AD A066110

LEVEL II

12

B.S

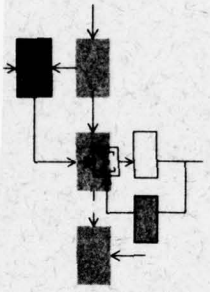
February 1979

LIDS-TH-885

Research Supported By:

*ARPA Contract ONR/N00014-75-C-1183*

DDC
RECEIVED
MAR 21 1979
C



# ROUTING TO MINIMIZE THE MAXIMUM CONGESTION IN A COMMUNICATION NETWORK

DDC FILE COPY

Francisco De Asis Ros Peran

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

Laboratory for Information and Decision Systems

Formerly

Electronic Systems Laboratory

MASSACHUSETTS INSTITUTE OF TECHNOLOGY, CAMBRIDGE, MASSACHUSETTS 02139

79 03 19 011

February 1979

LIDS-TH-885

(12)

# ROUTING TO MINIMIZE THE MAXIMUM CONGESTION IN A COMMUNICATION NETWORK

by

Francisco De Asis Ros Peran

DDC
RECEIVED
MAR 21 1979
C

This report is based on the unaltered thesis of Francisco De Asis Ros Peran, submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at the Massachusetts Institute of Technology in January 1979. This research was conducted at the M.I.T. Laboratory for Information and Decision Systems with partial support provided by the Advanced Research Project Agency of the Department of Defense under Contract ONR/ N00014-75-C-1183.

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

Laboratory for Information and Decision Systems
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

79 03 19 011

# REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

| | |
|---|---|
| 1. REPORT NUMBER (6) | 2. GOVT ACCESSION NO. (9) |
| 3. RECIPIENT'S CATALOG NUMBER: Doctoral Thesis | |
| 4. TITLE (and Subtitle): ROUTING TO MINIMIZE THE MAXIMUM CONGESTION IN A COMMUNICATION NETWORK. | 5. TYPE OF REPORT & PERIOD COVERED: Report (Thesis) |
| | 6. PERFORMING ORG. REPORT NUMBER (14): LIDS-TH-885 |
| 7. AUTHOR(s) (10): Francisco De Asis Ros Peran | 8. CONTRACT OR GRANT NUMBER(s) (15): ARPA Order No. 3045/5-7-75; ONR N00014-75-C-1183; ARPA Order-3045 |
| 9. PERFORMING ORGANIZATION NAME AND ADDRESS: Massachusetts Institute of Technology, Laboratory for Information and Decision Systems, Cambridge, Massachusetts 02139 | 10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: Program Code No. 5T10; ONR Identifying No. 049-383 |
| 11. CONTROLLING OFFICE NAME AND ADDRESS: Defense Advanced Research Projects Agency, 1400 Wilson Boulevard, Arlington, Virginia 22209 | 12. REPORT DATE (11): February 1979 |
| | 13. NUMBER OF PAGES: 332 |
| 14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): Office of Naval Research, Information Systems Program, Code 437, Arlington, Virginia 22217 (12) 336p. | 15. SECURITY CLASS. (of this report): UNCLASSIFIED |
| | 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE |

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Some new results in the area of routing algorithms for communication networks are presented. The aim has been to take as much advantage as possible of the inherent structure of the routing problem, as given by the graph supporting the network. It seems reasonable to believe that new insights in topological design can be developed as one gains new understanding of how performance varies with routing policy when topology is held fixed. For the purpose of making the topological structure of the network to explicitly appear in the routing policy, an optimization criteria that divides into a hierarchy of linear programs with a natural hierarchica

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

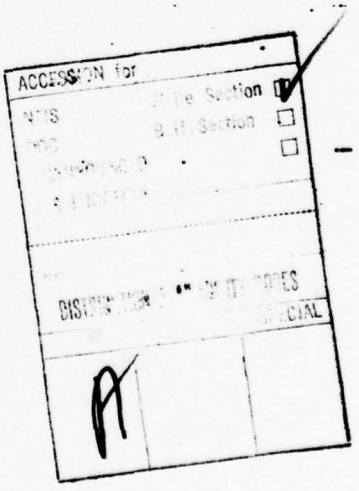
410950

20. (Continued)

partition of the network, has been chosen. The power of the criteria lies in the fact that, given the network state, it identifies the "worst" congestion problem and alleviates it with maximum resources. Furthermore the linear programming formulation allows one to express all results in terms of graph related matrices. This fact contributes, not only to reduce the amount of computation needed for finding, and dynamic tracking, the optimal solution, but it also provides with properties and simplifications that can be used for gaining a better understanding of the routing policy, or optimal routing policies in general if one assumes that optimal policies do not differ much from one another. This can also be viewed as a first step towards directly relating routing and topology, and may ultimately provide insight into the question of what is a good topology in the sense of simplifying the complexity of a routing algorithm.

A result of this work is the fact that one can describe the set of allowable paths, corresponding to an optimal solution, by means of a nonnegative distance assignment to links and the shortest route criterion. This distance assignment turns out to be highly non-unique, with a number of degrees of freedom equal to the number of problems included in the hierarchy of linear programs. An interesting property of these distance assignments is that one is always able to find an assignment with integer numbers, no matter if the capacities of the links are irrational.

An algorithm that makes use of all simplifying properties to solve the optimization problem is proposed. Closely related to it, another algorithm to track the optimal solution when changes in the requirement are present is also included. This reoptimization algorithm can work in a decentralized way whenever the requirement changes are such that no change of optimal basis is needed for any of the linear programs.



# ROUTING TO MINIMIZE THE MAXIMUM CONGESTION IN A COMMUNICATION NETWORK

by

FRANCISCO DE ASIS ROS PERAN

Ingeniero de Telecomunicación
Universidad Politécnica, Madrid (Spain)
(1972)

S.M., Massachusetts Institute of Technology
(1976)

SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE
DEGREE OF

DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

(DECEMBER, 1978)

Signature of Author.....
Department of Electrical Engineering and
Computer Science

Certified by.....
Thesis Supervisor

Accepted by.....
Chairman, Departmental Committee on Graduate Students

# ROUTING TO MINIMIZE THE MAXIMUM CONGESTION IN A COMMUNICATION NETWORK

by

Francisco A. Ros Perán

Submitted to the Department of Electrical Engineering and Computer Science on January 12, 1979, in partial fulfillment of the requirements for the Degree of Doctor of Philosophy.

## ABSTRACT

Some new results in the area of routing algorithms for communication networks are presented. The aim has been to take as much advantage as possible of the inherent structure of the routing problem, as given by the graph supporting the network. It seems reasonable to believe that new insights in topological design can be developed as one gains new understanding of how performance varies with routing policy when topology is held fixed. For the purpose of making the topological structure of the network to explicitly appear in the routing policy, an optimization criteria that divides into a hierarchy of linear programs with a natural hierarchical partition of the network, has been chosen. The power of the criteria lies in the fact that, given the network state, it identifies the "worst" congestion problem and alleviates it with maximum resources. Furthermore the linear programming formulation allows one to express all results in terms of graph related matrices. This fact contributes, not only to reduce the amount of computation needed for finding, and dynamic tracking, the optimal solution, but it also provides with properties and simplifications that can be used for gaining a better understanding of the routing policy, or optimal routing policies in general if one assumes that optimal policies do not differ much from one another. This can also be viewed as a first step towards directly relating routing and topology, and may ultimately provide insight into the question of what is a good topology in the sense of simplifying the complexity of a routing algorithm.

A result of this work is the fact that one can describe the set of allowable paths, corresponding to an optimal solution, by means of a nonnegative distance assignment to links and the shortest route criterion. This distance assignment turns out to be highly non-unique, with a number of degrees of freedom equal to the number of problems included in the hierarchy of linear programs. An interesting property of these distance assignments is that one is always able to find an

assignment with integer numbers, no matter if the capacities of the links are irrational.

An algorithm that makes use of all simplifying properties to solve the optimization problem is proposed. Closely related to it, another algorithm to track the optimal solution when changes in the requirement are present is also included. This reoptimization algorithm can work in a decentralized way whenever the requirement changes are such that no change of optimal basis is needed for any of the linear programs.

Thesis Supervisor: John M. Wozencraft

Title: Professor of Electrical Engineering and Computer Science

ACKNOWLEDGEMENTS

I wish to express by sincere gratitude to Professor John M. Wozencraft, my thesis supervisor, who suggested the topic of this thesis, generously contributed his time during this research, and also made a careful reading of the manuscript. It has been an honor and a pleasure to be associated with him, and his comments and suggestions are reflected throughout this thesis. His advice and support made the time I worked with him one of the most rewarding part of my education.

I would like to thank the members of my Doctoral Committee for invaluable counsel and personal support. To Professor Robert G. Gallager for many stimulating discussions and comments. To Professor James L. Massey for his helpful suggestions. And to Professor Gabriel Bitran for reviewing many of the results.

I am also grateful to Professor Jeffrey H. Shapiro, my faculty counselor, Professor Sanjoy K. Mitter and Professor Francis J. Reintjes for their advice during my graduate studies at M.I.T.

I am thankful to the Fundación del Instituto Tecnológico para Postgraduados (ITP), Spain, for its total financial support during my stay at M.I.T.

I thank Ms. Margaret Flaherty for her skillful typing of this thesis. This work has been partially supported by the Advanced Research Project Agency of the Department of Defense under Contract ONR/N00014-75-C-1183.

I would like to thank my parents for their confidence and encouragement. They have always been an inspiration and one of my greatest incentive.

TO

CLEMEN

TABLE OF CONTENTS


| | Page |
|---|---|
| ABSTRACT | 2 |
| ACKNOWLEDGEMENTS | 4 |
| TABLE OF CONTENTS | 6 |
| CHAPTER I. INTRODUCTION | |
| 1.1 Motivation | 9 |
| 1.2 Survey of Previous Work | 12 |
| 1.3 Thesis Organization | 18 |
| CHAPTER II. DESCRIPTION OF THE MODEL | |
| 2.1 Introduction | 22 |
| 2.2 The Basic Model | 22 |
| 2.3 A Linear Objective Function: Min-Max Routing | 27 |
| 2.4 Higher Order Criteria for the Linear Objective Function. The Successive Saturation Problem | 30 |
| CHAPTER III. THE FIRST SATURATION PROBLEM | |
| 3.1 Linear Programming Formulation. Primal and Dual Problems | 39 |
| 3.2 Optimal Solutions. Uniqueness and Degeneracy | 45 |
| 3.3 Solving the First Saturation Problem | 54 |
| 3.4 Distance Interpretation of Dual Variables. Optimal Solution as a Shortest Route Problem | 55 |
| 3.5 Graph Theoretic Properties and Bounds Related to an Optimal Solution | 63 |
| 3.6 Alternate Paths, Decision Variables and Canonic Equations | 69 |
| CHAPTER IV. MATRIX ANALYSIS OF THE FIRST SATURATION PROBLEM | |
| 4.1 Introduction | 80 |
| 4.2 The Inverse of an Optimal Basis and Other Program Matrices | 80 |
| 4.3 Properties and Graph Simplifications of Program Matrices | 109 |
| 4.4 Properties of Saturation Matrix Q | 118 |

| | | Page |
|---|---|---|
| 4.5 | The Capacity Normalized Network | 123 |
| 4.6 | The Inversion of the Saturation Matrix | 130 |
| | Method I (Gauss-Jordan) | 131 |
| | Method II | 134 |
| | Method III | 136 |
| CHAPTER V. | THE SUCCESSIVE SATURATION PROBLEM | |
| 5.1 | Introduction | 146 |
| 5.2 | The Second Saturation Level | 151 |
| 5.3 | Matrix Analysis of the Second Saturation Problem | 162 |
| 5.4 | Dual Variables Properties of the Second Level Problem. Distance Assignment. | 181 |
| 5.5 | The Third and Lower Saturation Problems | 202 |
| 5.6 | Other Results and Comments Relative to the Successive Saturation Problem | 217 |
| CHAPTER VI. | DYNAMIC ANALYSIS OF THE SUCCESSIVE SATURATION PROBLEM | |
| 6.1 | Introduction | 233 |
| 6.2 | Optimal Control and Information Flow Within a Linear Region | 235 |
| 6.3 | Change of Optimal Basis | 245 |
| 6.3.1 | A Slack Between Consecutive Saturation Levels Becomes Negative (I) | 245 |
| | Case Ia | 255 |
| | Case Ib | 257 |
| 6.3.2 | Further Comments on Case I | 259 |
| 6.3.3 | A Routing Variable Becomes Negative (II) | 265 |
| | Case IIa | 268 |
| | Case IIb | 269 |
| | Case IIc | 270 |
| 6.3.4 | Further Comments on Case II | 270 |
| 6.3.5 | A Spanning Tree Flow Becomes Negative (III) | 271 |
| | Case IIIa | 272 |
| | Case IIIb | 273 |

| | | Page |
|---|---|---|
| 6.4 | An Algorithm to Solve the Successive Saturation Problem | 273 |
| CHAPTER VII. | CONCLUSIONS | |
| 7.1 | Summary of Important Results | 278 |
| 7.2 | Areas of Future Work | 282 |
| APPENDIX I. | SOME MATRIX RESULTS | 283 |
| APPENDX II. | ON LINK INTERGER DISTANCE ASSIGNMENTS TO DESCRIBE AN OPTIMAL SOLUTION OF THE FIRST SATURATION PROBLEM | |
| AII.1 | Introduction | 287 |
| AII.2 | On the Largest Set of Integers for a Network with L Links. | 288 |
| AII.3 | On Valid Sets of Integer Distances and Corresponding Saturation Matrices | 294 |
| AII.4 | A Synthesis Procedure for the First Saturation Level | 300 |
| APPENDIX III. | EXAMPLES | 305 |
| REFERENCES | | 325 |
| BIOGRAPHICAL NOTE | | 329 |

# CHAPTER I

# INTRODUCTION

## 1.1 Motivation

Data communication networks have become an egineering reality. The literature in this area is already extensive [1]-[4] and can be expected to continue its rapid growth in the near future, since many important problems remain unsolved and others have to be reconsidered and studied under different perspectives. Experimentation with existing networks is a continuous source of new problems and ideas [5]-[6], while the availability of low cost and reliable microprocessors presents new opportunities for the design of sophisticated control systems [7]-[9]; furthermore, the price of storage devices has decreased rapidly and larger buffers are changing some aspects of data-network philosophy [5].

The overall subject of data networks can be looked upon from so many different points of view that it is worthwhile to impose some structure in terms of which specific subproblems can be isolated. Following the taxonomy adopted in [10], we can decompose an overall computer-communication network into four layers, namely

the computer subnet

the computer-communication interface

the communication subnet

the communication links,

each of which can be addressed with some degree of independence. Here we will concentrate on the communication subnet.

The main responsibilities of the communication subnet are twofold: first, to accept messages from their sources and deliver them expeditiously to their destinations; and if unable to do so, then, second to institute recovery procedures which maintain message integrity at a minimum cost. This second responsibility will not be fully considered here. For the acceptance and delivery of messages two aspects are of prime importance: (1) network topology and (2) the algorithms used to determine the routes which messages follow from their source to their destination.

The problem of routing assignments has been one of the most intensively studied areas in the field of data networks in recent years [11], [34], [35]. Desirable properties for a routing algorithm include that it must be adaptive, distributed, stable and deadlock-free. On the other hand, it must also be efficient, that is, easy to compute and with an information flow (updating traffic) as small as possible. Clearly there are difficulties in achieving all these attributes simultaneously and in quantifying the trade-offs among them.

Likewise there are multiple difficulties in the problem of network topology: First, we have the task of finding meaningful performance criteria that, subject to reasonable assumptions, will give rise to a tractable mathematical formulation. Second, we have the task of devising useful measure of performance or complexity as a function of topology. Finally, we must cope with the fact that, with adaptive routing, a single network topology corresponds to a large class of "virtual topologies", one for each possible routing policy.

In spite of the progress made in some of these areas, there still are many unanswered questions. For example, current methods are lacking

in some of the desired properties (adaptivity, etc.) mentioned before. In addition, unanticipated problems arising in practice show that many underlying issues are not yet well understood. Thus, it appears that new analytical tools and conceptual insights are needed for appropriately dealing with these problems. This present work is intended to be a step in that direction. We believe that new insights in topological design can be developed as we gain new understanding of how performance varies with routing policy when the topology is held fixed. Thus, the present work seeks to inter-relate routing and topology as much as possible. For this purpose we have chosen an optimization criteria that generates a natural hierarchical partition of the network, since it seems likely that, partitioning a network into hierarchies will be important to the efficient use of large scale networks.

Finally, we point out that in our study we will not differentiate between line-switching and packet-switching as a mode of operation. The reasons are that, on one hand, we do not see a clear distinction between these two modes and, on the other hand, that the model can be appropriately used for any of them, see [5], [10]. For obvious reasons we will not include in our model the excess traffic due to overheads, that is different for the various modes, see [6], nor the switching complexity of the nodes. Clearly, the switching operation involved in a packet network is much more complicated than that involved in a line or circuit network; for large scale networks this can be a decisive factor, although no formal study has been done to include the switching complexity as a constraint in data network design. A good starting point in this latter direction could be references [12] and [13].

## 1.2 Survey of Previous Work

Routing algorithms can be roughly classified as static, quasistatic and dynamic. Static algorithms are usually designed to determine fixed routing tables or network flows to minimize some statistic (usually the average) of network delay, assuming stationary input requirements. Such algorithms are used both in network design and in setting up fixed routing tables. A typical problem of this kind is as follows: For a given routing policy and traffic matrix, one can calculate the traffic on each link. Approximating the queueing delays on each link as a function of link traffic, one can calculate the expected delay per message in the network. The problem then is to choose the routing policy that will minimize the expected delay. This is a multicommodity flow problem and a recent algorithm for its solution due to Cantor and Gerla [14] is especially significant since it provides a mathematically elegant procedure for solving static problems with a computational efficiency comparable to the best available heuristic methods. It is worth noting, as pointed out in [10], that the mathematical problem formulation involves steady flow requirements rather than packets, and that the solution for each commodity is therefore "filamentary" in nature, by which we mean a lattice of flows with the property that flow is conserved at every intermediate node.

Two significant advances, in terms of computational efficiency, have been made by Defenderfer, in [11], beyond the original algorithm proposed in [14]. The first is an improved algorithm for determining the set of shortest distances between all pairs of nodes in a sparse network. The second is related to the significant advantages that may

be obtained in cases of practical interest by decomposing the flow into trees (one for each destination) rather than into extremal flows, as in the original Cantor-Gerla algorithm. For an excellent discussion and comparison of all these algorithms, see [11].

Quasi-static algorithms are routing procedures, in which changes of routes are allowed only at given intervals of time and/or whenever extreme changes occur. One approach to this problem is to have a special node in the network that periodically receives all the information related to the network and, at a given instant of time, reoptimizes the routes using any of the static algorithms mentioned before. These control algorithms are termed centralized. It seems clear that this quasi-static centralized policy entails a large communication problem since "state of the network" information must be transfered to a central location and then, after a calculation of a desirable routing policy has taken place at the central node, the corresponding new routing information must be distributed throughout the network. It seems possible that a distributed algorithm could reduce this communication problem at a small cost in increased nodal computer complexity. Some important contributions have appeared recently in the literature for this area; see [9] and [15] for problems and tradeoffs related to distributed routing algorithms.

There are many reasons, in addition to the communication problem mentioned before, why a distributed algorithm may be preferable to a centralized one. It can reduce reliance on a central node that might fail. It can also mitigate the effects of node or link failure since with a centralized algorithm the routes by which notification of such

failures are sent to the control node might in fact be destroyed by such failure. Further, it can reduce the protocol information flow required with a centralized algorithm, namely that from the network to the central node and from the central node to the network. All these potential advantages justify the need for research in the area of distributed algorithms. For more information of these points, see [7], [8].

This report will also investigate some aspects of decentralized routing. There are two properties that a distributed algorithm should have. On the one hand, a distributed algorithm should converge to an optimal solution which could be obtained from a corresponding centralized algorithm. This is true for the algorithms described in [9] and [15], but it is not for the ARPANET distributed algorithm [16]; the principal distinction, see [9], is that the ARPANET strives for what is called "user optimization," rather than for "system optimization." On the other hand another important characteristic that a distributed algorithm should have, as first pointed out by Gallager in [9], is that of being "loop free" at every iteration. To quote Gallager [9], "aside from reducing delay, it appears that loop freedom can be important in simplifying higher lever protocols. In fact, the major reason for building loop freedom into the algorithm was to prevent a potential deadlock in the protocol for communicating update information between the nodes." More recent work by Gallager [17] has extended the resuls on loop freedom. As we will see later some of those results generalize other results obtained independently in our research and give a clean and elegant treatment to the important problem of loops in multicommodity

flow.

The last type of routing algorithm, according to our classification, is that of dynamic routing which allows continuous changing of the routes. The routing of a particular message is not determined when it enters the network; instead, each node that receives the message selects the next node, to which the message is routed on its path to the destination, based on the present state of the network.

