m+1} we can express $-A_1$ as a positive linear combination of the other columns of $\bar{A}$. Hence $-A_1$ lies in the cone defined by $A_2, A_3, \ldots, A_{m+1}$.

ii) Finally, suppose $(v, 0)$ is a solution for matrix B with $B_1^T v > 0$.

We show that

$$(\bar{y} = v + \mu y, \quad \bar{x} = 0)$$

where

$$\mu = - \frac{A_{m+1}^T v}{A_{m+1}^T y}$$

is a solution for $\bar{A}$. Note that

$$A_{m+1}^T \bar{y} = A_{m+1}^T v - \frac{A_{m+1}^T v}{A_{m+1}^T y} A_{m+1}^T y = 0$$

Then certainly

$$\bar{A}\bar{x} = 0 , \quad x \geq 0 .$$

But

$$A_j^T \bar{y} = A_j^T \bar{y} + A_{m+1}^T = B_j^T \bar{y} \geq 0$$

$$= B_j^T v + \mu B_j^T y = B_j^T v \geq 0 ,$$

$$j = 1, 2, \ldots, m$$

so

$$\bar{A}^T \bar{y} \geq 0 .$$

Similarly

$$A_1^T \bar{y} = A_1^T \bar{y} + A_{m+1}^T \bar{y} = B_1^T \bar{y}$$

$$= B_1^T v > 0$$

so

$$A_1^T \bar{y} > 0 .$$

The last situation is illustrated in Figure B - 2b. There is a hyperplane with normal vector v which separates the vector $- B_1$ from the cone spanned by B_2, B_3, ..., B_m. With this we can construct a hyperplane which separates $-A_1$ from the cone spanned by A_2, A_3, ..., A_{m+1} as follows: Take as its normal vector $\bar{y}$, the sum of v and a sufficient multiple of y so that the result is orthogonal to A_{m+1}. Then, each B_j has the same component along $\bar{y}$ as it had along v because each B_j is orthogonal to y. Also the component of A_j along $\bar{y}$ is the same as the component of B_j along $\bar{y}$. Therefore all of the A_j are on the same side of the hyperplane whose normal vector is $\bar{y}$, A_1 strictly so. Hence the hyperplane separates the vector $-A_1$ from the cone as required.

Having shown the theorem valid for $m = 1$, and for $m + 1$ if it is true for m, the theorem must hold for all positive integers and the proof is complete.

APPENDIX C

The Theory of Non-Linear Programming

In this section we will consider problems of the form

minimize

$$\phi(x)$$

subject to

$$g(x) \geq 0 \qquad (C - 1)$$

$$x \geq 0 \qquad (C - 2)$$

where x is an n-vector to be determined and g(x) is a transformation from n-space into m-space

$$g(x) = \begin{bmatrix} g_1(x) \\ g_2(x) \\ - \; - \\ g_m(x) \end{bmatrix}$$

It will be assumed that ϕ and g are differentiable.

<u>Definition</u>: The set of points satisfying (C - 1) and (C - 2) is called the <u>constraint set</u> of the programming problem.

The constraint set will be identified by the letter C. It is a convex set if each coordinate function g_i is convex.

<u>Lemma C – 1</u>: Let C be the set of points which satisfy

$$g(x) \geq 0 , \quad x \geq 0 .$$

Then C is a convex set if the coordinate functions

$$g_1, \ldots, g_m \text{ of } g \text{ are convex.}$$

<u>Proof</u>: Let $\bar{x}$ and $\mathring{x}$ be any two points in C. Let $x = (1 - \alpha)\bar{x} + \alpha\mathring{x}$, $0 \leq \alpha \leq 1$, be any point on the line segment joining $\bar{x}$ and $\mathring{x}$. Obviously $x \geq 0$. Also

$$g_i(x) = g_i[(1 - \alpha)\bar{x} + \alpha\mathring{x}] \geq (1 - \alpha)\ g_i(\bar{x}) + \alpha g_i(\mathring{x}) \geq 0.$$

by the convexity of g_i.

An important question regarding programming problems asks what conditions will guarantee that a local minimum is an optimal solution. A sufficient condition is that the objective function be concave, and that the constraint set be convex as is proved below.

<u>Theorem C – 1</u>: If $\emptyset$ is a concave function and $g_1, \ldots, g_m$ are convex functions, and x minimizes $\emptyset(x)$ subject to $g(x) \leq 0$ and $x \leq 0$ in some neighborhood of $\bar{x}$, then x minimizes $\emptyset(x)$ in C.

<u>Proof</u>: Suppose $\bar{x}$ does not minimize $\emptyset(x)$ in C. Then let $\mathring{x}$ be some point in C for which $\emptyset(\mathring{x}) < \emptyset(\bar{x})$. The line segment $x = (1 - \alpha)\bar{x} + \alpha\mathring{x}$, $0 \leq \alpha \leq 1$, is entirely contained in C because C is convex. The objective function evaluated on this line segment is

$$\emptyset(\alpha) = \emptyset[(1 - \alpha)\bar{x} + \alpha\mathring{x}\ ] \geq (1 - \alpha)\ \emptyset(\bar{x}) - \alpha\emptyset\ (\mathring{x}).$$

Now there is a point on this line segment which is contained in the neighborhood of $\bar{x}$ but is not identical to $\bar{x}$. Let this point be

$x = (1 - \delta)\bar{x} + \delta\overset{\circ}{x}$ where $0 < \delta \leq 1$. Then

$$\phi(x) = \phi[(1 - \delta)\bar{x} + \delta\overset{\circ}{x}] \leq (1 - \delta)\,\phi(\bar{x}) + \delta\phi(\overset{\circ}{x})$$

by the concavity of $\phi(x)$. Since $\phi(\overset{\circ}{x}) < \phi(\bar{x})$, one has

$$\phi(x) \leq \phi(\bar{x}) - \delta[\phi(\bar{x}) - \phi(\overset{\circ}{x})]$$

hence

$$\phi(x) < \phi(\bar{x}) \ .$$

This contradicts the assumption that $\bar{x}$ is a local minimum.

In order to prove the fundamental theorem of non-linear programming (Theorem C – 2) it is necessary to impose a condition on the constraint set: for each boundary point of the constraint set there must exist a smooth curve terminating on the point and lying wholly within the constraint set. After proving the theorem, an example will be given which shows that the condition is necessary. A precise statement of the condition follows.

<u>Definition</u>: A set C of points satisfying a set of constraints

$$g(x) \leq 0 \ , \qquad x \leq 0$$

will be said to satisfy the <u>constraint qualification</u> if the following condition is satisfied for each point $\bar{x}$ of C.

Partition the constraints into two groups E and I such that

$$g_E(\bar{x}) = 0 \ , \qquad g_I(\bar{x}) > 0$$

$$g(x) = \begin{bmatrix} g_E(\bar{x}) \\ g_I(\bar{x}) \end{bmatrix} \ ,$$

and the components of $\bar{x}$ into two groups such that

$$\bar{x} = \begin{bmatrix} \bar{x}_P \\ \bar{x}_Z \end{bmatrix}, \quad \bar{x}_P > 0, \quad \bar{x}_Z = 0 .$$

Let $\overset{*}{x}$ be any vector such that

$$[dg_E(x), \overset{*}{x}] \geq 0 , \quad \overset{*}{x}_Z \geq 0 , \qquad (C - 3)$$

that is, $\overset{*}{x}$ points inward from the boundary of C. The condition is that for each $\overset{*}{x}$ satisfying (C - 3) there exists a differentiable curve

$$x = h(\alpha) , \quad 0 \leq \alpha \leq 1$$

such that

$$h(0) = \bar{x}$$

and

$$dh(0) = \lambda \overset{*}{x} \quad \text{for some} \quad \lambda > 0.$$

<u>Theorem C - 2</u>: (Fundamental Theorem of General Programming)

Consider the general programming problem

minimize

$$\phi(x)$$

with

$$g(x) \geq 0$$
$$x \geq 0$$

in which ϕ and g are differentiable and the constraint set C satisfies the constraint qualification. A necessary condition that $\bar{x}$ be an optimal vector of the general program is the existence of a $\bar{y}$ such that

$$\partial\phi(\bar{x}) - dg(\bar{x})^T \bar{y} \geq 0 \qquad g(\bar{x}) \geq 0$$

$$\bar{y} \geq 0 \qquad \bar{x} \geq 0 \qquad (C - 4)$$

$$\bar{y}^T g(\bar{x}) = 0 \qquad \bar{x}^T [\phi(\bar{x}) - dg(\bar{x})^T y] = 0$$

A sufficient condition that $\bar{x}$ be optimal in a neighborhood N of $\bar{x}$ is the existence of a y satisfying (C - 4) and that ϕ be concave and g convex in N.

Proof: Let $\bar{x}$ be some point in the constrained set C. Partition the constraints and variables so that

$$x = \begin{bmatrix} \bar{x}_P \\ \bar{x}_Z \end{bmatrix} . \qquad \begin{matrix} \bar{x}_P > 0 \\ \bar{x}_Z = 0 \end{matrix}$$

and

$$g(x) = \begin{bmatrix} g_E(\bar{x}) \\ g_I(\bar{x}) \end{bmatrix} \qquad \begin{matrix} g_E(\bar{x}) = 0 \\ g_I(\bar{x}) > 0 \end{matrix}$$

Now consider the system of inequalities

$$\begin{bmatrix} - \partial\phi(x) \; dg_E(x)^T & \begin{matrix} 0 \\ I_Z \end{matrix} \end{bmatrix}^T \overset{*}{x} \geq 0 \qquad (C - 5a)$$

$$\begin{bmatrix} - \partial\phi(\bar{x}) \; dg_E(\bar{x})^T & \begin{matrix} 0 \\ I_Z \end{matrix} \end{bmatrix} \begin{bmatrix} \tau \\ u_E \\ v_Z \end{bmatrix} = 0 , \qquad \begin{bmatrix} \tau \\ u_E \\ v_Z \end{bmatrix} \geq 0 \qquad (C - 5b)$$

$$[- \partial\phi(\bar{x}), \overset{*}{x}] + \tau > 0 \qquad (C - 5c)$$

According to Theorem B - 1 and Property B - 1 either this system possesses a solution for which $\tau > 0$, or it has a solution for which $[\partial\phi(\bar{x}), \overset{*}{x}] < 0$, but not both. If the system (C - 5) has a solution

$(\overset{*}{x}, \tau, u_E, v_Z)$ with $\tau > 0$, then

$$\bar{y} = \begin{bmatrix} u_E/\tau \\ \\ 0 \end{bmatrix}$$

satisfies the conditions of (C - 4). Conversely, if $\bar{y}$ satisfies (C - 4) then $(x = 0,\ \tau = 1,\ u_E = y_E,\ v_Z = \partial\phi(\bar{x}) - dg(\bar{x})^T\,\bar{y})$ is a solution of (C - 5) with $\tau > 0$.

To demonstrate the necessity assertion of the theorem, suppose that $\bar{x}$ is an optimal vector of the general programming problem. Suppose also that there is no vector y satisfying (C - 4), and a contradiction will be obtained. It follows from this last assumption that the system (C - 5) has no solution with $\tau > 0$. By Theorem D - 1, it must have a solution with $[\partial\phi(\bar{x}), \overset{*}{x}] < 0$. From (C - 5a) it is seen that x satisfies the relations

$$[dg_E(\bar{x}), \overset{*}{x}] \geq 0 , \qquad \overset{*}{x}_Z \geq 0 .$$

Thus the vector $\overset{*}{x}$ points "into" the constraint set. According to the constraint qualification a differentiable curve $f(\alpha)$ terminating at x exists which is contained entirely in C for which

$$df(0) = \lambda\,\overset{*}{x} , \quad \text{for some} \quad \lambda > 0 .$$

The derivative of ϕ along this curve is then

$$\frac{d\phi}{d\alpha}(0) = [\partial\phi(x), \lambda\,\overset{*}{x}] < 0 .$$

Therefore there are points in any neighborhood of $\bar{x}$ for which $\phi(\bar{x}) < \phi(\bar{x})$ denying that $\bar{x}$ is optimal.

To demonstrate the sufficiency assertion, suppose that $\bar{x}$ is not optimal and it will be shown that no $\bar{y}$ exists satisfying (C - 4). If

$\bar{x}$ is not optimal then there is a distinct $\overset{\circ}{x}$ in C such that $\phi(\overset{\circ}{x}) < \phi(\bar{x})$. Under the convexity assumption, the line segment

$$f(\alpha) = (1 - \alpha)\bar{x} + \alpha\overset{\circ}{x}, \quad 0 \leq \alpha \leq 1$$

joining $\overset{\circ}{x}$ and $\bar{x}$ is contained in C. Since ϕ is concave the derivative of ϕ along this segment evaluated at $\bar{x}$ must be negative

$$\frac{d\phi}{d\alpha}(0) = [\partial\phi(x), \overset{*}{x}] < 0$$

where $\overset{*}{x} = \overset{\circ}{x} - \bar{x}$. Then $\overset{*}{x}$, together with $\tau = 0$, $u_E = 0$, $v_Z = 0$, form a solution of (C - 5). It follows from Theorem B - 1 that (C - 5) has no solution with $\tau > 0$. Therefore (C - 4) has no solution unless $\bar{x}$ is optimal.