All the algorithms mentioned so far are based on the assumption of a convex objective function, for example, in an analytical model first introduced by Kleinrock [18] the steady state delay in each link is calculated explicitly as

$$D_\ell(f_\ell) = \frac{f_\ell}{c_\ell - f_\ell}$$

where

$f_\ell$ is the flow in link $\ell$ is messages/second

$c_\ell$ is the capacity of link $\ell$ in messages/second

$D_\ell$ is the total delay/unit time experienced by all messages in link

The total delay to be minimized by the routing procedure is then

$$D_T = \sum_\ell D_\ell(f_\ell).$$

Two important assumptions for this model are:

1) Queueing delays are the only nonnegligible source of delay in the network.

2) Each link's traffic can be modeled as Poisson message arrivals with independent exponentially distributed message lengths.

Assumption 2) is the result of the famous independence assumption requiring the messages to "lose their identity" at each node and be assigned new independent lengths. Assumption 1) has been generalized by Kleinrock himself [19], to account for overhead and propagation delays. On the other hand, in [9], [11] and [15], $D_\ell(f_\ell)$ is not required to have the form given above but only to be increasing and convex in $f_\ell$; nevertheless, the model used in these references is basically the same as that of Kleinrock.

Kleinrock's steady state and independence assumption render his model less appropriate for other than static of quasi-static types of situations, such as transients and dynamic strategies. Because of those drawbacks several attempts have been made to generalize this model. One of the most important efforts in this direction is the model proposed by Segall [20], where the contents of the queues at the nodes are viewed as continuous quantities, rather than as integer numbers of messages or bits. This macroscopic point of view not only provides a model that is analytically simpler than the classic finite-state models, but also is compatible with the fact that the effect of any single message on the total system performance should be minimal.

With this continuous model, the routing problem is formulated in [20] as a linear optimal control problem with state and control variable linear inequality constraints. The solution to this problem is approached in two ways. One way is via a feedback form obtained by means of Pontryagin's minimum principle and dynamic programming, since it turns out that for this problem the necessary conditions for optimality given

by the minimum principle are also sufficient. The other way is by replacing some constraints with penalty functions. This formulation has been used to investigate how to minimize the average message delay while disposing of whatever backlogs may exist in the network at any particular point in time. A comprehensive feedback solution has been obtained [21] for the case in which all backlog messages have the same destination and the inputs are constant in time. The approach runs into difficulties when the general case is considered and no general solution has yet been obtained.

To cope with these difficulties, while keeping the same continuous model, a second approach to the problem of dynamic routing was proposed by Wozencraft in [10]. This approach involves a change in the objective function. Rather than trying to minimize the average delay, the idea is to minimize the maximum delay. A wide variety of related network models and performance criteria can be approached from this point of view. The power of this "minimax policy" lies in the fact that, given the network state, it identifies the "worst" congestion problem and alleviates it with maximum resources.

It turns out that the minimax policy presents some canonical and structural properties that bring new insight to the problem of routing and topology for communication networks. The purpose of this research is to study in detail many of these properties and provide some general understanding of routing. The aim is to take as much advantage as possible of the inherent structure of the problem, as given by the graph supporting the network. By considering which links must be heavily loaded no matter what routing strategy is used, we can divide

the network naturally into subnetworks and study the interaction among these subnetworks. This can be viewed as a first step towards directly relating routing and topology, and may utilimately provide insight into the question of what is a good topology in the sense of simplifying the complexity of a routing algorithm.

## 1.3 Thesis Organization

The main object of this research is to obtain new insight into the problem of dynamic routing in data networks and, in particular, to study how one can effectively exploit the structure of the network as determined by the graph supporting the network.

In order to achieve these goals, we propose to work with an objective function that is simple and structured enough to allow a theoretical tracking of the optimal solution. It turns out that the minimax objective function described in Chapter II has this characteristic. The reason is that the minimax optimization problem divides into a hierarchy of linear programs that partition the set of links and requirements into subsets with a hierarchical relation among them. In Chapter II, we give a description of this objective function, with emphasis on its hierarchical structure. In Chapter II, we also introduce the basic model and approximations that will be used throughout the thesis.

In Chapter III, we describe the first linear program of the hierarchy of linear programs mentioned before. This linear program is for, what is called, the first saturation level problem. Writing the dual problem and applying some of the concepts and results on linear programming, we are able to find several properties that characterize

an optimal solution. One key result is the characterization of the optimal paths by means of an integer distance assignment to links, and the use of a shortest route criterion. We also give a set of equal-valued ratios, which we called the canonic equations, that contain all the essential information related to an optimal solution. These canonic equations allow us to define an equivalent capacity per allowable path that proves to be very useful for decentralized control of the optimal solution, within the region corresponding to an optimal basis of the linear program.

In Chapter IV we make a matrix analysis of the first saturation problem. The first major result of this section is a procedure to reconstruct the updated constraint matrix and the basis inverse using only the information in the sets of saturated arcs and nonzero flows. This result allows us to avoid the general inversion of the entire basis and to compute only those elements that are needed. For this computation, all the operations, except for the inverse of a saturation matrix whose dimension is the number of saturated links, are graph-theoretic operations. Furthermore, the partitioning we make of all matrices and vectors involved in the linear program allows us to compute, at any time, only the minimum number of elements needed to reoptimize.

Since the nature of the matrix we must invert is very important for the complexity of all these calculations, we devote part of Chapter IV to studying some properties of this matrix. Another key result of this section is the relation between the matrices of a capacity normalized network (all capacities equal to one) and a general network. It turns out that these networks are very simply related, and we can move from

one problem to the other very easily. This can be important for simplifying computation. We end the section by discussing attractive methods for inverting the saturation matrix.

In Chapter V we extend the results of Chapters III and IV to lower saturation problems. In particular we broaden the interpretation of dual variables as distance assignments and present an algorithm that, given the dual variables of every saturation problem up to level M (any M), produces a distance assignment to links that describes the optimal solutions of the global problem up to level M. As for the first level problem we can easily derive, from previous distance assignment, a valid integer assignment.

Chapter V also generalizes the matrix analysis of Chapter IV to lower levels and presents a procedure to reconstruct the updated constraint matrix and the basis inverse, from the saturated arcs and nonzero flows, by means of graph theoretic operations and the inversion of the corresponding saturation matrices. These saturation matrices are graph matrices, one corresponding to each level, whose dimension is the number of saturated links of the corresponding level. We conclude Chapter V with some comments and results relative to nonuniqueness and degeneracy of the global problem.

In Chapter VI we deal with the problem of the dynamic tracking of the optimal solution when changes in the requirement vector are present. This problem is divided in two cases, depending upon whether or not we need a change of optimal basis to accommodate the changes. For the case when a change of basis is not required we also present a decentralized procedure in which, knowing the requirement and the amount of the

change each individual node is able to compute the changes it has to introduce in the flow along the allowable paths leaving that node. For the case when a change of optimal basis is needed we study how the change can be accomplished without having to know the whole basis inverse, but only a very small part of it. It turns out that, using a modified version of the revised simplex method, we usually need to compute only a few elements of the inverse for each change. Chapter VI concludes with an algorithm to solve the successive saturation problem. Based on previous results most of the operations involved in this algorithm are graph theoretic operations.

Finally, in Chapter VII we draw together the most important conclusions obtained from this research, and suggest some of the work that still remains to be done as the object of further investigation. The appendices include examples and a few peripheral results, except for Appendix I, which is a summary of some matrix results that are used throughout the report.

# CHAPTER II

# DESCRIPTION OF THE MODEL

## 2.1 Introduction

In this section, the model, constraints and objective function that will be used throughout this research are introduced. It is a truism that, in the process of modeling a system, one must compromise between accuracy and mathematical tractability. The more detailed the model is, the more accuracy can presumably be achieved but at the cost of analytical complication. In a data-communication network the usual approach is to look at each message or packet as an entity, so that the state of the network is describable by the number and destination of the messages at each node. The difficulty with this approach is that the number of states becomes enormous, even for small networks.

The approach that we follow [20] to circumvent this difficulty is to realize that any individual packet contributes very little to the overall behavior of the network. Thus it is not really necessary to look separately at each of these messages. In our approach we will work with continuous variables, rather than with integer number of messages, bits, etc. This macroscopic point of view provides a model that is analytically simpler than the classical finite-state models.

## 2.2 The Basic Model

Consider a directed communication network consisting of N nodes and let

$$N = \{1,2,\ldots,N\}; \quad |N| = N \quad (*)$$

---

*$|A|$ means cardinality of set A.

be the collection of nodes. Similarly let

$$L = \{1,2,\ldots,L\}$$

where

$$|L| = L$$

is the collection of links. If the origin and termination of link $\ell$ are $o(\ell) = i$, and $t(\ell) = j$ respectively, then link $\ell$ can also be denoted as arc(i,j). The networks, containing these nodes and links, with which we shall be concerned, are assumed to be connected in the sense that for each pair of nodes i,j there is some, not necessarily directed, path going from i to j. If none of the paths from i to j is directed then i cannot communicate with j.

Define:

$R$ = set $\{i,j\}$ of node pairs such that node i can communicate to node j, $i \neq j$.

$r_i(k)$ = amount of input traffic at node i whose destination is node k, for every (i,k) in $R$.

$r(k)$ = set $\{r_i(k)\}$ of inputs over all $i \neq k$.

$r$ = vector whose components are the elements $r_i(k)$, all (i,k) in $R$, in some prescribed order.

For simplicity r(k) will not only denote the set defined above, but also the column vector whose components are the elements of the set. Although the ordering of the componenets in unimportant, it will be convenient, for simplicity of notation, to consider them ordered by

increasing subindex i. Note that $r_i(k)$ is not defined for pairs of nodes i,k such that i cannot communicate to k, that is, such that there exists no directed path from i to k.

If every pair of nodes can communicate then set $R$ will have $N(N-1)$ elements. Without loss of generality, and in order to simplify notation, we will assume, unless otherwise specified, that the network is strongly connected and therefore $|R| = N(N-1)$.

We will say that any traffic in the network is commodity n traffic if the final destination of that traffic is node n. We shall also, sometimes, say that the traffic is commodity (m,n) traffic if the origin was node m and the final destination is node n. Thus, $R$ is the set of all commodities (m,n) that can flow in the network.

We also define

$f_\ell(k)$ = commodity k traffic flowing over link $\ell$ [we shall also sometimes write $f_\ell(k)$ as $f_{ij}(k)$ when $\ell$ = arc(i,j)],

$f_\ell = \sum_k f_\ell(k)$ = total or aggregate traffic flowing over link $\ell$,

$f(k)$ = vector corresponding to the set $\{f_\ell(k)$, all $\ell \in L\}$

$F$ = $\{f_{ij}(k)$, all arc(i,j) $\in L$, all (i,k) $\in R\}$

= set of all existing flows,

$f$ = vector corresponding to the set $F$,

$c_\ell$ = capacity of link $\ell$.

We assume that all traffic defined for k, i.e., all commodity k traffic, is removed from the network when it reachs k. Thus, no traffic for k flows outward from k, and $f_\ell(k)$ is defined only for all $\ell \in L$ such that

$o(\ell) \neq k$. Many other situations can be reduced to the model described just by adding or substracting nodes and/or links.

The reader should not be disturbed by the fact that we have not assigned units to any of the variables defined above. For our problem, the units are immaterial as long as it is understood that all quantities and variables are measured consistently (e.g., by number of messages, by bits/seconds, by number of calls, etc.). The only restriction we make is that, as mentioned in section 2.1, the units are such that, after appropriate normalization, the amount of traffic can be well approximated by continuous variables.

The flows just defined must satisfy three basic constraints: nonnegativity, conservation, and capacity-consistency. The nonnegativity constraint states that

$$f_\ell(k) \geq 0, \qquad \text{all } \ell \in L, \quad k \in N. \tag{2.2.1}$$

The conservation constraint accounts for the fact that all traffic that comes into each node must go out again unless this node is its final destination. This constraint can be written as

$$\sum_j f_{ij}(n) - \sum_k f_{ki}(n) = \begin{cases} r_i(n) & i \neq n \\ -\sum_\ell r_\ell(n) & i = n \end{cases} \tag{2.2.2}$$

Summing over i in (2.2.2), we see that the equation corresponding to i=n is redundant. In other words (see [22], page 123), the balance of flow at N-1 nodes, and the knowledge that input equals output, together imply the balance of flow at the remaining node.

Finally, the capacity constraint is that

$$f_\ell \leq c_\ell \qquad \text{all } \ell \in L. \tag{2.2.3}$$

Definition: A set of flows f is a feasible multicommodity flow for the inputs $r \geq 0$ if (2.2.1) and (2.2.2) are satisfied.

Each feasible multicommodity flow represents a particular way of routing the inputs r to their destination, ignoring the capacity constraint. Such routings can be implemented, in the presence of dynamic variations on the inputs, by taking the dynamically varying traffic arriving at i destined for k and allocating it to outgoing links in the proportions given by flows $f_{ij}(k)$.

It will prove convenient to write equations (2.2.2) in matrix form. For each commodity, we have

$$E\, f(n) = b(n), \tag{2.2.4}$$

where

$$b(n) = [b_1(n), b_2(n), \ldots, b_N(n)]$$

$$b_i(n) \triangleq \begin{cases} r_i(n) & i \neq n \\ -\sum_\ell r_\ell(n) & i = n \end{cases}$$

and E is an N x L matrix called the node-arc incidence matrix of the network, [23], that is independent of destination node n. If $e_{i\ell}$ is the $(i,\ell)^{th}$ element of E, where i refers to a node and $\ell$ to a link, we can define E as follows

$$e_{i\ell} = \begin{cases} 1 \text{ if } o(\ell) = i \\ -1 \text{ if } t(\ell) = i \\ 0 \text{ otherwise .} \end{cases} \tag{2.2.5}$$

We saw that in (2.2.2) there is a redundant equation; this translates into a reduction by 1, from full rank, of the row rank of matrix E. If we assume, as is the case for nonpathological networks, that L > N, then, using the redundant equation argument, one can prove [22], [23], that

$$\text{rank } E = N-1 \ .$$

In order to eliminate redundancies we must eliminate one equation from (2.2.2) and one row from E. It turns out that, in order to simplify further calculations, it is convenient to eliminate the equation, and row, corresponding to i=n. If $E(\overline{rn})$ represents matrix E without row $n$[(*)], then (2.2.4) reduces to

$$E(\overline{rn})\ f(n) = r(n), \quad n \in N \ . \tag{2.2.6}$$

It is well known that the set of flows satisfying (2.2.6) and (2.2.1), i.e. the set of feasible multicommodity flows according to the definition above, is a convex polyhedron [24].

## 2.3 A Linear Objective Function: Min-Max Routing

The attractiveness that a linear objective function brings to a problem that already has linear constraints is obvious. In particular the linear dependence of the optimal solution to changes, without certain bounds, in the requirements makes the problem specially appealing from

(*) Throughout this report, if D is some matrix, then D(rn) will denote the n-th row of matrix D and $D(\overline{rn})$ will denote matrix D without row n. Similarly, we shall write D(cn) and $D(\overline{cn})$ with respect to columns. D(m,n) will denote the element in row m column n.

the dynamic point of view. On the other hand the availability of potent computational tools (i.e., the simplex method, [24]) has made linear problems highly popular. It is well known that large-scale programming problems can be solved if they possess special structural characteristics that can be exploited by a solution method. It is at this level, i.e., the development of techniques that exploit these special structures, where we want the linearity of the problem to appear in all its power. Our purpose is to use the known results of linear programming to determine some of the structural properties enclosed within a network problem. Since the actual form of the objective function in such problems is not so important, so long as it can be considered a "reasonable" function, see [15], [49], we anticipate that many of the properties evidenced by the linear problem will give some light to the problem of optimal routing and topologies as a whole.

We define the utilization factor or saturation level of a link as the quotient $f_\ell/c_\ell$. For a given routing strategy, there will be a set of links for which this saturation level is maximum. This set of links will form the "bottleneck" of the network for that routing strategy. The idea is to alleviate, with maximum resources, the traffic load on this bottleneck, by choosing an appropriate routing strategy. In other words, the objective is to minimize the greatest saturation level among the links over all possible routing policies. Therefore the problem is

$$\min_{f \text{ feasible}} \left( \max_{\ell \in L} \frac{f_\ell}{c_\ell} \right) \tag{2.3.1}$$

The standard way to treat these problems [25] is to consider, say $\alpha^1$, as the maximum saturation level and minimize the value of $\alpha^1$ over the set of feasible flows; that is

$$\begin{aligned} &\min \quad \alpha^1 \\ &\text{s.t.} \quad f_\ell \leq \alpha^1 c_\ell \qquad \text{all } \ell \in L \\ &\qquad f \text{ feasible} \end{aligned} \tag{2.3.2}$$

Stated in this way, the problem is equivalent to considering a network with the same topology as before but with new link capacities equal to $\alpha^1 c_\ell$ instead of $c_\ell$, where $\alpha^1 > 0$ is a scaling factor, and asking for the minimum value of $\alpha^1$ for which there exists a feasible flow that is consistent with the new capacities.

It is easy to see that as $\alpha^1$ reduces to some value, say $\alpha_0^1$, a subset of the links is bound to saturate, so that the requirements matrix cannot be satisfied for $\alpha^1 < \alpha_0^1$. For these saturated links $\ell$, (2.3.2) holds with equality: this implies that there must exist a subset of commodities that must flow through the bottleneck, or saturation set, in the sense that all the flow of these commodities must pass through the saturation set.

We will call

$S^1$ = the subset of links that saturate at $\alpha^1 = \alpha_0^1$, i.e., the first saturation set of links

$R^1$ = subset of commodities that must flow through $S^1$

These subsets will be defined more precisely in the next chapter.

Therefore by solving (2.3.2), we will have a triple $(\alpha_0^1, S^1, R^1)$ with the property that no feasible flow consistent with the capacity

constraint exists if either

i) any link $\ell \in S^1$ has its capacity reduced beneath $\alpha_0^1 c_\ell$, or if

ii) any requirement $r_m(n)$, s.t. $(m,n) \in R^1$, is increased while all other requirements are held fixed, and the link capacities remain at the value $\alpha_0^1 c_\ell$

There are many problems that can be formulated in a way mathematically similar to the one just described. For instance the problem of emptying a network of an accumulated backlog of traffic as fast as possible is of this type. Furthermore, it can be shown that the minimax formulation can be extended to random inputs. For all these related problems see [10].

## 2.4 Higher Order Criteria for the Linear Objective Function: The Successive Saturation Problem

It can be expected that many feasible flows will solve (2.3.2), or, in other words, that there will exist many possibilities to accommodate, or route, the commodities not in $R^1$. Because of this circumstance, we may ask for the "best" among the many solutions. One way to approach this question will be to apply the same minimax criterion to the part of the network not bound by the solution of the first saturation problem. We may therefore seek next to minimize a parameter $\alpha^2$ which characterizes the network obtained by clamping the capacities in $S^1$ and scaling the rest of the network. The new problem can then be stated as

$$\begin{aligned} &\min\ \alpha^2 \\ &\text{s.t.} \quad f_\ell \le \begin{cases} \alpha_0^1 c_\ell & \ell \in S^1 \\ \alpha^2 c_\ell & \text{otherwise} \end{cases} \qquad \text{all } \ell \in L \\ &\qquad f \text{ feasible} \end{aligned} \tag{2.4.1}$$

If $\alpha_0^2$ is the minimum value of $\alpha^2$ for which a feasible flow exists, there must exist a set of links, $S^2$, that must saturate and an additional set, $R^2$, of commodities that must flow through $S^2$. It is clear that some commodities in $R^1$ may also utilize links of $S^2$, since they must go from their source to $S^1$ and from $S^1$ to their destinations in some way. But it is also clear that no commodity in $R^2$ will use links of $S^1$, by definition of the various sets. We have, therefore, divided the links and commodities into disjoint subsets and established a hierarchical order among them.

It may happen that the solution to (2.4.1) is not unique. The procedure could then be repeated by clamping the links in $S^2$ at capacity $\alpha_0^2 c_\ell$, and scaling (by $\alpha^3$) those links of the network that are neither in $S^1$ nor in $S^2$. Continuing to iterate in this way until we find a unique solution to the corresponding problem or we have exhausted all links, we ultimately generate a finite number of triples $(\alpha_0^i, S^i, R^i)$ with the following properties

i) $\alpha_0^1 > \alpha_0^2 > \ldots > \alpha_0^M \geq 0$

ii) $S^i \cap S^j = \phi$
$R^i \cap R^j = \phi$ $\quad i \neq j$

iii) $\bigcup_{i=1}^{M} S^i = L$ †

$\bigcup_{i=1}^{M} R^i = R$

† Remark: In order for this property to be satisfied we make the agreement that if after finding a unique solution there still remain links not included in any $S^i$, then each of these links is associated with a new saturation set and these new sets are ordered by decreasing saturation levels. The corresponding commodities are similarly treated.

iv) $c_\ell^* = \alpha_0^i c_\ell$ if $\ell \in S^i$, $i=1,\ldots,M$

The $\{c_\ell^*\}$ are the link capacities of the "smallest" network, obtainable from $[c_\ell]$ by scaling, that still satisfies r.

v) If the arrows in Figure 2.4.1 represent "the right to flow through," then this figure indicates the hierarchies established by the solution procedure

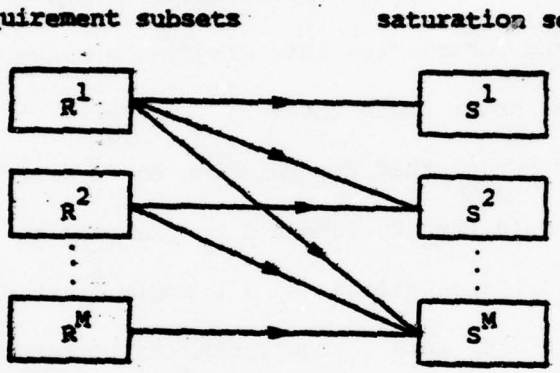



Figure 2.4.1

vi) If $f_0^i$ is the set of optimal solutions for the $i^{th}$ saturation level problem then we have

$$f_0^1 \supset f_0^2 \supset \ldots \supset f_0^M$$

All these properties, in particular v) and vi) give rise to a layer hierarchy, according to the terminology of [26]. This hierarchy of decision layers describes cases where there is a family of decision problems whose solution can be carried out optimally in a sequential manner, and yet which are coupled in the sense that the solution of any problem in the sequence fixes some parameters in the subsequent problems.

In this particular case the original or global problem can be states in several equivalent ways. We will consider two of them below:

- Global problem: form 1(*)

$$\min \sum_{\ell \in L} \alpha_\ell^K$$

$$f_\ell \leq \alpha_\ell c_\ell$$

f feasible

In this case $\alpha_\ell$ is the saturation level of link $\ell$ and K is a large constant. Since the objective function gives a bigger weight to larger saturation levels, then the optimum solution of the problem will minimize the larger levels as much as possible. The constant K has to be choosen large enough to guarantee that no pair of levels can exchange resources among themselves, in the sense that the objective function can never be reduced by decreasing some saturation level at the expense of the increment of some lower level saturation.

- Global problem: form 2

We can think of many possible ways to formulated an equivalent global problem with a linear objective function instead of a non-linear one. This counterpart is that some of the constraints will not be linear. One such global formulation which must also be solved in a sequential manner, is given below:

$$\min \sum_{i=1}^{M} (M-i)^K \alpha^i$$

$$f_\ell \leq [\sum_{i=1}^{M} x_\ell^i \alpha^i] c_\ell \quad \text{all } \ell \in L$$

---

*This form was suggested by R.G. Gallager.

$$x_{\ell}^{i} = 1,0$$

$$\sum_{i} x_{\ell}^{i} = 1$$

$$\alpha^{i} > \alpha^{j} \geq 0 \qquad \text{all } j > i$$

f feasible

The idea behind this form is to make explicit in the statement of the problem that there will be some links that will have a common saturation level. The x variables assign saturation levels to links. The constant M can be chosen equal to the number of links, L, since once all links have a given saturation level $\alpha$ assigned to them the remaining $\alpha^{i}$ can be choosen equal to zero. The linearity in the objective function is obtained by penalizing larger values of $\alpha$ more heavily by using a multiplying coefficient $(M-i)^{K}$ that decreases when i increases (that is, when the saturation level decreases). As in Form 1, K is a constant that has to be chosen large enough to guarantee that no pair of distinct levels can exchange resources among each other.

The purpose of writing the global problem in this form is to make the layer hierarchical structure of the problem more apparent. The manner in which the subproblems fit into the global problem is also clearer.

With this structure, we have a subproblem for each i, and each of these subproblems determines the parameters $x^{i}$ and $\alpha^{i}$ for lower subproblems. This subproblem is equivalent to the min-max problem, for the $i^{th}$ level, that we had before except for the fact that the objective function has been multiplied by a constant.

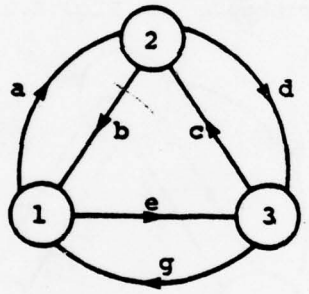
Now we present an example that illustrates the nature of the min-max solution for a three node network.

Example 2.4.1

Consider the network of Fig. E2.4.1a with the following requirement vector

$$r = [r_2(1),\ r_3(1),\ r_1(2),\ r_3(2),\ r_1(3),\ r_2(3)] =$$

$$= [4,\ 8,\ 5,\ 1,\ 1,\ 3]$$




$c_\ell = 10$, all $\ell$

Fig. E.2.4.1a

Suppose we decide to accomodate these requirements via the direct link between each pair of nodes, then, we have the flow distribution of Fig. E.2.4.1b

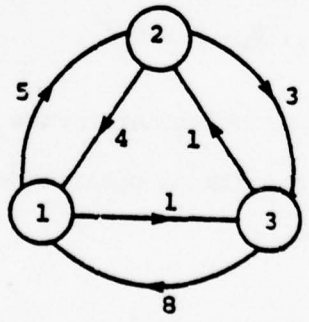



Fig. E.2.4.1b

For this flow distribution the link with the highest saturation level is link g. We can reduce this maximum saturation level if we route part of commodity (3,1) via node 2. Since by doing this operation we also increase the saturation level of both links b and c the maximum amount of commodity (3,1) that could be placed on the new path while minimizing the maximum saturation level of all the links of these two paths will be given by

$$\alpha_g = \max\{\alpha_b, \alpha_c\}$$

and the solution is represented in Fig. E.2.4.1c

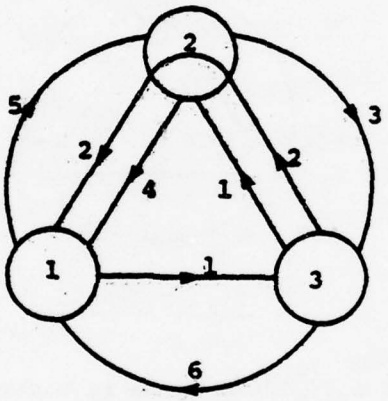



Fig. E.2.4.1c

Since for this solution

$$\alpha_g = \alpha_b = 0.6 > \alpha_a, \alpha_c, \alpha_d, \alpha_e$$

and, furthermore, one cannot further reduce $\alpha_g$ or $\alpha_b$ we can conclude that Fig. E.2.4.1c represents an optimal solution for the first saturation problem with

$$\alpha_0^1 = 0.6, \quad S^1 = \{b,g\}, \quad R^1 = \{(2,1), (3,1)\}$$

Observe also that the solution represented in Fig. E.2.4.1c is not the unique solution for the first saturation problem. In fact there are infinite ways, and thus infinite solutions, to arrange commodity (1,2), for instance, through both link a and via node 3 without increasing the saturation level of links b and g or bringing the saturation level of any other link up to the value $\alpha_0^1 = 0.6$. Because of these facts we can attempt the solution of a second saturation problem. According to (2.4.1) we will try to $\min\{\max(\alpha_a, \alpha_c, \alpha_d, \alpha_e)\}$ without increasing $\alpha_b$ or $\alpha_g$. Is very simple to see that the solution to this problem is that represented in Fig. E.2.4.1d,

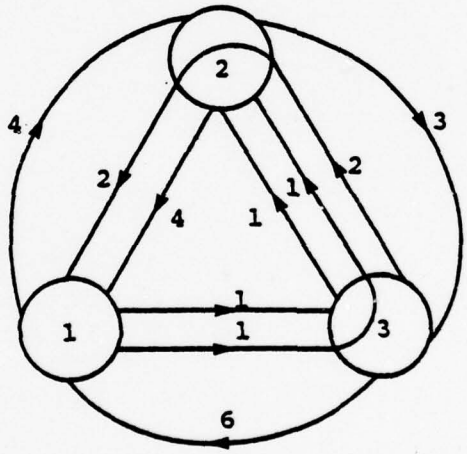



Fig. E.2.4.1d

where now

$$\alpha_0^2 = 0.4, \quad S^2 = \{a,c\}, \quad R^2 = \{(1,2), (3,2)\}$$

Note also that in this case Fig. E.2.4.1d represents the unique solution to the second saturation problem, and hence the unique solution to the successive saturation problem. Thus for this example M=2.

# CHAPTER III

## THE FIRST SATURATION PROBLEM

### 3.1 Linear Programming Formulation. Primal and Dual Problems

In Chapter II we formulated our optimization problem as a sequence of decision problems having the structure of a layer hierarchy, according to [26]. Here in this chapter we will concentrate in the first problem of this hierarchy. In this section, in particular, we will give a linear programming formulation to this first problem.

From (2.3.2) and (2.2.6) we have that the first saturation problem is given by

$$\begin{aligned} &\min \alpha^1 \\ &\text{s.t.} \\ &\sum_n f_\ell(n) - \alpha^1 c_\ell \leq 0 \qquad \text{all } \ell \in L \qquad (3.1.1) \\ &E(\overline{rn}) f(n) = r(n) \qquad \text{all } n \in N \\ &f \geq 0 \\ &\alpha^1 \geq 0 \end{aligned}$$

Note that the constraint $\alpha^1 \geq 0$ is not really needed since, by definition, $\alpha^1 = \max \frac{f_\ell}{c_\ell} > 0$, unless all requirements are null.*

Problem (3.1.1) can be written in the standard linear programming (LP) form as follows

---

*We could add the constraint $\alpha^1 \leq 1$, but we prefer instead to conclude, whenever $\alpha^1 > 1$, that some additional capacity is needed in order for the network to be able to accommodate the traffic flow.

(LP1):

$$\min x_1$$
$$A^1 x = b \qquad (3.1.2)$$
$$x \geq 0$$

where

$$x^T = (\alpha^1, s_1, s_2, \ldots, s_L, f_1(1), \ldots, f_L(1), f_1(2), \ldots, f_L(N)) =$$
$$= (\alpha^1, s^T, f^T(1), f^T(2), \ldots, f^T(N)) = (\alpha^1, s^T, f^T)$$

is an $1 + L + (N-1)\,L$ vector (recall that $f_{ij}(i)$ is not defined), with $s^T$ a vector of slack variables, one for each link, which give the amount of unused capacity to reach saturation for a given $\alpha^1$ and flow distribution.

The matrix

$$A^1 = \begin{array}{|c|c|c|c|c|c|} \hline -c(L) & I & I(\bar{1}) & I(\bar{2}) & \cdots & I(\bar{N}) \\ \hline & & E(\bar{r1}) & 0 & & 0 \\ & & 0 & E(\bar{r2}) & \cdots & 0 \\ 0 & 0 & \vdots & \vdots & & \vdots \\ & & 0 & 0 & & E(\bar{rN}) \\ \hline \end{array} \qquad (3.1.3)$$

($L$ rows; $N \times (N-1)$ columns)

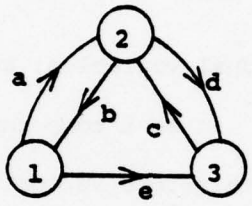
is an $L + N(N-1)$ by $NL+1$ matrix where:

- $c^T(L) = [c_1, c_2, \ldots, c_L]$ is the capacity vector
- $E(\bar{ri})$ is the incidence matrix for destination i as defined in section 2.2. Since as we said before $f_{ij}(i)$ is not defined we will consider that $E(\bar{ri})$ does not include the columns corresponding to the links leaving node i.
- I is an identity matrix of dimension L, and
- $I(\bar{k})$ is an identity matrix with the columns corresponding to links originating at node k deleted.

Note that not all $I(\bar{k})$, and therefore not all $E(\underline{\bar{r}k})$, are, in general, of the same dimension, because the number of columns of $I(\bar{k})$ and $E(\underline{\bar{r}k})$ is equal to $L - od(k)$, where $od(k)$ is the outdegree of node k, see [27]. Thus, unless $od(k) = od(i)$, all $i,k \in N$, not all $I(\bar{k})$ will have the same dimension.

Example 3.1.1

For the three node network below




we have

$$I(\bar{1}) = \begin{array}{c} \begin{matrix} b & c & d \end{matrix} \\ \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{matrix} a \\ b \\ c \\ d \\ e \end{matrix} \end{array} \quad ; \quad E(\underline{\bar{r}1}) = \begin{array}{c} \begin{matrix} b & c & d \end{matrix} \\ \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & -1 \end{pmatrix} \begin{matrix} 2 \\ 3 \end{matrix} \end{array}$$

$$I(\bar{3}) = \begin{array}{c} \begin{matrix} a & b & d & e \end{matrix} \\ \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{matrix} a \\ b \\ c \\ d \end{matrix} \end{array} \quad ; \quad E(\underline{\bar{r}2}) = \begin{array}{c} \begin{matrix} a & b & d & e \end{matrix} \\ \begin{pmatrix} 1 & -1 & 0 & 1 \\ 0 & 0 & -1 & -1 \end{pmatrix} \begin{matrix} 1 \\ 3 \end{matrix} \end{array}$$

Finally, the vector

$$b = \begin{pmatrix} \underline{0} \\ r(1) \\ r(2) \\ \vdots \\ r(N) \end{pmatrix} = \begin{pmatrix} \underline{0} \\ r \end{pmatrix} \qquad (3.1.4)$$

in an L + N(N-1) column vector.

The dual linear programming problem for (LP1) can be written as (DP1):

$$\max \lambda^1 \; b$$
$$\lambda^1 \; A^1 \leq \gamma^1 \tag{3.1.5}$$

where

$$\lambda^1 = (\pi^1, \sigma^1) = (\pi^1, \sigma^1(1), \ldots, \sigma^1(N))$$

is an L + N(N-1) vector of dual variables, and $\gamma^1 = (1,0,\ldots,0)$ is an NL+1 row vector. Vector $\pi^1$ is an L components row vector corresponding to the first L rows of $A^1$; therefore vector $\pi^1$ has one component per link of the network and $\pi^1_\ell$ will be the component associated with link $\ell$. Vector $\sigma^1(n)$ is an N-1 components row vector that correspond to the N-1 conservation equations associated with commodity n; component $\sigma^1_m(n)$ will be the component associated with commodity (m,n).

We will see later some physical and mathematical interpretations of these dual variables. For simplicity of notation, and because the first saturation problem will be the most extensively studied in this thesis, we will write the dual variables corresponding to this problem without the superscript 1 whenever we feel that there is no reason for confusion.

Constraint $\lambda A^1 \leq \gamma$ can be separated into the following set of inequalities:

$$-\sum_{\ell=1}^{L} \pi_\ell c_\ell \leq 1$$

$$\pi_\ell \leq 0, \qquad \text{all } \ell \in L$$

$$\sigma(n)\ E(\overline{\underline{rn}}) \leq -\pi\ , \qquad \text{all } n \in N$$

From these inequalities, the complementary slackness theorem of linear programming (see [28] p. 77) and (2.2.5) we have

Theorem 3.1.1

Let x and $\lambda$ be feasible solutions for the primal and dual problems, respectively. A necessary and sufficient condition that they both be optimal solutions is that for all $\ell \in L$ and all $n \in N$

$$\text{i)} \quad \sum_{\ell=1}^{L} \pi_\ell c_\ell = -1$$

$$\text{ii)} \quad s_\ell > 0 \Rightarrow \pi_\ell = 0 \qquad \pi_\ell < 0 \Rightarrow s_\ell = 0 \qquad (3.1.6)$$

$$\text{iii)} \quad f_\ell(n) > 0 \Rightarrow -\pi_\ell = \sigma_{o(\ell)} - \sigma_{t(\ell)}(n)$$

□

Now we will use the strong duality theorem of linear programming, see [28] page 72, to prove the following result that gives the first interpretation of the dual variables.

Theorem 3.1.2

If $\alpha_0^1$ is the optimal value of the first saturation level then

$$\alpha_0^1 = \sum_{i=1}^{N} \sum_{\substack{n=1 \\ n \neq i}}^{N} \sigma_i(n) r_i(n) \qquad (3.1.7)$$

$$\alpha_0^1 = -\sum_{\ell=1}^{L} \pi_\ell f_\ell = -\sum_{\ell \in S_0^1} \pi_\ell f_\ell$$

where $\{\pi_\ell\}$ and $\{\sigma_i(n)\}$ are values of an optimal solution of the dual problem, $f_\ell$ is an optimal aggregate flow, and

$$S_0^1 \triangleq \{\ell | \ell \in L,\ s_\ell = 0\} \tag{3.1.8}$$

Proof:

From Chapter II we know that, based on physical grounds, (LP1) has a finite optimal solution. Using now the duality theorem we conclude that correspondingly (DP1) has also a finite optimal solution and that, furthermore, the corresponding values of the objective functions are equal. We can write in this way

$$\alpha_0^1 = \lambda b = \sigma r = \sum_{n=1}^{N} \sigma(n) r(n) = \sum_{i=1}^{N} \sum_{\substack{n=1 \\ n \neq i}}^{N} \sigma_i(n) r_i(n)$$

that proves the first expression.

To prove the second set of equalities of (3.1.7) note that

$$s_\ell = 0 \Rightarrow \alpha_0^1 c_\ell = f_\ell ,$$

multiplying both sides of this equation by $\pi_\ell$ and summing over all links for which $s_\ell = 0$ we have

$$\alpha_0^1 \sum_{\ell \in S_0^1} c_\ell \pi_\ell = \sum_{\ell \in S_0^1} \pi_\ell f_\ell$$

The use of i) and ii) of Theorem 3.1.1 on previous equality gives the second relation in (3.1.7) .

Q.E.D.

An immediate consequence of this theorem is the following lemma.

Lemma 3.1.1*

If $\Delta r_i(n)$ is a small increment in commodity (i,n) (assuming all other requirements are constant), which does not cause the optimal basis to change, then the corresponding increment in the cost function is given by

$$\Delta\alpha_0^1 = \sigma_i(n)\ \Delta r_i(n) \tag{3.1.9}$$

□

For any problem that is able to accept small requirement increments without changing optimal basis we can easily interpret, from this Lemma, the dual variable $\sigma_i(n)$ as the sensitivity of saturation level $\alpha_0^1$ with respect to small changes in requirement $r_i(n)$.

## 3.2 Optimal Solutions. Uniqueness and Degeneracy

An optimal solution to (LP1) can be written, for a given requirement vector, as

$$x_B = b^0 = B^{-1}b$$
$$x_N = 0 \tag{3.2.1}$$

where $x_B$ are the basic variables, $x_N$ are the non basics and B is the basic matrix. Since the first L components of b are equal to zero, we can partition $B^{-1}$ into two blocks in such a way that

*Theorems and Lemmas will have independent numbering. Corollaries will have the same number as the corresponding Theorem or Lemma, except that the number will be preceeded by an L whenever the corollary corresponds to a Lemma.

$$x_B = B^{-1}b = [D_1\ D_2]\begin{bmatrix}0\\ r\end{bmatrix} = D_2 r \qquad (3.2.2)$$

If the requirement vector changes from r to $r + \Delta r$, an elementary sensitivity analysis, [24], [29], shows that the basic solution $(x_B, x_N) = (D_2(r+\Delta r), 0)$ remains optimal as long as $D_2(r+\Delta r) \geq 0$, since dual feasibility and complementary slackness are unaffected.

Considering r as a vector variable, we find that the range of r within which the optimal basis remains unchanged is given by the polyhedral cone, on the r-space, defined by

$$D_2 r \geq 0 . \qquad (3.2.3)$$

As long as r moves within the boundaries of this polyhedron, the values of the basic variables can be obtained by (3.2.2). Since this expression is a linear function of r, we will call the region described by (3.2.3) the <u>linear region</u> for basis B.

We said before that an optimal solution to (LP1) is, in general, highly non-unique since it leaves many degrees of freedom with respect to the routing of commodities not in $R^1$. On the other hand, we have not said anything yet regarding uniqueness or not of the dual problem (DP1). This is an important question that we will address in this section and for which the idea of degenerate solution is important.

A <u>degenerate solution</u> in linear programming is a basic solution that has a basic variable equal to zero. Since we have two types of variables that can be basics, flow variables and slack variables*, we

---

*Although $\alpha^1$ is a basic variable it will not be included in this reasoning because $\alpha_0^1$ cannot be equal to zero unless r = 0.

we will have to types of degeneracy:

link degeneracy, when a basic slack variable is equal to zero, $s_\ell = 0$;

flow degeneracy, when a basic flow variable is equal to zero, $f_\ell(n) = 0$.

If a problem has more than one solution, as will happen in general with (LP1), the fact that we have one degenerate solution does not necessarily imply that all solutions will be degenerate. If all optimal solutions are degenerate we will say that for the present value of r (LP1) is degenerate; otherwise, (LP1) will be non-degenerated. It is well known from linear programming (see [29], Chapter 6) that a non-degenerated optimal solution can always absorb a small increment (positive or negative) in its right-hand-side (RHS) without changing basis. For a degenerate optimal solution, on the other hand, this claim is not true because we can always find arbitrarily small changes in the RHS that force a change of basis.

Definition 3.2.1

For a given requirement vector r, problem (LP1) will be said to be stable, or to be at a stable point, if for any $\Delta r$, such that $r + \Delta r \geq 0$, we can find a positive real number $\varepsilon_1$ such that for any $\varepsilon < \varepsilon_1$ there exists an optimal solution to a modified problem, with $r + \varepsilon r$ as RHS, having the same optimal basis as an optimal solution of (LP1). Otherwise r will be an unstable point for (LP1)

□

An immediate consequence of this definition and previous discussion is:

Lemma 3.2.1

If (LP1) has a nondegenerate solution, then, (LP1) is at a stable point.