An example of a programming problem which does not satisfy the constraint qualification is the following

minimize $\quad - x_1$

with

$$\theta(x) = (1 - x_1)^3 - x_2 \geq 0$$

$$x_1 \geq 0 , \qquad x_2 \geq 0$$

It is evident that the optimal solution of the problem is the boundary point $x_1 = 1$, $x_2 = 0$. It is also true that at this point the direction $x = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ satisfies the conditions (C - 3) but is not directed into the constraint set. It is only in the case of such anomalous "cusps" that the applicability of the theorem breaks down.

To complete this appendix we will show that the constraint set determined by a set of linear inequalities always satisfies the constraint qualification.

Theorem: The convex polyhedral set

$$Ax \geq b, \quad x \geq 0$$

always satisfies the constraint qualification.

Proof: Let $\bar{x}$ be any point of the set. Partition the matrices A, b and the vector x so that

$$A_I \bar{x} > b_I$$
$$A_E \bar{x} = b_E$$
$$\bar{x}_P > 0$$
$$\bar{x}_Z = 0 .$$

Let x be any vector such that

$$A_E x \geq 0$$
$$x_Z \geq 0$$

Define the curve $h(\alpha)$ by

$$h(\alpha) = \bar{x} + \alpha x$$

Now

$$A_I h(\alpha) = A_I x + \alpha A_I x \geq b_I$$

for all positive α not greater than some $\alpha_1 > 0$, and

$$A_E h(\alpha) = A_E x + \alpha A_E x \geq b_E \quad \text{for all } \alpha \geq 0.$$

Also

$$[h(\alpha)]_P = x_P + \alpha x_P \geq 0$$

for all positive α not greater than some $\alpha_2 > 0$, while

$$[h(\alpha)]_Z = x_Z + \alpha x_Z \geq 0 \quad \text{for all } \alpha \geq 0.$$

Hence $h(\alpha)$ is contained in the constraint set for all α, $0 \leq \alpha \leq \text{Min}[\alpha_1, \alpha_2]$. Obviously $dh(0) = \lambda x$ for $\lambda = 1$.

APPENDIX D

The Fundamental Theorems of Quadratic Programming

The development here concerns the following pair of dual quadratic programs.

Primal quadratic program

Find an (x_Q, x_L) which

minimizes $\quad \phi(x_Q, x_L) = \frac{1}{2} x_Q^T P x_Q + c_Q^T x_Q + c_L^T x_L$

subject to $\quad A_Q x_Q + A_L x_L \geq b$

$$x_Q \geq 0 \qquad x_L \geq 0$$

Dual quadratic program

Find a (y_Q, y_L) which

maximizes $\quad \theta(y_Q, y_L) = -\frac{1}{2} y_Q^T P^{-1} y_Q + b^T y_L$

subject to $\quad A_Q^T y_L - y_Q \leq c_Q$

$$A_L^T y_L \leq c_L$$

y_Q unrestricted, $\qquad y_L \geq 0$.

In these formulations P is assumed to be a symmetric positive definite matrix. If follows that P^{-1} always exists and is also a symmetric positive definite matrix.

Note that the primal problem given here does not represent as general a class of quadratic programs as does the formulation given in section 2 - 2. Specifically only those problems inthe formulation of section 2 - 2 in which the matrix Q has the form

$$Q = \begin{bmatrix} P & 0 \\ 0 & 0 \end{bmatrix}$$

are encompassed by the primal formulation given above. The positive definite matrix P may have any order from zero to the order of Q — the former case is that of linear programming, the latter — that in which the objective function is a positive definite quadratic form.

First we will show that the objective functions of these quadratic programs are concave.

<u>Lemma D - 1</u>: The function

$$\phi(x) = x^T Qx + c^T x$$

is concave if and only if Q is positive semidefinite, and is strictly concave if and only if Q is positive definite.

<u>Proof</u>: By direct computation from the definition of a concave function

$$\begin{aligned} \phi[(1-\alpha)\overset{o}{x} + \alpha\bar{x}] = \\ &(1-\alpha)^2 \overset{o}{x}^T Q\overset{o}{x} + 2\alpha(1-\alpha)\overset{o}{x}^T Q\bar{x} + \alpha^2\bar{x}^T Q\bar{x} \\ &+ (1-\alpha)c^T\overset{o}{x} + \alpha c^T\bar{x} \\ = &(1-\alpha)\phi(\overset{o}{x}) + \alpha\phi(\bar{x}) \\ &+ \alpha(1-\alpha)(\overset{o}{x} - \bar{x})^T Q(\overset{o}{x} - \bar{x}) . \end{aligned}$$

The last term is greater than or equal to zero for all choices of $\overset{\circ}{x}$ and $\bar{x}$ if and only if Q is positive semidefinite. This term is greater than zero for all distinct $\overset{\circ}{x}$ and $\bar{x}$ if and only if Q is positive definite.

As is pointed out in Chapter Two, the quadratic programs stated above are considered duals because each yields the exact same Lagrangian problem, namely

<u>Quadratic programming Lagrangian problem</u>

Find (x_Q, x_L, y_Q, y_L) such that

$$A_Q x_Q + A_L x_L \geq b \qquad x_Q \geq 0 \qquad x_L \geq 0 \qquad \text{primal constraints} \qquad (D - 1)$$

$$A_Q^T y_L - y_Q \leq c_Q \qquad A_L^T y_L \leq c_L \qquad y_L \geq 0 \qquad \text{dual constraints} \qquad (D - 2)$$

$$x_Q = P^{-1} y_Q \qquad \text{primal-dual coupling} \qquad (D - 3)$$

$$y_L^T(A_Q x_Q + A_L x_L - b) = 0 \qquad x_Q^T(c_Q - A_Q^T y_L + y_Q) = 0 \qquad x_L^T(c_L - A_L^T y_L) = 0 \qquad \text{complementary slackness conditions} \qquad (D - 4)$$

Since the objective functions of the primal and dual quadratic programs are concave and convex, respectively, the fundamental theorem of non-linear programming (Theorem C - 2) and Theorem C - 1 gives the following result.

Theorem D - 1: A feasible vector (x_Q, x_L) of the primal quadratic program is an optimal vector if and only if there is a (y_Q, y_L) such that (x_Q, x_L, y_Q, y_L) is a solution of the quadratic programming Lagrangian problem. Likewise, a feasible vector (x_Q, x_L) of the dual quadratic program is optimal if and only if there is an (x_Q, x_L) such that (x_Q, x_L, y_Q, y_L) solves the Lagrangian problem.

Next, two simple properties are demonstrated which will be used in the proofs of the fundamental theorems of quadratic programming.

Property D - 1: If (x_Q, x_L) and (y_Q, y_L) are feasible solutions of the primal and dual quadratic programs, then $\phi(x_Q, x_L) \geq \theta(y_Q, y_L)$.

Proof: By multiplying the constraint inequalities of the primal problem by y_L one obtains

$$y_L^T A_Q x_Q + y_L^T A_L x_L \geq b^T y_L$$

Multiplying the constraints of the dual problem by x_Q and x_L, respectively, gives

$$x_Q^T A_Q^T y_L - x_Q^T y_Q \leq c_Q^T x_Q$$

$$x_L^T A_L^T y_L \leq c_L^T x_L .$$

The difference between the objective functions is

$$\phi(x_Q, x_L) - \theta(y_Q, y_L) =$$

$$\frac{1}{2} x_Q^T P x_Q + c_Q^T x_Q + c_L^T x_L$$

$$+ \frac{1}{2} x_Q^T P^{-1} y_Q - b^T y_L .$$

Using the above relations one finds

$$\phi(x_Q, x_L) - \theta(y_Q, y_L) \geq$$

$$\frac{1}{2} x_Q^T P x_Q - x_Q^T y_Q + \frac{1}{2} y_Q P^{-1} y_Q$$

$$= \frac{1}{2} [(x_Q - P^{-1} y_Q)^T P(x_Q - P^{-1} y_Q)] .$$

This quantity is a positive definite quadratic form and can never be negative.

Property D - 2: If (x_Q, x_L) and (y_Q, y_L) are feasible vectors for the primal and dual quadratic programs, and $\phi(x_Q, x_L) \leq \theta(y_Q, y_L)$, then (x_Q, x_L) and (y_Q, y_L) are optimal solutions of the primal and dual problems.

Proof: By Property D - 1, $\phi \geq \theta$, so $\phi = \theta$. If (x_Q, x_L) does not give the smallest value of ϕ the feasible vector which does and (y_Q, y_L) would violate Property D - 1. The same reasoning applies to (y_Q, y_L).

Theorem D - 2: (Duality Theorem) A feasible vector (x_Q, x_L) for the primal program is optimal if and only if there is a feasible vector (y_Q, y_L) for the dual program such that $\phi(x_Q, x_L) = \theta(y_Q, y_L)$. A feasible vector of the dual is optimal if and only if there is a feasible vector of the primal such that $\phi(x_Q, x_L) = \theta(y_Q, y_L)$.

Proof: Only the first assertion will be demonstrated here as the proof of the second is essentially the same. In view of Property D - 2, the sufficiency of the condition is clear. To show the necessity, suppose (x_Q, x_L) is an optimal vector of the primal quadratic program. Then, by Theorem D - 1, there exists a (y_Q, y_L) such that (x_Q, x_L, y_Q, y_L) solves the Lagrangian problem. This (y_Q, y_L) is then a feasible vector of the dual problem. Now consider the difference between the primal and dual objective functions.

$$\phi(x_Q, x_L) - \theta(y_Q, y_L) =$$

$$\frac{1}{2} x_Q^T P x_Q + c_Q^T x_Q + c_L^T x_L$$

$$+ \frac{1}{2} y_Q^T P^{-1} y_Q - b^T y$$

Using the complementary slackness relations (D – 4) this becomes

$$\phi - \theta = \frac{1}{2} x_Q^T P x_Q - x_Q^T y_Q + \frac{1}{2} y_Q^T P^{-1} y_Q .$$

Finally the coupling relation (D – 3) yields

$$\phi - \theta = 0 .$$

Theorem D – 3: (Existence Theorem) If both primal and dual quadratic programs possess feasible vectors, then both have optimal vectors. If either of the two programs has no feasible vector, then neither has an optimal vector.

Proof: The second assertion will be proved first. Suppose (x_Q, x_L) is a feasible vector of the primal quadratic program. By Theorem D – 1 (x_Q, x_L) can be optimal only if there is a (y_Q, y_L) that satisfies the dual restrictions. The same is true for a feasible vector of the dual problem. To show the sufficiency assertion, suppose (x_Q, x_L) and (y_Q, y_L) are feasible solutions of the primal and dual programs. It is sufficient to show that the objective function of the primal problem is bounded from below on any infinite ray in the primal constraint set. Let

$$\begin{bmatrix} x_Q \\ x_L \end{bmatrix} = \begin{bmatrix} \bar{x}_Q \\ \bar{x}_L \end{bmatrix} + \lambda \begin{bmatrix} \overset{*}{x}_Q \\ \overset{*}{x}_L \end{bmatrix}, \quad \lambda \geq 0$$

be any such ray. In order that (x_Q, x_L) be feasible for all non-negative values of λ we must have

$$A_Q\, x_Q + A_L\, x_L = \lambda(A_Q\, \overset{*}{x}_Q + A_L\, \overset{*}{x}_L) + A_Q\, \bar{x}_Q + A_L\, \bar{x}_L \geq b$$

and

$$x_Q \geq 0\,, \quad x_L \geq 0$$

for all $\lambda \geq 0$. This is true only if $(\bar{x}_Q, \bar{x}_L)$ is a feasible vector of the primal problem and

$$A_Q\, \overset{*}{x}_Q + A_L\, \overset{*}{x}_L \geq 0$$

$$\overset{*}{x}_Q \geq 0\,, \quad \overset{*}{x}_L \geq 0\,. \qquad (D - 5)$$

Now consider the value of the primal objective function on this ray as a function of λ.

$$\begin{aligned} \phi(\lambda) &= \frac{1}{2}\, \bar{x}_Q^{\,T}\, P\, \bar{x}_Q + \lambda\, \bar{x}_Q^{\,T}\, P\, \overset{*}{x}_Q + \frac{1}{2}\, \lambda^2\, \overset{*}{x}_Q^{\,T}\, P\, \overset{*}{x}_Q \\ &\quad + c_Q^{\,T}\, (\bar{x}_Q + \overset{*}{x}_Q) + c_L^{\,T}\, (\bar{x}_L + \overset{*}{x}_L) \\ &= \phi(\bar{x}_Q, \bar{x}_L) + \lambda^2\, \frac{1}{2}\, \overset{*}{x}_Q^{\,T}\, P\, \overset{*}{x}_Q \\ &\quad + \lambda(\bar{x}_Q^{\,T}\, P\, \overset{*}{x}_Q + c_Q^{\,T}\, \overset{*}{x}_Q + c_L^{\,T}\, \overset{*}{x}_L) \\ &= \alpha + \beta\, \lambda + \gamma\, \lambda^2 \end{aligned}$$

If any component of $\overset{*}{x}_Q$ is non-zero, the coefficient of λ^2,

$$\gamma = \frac{1}{2}\, \overset{*}{x}_Q^{\,T}\, P\, \overset{*}{x}_Q$$

will be greater than zero. In this case $\phi(\lambda)$ is certainly bounded from below. On the other hand, if there is a solution of (D - 5) with $\overset{*}{x}_Q = 0$, $\phi(\lambda)$ would be bounded below on the ray only if

$$\begin{aligned}\beta &= (\bar{x}_Q^T P \overset{*}{x}_Q + c_Q^T \overset{*}{x}_Q + c_L^T \overset{*}{x}_L) \\ &= c_L^T \overset{*}{x}_L\end{aligned}$$

is non-negative. We will show that this is true. Since (y_Q, y_L) is a feasible solution of the dual problem

$$A_L^T y_L \leq c_L , \quad y_L \geq 0 .$$

Hence

$$\overset{*}{x}_L^T A_L^T y_L \leq c_L^T \overset{*}{x}_L .$$

But also, from (D - 5)

$$y_L^T A_L \overset{*}{x}_L \geq 0 .$$

Therefore

$$c_L^T \overset{*}{x}_L \geq 0$$

as desired.

Theorem D - 4: (Complementary Slackness Principle and Uniqueness) i) If (x_Q, x_L) is an optimal solution of the primal program and (y_Q, y_L) is an optimal solution of the dual program, then (x_Q, x_L, y_Q, y_L) is a solution of the Lagrangian problem. In particular, the complementary slackness conditions are satisfied. ii) The values of the variables x_Q and y_Q are unique in optimal solutions of the primal and dual programs, respectively.

Proof: All requirements of the Lagrangian problem are satisfied automatically except the complementary slackness conditions and the relation $y_Q = P\, x_Q$. It remains to demonstrate that these conditions are satisfied. Let

$$\gamma = y_L^T (A_Q\, x_Q + A_L\, x_L - b)$$

$$\varepsilon = x^T (c_Q - A_Q^T\, y_L + y_Q) \qquad (D - 6)$$

$$\delta = x_L^T (c_L - A_L^T\, y_L)$$

represent the amount of dissatisfaction of the complementary slackness conditions. Under the constraints of the primal and dual problems γ, ε and δ must each be greater than or equal to zero. We have as the difference between the primal and dual objective functions

$$\phi(x_Q, x_L) - \theta(y_Q, y_L) = \frac{1}{2}\, x_Q^T\, P x_Q + c_Q^T\, x_Q + c_L\ \ x_L + \frac{1}{2}\, y_Q^T\, P^{-1} y_Q - b^T\, y_L .$$

Using the relations

$$b^T\, y_L = y_L\, A_Q\, x_Q + y_L^T\, A_L\, x_L - \gamma$$

$$c_Q^T\, x_Q = x_Q^T\, A_Q^T\, y_L - x_Q^T\, y_Q + \varepsilon$$

$$c_L^T\, x_L = x_L^T\, A_L^T\, y_L + \delta$$

obtained from (D - 6), we find

$$\phi(x_Q, x_L) - \theta(y_Q, y_L) = \frac{1}{2}\, [x_Q^T\, P x_Q - 2 x_Q^T\, y_Q + y_Q^T\, P^{-1}\, y_Q] + \ + \ +$$

$$= \frac{1}{2}\, (x_Q - P^{-1}\, y_Q)^T\, (P x_Q - y_Q) + \gamma + \varepsilon + \delta$$

$$= \frac{1}{2}\, (P x_Q - y_Q)^T\, P^{-1}\, (P x_Q - y_Q) + \gamma + \varepsilon + \delta$$

Theorem two states that this quantity must be identically equal to zero. Each term of it is non-negative; the first term because it is a positive definite form in the quantities $(Px_Q - y_Q)$, and the other terms as a result of the primal and dual constraints. Thus each term must vanish. This requires that the relation

$$y_Q = Px_Q$$

hold, and that the complementary slackness conditions be satisfied. Thus (x_Q, x_L, y_Q, y_L) is a solution of the Lagrangian problem.

To demonstrate the uniqueness property, let $(\mathring{x}_Q, \mathring{x}_L)$ and $(\bar{x}_Q, \bar{x}_L)$ be any two optimal solutions of the primal quadratic program. Theorem D – 2 shows that the dual problem must possess an optimal solution (y_Q, y_L). According to the first assertion of the present theorem, we must have

$$y_Q + P\,\mathring{x}_Q$$

$$y_Q + P\,\bar{x}_Q\,.$$

If follows that $\mathring{x}_Q = \bar{x}_Q$. In a similar manner the uniqueness of y_Q in optimal solutions of the dual may be shown.

Finally the selection of the particular formulation of the pair of dual quadratic program employed here will be justified. We will show that the application of the duality principle to the dual problem yields the original primal problem.

The starting point is the dual problem,

Maximize

$$\Theta(\bar{y}_Q, \bar{y}_L) = -\frac{1}{2}\,\bar{y}_Q^{\,T}\,P^{-1}\,\bar{y}_Q + b^T\,\bar{y}_L$$

subject to

$$\bar{A}_Q^T \bar{y}_L - \bar{y}_Q \leq \bar{c}_Q$$

$$\bar{A}_L^T \bar{y}_L \leq \bar{c}_L$$

$$\bar{y}_L \geq 0$$

in which we have barred the vectors and matrices to distinguish them in the following. By making the identifications

$$\begin{aligned} x_Q &= \bar{y}_Q & A_Q &= \begin{bmatrix} I \\ 0 \end{bmatrix} \\ x_L &= \bar{y}_L & & \\ c_Q &= 0 & A_L &= \begin{bmatrix} -\bar{A}_Q^T \\ -\bar{A}_L^T \end{bmatrix} \\ c_L &= -\bar{b} & & \\ P &= \bar{P}^{-1} & b &= \begin{bmatrix} -\bar{c}_Q \\ -\bar{c}_L \end{bmatrix}, \end{aligned} \qquad (D-7)$$

We may express this problem in terms of a primal problem in which x_Q is unrestricted, i.e.,

Minimize

$$\frac{1}{2} x_Q^T P x_Q + c_Q^T x_Q + c_L^T x_L$$

subject to

$$A_Q x_Q + A_L x_L \geq b$$

$$x_L \geq 0 .$$

Therefore the dual for this problem has equalities for those restrictions corresponding to the x_Q:

Maximize

$$-\frac{1}{2}\, y_Q\, P^{-1}\, y_Q + b^T y_L$$

subject to

$$A_Q^T y_L - y_Q = c_Q$$

$$A_L^T y_L \leq c_L$$

$$y_L \geq 0 .$$

Using the identifications (D – 7) we may write the dual of the dual problem:

Maximize

$$-\frac{1}{2}\, y_Q^T\, \bar{P} y_Q - \begin{bmatrix} \bar{c}_Q \\ \bar{c}_L \end{bmatrix}^T y_L$$

subject to

$$[I \quad 0]\; y_L - y_Q = 0 \qquad (D - 8)$$

$$[-\bar{A}_Q \; -\bar{A}_L]\; y_L \leq -\bar{b}$$

If we identify y_L with $\begin{bmatrix} \bar{x}_Q \\ \bar{x}_L \end{bmatrix}$, then (D – 8) requires that $y_Q = \bar{x}_Q$. We may rewrite the problem as

Minimize

$$\frac{1}{2}\, \bar{x}_Q\, \bar{P} \bar{x}_Q + \bar{c}_Q^T\, \bar{x}_Q + \bar{c}_L^T\, \bar{x}_L$$

subject to

$$\bar{A}_Q\, \bar{x}_Q + \bar{A}_L\, \bar{x}_L \geq \bar{b}$$

$$x_Q \geq 0 \qquad y_Q \geq 0$$

which is precisely our primal problem as was to be shown.

APPENDIX E

Some Properties of Concave Functions

The objective in the three theorem of this section is to show that to any strictly concave function ϕ there is a corresponding strictly convex function θ such that

$$-\partial\theta = (\partial\phi)^{-1},$$

that is, the differential of the second is the negative inverse of the map which is the differential of the first.

The notation [x, y] will be used here to denote the dot product of the vectors x and y; $\phi \ \theta$ will denote the composition of the maps (functions) ϕ and θ. R^n is the Euclidian space of n dimensions. Page references in brackets are to Buck's treatise* on which these results are based. First we define the notions of concave and convex functions.

__Definition__: A mapping $\phi: D \longrightarrow R$ defined on a region D in R^n is __concave__ if and only if for every x, y in D,

$$\phi[(1-\alpha)x + \alpha y] \leq (1-\alpha)\phi(x) + \alpha\phi(y) \qquad (E-1)$$

---

* R.C. Buck, __Advanced Calculus__, McGraw-Hill, 1956.

for all α, $0 < \alpha < 1$. Such a map is strictly concave if and only if (E - 1) holds as a strict inequality for all pairs of points in D. It is convex (strictly convex) if and only if $-\phi$ is concave (strictly concave).

A generalization of the strictly increasing property [p. 20] of a function of a single variable will be used in this development.

Definition: Let $t:D \longrightarrow R^n$ be a continuous transformation [p. 165] defined on a region D in R^n. We will say t has the strict increase property if and only if for every pair of distinct points x and y in D the following relation is satisfied:

$$[t(x), (y - x)] < [t(y), (y - x)]$$

t has the strict decrease property if and only if $-t$ has the strict increase property.

The following two lemmas will be used in the proof of Theorem E - 1. Since the proofs of the lemmas are fairly long, although straightforward, and their assertions are reasonably obvious, these proofs will be omitted for the sake of brevity.

Lemma E - 1: Let [0, 1] be the closed unit interval in R. Let $\phi:I \longrightarrow R$ be a function defined on some open interval I containing [0, 1] and having a continous derivative ϕ'. Suppose $\phi'(1) \leq \phi'(0)$. Then for some α, $0 < \alpha < 1$,

$$\phi(\alpha) \geq (1 - \alpha)\, \phi(0) + \alpha\, \phi(1) .$$

Lemma E - 2: Let $\phi:I \longrightarrow R$ be a differentiable function defined on the unit interval [0, 1]. Suppose for some α, $0 < \alpha < 1$,

$$\phi(\alpha) \geq (1 - \alpha)\, \phi(0) + \alpha\, \phi(1) \text{ .}$$

Then there are two distinct points x and y, $0 < x < \alpha < y < 1$, such that

$$\phi'(x) \geq \phi'(y) \text{ .}$$

Theorem E - 1: Let D be a convex region in R^n, and let $\phi : D \longrightarrow R$ be a function with continuous first derivatives. Let $t : D \longrightarrow R^n$ be the differential [pp. 180-184] of ϕ. Then ϕ is strictly concave (convex) if and only if t has the strict increase (decrease) property.

Proof: We will demonstrate the theorem for the case in which ϕ is strictly concave. The other case follows from this by considering $-\phi$.

1) ϕ strictly concave implies t has the strict increase property. Suppose t does not have the strict increase property. Then there exist distinct points x and y in D such that

$$[t(x), (y - x)] \geq [t(y), (y - x)] \text{ .}$$

Define the real-valued function $\eta(\xi)$ on some open interval containing [0, 1] by

$$\eta(\xi) = f \circ \eta h(\xi)$$

where

$$h(\xi) = x + (y - x)\xi$$

The derivative of $\eta(\xi)$ may be found by use of the chain rule [p. 191] as

$$\eta'(\xi) = [t \circ h(\xi), h'(\xi)]$$

$$= [t \circ h(\xi), (y - x)]$$

Thus

$$\eta'(0) = [t(x),\ (y - x)]$$

$$\eta'(1) = [t(y),\ (y - x)]$$

and it follows that $\eta'(0) \geq \eta'(1)$. From Lemma E - 1 there is a point α in the open interval (0, 1) such that

$$\eta(\alpha) \geq (1 - \alpha)\,\eta(0) + \alpha\,\eta(1)\ .$$

However, $h(0) = x$, $h(1) = y$, so that

$$\eta(0) = \phi(x),\ \eta(1) = \phi(y)$$

$$\eta(\alpha) = \phi \circ h(\alpha) = \phi[(1 - \alpha)x + \alpha y]\ .$$

Thus

$$\phi[(1 - \alpha)x + \alpha y] \geq (1 - \alpha)\ \phi(x) + \alpha\ \phi(y)$$

in contradiction to the strict concavity of ϕ.