□

Now we want to relate stability and degeneracy for (LP1). All degeneracy problems arise from the nature of the RHS vector. Suppose that, for (LP1), the requirement vector r is such that some of its components are equal to zero. In this case it might happen that every optimal solution would contain some basic flows, the same for all solutions, which are equal to zero. In the standard linear programming terminology the variables that are zero in every solution, no matter whether they are basics or not, are called null variables, (see [28] page 91) and they can be treated by deleting the corresponding columns of matrix A and adjoining the equalities $x_i = 0$ for these variables. When we are interested in the dynamic evolution of the system to changes in the RHS, as will be the case here, the standard treatment given to null variables is not quite adequate, because the information carried on by deleted columns will be needed at subsequent stages. We will, therefore, keep these columns in our problem formulation. This fact forced us to give special treatment to these basic null variables. The following theorem is a known result to detect null variables.

Theorem 3.2.1 ([28], page 92)

If (LP1) has at least one solution, the variable $x_i$ is a null variable for this problem if and only if there is a nonzero vector $\lambda$

such that

$$\lambda A^1 \geq 0$$

$$\lambda b = 0$$

and the i-th component of $\lambda A^1$ is strictly positive.

□

Suppose that for a given vector r we have detected all null variables and let $x_{B0}$ be the subset of such variables that are basic for a given optimal solution. If $x = (x_b, 0)$ is such an optimal solution and B is the optimal basis we can write, from (3.2.2) and after appropriate reordering,

$$x_B = \begin{bmatrix} x_{B0} \\ \hline x_{B1} \end{bmatrix} = D_2 r \tag{3.2.4}$$

Assume now that r has some zero components and that by reordering equations we can write

$$x_B = \begin{bmatrix} x_{B0} \\ \hline x_{B1} \end{bmatrix} = [D_{20}\ D_{21}] \begin{bmatrix} r_0 \\ \hline r_1 \end{bmatrix} = D_{20} r_0 + D_{21} r_1$$

where $r_0 = 0$ and $x_{B0} = 0$.

Partitioning $D_{20}$ and $D_{21}$ we can write

$$\begin{bmatrix} x_{B0} \\ \hline x_{B1} \end{bmatrix} = \begin{bmatrix} D^0_{20} \\ \hline D^1_{20} \end{bmatrix} r_0 + \begin{bmatrix} D^0_{21} \\ \hline D^1_{21} \end{bmatrix} r_1 \tag{3.2.5}$$

According to the definition of stability the only part of this last expression that can cause trouble is $x_{B0}$, where

$$x_{B0} = D^0_{20} r_0 + D^0_{21} r_1$$

Now consider the case in which $x_{BO} = 0$ iff $r_0 = 0$. It follows that we must have $D^0_{21} = 0$ and

$$x_{BO} = D^0_{20} r_0 \tag{3.2.6}$$

If we make a change in $r_0$ then

$$\Delta x_{BO} = D^0_{20} \Delta r_0 \quad , \tag{3.2.7}$$

but since $r_0 = 0$, $\Delta r_0$ must be positive because we cannot have negative requirements. Therefore, the new value of the variables in $x_{BO}$ (namely, $x_{BO} + \Delta x_{BO} = \Delta x_{BO}$) will be positive for any positive $\Delta r_0$ iff $D^0_{20}$ is a matrix of positive elements. Since, by hypothesis, $D^0_{21} = 0$ and $D^0_{20}$ is a positive matrix, we have that $r^T = [r_0^T \;\; r_1^T] = [0 \; r_1^T]$ is a stable point for (LP1) in spite of the problem being degenerate. Note that this is possible only because $\Delta r_0$ is not allowed to be negative.

From these results we can write

Lemma 3.2.2

If (LP1) has a degenerate solution where all zero basic variables are null variables due to some requirement being equal to zero, then (LP1) is stable.

□

Because of this result, and since many requirements can be zero throughout a period of operation of the network, it is obvious that we should be more interested in stability than in degeneracy. We begin by summarizing some important results of duality in linear programming.

Theorem 3.2.2 ([24] page 144, [29] page 268)

i) A non-degenerate optimal solution for the primal problem => A unique optimal solution for the dual.

ii) A unique degenerate optimal primal => A non-unique optimal dual

iii) A non-unique optimal primal => A degenerate optimal dual

iv) A non-unique optimal dual => A degenerate optimal primal

□

Now we will prove an important theorem concerning the stability of (LP1):

Theorem 3.2.3

A necessary and sufficient condition for (LP1) to be at a stable point is that (DP1) has a unique solution.

Proof:

Consider, first, that (DP1) has a non-unique solution and that $\sigma_i(n)$ and $\tilde{\sigma}_i(n)$ correspond to two different solutions. Then, from (3.1.9), we can write

$$\Delta\alpha_0^1 = \sigma_i(n)\ \Delta r_i(n)$$

$$\tilde{\Delta}\alpha_0^1 = \tilde{\sigma}_i(n)\ \Delta r_i(n)$$

where $\Delta r_i(n)$ is an arbitrarily small increment in requirement $r_i(n)$, assuming the remaining requirements are held constants. Thus, since $\Delta\alpha_0^1 \neq \tilde{\Delta}\alpha_0^1$ we have that $\alpha_0^1 + \Delta\alpha_0^1$ and $\alpha_0^1 + \tilde{\Delta}\alpha_0^1$ must correspond to different basic solutions and, therefore, the point is unstable.

Now we will show that an unstable point for (LP1) implies a non-unique solution for (DP1). Consider that r is an unstable point for (LP1) then we can find, at least, two directions of change $\Delta r$ and $\tilde{\Delta r}$ for which, for any arbritrarily small positive number $\varepsilon$, $r + \varepsilon\Delta r$ and $r + \varepsilon\tilde{\Delta r}$ will require two different optimal basis in the solution of their corresponding (LP1) problems. If B and $\tilde{B}$ are these two optimal basis we know, from duality theory, that to each one of them will correspond a different dual solution

$$\lambda = \gamma^{1}{}_{B}{}^{-1}$$

$$\tilde{\lambda} = \gamma^{1}{}_{\tilde{B}}{}^{-1}$$

Since r can be interpreted as the limit as $\varepsilon \to 0$ of $r + \varepsilon\Delta r$ and $r + \varepsilon\tilde{\Delta r}$, from the linearity of the problem we conclude that $\lambda$ and $\tilde{\lambda}$ are both optimal solutions for (DP1) at point r. Hence (DP1) is non-unique.

Thus we have shown that (LP1) is unstable iff (DP1) is non-unique, which is logically equivalent to the statement of the theorem.

Q.E.D.

Suppose now that we have a stable optimal solution with a non-basic slack variable $s_\ell$ having a zero dual variable. Then from duality, [24] Chap. 6, we can find another optimal solution where $s_\ell$ is basic and, from stability, nonnull. Thus, link $\ell$, in spite of the fact that $s_\ell = 0$ for the first basic solution, was improperly saturated. This fact lead us to the following formal definition of the first saturation set $S^1$.

Definition 3.2.2

For a stable optimal solution link $\ell$ belongs to the first saturation set $S^1$ iff $\pi_\ell < 0$

□

From (3.1.8) and this definition we have that, for a given solution, $S_0^1 - S^1$ will be the set of unproperly saturated links, i.e., links with $s_\ell = \pi_\ell = 0$. Also, from this definition an obvious corollary of Theorem 3.2.3 is:

Corollary 3.2.3

At a stable point the set $S^1$ and the optimal value of variables $\{\pi_\ell\}$ and $\{\sigma_i(n)\}$ are unique.

□

An almost immediate consequence of Definition 3.2.1 and Theorem 3.2.3 is

Lemma 3.2.3

At a stable point sets $S^1$ and $R^1$ cannot be divided in subsets $\{S_i^1\}$, $\{R_j^1\}$ in such a way that commodities in $R_j^1$ only flow through links of $S_i^1$, and links of $S_i^1$ are only used by commodities of $R_j^1$, for some i and j.

Proof: Obvious.

With respect to unstable points note that, from Theorems 3.2.2-3 and the fact that any degenerate point can be interpreted as the limit of a cluster of non-degenerate points (see [29] Chap. 6), an unstable point can be treated as the limit of a stable point. Thus, these unstable points will appear at the boundary line separating regions of stable points.

Later we shall be interested in the dynamic evolution of the optimal solution as the requirement vector changes. The fact that the limit of any solution as it approaches an unstable point remains a solution at the boundary itself means that we need not worry about the direction from which the boundary is approached. Therefore, for simplicity, most of our future results will be stated for stable points without being concerned about loss of generality.

## 3.3 Solving the First Saturation Problem

The first saturation problem can be solved by any of the standard algorithms used to solve linear programs, e.g., the simplex method [24]. The special structure of (LP1), on the other hand, suggests that special programs or algorithms might take advantage of this structure to reduce the computation for an optimal solution. Fortunately, there already exist techniques that have produced very efficient algorithms to solve problems of this type, see [34] and [36] for a survey. These techniques have been developed along two main lines: decomposition techniques and generalized upper bounding approaches.

Two main types of decomposition are currently in use: one based on the Dantzig-Wolfe decomposition techniques [24], and usually called price-directive algorithms; and the other, based on subgradient optimization [37], and called resource directive algorithms. These techniques have been studied by Assad in [35] for general multicommodity problems and by Defenderfer in [11] for routing algorithms in computer networks. Defenderfer, in particular, has applied these techniques to this same first saturation problem and has an excellent

discussion and comparison of the various decompositions used. He gives the advantages and disadvantages of decomposition by trees, extremal flows or chains depending on the size and solution stage of the problem.

Generalized upper bounding approaches are based on techniques devised by Dantzig and Van Slyke [30]. There are two main formulations, namely the node-arc formulation [31], [32] and the arc-chain formulation [33]. These techniques appear very promising since they exploit the special structure of the linear program solution and make use of whatever simplification may arise due to the network structure of the matrices involved. Because of these facts they make extensive use of some graph-theoretic results related to the graph supporting the network. Our work in the following chapters will be much within this same spirit, i.e., make as much use as possible of the special structure and topology of the problem to develop a set of properties that not only simplify the actual computation of the solution, but also contributes to the understanding of the physical behaviour of the optimal solution of these type of problems. In fact this second purpose will be our major concern in this thesis.

Also deserving of mention is a nonlinear approach that converges to an optimal solution of (LP1) and was suggested by Defenderfer [11]. It has the attractive features of being quite efficient and allowing the program to stop when it approximates the optimal solution to a specified tolerance.

## 3.4 Distance Interpretation of Dual Variables. Optimal Solution as a Shortest Route Problem

In this section we will use many of the results already obtained

in this chapter to give a distance interpretation of the dual variables. From this interpretation we will show that a shortest route criterion could be used to determine all allowable paths corresponding to an optimal solution. The main tool of this section will be Theorem 3.1.1. Following the reasoning of Section 3.2, we will be dealing only with stable points. From Theorem 3.1.1 and Definition 3.2.2 we have:

Lemma 3.4.1

At a stable point the following relations are verified:

$$\sum_{\ell \in S^1} \pi_\ell c_\ell = -1 \tag{3.4.1}$$

$$f_\ell(n) > 0, \quad \ell \in \bar{S}^1 \Rightarrow \sigma_{o(\ell)}(n) = \sigma_{t(\ell)}(n) \tag{3.4.2}$$

where

$$\bar{S}^1 \triangleq L - S^1 \tag{3.4.3}$$

□

Consider that for commodity (i,n), with $r_i(n) > 0$, the directed chain $\ell_1, \ell_2, \ldots \ell_k$, originated at node i and terminated at node n, has the property that $f_{\ell_j}(n) > 0$ for all j = 1,2, ... k, then, we will say that this directed chain is an allowable path for commodity (i,n) and will be represented by P(i,n). If commodity (i,n) has more than one allowable path we will differentiate them by means of subindexes, e.g., $P_1(i,n)$, $P_2(i,n)$ etc. From iii) of Theorem 3.1.1 we can write, for any of these allowable paths, that

$$\sigma_i(n) = -\pi_{\ell_1} + \sigma_{t(\ell_1)}(n) = -\pi_{\ell_1} - \pi_{\ell_2} + \sigma_{t(\ell_2)}(n) = \ldots = -\sum_{\ell \in P(i,n)} \pi_\ell . \tag{3.4.4}$$

Since $\sigma_i(n)$ is independent of the path we have

Lemma 3.4.2

The sum of the link dual variables, corresponding to an optimal solution of (LP1), along every allowable path of any commodity (i,n) is equal to the negative of the commodity dual variable $\sigma_i(n)$.

□

From (3.4.4) and Lemma 3.4.1 we can easily prove, using P(i,n) to represent any allowable path from i to n,

Lemma 3.4.3

$$\sigma_i(n) = - \sum_{\ell \in S^1 \cap P(i,n)} \pi_\ell \geq 0, \quad \text{all } (i,n) \in R$$

$$\sigma_i(n) = 0 \iff S^1 \cap P(i,n) = \phi$$

where $\phi$ is the empty set.

□

From this lemma we can write the following formal definition for $R^1$.

Definition 3.4.1

Commodity (i,n) will belong to subset $R^1$ iff $\sigma_i(n) \neq 0$.

□

If we interpret $-\pi_\ell$ as the length of link $\ell$ and define the distance of any path between two nodes as the sum of the lengths of its links, then we can summarize all previous results in the following theorem.

Theorem 3.4.1

Every allowable path for a given commodity (i,n), as given by any optimal solution to (LP1) (or a convex combination of solutions), is a minimum distance path between the two nodes. This minimum distance is equal to the optimal value of the dual variable $\sigma_i(n)$. The total distance for commodities not in $R^1$ is zero, and therefore all paths for these commodities are through links of zero distance.

□

A more general version of this result that applies to any loop-free multicommodity flow problem was independently, and almost simultaneously, obtained by Gallager, see [17]. Since the solution of the first saturation problem leaves many degrees of freedom for commodities not in $R^1$, we have no guarantee that an optimal solution as obtained from (LP1) is completely loopfree. In fact, this is one of the main justifications for lower saturation level analysis. Nevertheless, we can claim at this point that the solution has no loops containing links of $S^1$. This result will be generalized, in Chapter V, to include complete loop-freedom for all commodities when lower saturation problems are considered.

For the remaining part of this section we will show how to obtain an integer assignment of distances from the one just described in Theorem 3.4.1. Theorem 3.1.1 gives a set of conditions that vectors $\pi$ and $\sigma$ have to satisfy in order to be optimal solutions of the dual problem (DP1). From (3.4.4), it is easy to see that if $\pi$ and $\sigma$ satisfy conditions ii) and iii), and if k is a positive real number, then $k\pi$, $k\sigma$ also satisfy ii) and iii). Thus we have

Lemma 3.4.4

Conditions ii) and iii) of Theorem 3.1.1 define a ray* on the $\lambda$-space. The optimal solution of (DP1) is the point $(\pi,\sigma)$ of that ray that satisfies

$$\sum_{\ell \in S^1} \pi_\ell c_\ell = -1 \ .$$

□

From this Lemma we see that the only role of the capacity vector c in fixing the optimal point for (DP1) is through equation (3.4.1). From (3.4.4) and iii) of Theorem 3.1.1 we can see that the remaining equations, beside (3.4.1), that the $\pi$ variables have to satisfy is a set of homogeneous equations with 0, $\pm$ 1 coefficients. Thus, if C is the matrix with these 0, $\pm$ 1 coeffcients, we can write

$$\pi(S^1) \left[ c\left(S^1\right) \middle| C \right] \triangleq \pi(S^1)\, Q = \begin{bmatrix} -1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \qquad (3.4.6)$$

where Q must be a square full rank matrix. (We will give a physical interpretation for this matrix in Chapter IV, and we will see that it plays a very important role in the optimization process.) From (3.4.6) we can write

$$\pi(S^1) = \begin{bmatrix} -1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} Q^{-1} = Q^{-1}(\underline{r}1) \ .$$

Thus, from Cramer rule.

$$\pi_\ell = \frac{\det C(\overline{\underline{r}\ell})}{\det Q} = \frac{\det C(\overline{\underline{r}\ell})}{\sum_{\ell \in S^1} c_\ell \det C(\overline{\underline{r}\ell})}, \quad \text{all } \ell \in S^1, \qquad (3.4.7)$$

*We will say that $\rho$ is a ray in $\mathbb{R}^n$ if $\rho = \{x \mid x = a\lambda,\ x,\lambda \in \mathbb{R}^n,\ a \in \mathbb{R},\ a \geq 0\}$

but from the structure of matrix C, det $C(\overline{\underline{r\ell}})$ will be an integer number. From these facts we claim that $\pi_\ell$ and $\sigma_i(n)$ will have the form

$$\pi_\ell = -\frac{\theta_\ell}{D^1}$$

$$\sigma_i(n) = \frac{d_i(n)}{D^1} \qquad (3.4.8)$$

$$D^1 = \sum_{\ell \in S^1} \theta_\ell c_\ell$$

where $\theta_\ell$ and $d_i(n)$ are non negative integers. It is very easy to see, by sustitution, that (3.4.8) in fact do verify conditions i) - iii) of Theorem 3.1.1. If $\theta$ and d are the vectors with components $\theta_\ell$ and $d_i(n)$ respectively, it is also very easy to see, for any real coefficient $k > 0$, that $k(-\theta,d)$, also satisfies the same conditions. Thus $(-\theta,d)$ is a direction with integer components that describes the same ray as $(\pi,\sigma)$. If $D^1$ of (3.4.8) is the minimum common multiple of the denominators of the $\pi_\ell$'s then, $(-\theta,d)$ will be the smallest integer point of the ray*. In the future whenever we refer to vector $(-\theta,d)$ we will assume that it corresponds to the smallest integer point of the ray. Thus $S^1$ and the set of nonzero flows define, for a given network, a ray in the $\lambda$-space that is independent of the capacity and uniquely determined by the integer vector $(-\theta,d)$. We can summarize these results in the following

---

*Point x is the smallest integer point of the ray if all components of x are integer and kx is not integer for any $0 < k < 1$.

Lemma 3.4.5

$S^1$ and the set of allowable paths define, for a given network, a ray that can be characterized by an integer vector $(-\theta,d)$, defined by (3.4.8). The smallest such integer vector is obtained when $D^1$ = minimum common multiple of the denominators of the $\pi$'s. Given the information above vector $(-\theta,d)$ is independent of the actual value of the link capacities.

□

The independence of $(\theta,d)$ with respect to the capacities will simplify many arguments and calculations, since once $S^1$ and the set of allowable paths corresponding to an optimal solution are given, a variation on the value of the capacities will only affect the actual value of the dual variables by a scaling factor. Thus many computations can be carried through by normalizing all capacities to one. The actual values are, then, obtained by appropriate scaling whenever they are needed. We will use this fact in section 4.5.

An immediate consequence of Theorem 3.4.1 and Lemmas 3.4.4 and 3.4.5 is

Theorem 3.4.2

To any point of the ray defined by Lemmas 3.4.4 or 3.4.5 there correspond a different valid distance aassignment to describe the allowable paths according to Theorem 3.4.1. In particular the integers numbers $\{\theta_\ell\}$ defined in (3.4.8) constitute a valid assignment. For this assignment the total distance for any allowable path of commodity (i,n) will be $d_i(n)$ given in (3.4.8).

□

From now on $\theta_\ell$ will be called the <u>length or weight of link $\ell$;</u> $d_i(n)$ will be called the distance or <u>weight of commodity (i,n)</u> and $D^1$ will be called, from (3.4.8), the <u>total equivalent capacity</u> of $S^1$.

Given a set $S^1$ one might suspect the existance of a relation between the value of the largest weight in $\{\theta_\ell\}$ and the number of links in $S^1$. One might conjecture, for instance, that the largest weight need not be larger than $|S^1|$. This conjecture is false and we have found a family of examples, see Appendix II, where the maximum weight grows exponentially with the number of saturated links. Furthermore, these examples led us to a way of synthesizing a network with any given set of link weights provided that these weights are possible weights for a first saturation set.