2) t having the strict increase property implies ϕ is concave. Suppose ϕ is not concave. Then there are distinct points x, y in D such that for some α, $0 < \alpha < 1$,

$$\phi[(1 - \alpha)\ x + \alpha y] \geq (1 - \alpha)\ \phi(x) + \alpha\ \phi(y)\ . \qquad (E - 2)$$

Again define $\eta(\xi)$ as before by

$$\eta(\xi) = \phi \circ h(\xi)$$

where $h(\xi) = x + (y - x)t$ and we may write in place of (E - 2)

$$\eta(\alpha) \geq (1 - \alpha)\,\eta(0) + \alpha\,\eta(1)\ .$$

According to Lemma E - 2 there exist ξ_1 and ξ_2, $0 < \xi_1 < \alpha < \xi_2 < 1$, such that

$$\eta'(\xi_1) \geq \eta'(\xi_2)\ .$$

However, we have as before

$$\eta'(\xi) = [t \circ h(\xi), (y - x)].$$

Thus

$$[t \circ h(\xi_1), (y - x)] \geq [t \circ h(\xi_2), (y - x)].$$

If we let $x_1 = h(\xi_1)$, $y_1 = h(\xi_2)$, and note that $(y_1 - x_1)$ has the same direction as $(y - x)$, we have

$$[t(x_1), (y_1 - x_1)] \geq [t(y_1), y_1 - x_1] .$$

Thus t does not have the strict increase property, and the assertion is verified.

<u>Theorem E – 2</u>: Let D be a region in R^n and let $\emptyset : D \longrightarrow R$ be a function with continuous first derivatives. Let t be the differential of $\emptyset$. Then, if $\emptyset$ is either strictly concave or strictly convex, $t : D \longrightarrow R^n$ is a one-to-one map.

<u>Proof</u>: We give the proof for $\emptyset$ strictly concave. By theorem two, t has the strict increase property. Suppose t is not one-to-one. Then there are distinct points x and y in D such that

$$t(x) = t(y) .$$

Consider

$$\eta(\xi) = \emptyset \circ h(\xi)$$

where

$$h(\xi) = x + (y - x)\xi .$$

Then

$$\eta'(0) = [t(x), (y - x)] =$$

$$\eta'(1) = [t(y), (y - x)] ,$$

contradicting the assumption that t has the strict increase property.

<u>Theorem E - 3</u>: Let D be a region in R^n, and let $t:D \longrightarrow R^n$ be a continuous transformation having the strict increase property. Let E be a region in the range of t. Then t^{-1} exists and has the strict increase property on E.

<u>Proof</u>: Theorems E - 1 and E - 2 show that t is a one-to-one transformation. Hence, t^{-1} exists. Suppose t^{-1} does not have the strict increase property. Then there are distinct points u and v in E such that

$$[t^{-1}(u), (v - u)] \geq [t^{-1}(v), (v - u)] . \qquad (E - 3)$$

Let $x = t^{-1}(u)$, $y = t^{-1}(v)$, so that x and y are in D. Since t has the strict increase property

$$[t(x), (y - x)] < [t(y), (y - x)]$$

or

$$[u, y] - [u, x] < [v, y] - [v, x] .$$

From (E - 3) we have

$$[x, v] - [x, u] \geq [y, v] - [y, u]$$

which is a contradiction.

The preceding development has demonstrated the following facts.

Any strictly concave function $\emptyset$ has a differential $\partial\emptyset$ which is one-to-one and has the strict increase property. The negative inverse of this map $(\partial\emptyset)^{-1}$ has the strict decrease property and is the differential of some convex function Θ. That is $\partial\Theta = -(\partial\emptyset)^{-1}$.

For any strictly concave function $\emptyset$, the function Θ defined by the formula

$$\Theta(y) = \emptyset\,(\partial\emptyset)^{-1}(y) - [y,\ (\partial\emptyset)^{-1}(y)]$$

is called the Legendre transformation of $\emptyset$. Its differential may be computed by the use of the chain rule [p. 191].

$$\partial\Theta(y) = [\partial\emptyset \circ (\partial\emptyset)^{-1}(y),\ d(\partial\emptyset)^{-1}(y)] - [y,\ d(\partial\emptyset)^{-1}(y)] - (\partial\emptyset)^{-1}(y)$$

$$= [(\partial\emptyset) \circ (\partial\emptyset)^{-1}(y) - y),\ d(\partial\emptyset)^{-1}(y)] - (\partial\emptyset)^{-1}(y) = - (\partial\emptyset)^{-1}(y$$

Thus the differential of Θ is precisely the negative inverse of the differential of $\emptyset$. Theorems E - 1, E - 2, and E - 3 then show that, since $\emptyset$ is strictl concave, $\partial\emptyset$ has the strict increase property and therefore that $-(\partial\emptyset)^{-1}$ exists and has the strict decrease property. Thus Θ is a strictly convex function.

APPENDIX F

Duality Relations in Concave Programming

Based on the properties of the Legendre transform developed in Appendix E, the duality principle will be extended to a class of concave and convex programming problems. Specifically the discussion will concern a primal problem in which the objective function is a strictly concave function ϕ of variables x_C plus a linear combination of a distinct set of variables x_L

Primal concave program

Find an (x_C, x_L) which

minimizes

$$\phi(x_C) + c^T x_L$$

with

$$A_L x_L + A_C x_C \geq b$$

$$x_L \geq 0 , \quad x_C \geq 0 .$$

and a dual problem in which the Legendre transform of ϕ appears.

Dual convex program

Find the (y_C, y_L) which

maximizes

$$\theta(y_C) + b^T y_L$$

subject to

$$A_C^T y_L - y_C \leq 0$$

$$A_C^T y_L \geq c$$

y_C unrestricted, $y_L \geq 0$.

Here the Legendre transform Θ is defined by

$$\Theta(y_C) = \phi \circ (\partial\phi)^{-1} (y_C) - [y_C, (\partial\phi)^{-1}(y_C)] \qquad \text{(F - 1)}$$

and has the property that

$$\partial\Theta(y_C) = -(\partial\phi)^{-1}(y_C) ,$$

as was shown in Appendix E. In view of this property both the concave and the convex programming problem given above have the same corresponding Lagrangian problem.

Concave programming Lagrangian problem

Find (x_L, x_C, y_L, y_C) such that

$$\left.\begin{aligned} A_C x_C + A_L x_L &\geq b \\ x_L \geq 0, \quad x_C &\geq 0 \end{aligned}\right\} \text{primal constraints} \qquad \text{(F - 2)}$$

$$\left.\begin{aligned} A_C^T y_L - y_C &\leq 0 \\ A_L^T y_L &\leq c \\ y_L &\geq 0 \end{aligned}\right\} \text{dual constraints} \qquad \text{(F - 3)}$$

$$x_C = -\partial\Theta(y_C) = (\partial\phi)^{-1}(y_C) \quad \text{primal-dual coupling} \qquad \text{(F - 4)}$$

$$\left.\begin{aligned} y_L^T(A_L x_L + A_C x_C - b) &= 0 \\ x_C^T(y_C - A_C^T y_L) &= 0 \end{aligned}\right\} \text{complementary slackness} \qquad \text{(F - 5)}$$

$$x_L^T(c - A_L^T y_L) \quad = 0 \qquad \text{conditions}$$

Application of the fundamental theorem of non-linear programming to the dual pair of concave programs gives the following theorem which forms the basis for the proofs of this appendix.

<u>Theorem F - 1</u>: A feasible vector (x_C, x_L) of the primal and concave program is an optimal vector if and only if there is a (y_C, y_L) such that (x_C, x_L, y_C, y_L) is a solution of the concave programming Lagrangian problem. Likewise, a feasible vector of the dual convex program is optimal if and only if there is an (x_C, x_L) such that (x_C, x_L, y_C, y_L) solves the Lagrangian problem.

Two properties of the pair of concave programs will be demonstrated which are analogous to those proved for the dual pair of quadratic programs. First, however, a lemma is necessary.

<u>Lemma F - 1</u>: If $\emptyset(x)$ is a strictly concave function and the strictly convex function $\theta(y)$ is its Legendre transform, then the quantity

$$\psi(x, y) = \emptyset(x) - \theta(y) - y^T x$$

is never negative and is zero if an only if

$$x = -\partial\theta(y) = (\partial\emptyset)^{-1}(y) \ .$$

<u>Proof</u>: Employing the defining formula (F - 1) for the Legendre transformation

$$\psi(x, y) = \emptyset(x) - \emptyset \circ (\partial\emptyset)^{-1}(y) + [y, (\partial\emptyset)^{-1}(y)] - y^T x \ .$$

Let

$$w = (\partial\emptyset)^{-1}(y) \quad \text{so that} \quad y = \partial\emptyset(w).$$

Then

$$(x, w) = \phi(x) - \phi(w) - [\partial\phi(w), (x - w)].$$

If $x = (\partial\phi)^{-1}(y) = w$, then ψ is obviously zero. Now suppose x is distinct from w. According to the mean value theorem for a function of several variables [Buck,[*] p. 199] there is a point z on the line segment joining x and w such that

$$[\partial\phi(z), (x - w)] = \phi(x) - \phi(w) \ . \qquad (F - 6)$$

Also, since $\partial\phi$ has the strict increase property,

$$[\partial\phi(z), (z - w)] > [\partial\phi(w), (z - w)] \ ,$$

or, since (z – w) is parallel to (x – w),

$$[\partial\phi(z), (x - w)] - [\partial\phi(w), (x - w)] > 0$$

Using (F – 6) this becomes

$$\phi(x) - \phi(w), (x - w)] = \psi(x, w) > 0 \ .$$

Hence, $\psi(x, y)$ is strictly greater than zero except when $x = w = (\partial\phi)^{-1}(y)$.

<u>Property F – 1</u>: If (x_C, x_L) and (y_C, y_L) are feasible solutions of the primal concave and the dual convex programs, respectively, then

$$\phi(x_C) + c^T x_C \geq \theta(y_C) + b^T y_L \ .$$

<u>Proof</u>: By multiplying the constraints of the primal concave program by y_L one obtains

$$y_L^T A_C x_C + y_L^T A_L x_L \geq b^T y_L \ .$$

---

[*] R. C. Buck, <u>Advanced Calculus</u>, McGraw-Hill, 1956.

Similarly, from the constraints of the dual problem we have

$$x_C^T A_C^T y_L - x_C^T y_C \leq 0$$

$$x_L^T A_L^T y_L \leq c$$

The difference between the objective functions of the primal and dual programs is

$$\triangle = \phi(x_C) + c^T x_L - \theta(y\ ) - b^T y_L$$

or, using the above relations,

$$\triangle \geq \phi(x_C) - \theta(y_C) - y_C^T x_C$$

This quantity is never negative by Lemma F - 1.

<u>Property F - 2</u>: If (x_C, x_L) and (y_Q, y_L) are feasible vectors for the primal and dual programs, and $\phi(x_C) + c^T x_L \leq \theta(y_C) + b^T y_L$, then (x_C, x_L) and (y_C, y_L) are optimal solutions of the primal and dual problems.

<u>Proof</u>: By Property F - 1 $\phi(x_C) + c^T x_L \geq \theta\ (y_C) + b^T y_L$. Therefore the objective functions must be equal. If (x_C, x_L) is not optimal, then some other feasible vector must give a lower value of the primal objective function and this would violate Property D - 1. The same reasoning applies to (y_C, y_L) in relation to the dual problem.

The duality theorem and complementary slackness principle for concave programming may now be demonstrated.

<u>Theorem F - 2</u>: (Duality Theorem) A feasible vector (x_C, x_L) for the primal concave program is optimal if and only if there is a feasible vector of the dual program such that the objective functions are equal.

$$\phi(x_C) + c^T x_L + \theta(y_C) + b^T y_L$$

A feasible vector of the dual is optimal if and only if there is a feasible vector of the primal such that the objectives are equal.

Proof: Only the first assertion will be demonstrated here as the proof of the second is essentially the same. In view of Property F – 2 the sufficiency of the condition is clear. To show the necessity, suppose (x_C, x_L) is an optimal vector of the primal concave program. Then by theorem F – 1 there exists a (y_C, y_L) such that (x_C, x_L, y_C, y_L) is a solution of the concave programming Lagrangian problem. This (y_C, y_L) is then a feasible vector of the dual convex program. The difference between the primal and dual objective functions is

$$\Delta = \phi(x_C) + c^T x_L - \theta(y_C) - b^T y_L .$$

Using the complementary slackness conditions (F – 5), this becomes

$$\Delta = \phi(x_C) - y_C^T x_C - \theta(y_C) .$$

The coupling relation and Lemma F – 1 show that $\Delta = 0$.

Theorem F – 3: (Complementary Slackness Principle and Uniqueness) i) If (x_C, x_L) is any optimal vector of the primal program and (y_C, y_L) is any optimal vector of the dual program, the (x_C, x_L, y_C, y_L) is a solution of the concave programming Lagrangian problem. In particular the complementary slackness conditions are satisfied. ii) The values of the variables x_C and y_C are unique in the optimal vectors of the primal and dual programs, respectively.

Proof: All requirements of the Lagrangian problem are satisfied automatically except the complementary slackness conditions F – 5 and the coupling relations $x_C = (\partial\phi)^{-1}(y_C)$. It remains to demonstrate that these conditions are satisfied. Let

$$\gamma = \bar{y}_L^T(A_L \bar{x}_L + A_C \bar{x}_C - b)$$

$$\varepsilon = \bar{x}_L^T(\bar{y}_C - A_C^T \bar{y}_L)$$

$$\delta = \bar{x}_L^T(c - A_L^T \bar{y}_L)$$

represent the amount of dissatisfaction of the complementary slackness conditions. Under the constraints of the primal and dual problems γ, ε and δ must each be greater than or equal to zero. The difference between the primal and dual objective functions is then

$$\begin{aligned} &\phi(\bar{x}_C) + c^T x_L - \theta(\bar{y}_C) - b^T \bar{y}_L \\ &\quad = [\phi(\bar{x}_C) - \theta(\bar{y}_C) - y_C^T \bar{x}_C] + \gamma + \varepsilon + \delta \end{aligned} \qquad (F - 7)$$

Each term of the above expression is non-negative—the first as a consequence of Lemma F - 1 and the others by construction. Theorem F - 2, however, requires that the primal and dual objective functions be equal and, hence, that the expression (F - 7) vanish. Since each term is non-negative, they all must vanish separately. In the case of the first term, Lemma F - 1 shows that the coupling relation (F - 4) must be satisfied and the other terms require that the complementary slackness conditions be met. Thus (x_C, x_L, y_C, y_L) is a solution of the Lagrangian problem.

To demonstrate the uniqueness property, let $(\mathring{x}_C, \mathring{x}_L)$ and $(\bar{x}_C, \bar{x}_L)$ be any two optimal vectors for the primal concave program. Theorem F - 2 shows that the dual problem must possess an optimal vector (y_C, y_L). According to the first assertion of the present theorem we must have both

$$\mathring{y}_C = \partial\phi(x_C)$$
$$\bar{y}_C = \partial\phi(x_C)$$

It follows that $\overset{\circ}{x}_C = \bar{x}_C$ because $\partial\emptyset$ is a one-to-one map. In a similar manner the uniqueness of y_C in optimal vectors of the dual program may be shown.

A P P E N D I X G

Properties of Terminal Pair Solutions

In this section the properties of the terminal pair system

η	x	ξ	y	v	
	A	e			b
	$-Q$		A^T	I	c
1			e^T		0

will be studied. In particular, the purpose of this development is to demonstrate that the set of all pairs (ξ, η) which correspond to solutions of the system is a breakpoint curve in the plane. Writing the system out one has the three relations

$$A x + e \xi = b \tag{G - 1a}$$

$$A^T y - Q x + v = c \tag{G - 1b}$$

$$\eta = - e^T y \tag{G - 1c}$$

and the complementary slackness conditions

$$x \geq 0, \quad v \geq 0, \quad x^T v = 0. \tag{G - 1d}$$

The matrices in (G - 1) are supposed to have the following properties.

A - arbitrary rectangular matrix

b, c, e - arbitrary column matrices

Q - symmetric, positive semidefinite

First, a precise definition of what is meant by a breakpoint curve must be given.

Definition: A breakpoint curve is a sequence of k line segments in the plane with the following properties:

1) The first segment is a ray with non-zero direction vector (ξ, η) where $\xi \leq 0$, $\eta \leq 0$.

2) The k^{th} segment (not necessarily distinct from the first) is a ray with non-zero direction vector (ξ, η) where $\xi \geq 0$, $\eta \geq 0$.

3) Let $(\overset{o}{\xi}_i, \overset{o}{\eta}_i)$ be the lower endpoint of the i^{th} segment (except for the first) and let $(\bar{\xi}_i, \bar{\eta}_i)$ be the upper end point of the i^{th} segment (except for the last). Then

$$\overset{o}{\xi}_{i+1} = \bar{\xi}_i , \quad \overset{o}{\eta}_{i+1} = \bar{\eta}_i, \quad i = 1, \ldots, k - 1,$$

that is, succeeding segments must join in a common point.

4) Each segment must not have a negative slope.

$$\bar{\xi}_i - \overset{o}{\xi}_i \geq 0, \qquad i = 2, \ldots, k - 1$$

$$\bar{\eta}_i - \overset{o}{\eta}_i \geq 0$$

Definition: A terminal solution of a terminal pair system is a pair (ξ, η) which corresponds to some solution of the system.

The following property of the terminal solutions will be frequently used in the proofs:

Lemma G - 1: If $(\bar{\eta}, \bar{x}, \bar{\xi}, \bar{y}, \bar{v})$ and (η, x, ξ, y, v) are any two solutions of the system, then

$$\xi \, \bar{\eta} = x^T c - \bar{y}^T b + x^T Q \, \bar{x} - \bar{v}^T x$$

Proof: By straight forward calculation using relations (G - 1c), (G - 1a) and (G - 1b) in succession one finds

$$\xi \bar{\eta} = - \xi e^T \bar{y}$$

$$= y^T A x - y^T b$$

$$= x^T c - \bar{y}^T b + x^T Q \bar{x} - \bar{v}^T x$$

First it will be shown that the set of terminal solutions of the system has a "monotone increase property", namely, if ξ is increased η does not decrease and vice versa.

Lemma G - 2: Let (ξ, η) and (ξ, η) be terminal solutions of the system (G - 1). Then

$$(\bar{\xi} - \mathring{\xi}) (\bar{\eta} - \mathring{\eta}) \geq 0$$

Proof: Let $(\bar{\eta}, \bar{x}, \bar{\xi}, \bar{y}, \bar{v})$ and $(\mathring{\eta}, \mathring{x}, \mathring{\xi}, \mathring{y}, \mathring{v})$ be any solutions of (G - 1) corresponding to the given terminal solutions. Applying Lemma G - 1 one obtains

$$(\bar{\xi} - \mathring{\xi}) (\mathring{\eta} - \mathring{\eta}) = \bar{\xi} \bar{\eta} + \xi \eta - \xi \eta$$

$$= \bar{x}^T c + \mathring{x}^T c - \bar{x}^T c - \mathring{x}^T c$$

$$+ \bar{y}^T c + \mathring{y}^T c - \mathring{y}^T c - \bar{y}^T c$$

$$+ \bar{x}^T Q \bar{x} + \mathring{x}^T Q \mathring{x} - \bar{x}^T Q \mathring{x} - \mathring{x}^T Q \bar{x}$$

$$- \bar{v}^T \bar{x} - \mathring{v}^T \mathring{x} + \mathring{v}^T \bar{x} + \bar{v}^T \mathring{x}$$

$$- (\bar{x} - \mathring{x})^T Q (\bar{x} - \mathring{x}) + \bar{v}^T \mathring{x} + \mathring{v}^T \bar{x} \geq 0$$

This expression is never less than zero because it is the sum of a positive definite form and products of non-negative vectors.

Next the existence of line segments in the plane all points of which are terminal solutions will be shown. In any solution of (G - 1), some components of x and some components of v must be zero in order that the complementary slackness condition be satisfied. Define a primal variable set P and a dual variable set D such that i not in P implies $x_i = 0$,

i not in D implies $v_i = 0$.

A choice of the sets P and D is <u>allowable</u> if each i is contained in one and only one of P and D.

<u>Lemma G - 3</u>: Let P and D be an allowable pair of primal and dual variable sets. Let S be the set of all terminal solutions (ξ, η) which correspond to solutions of the terminal pair system consistent with P and D. Then either S is empty, S is a single point, or S is a line segment (possibly semi-infinite or infinite).

<u>Proof</u>: Let $(\bar{\xi}, \bar{\eta})$ and $(\mathring{\xi}, \mathring{\eta})$ be two distinct members of S and let $(\bar{\eta}, \bar{x}, \bar{\xi}, \bar{y}, \bar{v})$ and $(\mathring{\eta}, \mathring{x}, \mathring{\xi}, \mathring{u}, \mathring{v})$ be corresponding solutions of (G - 1) consistent with P and D. Then, by direct substitution in (G - 1)

$$\begin{bmatrix} \eta \\ x \\ \xi \\ y \\ v \end{bmatrix} = (1 - \alpha) \begin{bmatrix} \bar{\eta} \\ \bar{x} \\ \bar{\xi} \\ \bar{y} \\ \bar{v} \end{bmatrix} + \alpha \begin{bmatrix} \mathring{\eta} \\ \mathring{x} \\ \mathring{\xi} \\ \mathring{y} \\ \mathring{v} \end{bmatrix}$$

is also such a solution for $0 \leq \alpha \leq 1$. Hence all points on the line segment joining $(\bar{\xi}, \bar{\eta})$ and $(\mathring{\xi}, \mathring{\eta})$ are terminal solutions and members of S.

Now let $(\overset{*}{\xi}, \overset{*}{\eta})$ be any member of S distinct from $(\bar{\xi}, \bar{\eta})$ and it will be shown that this point lies on the line determined by $(\bar{\xi}, \bar{\eta})$ and $(\mathring{\xi}, \mathring{\eta})$. Specifically it will be shown that

$$(\overset{*}{\xi} - \bar{\xi})(\mathring{\eta} - \bar{\eta}) = (\mathring{\xi} - \bar{\xi})(\overset{*}{\eta} - \bar{\eta}) .$$

This is true only if the line joining $(\bar{\xi}, \bar{\eta})$ and $(\overset{*}{\xi}, \overset{*}{\eta})$ is parallel to and hence identical with the line joining $(\bar{\xi}, \bar{\eta})$ and $(\overset{o}{\xi}, \overset{o}{\eta})$. Using the same method as in Lemma G - 2, one finds

$$(\overset{*}{\xi} - \bar{\xi})(\overset{o}{\eta} - \bar{\eta}) - (\overset{*}{x} - \bar{x})^T Q(\overset{o}{x} - \bar{x})$$

and

$$(\overset{o}{\xi} - \bar{\xi})(\overset{*}{\eta} - \bar{\eta}) = (\overset{o}{x} - \bar{x})^T Q(\overset{*}{x} - \bar{x}).$$

Since the right hand sides are identical, the proof is complete.

The next property shows that there is a unique intersection between the set of terminal solutions and any 45 degree line of negative slope.

Lemma G - 4: If the system (G - 1) has any solution then for every value of a scalar α there is a unique terminal solution (ξ, η) of the system which satisfies the condition

$$\eta + \xi = \alpha. \qquad (G - 2)$$

Proof: Consider the system of relations formed by appending the above condition to the terminal pair system:

$$\begin{aligned} A\,x + e\,\xi &= b \\ x &\geq 0 \\ A^T y - Q\,x + v &= c \\ \eta &= -e^T y \\ \xi + \eta &= \alpha \\ v &\geq 0 \\ v^T x &= 0 \end{aligned} \qquad (G - 3)$$

These relations comprise the Lagrangian problem associated with the following quadratic programming problem.

Minimize

$$\frac{1}{2} x^T Q\,x + c^T x + \frac{1}{2}\xi^2 - \alpha\,\xi \qquad (G - 4a)$$

with

$$A\,x + e\,\xi = b \qquad (G - 4b)$$
$$x \geq 0$$

By assumption the terminal pair system has some solution (η, x, ξ, y, v). If we take $\alpha = \xi - e^T y$, then (η, x, ξ, y, v) is a solution of the transformed system (G - 3) for this value of α. Hence (x, ξ) is an optimal solution of the quadratic program (G - 4). It will be shown that the objective function (G - 4a) is bounded below on any ray in the constraint set (G - 4b) regardless of the value of α. This implies that (G - 4) has an optimal vector for each α and, hence, that (G - 3) has a solution for every α.

A ray in the constraint set (G - 4b) is described by

$$\left.\begin{aligned} x &= \overset{o}{x} + \lambda\,\bar{x} \\ \xi &= \overset{o}{\xi} + \lambda\,\bar{\xi} \end{aligned}\right\},\ \lambda \geq 0$$

where

$$A\,\overset{o}{x} + e\,\overset{o}{\xi} \geq b$$
$$\overset{o}{x} \geq 0$$
$$A\,\bar{x} + e\,\bar{\xi} \geq 0$$
$$\bar{x} \geq 0$$

The objective function (G - 3a) evaluated on this ray is

$$\begin{aligned} \phi(\lambda) &= \frac{1}{2}\,(x + \lambda\,\bar{x})^T\,Q(\overset{o}{x} + \lambda\,\bar{x}) + c^T(\overset{o}{x} + \lambda\,\bar{x}) + \frac{1}{2}(\overset{o}{\xi}+\bar{\xi})^2 - \alpha(\overset{o}{\xi}+\bar{\xi}). \\ &= \phi(0) + \frac{1}{2}\,\lambda^2(\bar{x}^T Q\,\bar{x} + \bar{\xi}^2) + \lambda(\overset{o}{x}^T Q\,\bar{x} + c^T\bar{x} + \overset{o}{\xi}\,\bar{\xi} - \alpha\,\bar{\xi}) \end{aligned}$$

This is certainly bounded below if the coefficient of λ^2 is strictly positive. On the other hand, for a ray in which the coefficient of λ^2 is zero (which requires $\bar{\xi} = 0$), the objective function is

$$\phi\,(\lambda) = \lambda(\overset{o}{x}^T\ Q\,\bar{x} + c^T\,\bar{x}).$$

Since this is bounded below for some α, it must be bounded below for all α. This establishes the existence of a solution to (G - 3) for every α. It remains to show that the corresponding terminal solution is unique.