Finally, as an application of the independence of $\theta$ with respect to the capacity values we will obtain the change in the value of $\alpha_0^1$ due to a small change in capacity $c_\ell$, $\ell \in S^1$, assuming that the optimal basis does not change. Substituting $\sigma_i(n)$ from (3.4.8) in (3.1.7) and using (3.4.5), we have

$$\alpha_0^1 = \frac{1}{D^1} \sum_{(i,n) \in R^1} d_i(n) r_i(n) \quad . \tag{3.4.9}$$

If we have an increment $\Delta c_\ell$ in $c_\ell$, that does not change the optimal basis, the new value of the saturation level will be given by

$$\alpha_0^1 + \Delta\alpha_0^1 = \frac{1}{D^1+\Delta D^1} \sum_{(i,n) \in R^1} d_i(n) r_i(n) \tag{3.4.10}$$

where

$$D^1 + \Delta D^1 = D^1 + \theta_\ell \Delta c_\ell \ . \qquad (3.4.11)$$

Substituting (3.4.11) and the sum from (3.4.9) in (3.4.10) and using (3.4.8) we have the following result:

Lemma 3.4.6

The variation in the optimal value of the saturation level due to a small change in the capacity of link $\ell$, that does not induce a change of optimal basis, is given by

$$\Delta\alpha_0^1 = \alpha_0^1 \frac{\pi_\ell \Delta c_\ell}{1-\pi_\ell \Delta c_\ell}$$

or (3.4.12)

$$\frac{\Delta\alpha_0^1}{\alpha_0^1 + \Delta\alpha_0^1} = \pi_\ell \Delta c_\ell \ . \qquad \square$$

## 3.5 Graph Theoretic Properties and Bounds Related to an Optimal Solution

Consider an optimal solution to (LP1) and assume for simplicity that the solution is stable and $r>0$[*]. The basis corresponding to this optimal solution will be composed of the following variable:

i) nonzero slack variables, i.e., $\{s_\ell | \ell \notin S_0^1\}$

ii) nonzero flows, i.e., $F_B \triangleq \{f_\ell(n) | f_\ell(n) \geq 0\}$

iii) objective function $\alpha^1$

Since rank $A^1 = L + N(N-1)$ any basis must have dimension $L + N(N-1)$ and therefore the number of the nonzero flows has to be

$$|F_B| = L + N(N-1) - (L - |S_0^1|) - 1 = N(N-1) + |S_0^1| - 1$$

---

[*] If some $r_i(n) = 0$ we can make $r_i(n) = \varepsilon$ and then make $\varepsilon \to 0$. If the optimal solution gives $f_\ell(n) = g(\varepsilon)^i \to 0$ as $\varepsilon \to 0$ we will consider $f_\ell(n) > 0$ in ii).

where the -1 accounts for the fact that $\alpha^1$ is a basic variable.

On the other hand, we have the following general result for multicommodity flow problems in which r>0.

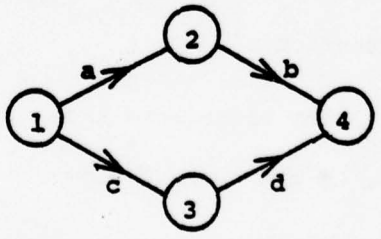
Theorem 3.5.1 ([24], pp. 356-7) ([17])

Any feasible multicommodity flow must have in its basis, for each destination (commodity) of the network, the flow components of, at least, the links of a directed spanning tree with such destination as sink (assuming a capacitiated network.)

□

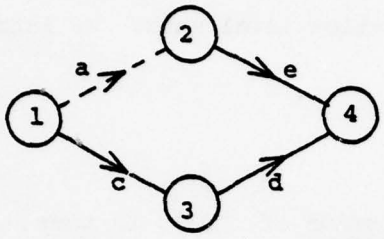
This means that, using the links $\ell$ such that $f_\ell(n) > 0$, we must be able to form at least N spanning trees, one for each commodity. Since any spanning tree must have N-1 link components (recall that we assumed that any pair of nodes can communicate in both directions(*)), we need N(N-1) links, and thus $f_\ell(n) > 0$ variables, to form those trees. Hence, we still have $|S_0^1| - 1$ flow variables, $f_\ell(n) > 0$, that have not been included in the spanning trees. It is not difficult to see that these additional flows will create alternate paths for some commodities.

For instance if, in the network below, $f_\ell(4) > 0$ for all $\ell$, then, we have two allowable paths for commodity (1,4). Of these two allowable




*Although this assumption does not subtract must generality from the problem and many results are unaffected by it, we will relax it later in this section to obtain some results that are, in fact, affected.

paths one will be in the spanning tree rooted at 4 and the other will




constitute an alternate path.

Lemma 3.5.1

The number of alternate paths given by an optimal basic solution of (LP1) is equal to $|S_0^1| - 1$.

□

From Definition 3.2.2 we know that if $|S_0^1| > |S^1|$ then the elements of $S_0^1 - S^1$ are elements with $s_\ell = 0$ and $\pi_\ell = 0$ and constitute a set of unproperly saturated links. If we consider problem (LP1) immersed within the collection of successive saturation problems where, loosely speaking, the global goal is to minimize the maximum saturation level of all links in the network, it seems clear that a solution with a smaller $|S_0^1|$ will be, in principle, preferable to a solution with a larger $|S_0^1|$. On the other hand if we consider problem (LP1) in isolation we may think that a solution with a larger number of alternate paths, and consequently a larger $|S_0^1|$, could be desirable because it will give to the problem a stronger robustness against failures. For this thesis, nevertheless, we will be more concerned about the global problem so we will be looking for $|S_0^1|$ as small as possible, that is $|S_0^1| = |S^1|$. It may happen, though, that none of the possible basic optimal solutions of (LP1) could

reach this lower bound for $|S_0^1|$ (we give one such example in Section 5.6.) In this case, as we will see in Chapter V, those extra paths will become part of lower saturation level sets. We introduce the following definition.

Definition 3.5.1

A minimal optimal solution of (LP1), if there exists one, is a basic optimal solution for which $|S_0^1| = |S^1|$.

□

An immediate consequence of this definition, Lemma 3.5.1 is

Lemma 3.5.2

The number of alternate paths given by a minimal optimal solution is $|S^1| - 1$.

□

For the remaining of this section all results will be related to minimal solution, although their extension to general basic solutions is obvious from Lemmas 3.5.1 and 3.5.2. To relate the number of alternate paths to the number of nodes and links in the network we will use the following result of graph theory:

Theorem 3.5.2 ([27], page 39)

The maximum number of independent circuits in a connected graph with L links and N nodes is L - N + 1. The maximum number of independent cut-sets is N-1.

□

Therefore if M is the circuit matrix of the network the first part of Theorem 3.5.2 says that rank M = L- N + 1. Since any alternate path

forms a circuit with its corresponding spanning tree we have that the number of circuits in the basis will be given by the number of alternate paths. Suppose that $M_B$ is the circuit matrix, i.e., the columns corresponding to flows creating alternate paths for an optimal basis of (LP1); then since (according to Theorem 3.4.2) the sum of the distances along any circuit must be zero, we have that if $\theta$ is the vector of link distances

$$M_B\theta = 0$$

If $M_B$ is full rank, i.e., rank $M_B = L - N + 1$, then, from the orthogonality of the circuit and cut-set matrices of a graph (see [48] page 97] $\theta$ must denote a cut-set or a linear combination of cut-sets. If the network is strongly connected then no cut-set exists having all its links pointing in the same direction. Thus for a strongly connected network if $M_B$ is full rank then $\theta$ cannot have all its components of the same sign which contradicts the definition of $\theta$.

Consider now the weakly connected networks of Fig. 3.5.1. For the network of Fig. a)

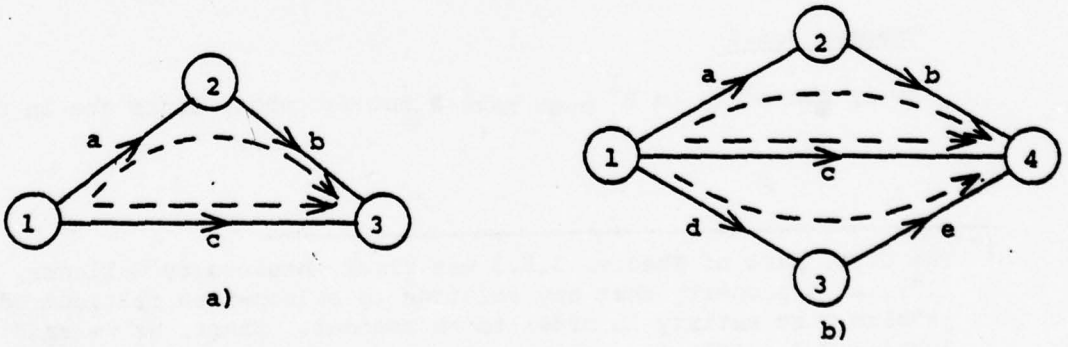



Fig. 3.5.1

the maximum number of alternate path is, obviously, one, but, $L - N + 1 = 1$ for this network. Similarly for the network of Fig. b) the maximum number of alternate paths is two, where $L - N + 1 = 2$.

Hence we conclude:

Theorem 3.5.3

For a strongly connected network the maximum number of alternate paths required by a minimal optimal solution of (LP1) is given by $L - N$. For a weakly connected network this maximum number can be equal to $L - N + 1$.(*)

□

From Lemma 3.5.2 and Theorem 3.5.3 we have

Lemma 3.5.3

For a network with L links and N nodes

$$|S^1| \leq L - N + 1$$

□

We now end this section with an important but rather obvious result.

Theorem 3.5.4

Every commodity in $R^1$ must have a cut-set whose links are in $S^1$.

---

(*) The first part of Theorem 3.5.3 was first obtained by Gallager, in [17], as a property that any solution to a loop-free multicommodity problem must satisfy in order to be optimal. Since, as we said in Section 3.4, (LP1) is only loop-free with respect to loops containing links of $S^1$, we have to restrict our previous results to basic solutions. In Chapter V, these results will be generalized to include all saturation levels.

Proof:

Suppose not. Then there is a commodity in $R^1$ that has a path, from source to destination, that does not have to go through links of $S^1$. But if this is true then that commodity is not in $R^1$, by definition of $R^1$, and we have a contradiction.

Q.E.D.

## 3.6 Alternate Paths, Decision Variables and Canonic Equations

Given an optimal solution to (LP1), take a node, say m, and define, for that solution, the sets

$$\Gamma_m(n) = \{\ell \mid o(\ell) = m, \text{ and } f_\ell(n) > 0\} \tag{3.6.1}$$

$$\Gamma_{\overline{m}}(n) = \{\ell \mid t(\ell) = m, \text{ and } f_\ell(n) > 0\}$$

From these definitions the traffic going from m to n is properly divided among all links in $\Gamma_m(n)$, so

$$\sum_{\ell \in \Gamma_m(n)} f_\ell(n) = r_m(n) + \sum_{\ell \in \Gamma_{\overline{m}}(n)} f_\ell(n) \triangleq r_m^1(n) \tag{3.6.2}$$

Note that, since we are characterizing commodities by destination, the amount of traffic going from m to n is not only $r_m(n)$, but also all traffic going to n that uses m as an intermediate node.

Suppose now that $|\Gamma_m(n)| > 1$; then from Theorem 3.5.1 we know that one of the flows, say $f_\tau(n)$, $\tau \in \Gamma_m(n)$, will be in the spanning tree for commodity n and the remaining flows will constitute alternate paths, from node m, for that commodity. Now suppose that we have a small increment $\Delta r_m(n)$ and that for the new requirement the present

basis is still optimal. The increment will then be apportioned among the links in $\Gamma_m(n)$ so that, if $\tau$ is the link in the spanning tree,

$$\Delta f_\ell(n) = \Delta y_\ell(n) \; ; \qquad \ell \neq \tau$$
$$\Delta f_\tau(n) = \Delta r_m(n) - \sum_{\substack{\ell \in \Gamma_{m(n)} \\ \ell \neq \tau}} \Delta y_\ell(n) \; , \tag{3.6.3}$$

where the values of the $\Delta y_\ell(n)$ have to be obtained from the present (LP1) optimal solution or from the canonic equations of the optimal solution as we will explain later on in this section. If we call $y_\ell(n) \equiv y_{mj}(n)$ the <u>flow or routing variable</u> associated with the alternate path for commodity (m,n) whose first link is link $\ell$, we associate a variable with each of the nonzero flows not included in the spanning trees. For an increment in r within the linear region for the present optimal basis, these variables will indicate how this increment should be apportioned among the various allowable paths.

Consider now that $|S^1| = k$ and that $\ell 1, \ldots, \ell k$, are the elements of $S^1$, then, we can write

$$\alpha_0^1 = \frac{f_{\ell 1}}{c_{\ell 1}} = \frac{f_{\ell 2}}{c_{\ell 2}} = \ldots = \frac{f_{\ell k}}{c_{\ell k}} \; . \tag{3.6.4}$$

This set of equations will be called the set of <u>canonical equations</u> for the first saturation problem. Since the aggregate flows are an algebraic sum of requirements and flow variables, these equations will be a collection of ratios whose numerators are algebraic sums of requirements and flow variables and whose denominators are the capacities of the saturated links.

We can apply the law of proportions(*) to this collection of ratios in order to find an equivalent ratio without decision variables. We will call this ratio the canonical ratio. If we consider, for simplicity, a minimal solution or a convex combination of minimal solutions, then, since the set $S^1$ is unique, we can use (3.4.9) to prove

Lemma 3.6.1

The canonical ratio for a minimal solution or a convex combination of minimal solutions will be given by

$$\alpha^1 = \frac{\sum_{(i,j) \in R^1} d_i(j) r_i(j)}{\sum_{\ell \in S^1} \theta_\ell c_\ell} \triangleq \frac{r^1}{D^1} \quad . \tag{3.6.5}$$

□

Considering now the previous $\Delta r_m(n)$ increment applied to this ratio, and assuming that the remaining requirements are held constant, we have, using (3.1.8)

$$\Delta\alpha^1 = \frac{d_m(n)\ \Delta r_m(n)}{D^1} = \sigma_m(n)\ \Delta r_m(n) \ , \tag{3.6.6}$$

which gives the value of dual variables $\sigma_m(n)$. In a similar way we could have obtained the values of all dual variables. Thus we have

Lemma 3.6.2

If the only information we have about a minimal optimal solution

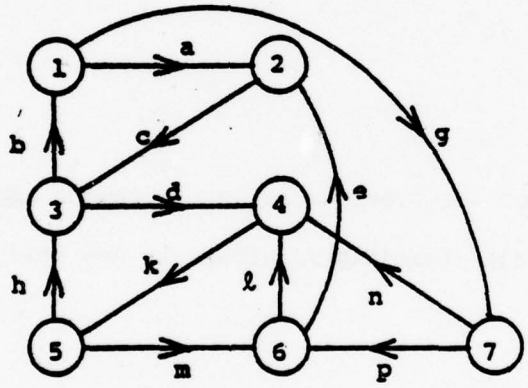
(*) $\frac{a}{b} = \frac{c}{d} = \frac{a+c}{b+d} = \frac{a-c}{b-d}$

are the sets of allowable paths and saturated links, we can obtain from them, via canonic equations and canonical ratio, the set of dual variables using (3.6.5), (3.4.5) and (3.4.8). □

To illustrate these facts and to show other properties of the canonic equations we will use the example below. Through the example we will be able to present the methodology of obtaining parameters via canonic equations without falling into cumbersome notation and without lossing generality.

Example 3.6.1

Consider the network




and suppose that, from an optimal solution of the (LP1) problem for a given requirement vector r, we know that the sets of saturated links and allowable paths are as follow:

$$S^1 = \{a,g,d,m\}$$

$$P_1(1,6) = (g,p) \qquad P_2(1,6) = (a,c,d,k,m)$$

$$P_1(3,2) = (b,a) \qquad P_2(3,2) = (d,k,m,e)$$

$$P_1(5,4) = (m,\ell) \qquad P_2(5,4) = (h,d) \qquad (3.6.7)$$

and assume, for simplicity, that r is such that all other requirements are equal to zero. If we associate a variable $y_i(n)$ to each one of the three alternate paths of (3.6.7) where i is the first link of the alternate path, the canonic equations will be

$$\alpha_0^1 = \frac{r_1(6) - y_a(6)}{c_g} = \frac{r_3(2) - y_d(2) + y_a(6)}{c_a} \tag{3.6.8}$$

$$= \frac{r_5(4) - y_h(4) + y_d(2) + y_a(6)}{c_m} = \frac{y_h(4) + y_d(2) + y_a(6)}{c_d}$$

Adding the first ratio to the last three we eliminate variable $y_a(6)$, and we can then add the new second ratio to the new last two ratios to eliminate variable $y_d(2)$. Finally adding the resulting last two ratios together we would obtain the canonical ratio

$$\alpha_0^1 = \frac{4r_1(6) + 2r_3(2) + r_5(4)}{4c_g + 2c_a + c_d + c_m} \tag{3.6.9}$$

From this ratio and (3.6.5) we can immediately identify

$$\theta_g = 4, \quad \theta_a = 2, \quad \theta_d = \theta_m = 1$$

$$d_1(6) = 4, \quad d_3(2) = 2, \quad d_5(4) = 1 \tag{3.6.10}$$

and if $c_i = 1$ all $i \in L$.

$$\pi_g = -\frac{1}{2}, \quad \pi_a = -\frac{1}{4}, \quad \pi_d = \pi_m = -\frac{1}{8}$$

$$\sigma_1(6) = \frac{1}{2}, \quad \sigma_3(2) = \frac{1}{4}, \quad \sigma_5(4) = \frac{1}{8} \tag{3.6.11}$$

Suppose now we have an increment $\Delta r_3(2)$ in commodity (3,2). From the canonic equation we can write that the corresponding increment in $\alpha_0^1$ will be

$$\Delta\alpha_0^1 = \frac{2\Delta r_3(2)}{D^1} \tag{3.6.12}$$

where from (3.4.8)

$$D^1 = 4c_g + 2c_a + c_d + c_m \ . \tag{3.6.13}$$

To obtain the increment caused by $\Delta r_3(2)$ in variable $y_3(2)$ we need a ratio where $y_3(2)$ is the unique flow variable. This ratio can be obtained by summing the first two ratios of (3.6.8)

$$\alpha_0^1 = \frac{r_1(6) + r_3(2) - y_d(2)}{c_g + c_a} \tag{3.6.14}$$

from which

$$\Delta\alpha_0^1 = \frac{\Delta r_3(2) - \Delta y_d(2)}{c_g + c_a} \tag{3.6.15}$$

and solving for $\Delta y_d(2)$

$$\Delta y_d(2) = \Delta r_3(2) - (c_g + c_a)\ \Delta\alpha_0^1 \ ,$$

using (3.6.12) we have

$$\Delta y_d(2) = \frac{2c_g + c_d + c_m}{D^1}\ \Delta r_3(2) \tag{3.6.16}$$

$$\Delta y_d(2) = \frac{2c_g + c_d + c_m}{2}\ \Delta\alpha_0^1$$

In a similar way we could compute the change induced in the value of other flow variables by the increment $\Delta r_3(2)$

$$\Delta y_a(6) = -c_g \Delta\alpha_0^1 = -\frac{2c_g}{D^1}\Delta r_3(2)$$

$$\Delta y_h(4) = \frac{c_d - c_m}{2}\Delta\alpha_0^1 = \frac{c_d - c_m}{D^1}\Delta r_3(2) \qquad (3.6.17)$$

From equations (3.6.16), (3.6.17) we can define the quantities

$$\eta_3^{d2}(2) = \frac{2c_g + c_d + c_m}{D^1}\ ;\quad c_3^{d2}(2) = \frac{2c_g + c_d + c_m}{2}$$

and similarly for the expressions of (3.6.17). Thus, in general, we could write

$$\Delta y_a(i) = \eta_m^{ai}(n)\Delta r_m(n) = c_m^{ai}(n)\ \Delta\alpha_0^1 \qquad (3.6.18)$$

where

$$c_m^{ai}(n) = \frac{\eta_m^{ai}(n)}{\sigma_m(n)} \qquad (3.6.19)$$

The importance of $\eta_m^{ai}(n)$ and $c_m^{ai}(n)$ is that they are constant for a given linear region and will, therefore, control the value of the flow variable within that region.

By analogy with $D^1$ the term $c_m^{ai}(n)$ will be called the <u>equivalent (pseudo) capacity</u> of variable $y_a(i)$ for changes in $r_m(n)$, where the word "pseudo" is introduced to point out that this term, although a rational function of the actual link capacities, cannot be fully considered as a capacity since it can sometimes be negative, as we shall see below. Equivalently, $\eta_m^{ai}(n)$ will be called the <u>sensitivity factor</u> of variable

$y_a(i)$ for changes in commodity $r_m(n)$. These quantities constitute all the information that is needed to apportion all changes within a linear region. Furthermore they can be distributed throughout the network in such a way that the network will perform as a decentralized system. Specifically, node m will need to know $\{\eta_m^{\ell i}(n)\}$ for all $\ell \in \Gamma_m(i)$ and all n and i such that $(m,i)$, $(m,n) \in R^1$. Efficient representation of the sensitivity factors will be studied, together with other dynamic aspects of the problem, in Chapter VI.