Suppose there are two terminal solutions $(\mathring{\xi}, \mathring{\eta})$ and $(\bar{\bar{\xi}}, \bar{\eta})$ which satisfy (G - 3) for the same value of α. Lemma (G - 1) and condition (G - 2) yield

$$\bar{\bar{\xi}}\,\mathring{\xi} = -\bar{\eta}\,\mathring{\xi} + \alpha\,\mathring{\xi} = -c^T\mathring{x} + b^T\bar{y} - \mathring{x}^T Q\,\bar{x} + \bar{v}^T\mathring{x} + \alpha\,\mathring{\xi}$$

Now consider the product

$$\begin{aligned}\frac{1}{2}(\bar{\bar{\xi}} - \mathring{\xi})^2 &= \frac{1}{2}\mathring{x}^T Q\,\mathring{x} + c^T\mathring{x} + \frac{1}{2}\mathring{\xi}^{\,2} - \alpha\,\mathring{\xi}\\ &\quad - \frac{1}{2}\bar{x}^T Q\,\bar{x} - c^T\bar{x} - \frac{1}{2}\bar{\bar{\xi}}^{\,2} + \alpha\,\bar{\bar{\xi}}\\ &\quad - \frac{1}{2}(\bar{x} - \mathring{x})^T\ Q(\bar{x} - \mathring{x}) - \bar{v}^T\mathring{x}\end{aligned}$$

$$\frac{1}{2}(\bar{\bar{\xi}} - \mathring{\xi})^2 = \phi(\mathring{x}, \mathring{\xi}) - \phi(\bar{x}, \bar{\bar{\xi}}) - \frac{1}{2}(\bar{x} - \mathring{x}) - \bar{v}^T\mathring{x}. \qquad \text{(G - 6)}$$

Since $(\mathring{x}, \mathring{\xi})$ and $(\bar{x}, \bar{\bar{\xi}})$ must both be optimal solutions of (G - 4), they yield identical values of the objective function. The other terms on the right side of (G - 6) can never be greater than zero. The left side is never less than zero. Hence $\bar{\bar{\xi}} - \mathring{\xi}$ and (G - 2) shows that $\bar{\eta} = \mathring{\eta}$.

The above result allows us to show that there is a ray of terminal solutions extending infinitely toward positive coordinate values, and also a ray extending toward negative coordinate values.

<u>Lemma H - 5</u>: For some allowable P and D the corresponding set of terminal solutions is a ray

$$(\xi, \eta) = (\bar{\bar{\xi}}, \bar{\eta}) + \lambda\,(\overset{*}{\xi}, \overset{*}{\eta}),\ \lambda \geq 0$$

where $\overset{*}{\xi} \leq 0$, $\overset{*}{\eta} \leq 0$ but at least one of $\overset{*}{\xi}$ and $\overset{*}{\eta}$ is non-zero. Also, for some allowable P and D the corresponding set of terminal solutions is a ray with $\overset{*}{\xi} \geq 0$, $\overset{*}{\eta} \geq 0$ but at least one of $\overset{*}{\xi}$ and $\overset{*}{\eta}$ is non-zero.

<u>Proof</u>: If all segments were finite, one could choose α sufficiently negative so that

$$\alpha < \bar{\xi}_i + \bar{\eta}_i, \text{ all } i$$

and there could be no terminal solution satisfying (G - 2). The direction of the ray must be within the range specified so that Lemma G - 2 is satisfied. Taking α to be sufficiently positive demonstrates the second assertion.

In the remaining proofs the term line segment refers to the non-empty set of terminal solutions of (G - 1) for some allowable choice of P and D. A line segment may consist of only one point. Suppose there are k choices of the sets P and D for which there are associated line segments. The <u>lower</u> and <u>upper</u> end points of the i^{th} line segment will be denoted by $(\overset{o}{\xi}_i, \overset{o}{\eta}_i)$ and $(\bar{\xi}_i, \bar{\eta}_i)$, respectively where

$$\overset{o}{\xi}_i + \overset{o}{\eta}_i \leq \bar{\xi}_i + \bar{\eta}_i.$$

<u>Lemma G - 6</u>: The line segments comprising the set of terminal solutions can be arranged in a sequence $L_1, L_2, \ldots, L_k$ so that

$$\overset{o}{\xi}_{i+1} + \overset{o}{\eta}_{i+1} \geq \bar{\xi}_i + \bar{\eta}_i, \quad i = 1, \ldots, k - 1.$$

<u>Proof</u>: Take L_1 to be the line segment extending indefinitely toward negative η and/or ξ. Given L_1 through L_i, take L_{i+1} to be that one of the remaining segments for which $\overset{o}{\xi} + \overset{o}{\eta}$ is the smallest. Then

$$\overset{o}{\xi}_{i+1} + \overset{o}{\eta}_{i+1} \geq \bar{\xi}_i + \bar{\eta}_i \qquad (G - 7)$$

because

$$\mathring{\xi}_i + \mathring{\eta}_i \leq \mathring{\xi}_{i+1} + \mathring{\eta}_{i+1}$$

and if (G - 7) were not true there would be a point on L_i with

$$\xi + \eta = \mathring{\xi}_{i+1} + \mathring{\eta}_{i+1} .$$

This would contradict Lemma G - 4) for $\alpha = \mathring{\xi}_{i+1} + \eta_{i+1}$. The last segment L_k of the sequence is the segment extending indefinitely toward positive ξ and/or η. Otherwise succeeding segments would violate Lemma G - 4).

Theorem G - 1: The set of all terminal solutions of a terminal pair system is a breakpoint curve.

Proof: The preceeding lemmas have shown that the set of all terminal solutions is a sequence of straight line segments with non-negative slopes. It remains to demonstrate that these segments join to form a continuous curve, that is, that

$$\left.\begin{aligned} \bar{\xi}_i &= \mathring{\xi}_{i+1} \\ \bar{\eta}_i &= \mathring{\eta}_{i-1} \end{aligned}\right\} , \quad i = 1, \ldots, k - 1 \qquad \text{(G - 8)}$$

Suppose not. Then for some i (G - 8) is not satisfied. Lemma (G - 6) shows that

$$\bar{\xi}_i + \bar{\eta}_i \leq \mathring{\xi}_{i+1} + \mathring{\eta}_{i+1} . \qquad \text{(G - 9)}$$

If (G - 9) is satisfied as an equality both $(\bar{\xi}_i, \bar{\eta}_i)$ and $(\mathring{\xi}_{i+1}, \mathring{\eta}_{i+1})$ would be solutions of

$$\xi + \eta = \alpha \qquad \text{(G - 10)}$$

in violation of Lemma G - 4. If (G - 9) is satisfied as an inequality there is no terminal solution satisfying (G - 10) with

$$\alpha = \frac{1}{2} (\bar{\xi}_i + \bar{\eta}_i + \mathring{\xi}_{i+1} + \mathring{\eta}_{i+1})$$

in violation of Lemma G - 4.

To conclude this section two special classes of terminal pair systems are considered. In the first the matrix Q is taken to be null and it is shown that the corresponding breakpoint curve contains only horizontal and vertical segments. Secondly, Q and either b or c are taken to be zero. In the first instance all vertical segments collapse into the η-axis, in the second, all horizontal segments collapse into the ξ-axis.

<u>Theorem G - 2</u>: The breakpoint curve for a terminal pair system consists only of vertical and horizontal line segments when Q is the null matrix.

<u>Proof</u>: Let $(\overset{\circ}{\xi}\ \overset{\circ}{\eta})$ and $(\bar{\xi}\ \bar{\xi})$ be any two terminal solutions of (G - 1) for the same segment (having the same primal and dual active variable sets, P and D). Lemma G - 1 with Q set to zero yields

$$(\bar{\xi} - \overset{\circ}{\xi})(\bar{\eta} - \overset{\circ}{\eta}) = (\bar{v} - \overset{\circ}{v})^T (\bar{x} - \overset{\circ}{x})$$

which is zero by the complementary slackness condition. Hence either

$$(\bar{\xi} - \overset{\circ}{\xi}) > 0$$

or

$$(\bar{\eta} - \overset{\circ}{\eta}) > 0$$

but not in both. In the first case the segment must be horizontal; in the second, vertical.

<u>Theorem G - 3</u>: The breakpoint curve for a terminal pair system in which matrices Q and b are null contains no vertical segment in which ξ is not zero; if Q and c are null it contains no horizontal segment in which η is not zero.

<u>Proof</u>: Let $(\xi\ \overset{o}{\eta})$ and $(\xi\ \bar{\eta})$ be two terminal solutions of (G - 1) for the same vertical segment $(\bar{\eta} > \overset{o}{\eta})$. Then, using Lemma G - 1 with Q and c set to zero one has

$$\xi(\bar{\eta} - \overset{o}{\eta}) = \bar{v}^T x - \overset{o}{y}^T b + \bar{y}^T b + \bar{y}^T b = 0$$

hence $\xi = 0$ as required. The second assertion is proved in the same manner.

The following result concerning basic solutions of the terminal pair system is used in Chapter Five.

<u>Lemma G - 7</u>: Let $\bar{z}$ be a basic solution of the terminal pair system in which the conjugate variable pair (x_b, v_b) is at breakpoint. Let $\overset{o}{z}$ and $\overset{*}{z}$ be the unit incremental solutions associated with this basic solution and suppose $\overset{o}{x}_b > 0$ and $\overset{*}{v}_b > 0$, $\overset{o}{v}_b = \overset{*}{x}_b = 0$. Then

$$\overset{o}{\xi}\ \overset{*}{\eta} > \overset{*}{\xi}\ \overset{o}{\eta}$$

<u>Proof</u>: The incremental solutions $\overset{o}{z}$ and $\overset{*}{z}$ satisfy the terminal pair system with zero right hand side. Using Lemma G - 1 we have

$$\begin{aligned}\overset{o}{\xi}\ \overset{*}{\eta} - \overset{*}{\xi}\ \overset{o}{\eta} &= \overset{o}{x}^T Q\, \overset{*}{x} - \overset{*}{v}^T \overset{o}{x} \\ &\quad - \overset{*}{x}^T Q\, \overset{o}{x} + \overset{o}{v}^T \overset{*}{x} \\ &= \overset{*}{v}_b\, \overset{o}{x}_b\ 0\ .\end{aligned}$$

APPENDIX H

The Generalized Breakpoint Stepping Method

The algorithm developed in Chapter Five for tracing the breakpoint curve of the terminal pair system

η	x	ξ	y	v	
	A	e			b
	$-Q$		A^T	I	c
1			e^T		0

(H – 1)

fails if more than one conjugate variable pair is at breakpoint at some basic solution during the computation. In this appendix, theoretical developments will be made which show how to modify the method so that it will work without exception. In order to resolve the degeneracy problem the original syste of relations (H – 1) describing the breakpoint curve will be expanded so that in a sense, breakpoints that were superimposed in the terminal pair system become distinct in a <u>generalized system</u>. Each variable Z_i of the original sy tem is replaced by a vector $\mathbf{Z}_i$ of m + n + 2 components in the generalized sys tem. The first component of the vector is identified with the original variable. Similarly each component d_i of the right hand side is replaced by a vector D_i. Hence the original system is replaced by

$$P \; Z^T = D^T \qquad (H - 2)$$

where

$$P = \begin{bmatrix} 0 & A & e & 0 & 0 \\ 0 & -Q & 0 & A^T & I \\ 1 & 0 & 0 & e^T & 0 \end{bmatrix}, \qquad D = \begin{bmatrix} b^T & c^T & 0 \\ I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

and

$$Z = [f \quad X \quad g \quad Y \quad V] .$$

Using this generalized system, it will be shown that the quantity $f + g$ increases strictly on each step of the breakpoint stepping algorithm. However, this increase is not in the ordinary algebraic sense, but in a special sense which is defined below.

Definition: A vector x is greater than zero in the lexicographic sense and one writes

$$x \overset{*}{>} 0$$

if x is not zero and its first non-zero component is positive. The vector is less than zero and one writes

$$x \overset{*}{<} 0$$

if its negative is greater than zero. The relation

$$x \overset{*}{>} y$$

is true if

$$x - y \overset{*}{>} 0 .$$

Definition: A square matrix which is the product of two vectors

$$A = x \, y^T$$

is said to be greater than zero in the <u>lexicographic sense</u> and one writes

$$A \overset{*}{>} 0$$

if A is not identically zero and

$$a_{ij} > 0$$

where i is the first non-zero row and j is the first non-zero column of A.

The following property is an immediate consequence of the definitions.

<u>Property 1</u>: The relation $x\, y^T \overset{*}{>} 0$ is true if and only if

$$x \overset{*}{>} 0, \quad y \overset{*}{>} 0 \quad \text{or} \quad x \overset{*}{<} 0, \quad y \overset{*}{<} 0 .$$

The requirement that the variables x_i and v_i be non-negative in the original system is replaced by the requirements

$$X_i \overset{*}{\geq} 0 , \qquad V_i \overset{*}{\geq} 0$$

in the generalized system. The complementary slackness condition

$$x_i\, v_i = 0, \quad \text{all } i$$

becomes

$$X_{ji}\, V_{ji} = 0, \quad \text{all } i \text{ and } j,$$

that is, each component of the new vector variable must satisfy the same condition as the single variable in the original system.