Note that the elements $\eta_m^{\ell i}(n)$ (or $c_m^{\ell i}(n)$) can be considered as forming a matrix with as many rows as variables $y_\ell(i)$ appear in the solution, and as many columns as requirements are in $R^1$. From (3.6.3) we have

$$\Delta r_m(n) = \Delta f_\tau(n) + \sum_{\ell \in \Gamma_m(n)} \Delta y_\ell(n) \qquad (3.6.20)$$

where $\Gamma_m(n)$ was defined in (3.6.1). If by analogy with (3.6.17) we define

$$\Delta f_\tau(n) \triangleq \eta_m^{\tau n}(n)\ \Delta r_m(n) \qquad (3.6.21)$$

then substituting (3.6.18) and (3.6.21) in (3.6.20) and dividing by $\Delta r_m(n)$ we have the following important relation:

Lemma 3.6.3

The sensitivity factors defined by (3.6.18) and (3.6.21) satisfy

$$\sum_{\ell \in \Gamma_m(n)} \eta_m^{\ell n}(n) = 1 \qquad (m,n) \in R^1 \ . \qquad (3.6.22)$$

Similarly the equivalent pseudo capacities satisfy

$$\sum_{\ell \in \Gamma_m(n)} c_m^{\ell n}(n) = \frac{1}{\sigma_m(n)} \qquad (m,n) \in R^1 \tag{3.6.23}$$

□

Note furthermore that the sum of the variations of flows of commodities different from (m,n) has to be zero, since only $r_m(n)$ is changing. Following identical steps as before, we have

Lemma 3.6.4

The sensitivity factors and equivalent pseudo capacities are also related by

$$\sum_{\substack{a \in \bar{\Gamma}_m(n) \\ i \in \chi_a(n)}} \eta_m^{ai}(n) = \sum_{a,i} c_m^{ai}(n) = 0 \tag{3.6.24}$$

where

$$\Gamma \triangleq \{\ell \mid \ell \in L,\ f_\ell(j) > 0,\ \text{all } j \in N \ \text{ and } \ |\Gamma_{o(\ell)}(j)| > 1\} \tag{3.6.25}$$

$$\bar{\Gamma}_m(n) \triangleq \Gamma - \Gamma_m(n) \tag{3.6.26}$$

and

$$\chi_a(n) \triangleq \{i \mid i \in N,\ (j,i) \in R^1,\ \text{all } j \in N,\ |\Gamma_{o(a)}(i)| > 1\}. \tag{3.6.27}$$

□

The condition $|\Gamma_{o(\ell)}(j)| > 1$ in (3.6.25) and (3.6.27) is needed because the sensitivity factors are not defined for commodities without alternate paths. From (3.6.23) we can clearly see that some $c_m^{ai}(n)$ will be negative, so, the introduction of the word "pseudo" in the name of

these quantities is justified.

Following procedures similar to the one just described, we could obtain, by constructing the appropriate ratios from the canonic equations via the law of proportions, any existing dependence between any pair of variables or between a requirement and a variable. In this sense the term "canonic equations" becomes now more appropriate since they contain all the relevant information provided by an optimal solution and they are an excellent tool for finding characteristics of that solution.

To end this section we introduce the concept of equivalent links. Consider a commodity $(m,n) \in R^1$ and all its routing variables $y_\ell(n)$. Suppose that $r_m(n)$ and the variables $y_\ell(n)$ appear only in a subset of the ratios corresponding to the canonic equations, say in those corresponding to links in $S_m^1(n) \subset S^1$. Then we can reduce all ratios corresponding to the complement of $S_m^1(n)$ to a smaller set of ratios not containing internal variables of subset $\bar{S}_m^1(n) \triangleq S^1 - S_m^1(n)$. That is, we can eliminate the internal variables of $\bar{S}_m^1(n)$. As far as commodity $(m,n)$ is concerned the resulting ratios can be interpreted as corresponding to <u>equivalent links</u> with a capacity given by the denominator and an aggregate flow given by the numerator of the ratios. This concept can be extended to equivalence with respect to variables or paths and can be useful to obtain properties of the solution. In the example of Section 3.6 for instance, the canonic equations (3.6.8) can be reduced to

$$\alpha_0^1 = \frac{2r_1(6) + r_3(2) + r_5(4) - y_h(4)}{2c_g + c_a + c_m} = \frac{2r_1(6) + r_3(2) + y_h(4)}{2c_g + c_a + c_d}$$

as far as variable $y_h(4)$ is concerned. From these two ratios, we can easily obtain how every requirement or capacity change will affect this variable. These two ratios can be interpreted as belonging to two new links with capacities given by the denominators and total flow given by the numerators.

Appendix II presents more examples about the use of canonic equations. Note that this methodology is more useful for hand-calculation purposes than for computer calculations. In the next section, we present a different methodology more appropriate for computer use.

# CHAPTER IV

# MATRIX ANALYSIS OF THE FIRST SATURATION PROBLEM

## 4.1 Introduction

Whenever a basis reoptimization is needed in linear programming, matrix $B^{-1}$ plays a central role. Because of this fact one might argue that, even if we find a decentralized procedure to act within a linear region, as we anticipated in section 3.6, we gain little if the whole matrix $B^{-1}$ has to be stored anyway to provide the information needed for a change of linear regions. The rejoinder is that we do not need to store $B^{-1}$ either, because it can be immediately built from the information provided by the canonical equations. Furthermore, whenever optimality is lost, only a fraction of matrix $B^{-1}$ will have to be reconstructed. In this chapter, we will study properties of the linear-program-related matrices that will simplify the dynamic reoptimization process.

## 4.2 The Inverse of an Optimal Basis and other Program Matrices

Suppose that from an optimal solution of (LP1) we are given $S^1$ and $F_B$, where $F_B$ is the set of basic flows

$$F_B = \{f_\ell(n) \,|\, f_\ell(n) > 0, \quad \text{all } \ell \in L, \quad n \in N\} \tag{4.2.1}$$

For simplicity, we will assume that the optimal solution is minimal. (As we will see in Section 5.7, this need entail no loss of generality.) According to Theorem 3.5.1, we can partition the set $F_B$ into two sets, $F_{BA}$ and $F_{BT}$, where $F_{BT}$ will contain the flows corresponding to the

spanning trees and $F_{BA}$ will contain the remaining $|S^1| - 1$ flows, that is, the flows generating alternate paths, see Section 3.5. If we consider the process of finding an optimal solution to (3.1.2) we can divide this process into two steps:

step 1 - the elements of $F_{BT}$ are made basic

step 2 - $\alpha^1$ and the elements of $F_{BA}$ are made basic

Before analyzing these two steps we introduce some notation and definitions. If F is the set of all flows, then

$$F_{\bar{B}} = F - F_B \tag{4.2.2}$$

Similarly

$$\bar{S}^1 = L - S^1$$

where

$$|S^1| \triangleq q$$

and the capacity vector

$$c(L) = (c(S^1), c(\bar{S}^1)) = (c^1, \bar{c}^1) \tag{4.2.3}$$

In the analysis that follows, it will be convenient to reorder the elements of x and therefore the columns of $A^1$, as stated below

$$x^T \equiv (S^1, \bar{S}^1, \alpha^1, F_{BA}, F_{BT}, F_{\bar{B}}) \tag{4.2.4}$$

The new, reordered and partitioned, matrix $A^1$ will then be

$$A^1 = \begin{array}{c|c|c|c|c|c|c} S^1 & \bar{S}^1 & \alpha^1 & F_{BA} & F_{BT} & & F_{\bar{B}} \\ \hline I & 0 & -c^1 & J_B & J_{T1} & \cdots & J_{TN} \quad J_{\bar{B}} \\ \hline 0 & I & -\bar{c}^1 & \bar{J}_B & \bar{J}_{T1} & \cdots & \bar{J}_{TN} \quad \bar{J}_{\bar{B}} \\ \hline & & & E_{B1} & E_{T1} & 0 & E_{\bar{B}1} \\ & 0 & 0 & \vdots & & & \vdots \\ & & & E_{BN} & 0 & E_{TN} & E_{\bar{B}N} \end{array} \qquad (4.2.5)$$

where the J and E matrices can be easily obtained from (3.1.3); in particular if, for any of the J-matrices, $J(i,\ell n)$ is the element corresponding to flow $f_\ell(n)$ and link i we have

$$J(i,\ell n) = \begin{cases} 1 & i = \ell \\ 0 & i \neq \ell \end{cases} \quad \text{all } i \in L, \quad f_\ell(n) \in F \qquad (4.2.6)$$

Similarly if, for any of the E-matrices, $E(ij,\ell n)$ is the element corresponding to flow $f_\ell(n)$ and commodity (i,j) we have

$$E(ij,\ell n) = \begin{cases} 1 & i = o(\ell),\ j = n,\ i \neq n \\ -1 & i = t(\ell),\ j = n,\ i \neq n \\ 0 & \text{otherwise} \end{cases} \quad \begin{array}{l} \text{all } (i,j) \in R \\ \text{all } f_\ell(n) \in F \end{array} \qquad (4.2.7)$$

The finding of an optional basic solution then proceeds as follows:

Step 1

Since the RHS of (3.1.2) has the structure given in (3.1.4) is it not difficult to see that the N(N-1) pivots needed for this first step will take place in the N(N-1) bottom rows of $A^1$. After this first transformation, matrix $A^1$ of (4.2.5) will become

$A_I^1 =$

| $S^1$ | $\bar{S}^1$ | $\alpha^1$ | $F_{BA}$ | $F_{BT}$ | $F_{\bar{B}}$ |
|---|---|---|---|---|---|
| I | 0 | $-c^1$ | $P_B$ | 0 | $P_{\bar{B}}$ |
| 0 | I | $-\bar{c}^1$ | $\bar{P}_B$ | | $\bar{P}_{\bar{B}}$ |
| 0 | | | $M_{B1}$ | I | $M_{\bar{B}1}$ |
| | | | $M_{BN}$ | | $M_{\bar{B}N}$ |

(4.2.8)

where

$$P_\varepsilon = J_\varepsilon - J_T E_T^{-1} E_\varepsilon$$

$$\bar{P}_\varepsilon = \bar{J}_\varepsilon - J_T E_T^{-1} E_\varepsilon \qquad \varepsilon = B, \bar{B} \qquad (4.2.9)$$

and

$$M_\varepsilon = E_T^{-1} E_\varepsilon \qquad \varepsilon = B, \bar{B} \tag{4.2.10}$$

here $\varepsilon$ is an index that is introduced to avoid repetitions and that can be equal to B or $\bar{B}$. $J_T$ and $E_T$ are the upper and lower matrices below $F_{BT}$ in (4.2.5), respectively.

Now we will proceed to interpret (4.2.9) and (4.2.10) as graph matrices. The first task is to determine $E_T^{-1}$. Since $E_T$ is a block diagonal matrix it is clear that

$$E_T^{-1} = \begin{bmatrix} E_{T1}^{-1} & & 0 \\ & \ddots & \\ 0 & & E_{TN}^{-1} \end{bmatrix} \tag{4.2.11}$$

In order to compute $E_{Ti}^{-1}$ and subsequent matrix operations we introduce the following definitions:

Definition 4.2.1

Let T(n) denote the basic spanning tree rooted at node n. Since a tree defines a partial ordering of nodes we will say that node i is below node j, or succeeds node j, if node j is in the path from i to n at tree T(n). Similarly we will say that node i is above, or precedes node j. It may happen that node i neither succeeds nor precedes node j. In this case they belong to different branches of the tree.

$T_i(n)$ will denote the subtree of T(n) rooted at node i, that is, it will refer to the collection of nodes, and corresponding links, that preceeds node i.

$T^1(n)$, or $T_i^1(n)$, will denote tree $T(n)$, or subtree $T_i(n)$, with all nonsaturated links ($\ell \in \bar{S}^1$) shortcircuited.

$T(m,n)$ will denote the path from m to n on tree $T(n)$. $T^1(m,n) \triangleq T(m,n) \cap S^1$.

Definition 4.2.2

$C_\ell(n)$ will denote the circuit formed by the insertion of link $\ell$ into tree $T(n)$. The positive direction of this circuit will correspond to the direction of link $\ell$.

$C_\ell^1(n)$ will denote circuit $C_\ell(n)$ with all nonsaturated links ($\ell \in \bar{S}^1$) shortcircuited, and will be called the saturated or reduced circuit.

□

From Definition 4.2.1 it is clear that $E_{Tn}$ will be the incidence matrix for tree $T(n)$ without row n. The columns of $E_{Tn}$ will correspond to links and the rows will correspond to nodes of the tree, except for node n. It is very easy to see that

Lemma 4.2.1 ([24] page 357, [22])

The incidence matrix of a tree, and therefore $E_{Tn}$, is triangular and unimodular (i.e., has determinant = 1).

Lemma 4.2.2

$E_{Tn}$ can be transformed, by row reordering, into a triangular matrix with all its diagonal elements equal to 1 and, at most, one off diagonal element per column equal to -1.

Proof

Follows from Lemma 4.2.1, the definition of $E_{Tn}$, and the fact that since node n is eliminated from the incidence matrix all nodes of the tree have an out-degree equal to one.

Q.E.D.

To obtain $E_{Tn}^{-1}$ we will apply Theorem AI.1 of Appendix I; therefore, we have to eliminate all nonzero off diagonal elements by means of row transformations. Row i of $E_{Tn}$ will have a +1 in the position corresponding to the arc leaving i, arc (i,·), and -1 in positions corresponding to all arcs incident at i. If (j,i) is such an incident arc then the -1 corresponding to it can be eliminated by adding row j to row i, but in the process we introduce all the -1's of row j that correspond to the links incident at j, see Fig. 4.2.1; these

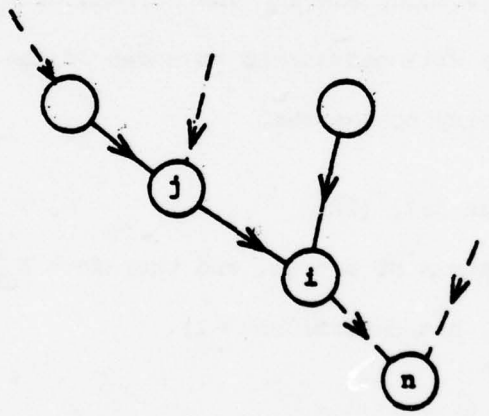



Fig. 4.2.1

new -1's can, nevertheless, be eliminated in the same way. At the end of the process we have that in order to eliminate all -1's from row i we have to sum to it the rows corresponding to all nodes, of

tree T(n), preceding i. If we assume the rows of $E_{Tn}$ are ordered according to Lemma 4.2.2, then, row i does not have to be transformed further in the process of moving from $E_{Tn}$ to I; therefore, if we apply the same sequence of row operations to I we obtain the i-th row of $E_{Tn}^{-1}$. The results of these row transformations into I will be a row with 1 in all places corresponding to nodes of $T_i(n)$. Note that in $E_{Tn}^{-1}$ the rows correspond to links and the columns to nodes. We have therefore proved

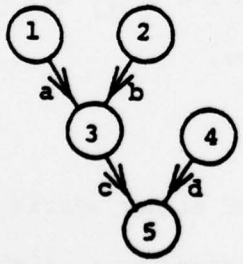
Theorem 4.2.1

$E_{Tn}^{-1}$ is a graph matrix in which the $(\ell,i)$-th element, corresponding to link $\ell$ and node i of T(n), is given by

$$E_{Tn}^{-1}(\ell,i) = \begin{cases} 1 & \text{if } i \in T_{o(\ell)}(n) \\ 0 & \text{otherwise} \end{cases}$$

Example

For the tree below, the incidence matrix $E_{T5}$ and its inverse are:




$$E_{T5} = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{cccc} a & b & c & d \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & -1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{array} ; \quad E_{T5}^{-1} = \begin{array}{c} \\ a \\ b \\ c \\ d \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{array}$$

Thus, the column of $E_{Tn}^{-1}$ corresponding to node i defines the path in tree $T_n$ from node i to node n.

Having determined $E_T^{-1}$, we shall next proceed with the computation of the products $J_T E_T^{-1}$ and $J_T E_T^{-1}$. We will call

$$J \triangleq J_T E_T^{-1}$$
$$\bar{J} \triangleq \bar{J}_T E_T^{-1} \qquad (4.2.12)$$

Obviously $J$ and $\bar{J}$ will have the structure

$$J = \left[ J_1 \,\vdots\, \cdots \,\vdots\, J_N \right] \quad ; \quad \bar{J} = \left[ \bar{J}_1 \,\vdots\, \cdots \,\vdots\, \bar{J}_N \right] \qquad (4.2.13)$$

where, from (4.2.11)

$$J_n = J_T E_{Tn}^{-1} = J_{Tn} E_{Tn}^{-1}$$
$$\bar{J}_n = \bar{J}_{Tn} E_{TN}^{-1} \qquad (4.2.14)$$

and from (4.2.6) and Theorem 4.2.1 we conclude

Lemma 4.2.3

Each row of matrix $J_n (\bar{J}_n)$ correspond to a link of $S^1 (\bar{S}^1)$ and each column correspond to a commodity (i,n), $i \neq n$. The row corresponding to link $\ell$, for $\ell \in T(n)$, (if $\ell \in S^1$ then row $\ell \in J_n$ and if $\ell \in \bar{S}^1$ then row $\ell \in \bar{J}_n$) will have a+1 in all positions i such that $i \in T_{o(\ell)}$ and zeroes otherwise. If $\ell \notin T(n)$ then row $\ell$ is a row of zeros.

Corollary L4.2.3

The column corresponding to commodity (i,n), of either matrix $J_n$ or matrix $\bar{J}_n$, will have a+1 in all positions corresponding to links of the path i to n, in the spanning tree T(n), and zeroes otherwise. If the link is in $S^1$ then the element will be in matrix $J_n$, and will be in $\bar{J}_n$ otherwise.

□

From Theorem 4.2.1 and Corollary L4.2.3 we see that the rows of $E_{Tn}^{-1}$ are assigned either to $J_n$ or $\bar{J}_n$. In addition $J_n(\bar{J}_n)$ will have rows of zeroes for all links in $S^1(\bar{S}^1)$ not included in the tree T(n).

Finally we need to obtain $JE_\varepsilon$ and $\bar{JE}_\varepsilon$. If $E_\varepsilon(\underline{c}\ell n)$ denotes the column of $E_\varepsilon$ corresponding to variable $f_\ell(n)$ then from (4.2.9)

$$P_\varepsilon(\underline{c}\ell n) = J_\varepsilon(\underline{c}\ell n) - JE_\varepsilon(\underline{c}\ell n)$$

and similarly for $\bar{P}_\varepsilon$. But from (4.2.7) and (4.2.13)

$$JE_\varepsilon(\underline{c}\ell n) = J_n E_\varepsilon(\underline{c}\ell n)$$

The interpretation of these quantities is evident in the following example.

Example 4.2.1

Consider the tree T(7) of Fig. E.4.2.1.

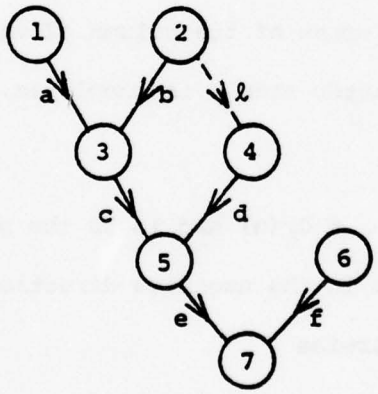



Fig. E. 4.2.1

we have

$$E_\varepsilon(\underline{c}\ell 7) = \begin{bmatrix} 0 \\ 1 \\ 0 \\ -1 \\ 0 \\ 0 \end{bmatrix} \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{matrix}$$

$$J_7(\underline{r}c) = [1 \quad 1 \quad 1 \quad 0 \quad 0 \quad 0]$$

$$J_7(\underline{r}d) = [0 \quad 0 \quad 0 \quad 1 \quad 0 \quad 0]$$

$$J_7(\underline{r}e) = [1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 0]$$

so that

$$J_7(\underline{r}c)\ E_\varepsilon(\underline{c}\ell 7) = 1$$

$$J_7(\underline{r}d)\ E_\varepsilon(\underline{c}\ell 7) = -1$$

$$J_7(\underline{r}e)\ E_\varepsilon(\underline{c}\ell 7) = 0$$

etc.

Thus, using (4.2.7), Lemma 4.2.3 and Definition 4.2.2, together with the interpretation provided by the example, we have

Theorem 4.2.2

If $p_I(a,\ell n)$ is the element of the column of variable $f_\ell(n) \in F_{BA} \cup F_{\bar{B}}$ corresponding to link a, after step 1 is completed, then this element will be given by

$$p_I(a,\ell n) = \begin{cases} 1 & \text{If } a \in C_\ell(n) \text{ and is in the positive direction} \\ -1 & \text{Idem in the negative direction} \\ 0 & \text{otherwise} \end{cases}$$

□

Since element $p_I(a,\ell n)$ belongs to one of the P-matrices of (4.2.8) we also have

Corollary 4.2.2

If $P_B(\underline{c}\ell n)$ is the column of $P_B$ corresponding to $f_\ell(n) \in F_{BA}$ then $P_B(\underline{c}\ell n)$ is a vector representing the saturated circuit of variable $f_\ell(n)$.

Similarly for $P_{\overline{B}}(\underline{c}\ell n)$ and $f_\ell(n) \in F_{\overline{B}}$

□

This concludes determination of (4.2.9). We can also prove an equivalent result for the M-matrices of (4.2.10) with the help of Corollary L4.2.3 and Example 4.2.1

Theorem 4.2.3

If $M_B(ij,\ell n)$ is the element of matrix $M_B$ corresponding to commodity (i,j) and variable $f_\ell(n) \in F_{BA}$ then

$$M_B(ij,\ell n) = \begin{cases} 1 & \text{if } j=n,\ \text{arc}(i,\cdot) \in C_\ell(n) \cap T(n) \text{ and is in the negative direction} \\ -1 & \text{Idem in the positive direction} \\ 0 & \text{otherwise} \end{cases}$$

where $\text{arc}(i,\cdot)$ denotes the (only) arc leaving i in T(n), and $C_\ell(n) \cap T(n) = C_\ell(n) - \ell$.

Similarly for $M_{\overline{B}}(ij,\ell n)$, $f_\ell(n) \in F_{\overline{B}}$ .

Corollary 4.2.3

$$M_B(in,\ell n) = \begin{cases} -P_B(a,\ell n) & \text{for } o(a)=i,\ a \neq \ell,\ a \in T(n) \\ 0 & \text{otherwise} \end{cases}$$

Similarly for $M_{\overline{B}}$ and $P_{\overline{B}}$ .

From Theorems 4.2.2 and 4.2.3 we see that M and $P_B$ are graph-related matrices and therefore can be easily obtained from the graph. The result

of Theorem 4.2.2 is a known result of networks, see [38], but we decided to prove it, here again, because we need the intermediate results to prove Theorem 4.2.3 and other results below. For instance we will need some of those results to obtain $B_I^{-1}$, the inverse of the current basis. For this purpose note that if we start the linear program with an extended matrix $(A^1, I)$ then $B_I^{-1}(A^1, I) = (A_I^1, B_I^{-1})$, see Theorem AI.1 of Appendix I, so $B_I^{-1}$ is the matrix which occupies the position of I after the first step is completed. Therefore

Theorem 4.2.4

The inverse of the current basis, after the first step is completed, is

$$B_I^{-1} = \begin{array}{c|c|c|c} \multicolumn{2}{c}{S^1 \quad \bar{S}^1} & \multicolumn{2}{c}{R} \\ \hline I & 0 & -J_1 \cdots & -J_N \\ \hline 0 & I & -\bar{J}_1 \cdots & -\bar{J}_N \\ \hline \multicolumn{2}{c|}{\multirow{2}{*}{0}} & E_{T1}^{-1} & 0 \\ \multicolumn{2}{c|}{} & \vdots & \\ \hline \multicolumn{2}{c|}{} & 0 & E_{TN}^{-1} \end{array} \qquad (4.2.15)$$

where the J-matrices are given by Lemma 4.2.3 and the $E_T^{-1}$-matrices by Theorem 4.2.1.

With respect to other vectors involved with the linear program, note that neither the vector of partial cost nor the value of $\alpha^1$, has changed during this first step. On the other hand the new RHS, $b_I$, is given by

$$b_I = B^{-1}b = B^{-1}\begin{bmatrix}0\\ \hline r\end{bmatrix} , \tag{4.2.16}$$

where if we partition $b_I$ into three parts corresponding to $S^1$, $\bar{S}^1$ and $F_{BT}$ respectively, we have

$$\begin{aligned} b_{I1} &= -Jr = -\sum_{n=1}^{N} J_n r(n) \\ b_{I\bar{1}} &= -\bar{J}r = -\sum_{n=1}^{N} \bar{J}_n r(n) \\ b_{IT} &= E_T^{-1} r = \sum_{n=1}^{N} E_{Tn}^{-1} r(n) \end{aligned} \tag{4.2.17}$$

and partitioning $b_{IT}$ further, into N subvectors, we have

$$b_{IT}(n) = E_{Tn}^{-1}\, r(n), \quad n \in N \tag{4.2.18}$$

From (4.2.17), (4.2.18), Lemma 4.2.3 and Theorem 4.2.1 we can write

Lemma 4.2.4

The element of $b_I$ corresponding to link $\ell$ will be given by

$$b_I(\ell) = -\sum_{n=1}^{N} \sum_{i \in T_{o(\ell)}(n)} r_i(n) \qquad \text{all } \ell \in L \tag{4.2.19}$$

Corollary L4.2.4

The value of $b_I(\ell)$ will be equal to minus the value of the aggregate flow that link $\ell$ would have if all commodities could flow only through links of spanning trees. For this reason we will oftenly call $-b_I(\ell) \equiv f_{T\ell}$

Lemma 4.2.5

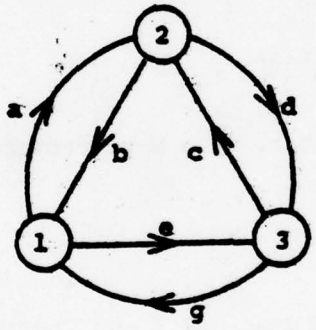
The subvector of $b_{IT}$ corresponding to commodity n will be given by (4.2.18). The component of this subvector corresponding to commodity (i,n) will be given by

$$b_{IT}(i,n) = \sum_{j \in T_i(n)} r_j(n) \quad \text{all } (i,n) \in R \tag{4.2.20}$$

The foregoing analysis of Step 1 is summarized in the following

Example 4.2.2

Consider the three nodes completely connected network below and assume that, from an optimal basic solution we are given the following

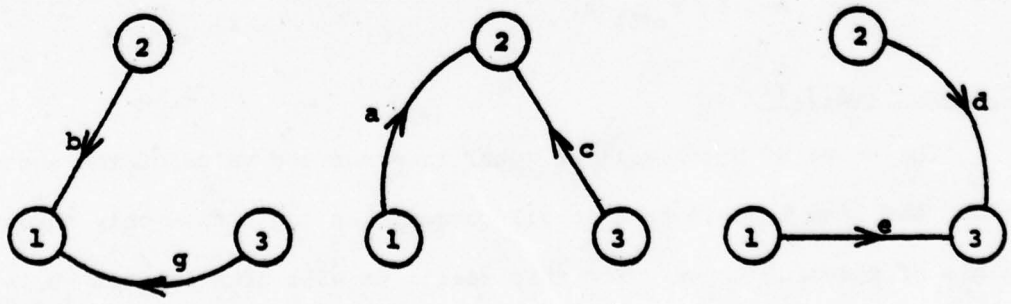



data

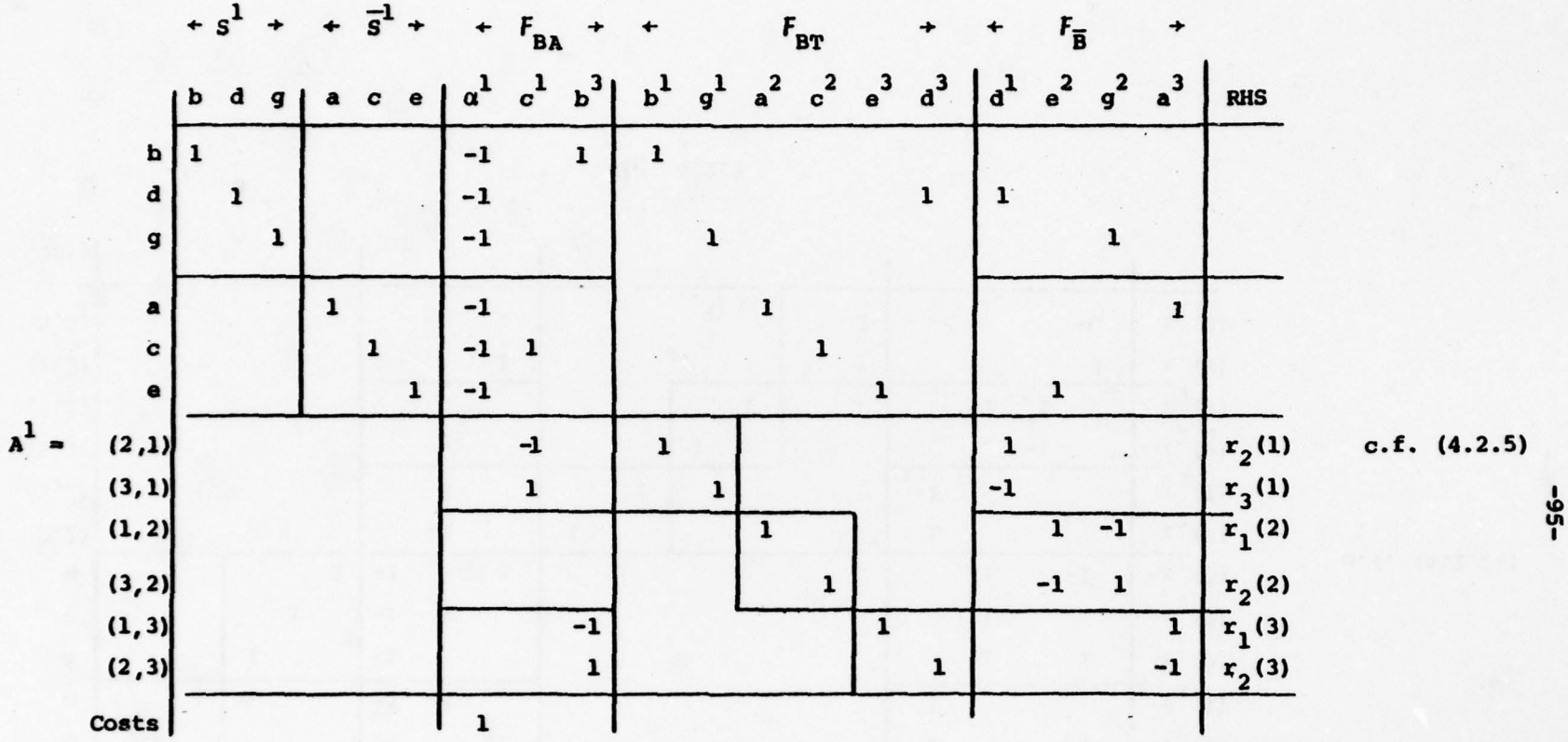
$$S^1 = \{b, d, g\}$$

$$F_B = \{f_b(1), f_c(1), f_g(1), f_a(2), f_c(2), f_b(3), f_e(3), f_d(3)\}$$

If we choose the trees below as the basic trees, then, $F_B$ is partioned




as follows




c.f. (4.2.5)

FIG. 4.2.2

$A_I^1$ =

| | ← $S^1$ → | | | ← $\bar{S}^{-1}$ → | | | ← $F_{BA}$ → | | | ← $F_{BT}$ → | | | | | | ← $F_{\bar{B}}$ → | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | b | d | g | a | c | e | $\alpha^1$ | $c^1$ | $b^3$ | $b^1$ | $g^1$ | $a^2$ | $c^2$ | $e^3$ | $d^3$ | $d^1$ | $e^2$ | $g^2$ | $a^3$ | RHS |
| b | 1 | | | | | | -1 | 1 | 1 | | | | | | | -1 | | | | $-r_2(1)$ |
| d | | 1 | | | | | -1 | | -1 | | | | | | | 1 | | | 1 | $-r_2(3)$ |
| g | | | 1 | | | | -1 | -1 | | | | | | | | 1 | | | | $-r_3(1)$ |
| a | | | | 1 | | | -1 | | | | | | | | | | -1 | | 1 | $-r_1(2)$ |
| c | | | | | 1 | | -1 | 1 | | | | | | | | | 1 | | | $-r_3(2)$ |
| e | | | | | | 1 | -1 | | 1 | | | | | | | | 1 | | -1 | $-r_1(3)$ |
| (2,1) | | | | | | | | -1 | | 1 | | | | | | 1 | | | | $r_2(1)$ |
| (3,1) | | | | | | | | 1 | | | 1 | | | | | -1 | | | | $r_3(1)$ |
| (1,2) | | | | | | | | | | | | 1 | | | | | 1 | -1 | | $r_1(2)$ |
| (3,2) | | | | | | | | | | | | | 1 | | | | -1 | 1 | | $r_3(2)$ |
| (1,3) | | | | | | | | | -1 | | | | | 1 | | | | | 1 | $r_1(3)$ |
| (2,3) | | | | | | | | | 1 | | | | | | 1 | | | | -1 | $r_2(3)$ |
| Costs | | | | | | | 1 | | | | | | | | | | | | | |

c.f. (4.2.28)

Fig. 4.2.3

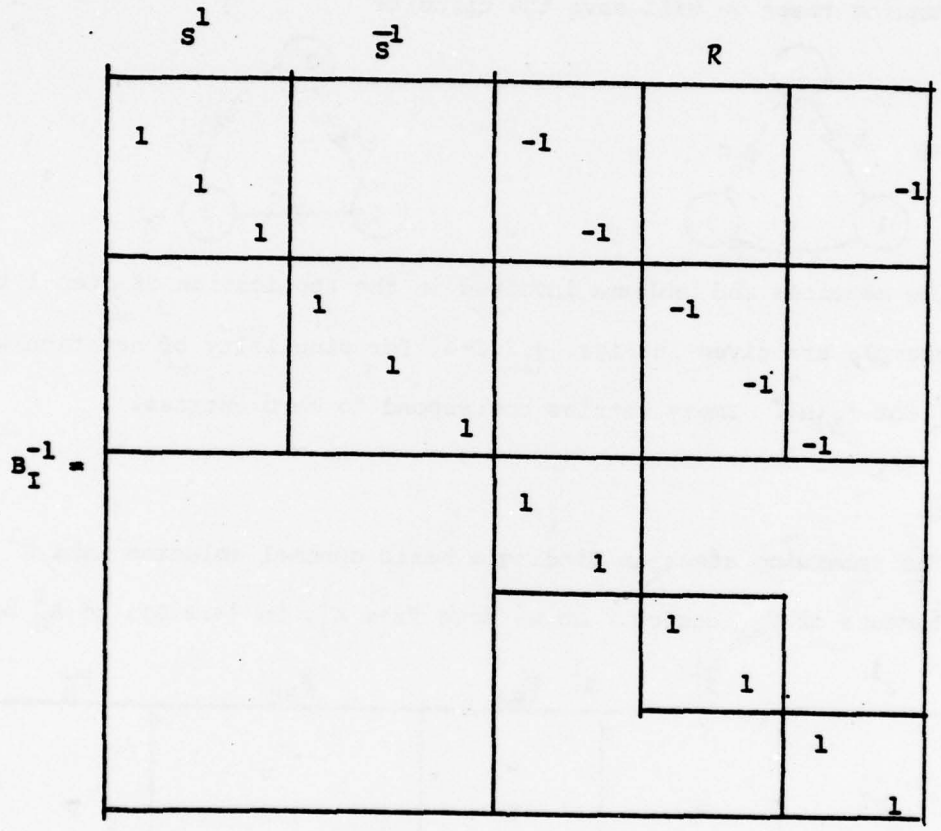
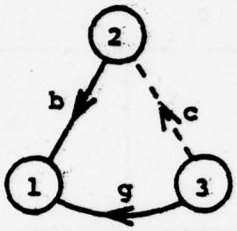


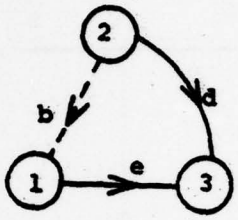

Fig. 4.2.4

$$F_{BT} = \{f_b(1),\ f_g(1),\ f_a(2),\ f_c(2),\ f_e(3),\ f_d(3)\}$$

$$F_{BA} = \{f_c(1),\ f_b(3)\}$$

Thus we have two alternate paths. Inserting the elements of $F_{BA}$ into the spanning trees we will have the circuits




The matrices and tableaus involved in the application of Step 1 to this example are given in Figs. 4.2.2-4, for simplicity of notation we will use $\ell^n$ for $f_\ell(n)$. Empty entries correspond to zero entries.

Step 2

The remaining steps in finding a basic optimal solution make $\alpha^1$ and the elements of $F_{BA}$ basics. So we move from $A_I^1$, in (4.2.8), to $A_0^1$ below

$A_0^1 =$

| $S^1$ | $\bar{S}^1$ | $\alpha^1$ | $F_{BA}$ | $F_{BT}$ | $F_{\bar{B}}$ |
|---|---|---|---|---|---|
| $Q^{-1}$ | 0 | I | | 0 | P |
| $\bar{G}$ | I | 0 | | | $\bar{P}$ |
| $G_1$ | 0 | | | I | $M_1$ |
| ⋮ | | | | | ⋮ |
| $G_N$ | | | | | $M_N$ |

(4.2.21)

where

$$Q = \left[ -c^1 \,\middle|\, P_B \right] \tag{4.2.22}$$

$$\begin{aligned} \bar{G} &= -\bar{\Phi} Q^{-1} \\ G_n &= -\Phi_n Q^{-1} \qquad \text{all } n \in N \\ P &= Q^{-1} P_{\bar{B}} \\ \bar{P} &= \bar{P}_{\bar{B}} - \bar{\Phi} P = \bar{P}_{\bar{B}} + \bar{G} P_{\bar{B}} \\ M_n &= M_{\bar{B}n} - \Phi_n P = M_{\bar{B}n} + G_n P_{\bar{B}} \end{aligned} \tag{4.2.23}$$

and

$$\bar{\Phi} = \left[ -\bar{c}^1 \,\middle|\, \bar{P}_B \right] \quad ; \quad \Phi_n = \left[ 0 \,\middle|\, M_{Bn} \right] \tag{4.2.24}$$

Thus

Theorem 4.2.5

The updated constraint matrix for the optimal solution(*) (4.2.1) is given by (4.2.21) - (4.2.24), where all matrices in the RHS of equations (4.2.22) - (4.2.24) are graph-related matrices and where the only matrix that has to be inverted is Q, a $|S^1| \times |S^1|$ matrix.

□

Because of the importance of the matrix Q we devote some sections of this chapter to the study of its structure and some of its properties

(*) Note that there will be many constraint matrices depending on how we choose the spanning trees and on how we order the elements of $F_{B1}$, $F_{BT}$ and $F_{\bar{B}}$. Nevertheless, all represent the same optimal solution, but written in different coordinates.

From Theorems 4.2.4 and 4.2.5 the next theorem follows

Theorem 4.2.6

The inverse matrix of an optimal basis(*) for the optimal solution defined by (4.2.1) is

$$B^{-1} = \begin{array}{c|c|ccc} S^1 & \bar{S}^1 & & R & \\ \hline Q^{-1} & 0 & V_1 & \cdots & V_N \\ \hline \bar{G} & I & \bar{V}_1 & \cdots & \bar{V}_N \\ \hline G_1 & & U_{11} & \cdots & U_{1N} \\ \vdots & 0 & \vdots & \cdots & \vdots \\ G_N & & U_{N1} & \cdots & U_{NN} \end{array} \qquad (4.2.25)$$

where the matrices under $S^1$ are given in (4.2.22) - (4.2.24) and

$$\begin{aligned} V_n &= -Q^{-1}J_n \qquad \text{all } n \in N \\ \bar{V}_n &= -\bar{J}_n - \bar{\Phi}V_n = -\bar{J}_n + GJ_n \qquad \text{all } n \in N \\ U_{in} &= -\Phi_i V_n = -G_i J_n \qquad \text{all } (i,n) \in R \\ U_{ni} &= -\Phi_n V_i = -G_n J_i \\ U_{nn} &= E_{Tn}^{-1} - \Phi_n V_n = E_{Tn}^{-1} - G_n J_n \qquad \text{all } n \in N \end{aligned} \qquad (4.2.26)$$

□

(*) See note for Theorem 4.2.5.

Now we will compute the optimal partial costs, dual variables and RHS. From standard linear programming analysis, see [28] [29], we have:

• for the dual variables

$$\lambda \triangleq (\pi,\sigma) = \gamma^1 B^{-1} \tag{4.2.27}$$

• for the optimal partial cost vector

$$z = \gamma^1 - \lambda A^1 \tag{4.2.28}$$

• for the optimal RHS, that is for the optimal value of the basic variables,

$$b_0 = B^{-1} b \tag{4.2.29}$$

and finally

• for the optimal objective function

$$\alpha_0^1 = \lambda b \tag{4.2.30}$$

Note that these formulas express the optimal values as a function of starting matrices of (LP1). To consider this process broken into two steps, as we are doing in this section, observe that the optimal values can also be obtained by considering $b_I$ of (4.2.16) as an additional column of $A_I^1$ and applying to it the same rules applied to the $A_I^1$ submatrices, see (4.2.23). We will apply here both procedures, depending on which one is easier or more relevant for each case.

From (3.1.5) we know that $\gamma^1$ has a 1 at the position corresponding to $\alpha^1$ and zeroes elsewhere. From (4.2.25) we see that $\alpha^1$ is the first element among the basic variables, so we conclude

Theorem 4.2.7

The optimal dual variables of (LP1) are given by the first row of $B^{-1}$, in (4.2.25). In particular

$$\pi(S^1) = Q^{-1}(\underline{r}1)$$

$$\pi(\bar{S}^1) = 0 \qquad (4.2.31)$$

$$\sigma(n) = V_n(\underline{r}1) \quad n \in N$$

Lemma 4.2.6

$$\sigma_i(n) = -\sum_{\ell \in T(i,n) \cap S^1} \pi_\ell \qquad (4.2.32)$$

$$T(i,n) \cap S^1 = \phi \Rightarrow \sigma_i(n) = 0$$

Proof: From (4.2.26), corollary L4.2.3 and (4.2.31).

Note that some of the results of this Theorem and Lemma were already obtained in Chapter III, see Section 3.4, by means of the complementary slackness condition of linear programming.