It is clear that for any matrix Z which is a solution of the generalized system, the first components of the columns of Z form a solution of the original system.

Next, the notion of a basis will be defined for the generalized system.

<u>Definition</u>: A <u>basic solution</u> of the generalized system is a solution Z in which n + 1 of the columns of X and V are identically zero. The

m+n+1 columns of P corresponding to columns of Z not required to be zero comprise the basis associated with the basic solution. A basis must always include the P_i corresponding to the vectors f, g and the columns of Y.

The following property of a basis of the generalized system is crucial in succeeding arguments.

Lemma H - 1: Let P_i i in B be the basis vectors corresponding to some basic solution Z. Then these vectors form a linearly independent set and

$$Z_i \overset{*}{>} 0 , \quad i \in B .$$

Proof: The vectors of the basis P_i, $i \in B$ form an m + n + 1 by m + n + 1 square matrix R. The corresponding columns of Z form an m + n + 2 by m + n + 1 matrix W such that

$$R W^T = D^T = \begin{bmatrix} b & I & 0 & 0 \\ c & 0 & I & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} .$$

The matrix D was constructed to have rank m + n + 1. Hence, both R and V must have rank m + n + 1. With respect to R, this shows that its columns are linearly independent; with respect to W this shows that none of its columns can be identically zero.

Thus each vector variable associated with a vector of the basis for any basic solution cannot be identically zero. On the other hand, the complementary slackness condition requires that one member of each pair of conjugate variables must be identically zero for each solution. Since n + 1 of the X_i and

V_i must be zero for a basic solution, precisely one pair of conjugate variables must have $X_i = 0$ and $V_i = 0$. The remaining X_i and V_i correspond to vectors in the basis and therefore cannot be zero by the above lemma.

Definition: For a basic solution the pair of conjugate variables (X_i, V_i) for which $X_i = 0$ and $V_i = 0$ will be said to be at breakpoint.

Since the vectors of the basis are linearly independent and form a square matrix, the inverse of this matrix exists. The row of the inverse corresponding to a column P_i of the basis will be represented by the column vector S_i. Hence

$$P_i^T S_j = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \qquad i, j \in B$$

because the product of a matrix with its inverse is the identity matrix.

The three theorems which follow show that, given a basic solution of the generalized system, either a new basic solution may be formed with increased f + g or a class of solutions is available on which the quantity f + g increases indefinitely. In Theorems H - 1 and H - 2 the existence of a class of solutions for which f + g increases and one for which f + g decreases is demonstrated as a preliminary for Theorem H - 3 where the main result is proved.

Theorem H - 1: Let $\bar{Z}$ be a solution of the generalized system in which the conjugate variable pair (X_j, V_j) is at breakpoint. Then a class of solutions $Z(t)$ can be constructed for which $X_j(t) \overset{*}{>} 0$ for $t \overset{*}{>} 0$, and one for which $V_j(t) \overset{*}{>} 0$ for $t \overset{*}{>} 0$.

Proof: Let P_h be the column of P associated with the variable vector X_j. P_h is not a member of the basis, but can be expressed in terms of the linearly independent basis vectors.

$$P_h = I\ P_y = \sum_{i \epsilon B} (P_i^T\ S_i)\ P_h$$

$$= \sum_{i \epsilon B} (S_i^T\ P_h)\ P_i$$

Since $\bar{Z}$ is a basic solution

$$\sum_{i \epsilon B} P_i\ \bar{Z}_i^T = D\ .$$

Then

$$\sum_{i \epsilon B} P_i\ \bar{Z}_i^T - P_h\ t^T + P_h\ t^T = D$$

or

$$\sum_{i \epsilon B} P_i(\bar{Z}_i - S_i^T\ P_h)\ t)^T = D$$

Therefore

$$Z_i(h) = \begin{cases} \bar{Z}_i - (S_i^T\ P_h)h, & i \epsilon B \\ h, & i = h \\ 0, & \text{other } i \end{cases}$$

is a solution of the generalized system for $h \overset{*}{\geq} 0$ as long as $(S_i + P_h)t \overset{*}{\leq} \bar{Z}_i$ for all i in B corresponding to vector variables in **X** or V.

In a similar manner let P_k be the column of **P** associated with the variable vector $\mathbf{V}_j$. P_k is not in the basis, but

$$P_k = \sum_{i \epsilon B} (S_i^T\ P_k)\ P_i\ .$$

and

$$\mathbf{Z}_i(k) = \begin{cases} \mathbf{Z}_i - (\mathbf{S}_i^T P_k)\, k & i \varepsilon B \\ k & i = k \\ 0 & \text{other } i \end{cases}$$

is a solution of the modified system for $t \overset{*}{\geq} 0$ as long as

$$(\mathbf{S}_i^T P_k) k \leq \bar{\mathbf{Z}}_i .$$

The proof of the next theorem requires two preliminary results.

Lemma H - 2: Let $\bar{\mathbf{Z}}$ be a basic solution of the generalized terminal pair system in which the pair of conjugate variable vectors (X_j, V_j) is at breakpoint. Let

$$\overset{\circ}{\mathbf{Z}} = \mathbf{Z}(h) , \qquad h \overset{*}{>} 0$$

and

$$\overset{*}{\mathbf{Z}} = \mathbf{Z}(k) , \qquad k \overset{*}{>} 0$$

be the two classes of solutions formed in Theorem H - 1. Then

$$(\overset{*}{f} - \bar{f})\ (\overset{*}{g} - \bar{g})^T \overset{*}{\geq} 0$$

and

$$(\overset{\circ}{f} - \bar{f})\ (\overset{\circ}{g} - \bar{g})^T \overset{*}{\geq} 0 .$$

Proof: It is clear that the difference between two solutions of the generalized system is a solution of the system with its right hand side replaced by zero.

$$A(\overset{*}{X} - \bar{X}) \quad + e(\overset{*}{f} - \bar{f}) \quad = 0$$

$$- Q(\overset{*}{X} - \bar{X})^T + A^T(\overset{*}{Y} - \bar{Y})^T + (\overset{*}{V} - \bar{V}) = 0$$

$$(\overset{*}{g} - \bar{g})^T + e^T(\overset{*}{Y} - \bar{Y})^T = 0$$

Multiplying these relations by $(\overset{*}{Y} - \bar{Y})$, $(\overset{*}{X} - \bar{X})$ and $(\overset{*}{f} - \bar{f})$, respectively, and combining yields

$$T = (\overset{*}{f} - \bar{f})\,(\overset{*}{g} - \bar{g})^T = (\overset{*}{X} - \bar{X})Q\,(\overset{*}{X} - \bar{X})^T + \overset{*}{V}\,\bar{X}^T + \bar{V}\,\overset{*}{X}^T .$$

The (i, j) element of the square matrix T is

$$T_{ij} = ((\overset{*}{X} - \bar{X})^T)_i^T \;\; Q((\overset{*}{X} - \bar{X})^T)_j + (\overset{*}{V}^T)_i^T\,(\bar{X}^T)_j + (\bar{V}^T)_i^T\,(\overset{*}{X}^T)_j .$$

The complementary slackness condition shows that the last two terms are zero. (The only variable vector which was zero in $\bar{Z}$ but non-zero in $\overset{*}{Z}$ is $\overset{*}{V}_j$. However $\overset{\circ}{X}_j$ was assumed to be zero.) The matrix T is symmetric because Q is symmetric. Let r be the index of the first non-zero row of T (assuming T is not null). Then r is also the index of the first non-zero column. Since T is the product of two vectors, it follows by definition that

$$T = (\overset{*}{f} - \bar{f})\,(\overset{*}{g} - \bar{g})^T \overset{*}{\geq} 0 ,$$

The same argument proves the second assertion.

<u>Lemma H – 3</u>: Let $\bar{Z}$ be a basic solution of the generalized terminal pair system and let

$$\overset{\circ}{Z} = Z(h) , \qquad h \overset{*}{>} 0$$

and

$$\overset{*}{Z} = Z(k) , \qquad k \overset{*}{>} 0$$

be the two classes of solutions formed in Theorem H – 1. Then

$$(\overset{*}{f} - \bar{f})\,(\overset{\circ}{g} - \bar{g})^T \overset{*}{>} (\overset{*}{g} - \bar{g})\,(\overset{\circ}{f} - \bar{f})^T .$$

<u>Proof</u>: Let j be the index of the conjugate variable pair which is at breakpoint for the basic solution $\bar{Z}$. Then $\overset{\circ}{V}_j = 0$ and $\overset{*}{X}_j = 0$ while

$\mathring{X}_j = h \overset{*}{>} 0$ and $\overset{*}{V}_j = k \overset{*}{>} 0$.

By manipulation similar to that used in the proof of Lemma H - 2, one finds

$$(\overset{*}{f} - \bar{f})\,(\mathring{g} - \bar{g})^T - (\overset{*}{g} - \bar{g})\,(\mathring{f} - \bar{f})^T$$

$$= (\overset{*}{X} - \bar{X})Q\,(\mathring{X} - \bar{X})^T$$

$$+ \overset{*}{V}\,\mathring{X}^T - \bar{V}\,\mathring{X}^T + \bar{V}\,\bar{X}^T - \overset{*}{V}\,\bar{X}^T$$

$$- (\overset{*}{X} - \bar{X})Q\,(\mathring{X} - \bar{X})^T$$

$$- \overset{*}{V}\,\mathring{X}^T + \bar{V}\,\mathring{X}^T - \bar{V}\,\bar{X}^T + \overset{*}{V}\,\bar{X}^T$$

$$= \overset{*}{V}\,\mathring{X}^T - \bar{V}\,\mathring{X}^T - \overset{*}{V}\,\bar{X}^T - \overset{*}{V}\,\mathring{X}^T + \bar{V}\,\mathring{X}^T + \overset{*}{V}\,\bar{X}^T$$

But $\bar{V}\,\bar{X}^T = 0$ by the complementary slackness condition. Also

$$\mathring{V}\,\bar{X}^T = 0$$

$$\bar{V}\,\mathring{X}^T = 0$$

$\mathring{X}$ $\bar{X}$

because the same columns are non-zero in $\mathring{V}$ as in $\bar{V}$ and the only new non-zero column in $\mathring{X}$ is $\mathring{X}_j$ but it was postulated that $\mathring{V}_j = 0$. In the same manner

$$\overset{*}{V}\,\bar{X}^T = 0$$

$$\bar{V}\,\overset{*}{X}^T = 0$$

and

$$\mathring{V}\,\overset{*}{X}^T = 0$$

On the other hand

$$\overset{*}{V}\,\mathring{X}^T = \overset{*}{V}_j\,\mathring{X}_j^{\,T} = k\,h^T .$$

Thus

$$(\overset{*}{f} - \bar{f})\,(\mathring{g} - \bar{g})^T - (\overset{*}{g} - \bar{g})\,(\mathring{f} - \bar{f})^T = k\,h^T \overset{*}{>} 0$$

verifying the assertion of the lemma.

<u>Theorem H – 2</u>: Given a basic solution $\bar{Z}$ of the generalized system, and the two classes of solutions $\overset{\circ}{Z}(t)$ and $\overset{*}{Z}(t)$ constructed in Theorem H – 1, one of these classes has

$$\bar{f} + \bar{g} \overset{*}{>} \overset{\circ}{f}(h) + \overset{\circ}{g}(h) , \quad h \overset{*}{>} 0 ,$$

and the other has

$$\overset{*}{f}(k) + \overset{*}{g}(k) \overset{*}{>} \bar{f} + \bar{g}, \quad k \overset{*}{>} 0 .$$

<u>Proof</u>: From the proof of Theorem H – 1 one has

$$\begin{aligned} \overset{\circ}{f}(h) &= \bar{f} + \overset{\circ}{\alpha} h , & \overset{\circ}{\alpha} &= s_f^{\,T} P_r \\ \overset{\circ}{g}(h) &= \bar{g} + \overset{\circ}{\beta} h , & \overset{\circ}{\beta} &= s_g^{\,T} P_r \\ \overset{*}{f}(k) &= \bar{f} + \overset{*}{\alpha} k , & \overset{*}{\alpha} &= s_f^{\,T} P_s \\ \overset{*}{g}(k) &= \bar{g} + \overset{*}{\beta} k , & \overset{*}{\beta} &= s_f^{\,T} P_s \end{aligned} .$$

In terms of this notation, Lemma H – 2 states that

$$\overset{\circ}{\alpha}\, h\, \overset{\circ}{\beta}\, h^T \overset{*}{\geq} 0 \quad \text{which requires} \quad \overset{\circ}{\alpha}\, \overset{\circ}{\beta} \geq 0 \qquad (H – 3)$$

and

$$\overset{*}{\alpha}\, k\, \overset{*}{\beta}\, k^T \overset{*}{\geq} 0 \quad \text{which requires} \quad \overset{*}{\alpha}\, \overset{*}{\beta} \geq 0 . \qquad (H – 4)$$

Also Lemma H – 3 states that

$$\overset{*}{\alpha}\, k\, \overset{\circ}{\beta}\, h^T \overset{*}{>} \overset{*}{\beta}\, k\, \overset{\circ}{\alpha}\, h^T$$

which requires

$$\overset{*}{\alpha}\, \overset{\circ}{\beta} > \overset{*}{\beta}\, \overset{\circ}{\alpha} \qquad (H – 5)$$

According to this relation neither both $\overset{\circ}{\alpha}$ and $\overset{\circ}{\beta}$ can be zero, nor can both $\overset{*}{\alpha}$ and $\overset{*}{\beta}$. It follows that the quantity f + g must either increase or decrease for each of the two classes of general system solutions — it cannot remain fixed.

We shall assume that

$$\overset{\circ}{f} + \overset{\circ}{g} \overset{*}{>} \bar{f} + \bar{g} \quad \text{for} \quad h \overset{*}{>} 0 \qquad (H-6)$$

and show that

$$\overset{*}{f} + \overset{*}{g} \overset{*}{<} \bar{f} + \bar{g} \quad \text{for} \quad k \overset{*}{>} 0 \,. \qquad (H-7)$$

Relation (H – 6) requires $\overset{\circ}{\alpha} + \overset{\circ}{\beta} > 0$. Suppose $\overset{*}{\alpha} + \overset{*}{\beta} > 0$ and we will obtain a contradiction. Consider the product

$$T = (\overset{*}{f} - \overset{\circ}{f})\ (\overset{*}{g} - \overset{\circ}{g})^T$$

which must be $\overset{*}{\geq} 0$ by Lemma H – 2. Substitution yields

$$T = (\overset{*}{\alpha}\, k - \overset{\circ}{\alpha}\, h)\ (\overset{*}{\beta}\, k - \overset{\circ}{\beta}\, h)^T \,.$$

Take some sufficiently small $h \overset{*}{>} 0$ and choose $k \overset{*}{>} 0$ such that

$$\frac{1}{\overset{\circ}{\alpha} + \overset{\circ}{\beta}}\, k = \frac{1}{\overset{*}{\alpha} + \overset{*}{\beta}}\, h = t \overset{*}{\geq} 0 \,.$$

Then

$$T = (\overset{*}{\alpha}\, \overset{\circ}{\beta} - \overset{\circ}{\alpha}\, \overset{*}{\beta})\ (\overset{*}{\beta}\, \overset{\circ}{\alpha} - \overset{\circ}{\beta}\, \overset{*}{\alpha})\ t\ t^T$$

and relation (H – 5) shows that $T \overset{*}{<} 0$ in violation of Lemma H – 2.

It follows that to satisfy (H – 4) and (H – 5) we must have $\overset{*}{\alpha} + \overset{*}{\beta} > 0$ and (H – 7) is an immediate consequence. Similarly, assuming

$$\overset{\circ}{f} + \overset{\circ}{g} \overset{*}{<} \bar{f} + \bar{g} \quad \text{for} \quad h \overset{*}{>} 0$$

leads to the conclusion that

$$\overset{*}{f} + \overset{*}{g} \overset{*}{>} \bar{f} + \bar{g} \quad \text{for} \quad k \overset{*}{>} 0 \,.$$

<u>Theorem H – 3</u>: Let $\bar{Z}$ be a basic solution of the generalized system for which the pair of conjugate variables (X_j, V_j) is at breakpoint. Then either a basic solution $\overset{\circ}{Z}$ can be constructed for which

$$\overset{\circ}{f} + \overset{\circ}{g} > \bar{f} + \bar{g} \,,$$

or an infinite class of solutions $Z(t)$ can be formed in which

$$\overset{\circ}{f}(t) + \overset{\circ}{g}(t) > \bar{f} + \bar{g}, \quad \text{for all} \quad t \overset{*}{>} 0 .$$

Proof: By Theorem H - 1 there are two classes of solutions of the generalized system which can be constructed and Theorem H - 2 shows that

$$f(t) + g(t)$$

is strictly increasing on one of these. Using the notation of the proof of Theorem H - 1 suppose that

$$Z_i(t) = \begin{cases} \bar{Z}_i - (S_i^T P_h)t, & i \in B \\ t, & i = h \\ 0, & \text{other } i \end{cases}$$

was the class for which

$$f(t) + g(t) \overset{*}{>} \bar{f} + \bar{g} \quad \text{for} \quad t \overset{*}{>} 0 .$$

If $S_i^T P_h \leq 0$ for all i in B corresponding to columns of X or V, then $Z(t)$ is a solution of the modified system for all $t \overset{*}{\geq} 0$.

If not, choose r so that

$$\overset{*}{t} = \frac{1}{S_r^T P_h} \bar{Z}_r \geq \frac{1}{S_i^T P_h} \bar{Z}_i$$

for each i in B for which $S_i^T P_h > 0$. Then $\overset{*}{Z} = Z(\overset{*}{t})$ is another basic solution in which P_h is in the basis and P_r is not, and

$$\overset{*}{f} + \overset{*}{g} \overset{*}{>} \bar{f} + \bar{g} .$$

With these results rules may be formulated which remove the difficulty caused by degeneracy in the breakpoint tracing method. A little consideration shows that the succession of bases and corresponding basic solutions may be computed using the same recursion formulae given in Chapter Five. The problem arises in deciding which vector should be dropped from the basis in any given step. It turns out that sufficient information to make this decision is available from a basic solution of the ungeneralized system (H - 1).

Given a basic solution Z for the generalized system and the vector P_h to be substituted in the basis, the problem is to find r such that

$$\frac{1}{S_r^T P_h} Z_r \overset{*}{\geq} \frac{1}{S_i^T P_h} Z_i \qquad (H-8)$$

for each i corresponding to columns of P_X of P_V in the basis with $S_i^T P_h > 0$. This may be done as follows: First consider only the first components of the vector equation (H - 8) and choose r such that

$$\frac{z_{1r}}{S_r^T P_h} \geq \frac{z_{1i}}{S_i^T P_h}$$

i as above.

If there is just one such r, P_r is the vector to be dropped from the basis. Otherwise consider the second component of (H - 9).

$$\frac{z_{jr}}{S_r^T P_h} \geq \frac{z_{ji}}{S_i^T P_h} \qquad (H-9)$$

i as above, j = 2

Inspection of the generalized terminal pair system shows that

$$\sum_{i \varepsilon B} P_i z_{ji} = E_j , \qquad j = 2, 3, \ldots, m + n + 2 ,$$

where E_j is the unit vector with a one as its j^{th} component and remaining components zero. Thus the quantities z_{ji}, j > 1 are the coefficients of the expansion of the unit vectors as linear combinations of the basis vectors. It follows that these quantities are themselves the components of the basis inverse vectors.

$$z_{ij} = s_{i,j-1}, \quad i \text{ in } B, \quad j = 2, 3, \ldots, m + n + 2$$

With this result (H - 9) becomes

$$\frac{s_{jr}}{S_r^T P_h} \geq \frac{s_{ji}}{S_i^T P_h} \qquad \text{(H - 10)}$$

$$i \text{ as in (H - 8)},$$

and involves only quantities available from a basic solution of the ungeneralized system.

If (H - 10) provides a unique r for j = 2, then this r will satisfy (H - 8) and P_r is to be dropped from the basis. Otherwise apply (H - 10) with j = 3, and so on, until a unique r is obtained. If any tie is not broken by the time j = m + n + 2, it would mean that at least two columns of Z are proportional to each other. This, however, is impossible as Lemma H - 1 requires the columns of Z to be linearly independent.

BIOGRAPHICAL NOTE

Jack Bonnell Dennis was born in Elizabeth, New Jersey on October 13, 1931. After graduating from Darien High School, Darien, Connecticut in 1949, he entered the Massachusetts Institute of Technology and participated in the cooperative program in Electrical Engineering. This program culminated in the award of the Bachelor and Master of Science Degrees in June 1954.

The industrial experience of the author includes cooperative work with the Air Force Cambridge Research Center on non-linear devices, electronic digital data processing equipment and continuous wave radar. The summer of 1954 was spent on pulse circuit design at the Bell Telephone Laboratories. Since then he has been on the teaching staff of the M.I.T. Electrical Engineering Department for one year and has subsequently participated in the research program of the Operations Research Group at M.I.T.

The author is a member of Eta Kappa Nu, Sigma Xi, Tau Beta Pi and the Institute of Radio Engineers. He has published a paper on "A High-Speed Computer Technique for the Transportation Problem," which appeared in the Journal of the Association for Computing Machinery in April 1958.

The author is married to the former Jane Eldredge Hodgson of Miami, Florida.

B I B L I O G R A P H Y

1. Arrow, K.J., Hurwicz, L., "A Gradient Method for Approximating Saddle Points and Constrained Maxima," RAND Report P-228 (June 1951).

2. Barankin, E.W., Dorfman, R., "Toward Quadratic Programming," unpublished, (1955).

3. Brown, R.R., "A Generalized Computer Procedure for the Design of Optimum Systems," parts I and II, A.I.E.E. Conference Papers 58-234 and 48-235, (1958).

4. Charnes, A., "Degeneracy and Optimality in Linear Programming," Econometrica, vol. 20, no. 2.

5. Courant, R., Hilbert, D., Methoden der Mathematischen Physik, vol. 1, Julius Springer, Berlin, 1931 (Interscience, N.Y., 1943), pp. 199-201.

6. Courant, R., Hilbert, D., Methoden der Mathematischen Physik, vol. 2, Julius Springer, Berlin, 1937 (Interscience, N.Y., 1943), pp. 26-31.

7. Curry, H.B., "The Method of Steepest Descent for Non-Linear Problems of Minimization," Quarterly of Applied Mathematics, vol. 2, (1949), pp. 460-468.

8. Dantzig, G.B., "Application of the Simplex Method to a Transportation Problem," Activity Analysis of Production and Allocation, T.C. Koopmans, Ed., Wiley and Sons, (1951), pp. 359-373.

9. Dantzig, G.B., Ford, L.R., Fulkerson, D.R., "A Primal Dual Algorithm for Linear Programming," Linear Inequalities and Related Systems, H.W. Kuhn and A.W. Tucker, Eds., Princeton Univ. Press, (1956), pp. 171-182.

10. Dantzig, G.B., Orden, A., Wolfe, P., "The Generalized Simplex Method for Minimizing a Linear Form under Linear Inequality Restraints," Pacific Journal of Mathematics, vol. 5, June 1955, pp. 183-195.

11. Dennis, J.B., Nease R.F., Saunders, R.M., "System Synthesis with the Aid of Digital Computers," A.I.E.E. Conference Paper 56-326, (1956).

12. Farkas, J., "Über die theorie der einfachen Ungleichungen," Journal fur die reine und angewandte Mathematik, vol. 124, (1902), pp. 1 - 27.

DISTRIBUTION LIST

Commanding Officer, Office of Ordnance Research, Box CM, Duke Station, Durham, North Carolina.

13. Feder, D.P., "Solution of Non-Linear Equations," Wayne State University Computation Laboratory Summer Program, (July, 1956).

14. Fenchel, W., Convex Cones, Sets and Functions, Office of Naval Research Logistics Project Report, Department of Mathematics, Princeton University, (1953).

15. Frank, Marguerite, Wolfe, Philip, "An Algorithm for Quadratic Programming," Naval Research Logistics Quarterly, vol. 3, nos. 1 and 2, March and June 1956, pp. 95-110.

16. Ford, L.R., Fulkerson, D.R., " A Primal Dual Algorithm for the Capacitated Hitchcock Problem," Naval Research Logistics Quarterly, vol. 4, no. 1, (1957), pp. 47-54.

17. Ford, L.R., Fulkerson, D.R., "Maximal Flow Through a Network," Canadian Journal of Mathematics, vol. i, no. 3, (1956) pp. 399-404.

18. Hildreth, C., "A Quadratic Programming Procedure," Naval Research Logistics Quarterly, vol. 4, no. 1, (March 1947), pp. 79-86.

19. Jewell, W.S., "Optimal Flow Through Networks," Interim Technical Report No. 8, M.I.T. Project Fundamental Investigations in Methods of Operations Research, (1958).

20. Kuhn, Harold W., Tucker, A.W., "Non-linear Programming," Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, (1951), pp. 481-492.

21. Lemke, C.E., "The Dual Method of Solving the Linear Programming Problem," Naval Research Logistics Quarterly, vol. 1, no. 1, pp. 36-47.

22. Markowitz, H., "The Optimization of a Quadratic Function Subject to Linear Constraints," Naval Research Logistics Quarterly, vol. 3, nos. 1 and 2, March and June 1956, pp. 111-133.

23. Maxwell, J.C., "A Treatise on Electricity and Magnetism," Clarendon Press, Oxford, (1873), pp. 336-337.

24. Tucker, A.W., "Dual Systems of Homogeneous Linear Relations," Linear Inequalities and Related Systems, H.W. Kuhn and A.W. Tucker, Eds., Princeton Univ. Press, (1956), pp. 3-18.

25. Wolfe, P., "The Simplex Method for Quadratic Programming," RAND Report P-1205, (October 1957).