Next we consider the vector of partial costs, which has the following structure:

$$z = [z(S^1), z(\bar{S}^1), z(\alpha), z(F_{B1}), z(F_{BT}), z(F_{\bar{B}})], \qquad (4.2.23)$$

Since all basic variables have a null cost, the only nonzero subvectors are $z(S^1)$ and $z(F_{\bar{B}})$. From the definition of $\gamma^1$, (3.1.5), and $\lambda$ we have

$$z(S^1) = -\pi(S^1),$$

since the submatrix of $A^1$ corresponding to $S^1$ was the identity, see (4.2.5). Similarly

$$z(F_{\bar{B}}) = -\lambda A^{1}(F_{\bar{B}}) = -(\pi,\sigma)\begin{bmatrix} J_{\bar{B}} \\ \hline \bar{J}_{\bar{B}} \\ \hline E_{\bar{B}} \end{bmatrix} =$$

$$= -\pi(S^{1})J_{\bar{B}} - \sigma E_{\bar{B}} \qquad (4.2.34)$$

From the definition of $J_{\bar{B}}$, Lemma 4.2.3, and of $E_{\bar{B}}$, (4.2.7), we can write, for the partial cost of variable $f_{\ell}(n)$

$$z_{\ell}(n) = -\pi_{\ell} - \sigma_{o(\ell)}(n) + \sigma_{t(\ell)}(n)$$

Summarizing, we have

Theorem 4.2.8

If z, given by (4.2.33), is the(*) optimal partial cost vector of (LP1) then the only nonzero subvectors are $z(S^{1})$ and $z(F_{\bar{B}})$, where

$$z(S^{1}) = -\pi(S^{1})$$

$$z_{\ell}(n) = \sigma_{t(\ell)}(n) - \sigma_{o(\ell)}(n) - \pi_{\ell}, \quad f_{\ell}(n) \in F_{\bar{B}} \qquad (4.2.35)$$

□

Finally, to study the nature of the optimal RHS we partition $b_0$ into three subvectors, as in (4.2.17), where

$$b_0 = \begin{bmatrix} b_{01} \\ \hline b_{0\bar{1}} \\ \hline b_{0T} \end{bmatrix} = \begin{bmatrix} \alpha^{1} \\ F_{\bar{B}\Delta} \\ \hline s(\bar{S}^{-1}) \\ \hline F_{BT} \end{bmatrix} \qquad (4.2.36)$$

(*) Under our assumptions of this section the dual problem is unique, see section 3.2.

and $s(\bar{S}^1)$ is the set of slack variables corresponding to the links of $\bar{S}^1$.

From (4.2.29) and Corollary L4.2.4 we have

$$b_{01} = -Q^{-1}Jr = Q^{-1}b_{I1} = -Q^{-1}\begin{bmatrix} f_{1t} \\ \cdot \\ \cdot \\ f_{qt} \end{bmatrix} \triangleq -Q^{-1}f_T(S^1) \qquad (4.2.37)$$

In particular

$$\alpha_0^1 = -\sum_{\ell \in S^1} \pi_\ell f_{T\ell} \qquad (4.2.38)$$

which gives another expression for $\alpha^1$, different from the two descriptions given in (3.1.7).

Similarly from corollary L4.2.4, (4.2.24) and (4.2.37)

$$b_{0\bar{1}} = b_{I\bar{1}} - \bar{\Phi}b_{01} = -f_T(\bar{S}^1) - \left[-\bar{c}^1 \middle| \bar{P}_B\right]\begin{bmatrix} \alpha^1 \\ \hline F_{BA} \end{bmatrix}$$

where $f_T(\bar{S}^1)$ is the vector of aggregate flows obtained from the spanning trees for links of $\bar{S}^1$. Thus, since $b_{0\bar{1}}$ corresponds to the value of the slack for links of $\bar{S}^1$,

$$s(\bar{S}^1) = -f_T(\bar{S}^1) + \alpha^1\bar{c}^1 - \bar{P}_B F_{BA}$$

$$= \alpha^1\bar{c}^1 - \left[f_T(\bar{S}^1) + \bar{P}_B F_{BA}\right]$$

Therefore

$$f(\bar{S}^1) = f_T(\bar{S}^1) + \bar{P}_B F_{BA}$$

is the vector of values of the aggregate flows for links of $\bar{S}^1$ corresponding to the present optimal solution.

From (4.2.29), (4.2.17) and (4.2.24) we have for the values of the basic flows on the spanning tree

$$b_{0T} = b_{IT} - \Phi b_{01} = b_{IT} - \left[0 \mid M_B\right]\begin{bmatrix}\alpha^1 \\ \hline F_{BA}\end{bmatrix}$$

$$= b_{IT} - M_B F_{BA}$$

Thus

$$b_{0T}(n) = b_{IT}(n) - M_{Bn} F_{BA}$$

where $M_{Bn} F_{BA}$ determines the amounts of flow taken out from links of the spanning trees by the alternate paths created by $F_{BA}$.

In summary,

Theorem 4.2.9

The optimal RHS for (LP1), partitioned as in (4.2.36), has the following values.

$$\alpha^1 = -\sum_{\ell \in S^1} \pi_\ell f_{T\ell} = -\pi(S^1) f_T(S^1)$$

$$F_{BA} = -Q^{-1}(\overline{r1}) f_T(S^1)$$

$$S(\bar{S}^{-1}) = \alpha^1 c^{\bar{1}-1} - f(\bar{S}^{-1}) \qquad (4.2.39)$$

where

$$f(\bar{S}^{-1}) = f_T(\bar{S}^{-1}) + \bar{P}_B F_{BA}$$

is the vector of optimal values for the aggregate flows of links of $\bar{S}^{-1}$. Finally,

$$F_{BT}(n) = b_{IT}(n) - M_{Bn}F_{BA} \quad . \qquad \text{all } n \in N$$

where $f_T(S^1)$ and $f_T(\bar{S}^1)$ are given by Corollary L4.2.4, $b_{IT}(n)$ is given by (4.2.20), and $\bar{P}_B$ and $M_{Bn}$ are given in (4.2.23) and Theorem 4.2.3, respectively.

□

Example 4.2.3

Applying Step 2 to the three node example of this section (Example 4.2.2) we have the matrices and tableaus of Fig. E.4.2.3. We also represent below the collection of allowable paths and their traffic according to the basic solution obtained in the tableaus

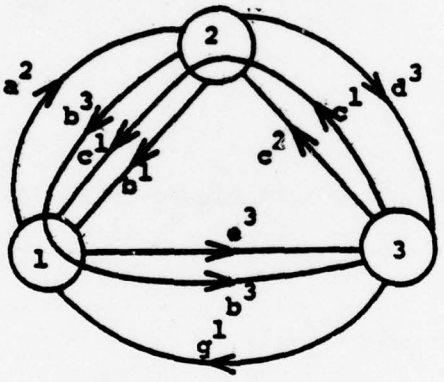



Fig. 4.2.3a

The optimal dual variables are

$$\pi_b = \pi_d = \pi_g = -\frac{1}{3} \; ; \quad \pi_a = \pi_c = \pi_e = 0$$

$$\sigma_2(1) = \sigma_3(1) = \sigma_2(3) = \frac{1}{3} \; ; \quad \sigma_1(2) = \sigma_3(2) = \sigma_1(3) = 0 \; ,$$

and for the requirement vector value

$$r = [r_2(1), r_3(1), r_1(2), r_3(2), r_1(3), r_2(3)] = [3,8,5,1,1,7]$$

$A_0^1 =$

| | b | d | g | a | c | e | $\alpha^1$ | $c^1$ | $b^3$ | $b^1$ | $g^1$ | $a^2$ | $c^2$ | $e^3$ | $d^3$ | $d^1$ | $e^2$ | $g^2$ | $a^3$ | RHS |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| b | -1/3 | -1/3 | -1/3 | | | | 1 | | | | | | | | | -1/3 | | -1/3 | -1/3 | $\alpha_0^1$ |
| d | 1/3 | 1/3 | -2/3 | | | | | 1 | | | | | | | | -2/3 | | -2/3 | 1/3 | $c_0^1$ |
| g | 1/3 | -2/3 | 1/3 | | | | | | 1 | | | | | | | -2/3 | | 1/3 | -2/3 | $b_0^3$ |
| a | -1/3 | -1/3 | -1/3 | 1 | | | | | | | | | | | | -1/3 | -1 | -1/3 | -1/3 | $s_a$ |
| c | -2/3 | -2/3 | 1/3 | | 1 | | | | | | | | | | | 1/3 | 1 | 1/3 | -2/3 | $s_c$ |
| e | -2/3 | 1/3 | -2/3 | | | 1 | | | | | | | | | | 1/3 | 1 | -2/3 | 1/3 | $s_e$ |
| (2,1) | 1/3 | 1/3 | -2/3 | | | | | | | 1 | | | | | | 1/3 | | -2/3 | 1/3 | $r_2(1) + c_0^1$ |
| (3,1) | -1/3 | -1/3 | 2/3 | | | | | | | | 1 | | | | | -1/3 | | 2/3 | -1/3 | $r_3(1) - c_0^1$ |
| (1,2) | | | | | | | | | | | | 1 | | | | | 1 | -1 | | $r_1(2)$ |
| (3,2) | | | | | | | | | | | | | 1 | | | | -1 | 1 | | $r_3(2)$ |
| (1,3) | 1/3 | -2/3 | 1/3 | | | | | | | | | | | 1 | | -2/3 | | -1/3 | 1/3 | $r_1(3) + b_0^3$ |
| (2,3) | -1/3 | 2/3 | 1/3 | | | | | | | | | | | | 1 | 2/3 | | -1/3 | -1/3 | $r_1(3) - b_0^3$ |
| Costs | 1/3 | 1/3 | 1/3 | | | | | | | | | | | | | 1/3 | | 1/3 | 1/3 | |

$$\alpha_0^1 = \tfrac{1}{3}[r_2(1) + r_2(3) + r_3(1)] \qquad s_a = \alpha_0^1 - r_1(2)$$

$$c_0^1 = \tfrac{1}{3}[-r_2(1) - r_2(3) + 2r_3(1)]; \qquad s_c = \alpha_0^1 - c_0^1 - r_3(2)$$

$$b_0^3 = \tfrac{1}{3}[-r_2(1) + 2r_2(3) - r_3(1)]; \qquad s_e = \alpha_0^1 - b_0^3 - r_1(3)$$

FIG. 4.2.5

$B^{-1}$ =

| $S^1$ | | | $\bar{S}^1$ | | | $R$ | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| -1/3 | -1/3 | -1/3 | | | | 1/3 | 1/3 | | | 1/3 |
| 1/3 | 1/3 | -2/3 | | | | -1/3 | 2/3 | | | -1/3 |
| 1/3 | -2/3 | 1/3 | | | | -1/3 | -1/3 | | | 2/3 |
| -1/2 | -1/3 | -1/3 | 1 | | | 1/3 | 1/3 | -1 | | 1/3 |
| -2/3 | -2/3 | 1/3 | | 1 | | 2/3 | -1/3 | | -1 | 2/3 |
| -2/3 | 1/3 | -2/3 | | | 1 | 2/3 | 2/3 | | | -1 -1/3 |
| 1/3 | 1/3 | -2/3 | | | | 2/3 | 2/3 | | | -1/3 |
| -1/3 | -1/3 | 2/3 | | | | 1/3 | 1/3 | | | -1/3 |
| | | | | | | | | 1 | | |
| | | | | | | | | | 1 | |
| 1/3 | -2/3 | 1/3 | | | | -1/3 | -1/3 | | | 1 2/3 |
| -1/3 | 2/3 | 1/3 | | | | 1/3 | 1/3 | | | -2/3 |
| 1/3 | 1/3 | 1/3 | | | | 1/3 | 1/3 | | | 1/3 ← Optimal Dual Variables |

$R = \{(2,1), (3,1), (1,2), (3,2), (1,3), (2,3)\}$

FIG. 4.2.6

we have the solution of Fig. E.4.2. 3b with $\alpha_0^1 = 6$

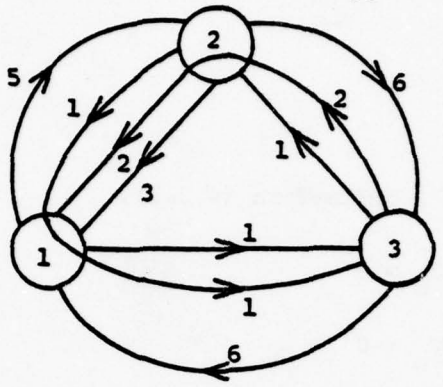



Fig. E.4.2.3b

## 4.3 Properties and Graph Simplification of Program Matrices

Now we procede to a microscopic study of the submatrices appearing in $A_0^1$ and $B^{-1}$, in order to bring further simplifications into these matrices(*). Consider, for instance, a destination n such that there are no commodities $(i,n) \in R^1$ and define $\bar{N}_1$ as the sets of destinations with this property, that is

$$\bar{N}^1 \triangleq \{n \mid (i,n) \notin R^1 \text{ for all } i \in N,\ i \neq n\} . \tag{4.3.1}$$

It is obvious that

$$n \in \bar{N}^1 \Rightarrow f_\ell(n) \notin F_{BA} \text{ for all } \ell \in L \tag{4.3.2}$$

so

$$n \in \bar{N}^1 \Rightarrow M_{Bn} = 0 . \tag{4.3.3}$$

Furthermore, from Lemma 4.2.3

$$n \in \bar{N}^1 \Rightarrow J_n = 0 . \tag{4.3.4}$$

(*) Many of the properties of this section are evidenced by the three node example of last section.

Substituting (4.3.3) and (4.3.4) in (4.2.22) - (4.2.24) and (4.2.26) we have

Lemma 4.3.1

For all $n \in \bar{N}^1$, defined in (4.3.1),

$$\left.\begin{array}{ll} M_{Bn} = 0 & J_n = 0 \\ \Phi_n = 0 & V_n = 0 \\ G_n = 0 & \bar{V}_n = -\bar{J}_n \\ M_n = M_{Bn} & U_{in} = 0 \\ U_{nn} = E_{Tn}^{-1} & U_{ni} = 0 \end{array}\right\} \quad \text{all } i \in N,\ i \neq n .$$

If, in accordance with our standard terminology, D denotes some matrix then $D(\underline{c}\ell n)$ represents the column of D corresponding to variable $f_\ell(n)$. The following lemmas are an immediate consequence of previous definitions and results.

Lemma 4.3.2

If $f_\ell(n) \in F_{BA}$ then

$$M_{Bi}(\underline{c}\ell n) = 0 \quad \text{for all } i \in N,\ i \neq n .$$

Similarly if $f_\ell(n) \in F_B$

$$M_{Bi}(\underline{c}\ell n) = 0 \quad \text{for all } i \in N,\ i \neq n .$$

Furthermore(*)

---

(*) $\phi$ is the empty set, and the circuit $C_\ell(n)$ has been defined in Definition 4.2.2

$$C_\ell(n) \cap \bar{S}^1 = \phi \Rightarrow \bar{P}_B(\underline{c}\ell n) = 0 \qquad \text{all } f_\ell(n) \in F_{BA}$$

$$\left.\begin{aligned} C_\ell(n) \cap S^1 = \phi &\Rightarrow P_{\bar{B}}(\underline{c}\ell n) = 0 \\ C_\ell(n) \cap \bar{S}^1 = \phi &\Rightarrow \bar{P}_{\bar{B}}(\underline{c}\ell n) = 0 \end{aligned}\right\} \quad \text{all } f_\ell(n) \in F_{\bar{B}}$$

Corollary L4.3.2

For $f_\ell(n) \in F_{\bar{B}}$

$$C_\ell(n) \cap S^1 = \phi \Rightarrow \begin{cases} \bar{P}(\underline{c}\ell n) = \bar{P}_{\bar{B}}(\underline{c}\ell n) \\ M_i(\underline{c}\ell n) = \begin{cases} M_{\bar{B}i}(\underline{c}\ell n) = 0 & \text{all } i \in N,\ i \neq n \\ M_{\bar{B}n}(\underline{c}\ell n) & i = n \end{cases} \end{cases}$$

□

If, in a similar way, $D(\underline{c}in)$ is the column corresponding to commodity (i,n) of matrix D, we can write

Lemma 4.3.3

If T(i,n) is the path from i to n on the spanning tree T(n), then

$$\left.\begin{aligned} T(i,n) \cap S^1 = \phi &\Rightarrow J_n(\underline{c}in) = 0 \\ T(i,n) \cap \bar{S}^1 = \phi &\Rightarrow \bar{J}_n(\underline{c}in) = 0 \end{aligned}\right\} \text{ all } i \in N,\ i \neq n .$$

Corollary L4.3.3

For all commodities such that $T(i,n) \cap S^1 = \phi$ we have

$$V_n(\underline{c}in) = 0$$

$$\bar{V}_n(\underline{c}in) = -\bar{J}_n(\underline{c}in)$$

$$U_{jn}(\underline{c}in) = 0 \qquad \text{all } j \in N,\ j \neq n$$

$$U_{nn}(\underline{c}in) = E_{Tn}^{-1}(\underline{c}in)$$

where the V and U matrices were defined in (4.2.26).

□

Note that Lemma 4.3.1 is included in Lemmas 4.3.2 and 4.3.3; nevertheless is worthwhile to have Lemma 4.3.1 as an independent result.

Applying some of these results to Theorem 4.2.8 we immediately obtain

Lemma 4.3.4

$$\ell \in \bar{S}^1 \Rightarrow z_\ell(n) = \sigma_{t(\ell)}(n) - \sigma_{o(\ell)}(n) \tag{4.3.5}$$

$$n \in N(\bar{I}) \Rightarrow z_\ell(n) = \begin{cases} -\pi_\ell & \ell \in S^1 \\ 0 & \ell \in \bar{S}^1 \end{cases} \tag{4.3.6}$$

$$\sigma_{t(\ell)}(n) \geq \sigma_{o(\ell)}(n) + \pi_\ell \quad \text{all } \ell, n \tag{4.3.7}$$

$$\ell \in \bar{S}^1 \text{ and } \sigma_{t(\ell)}(n) = 0 \Rightarrow \sigma_{o(\ell)}(n) = 0 \tag{4.3.8}$$

$$z_\ell(n) = 0 \Rightarrow \sigma_{t(\ell)}(n) = \sigma_{o(\ell)}(n) + \pi_\ell, \quad \text{all } \ell, n \tag{4.3.9}$$

□

With the help of figure 4.3.1 note that

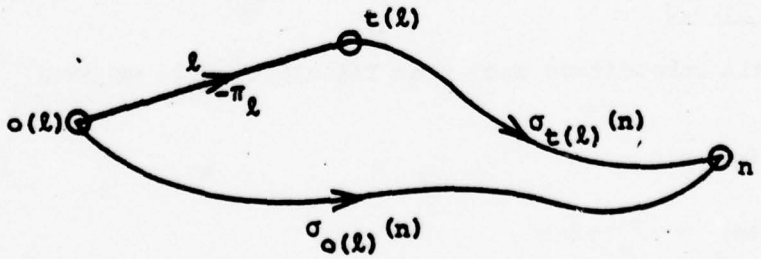



Fig. 4.3.1

equation (4.3.9) means that if both paths going to n have the same length, then we can use link $\ell$ to carry commodity $(o(\ell), n)$ and, therefore, make $f_\ell(n)$ nonzero and basic. This, of course, agrees with the interpretation that, for a nonbasic variable, $z_\ell(n) = 0$ means a non-unique solution, where any of the variables with zero cost can be made basic.

A slightly different, although totally equivalent, form of $z_\ell(n)$ can be obtained if we observe that from (4.2.34) and (4.2.27)

$$z(F_{\bar{B}}) = -\lambda A^1(F_{\bar{B}}) = -\gamma^1 B^{-1} A^1(F_{\bar{B}}) = \\ = -\gamma^1 \begin{bmatrix} P \\ \hline \bar{P} \\ \hline M \end{bmatrix} = -P(\underline{r}1) \tag{4.3.10}$$

where $P(\underline{r}1)$ means the first row of matrix P. On the other hand from (4.2.23) we can write

$$P(\underline{r}1) = Q^{-1}(\underline{r}1) P_{\bar{B}} = \pi(S^1) P_{\bar{B}} \tag{4.3.11}$$

But using the definition of $P_{\bar{B}}$ given in Theorem 4.2.2 we have

Lemma 4.3.5

$$z_\ell(n) = -\pi(S^1) C_\ell^1(n) \qquad \text{all } f_\ell(n) \in F_{\bar{B}}$$
$$= \sum_{a \in C_\ell^{1-}(n)} \pi_a - \sum_{a \in C_\ell^{1+}(n)} \pi_a$$

where $C_\ell^{1-}(n)$ is the set of links of $C_\ell^1(n)$ having negative direction and $C_\ell^{1+}(n)$ the set of links on the positive direction.

□

Comparing this result with Theorem 4.2.8 we clearly have

$$-\sum_{a \in C_\ell^{1+}(n)} \pi_a = \sigma_{t(\ell)}(n) - \pi_\ell \qquad (4.3.12)$$

$$\sum_{a \in C_\ell^{1-}(n)} \pi_a = -\sigma_{o(\ell)}(n)$$

and we can interpret the results of Lemma 4.3.4 in terms of empty or non-empty reduced circuits. In particular the following result gives a necessary and sufficient condition for a flow variable to have a zero partial cost; in this sense note that (4.3.9) gives a sufficient condition.

Theorem 4.3.1

A nonbasic flow variable $f_\ell(n)$ will have a zero partial cost at an optimal point of (LP1) iff either the saturated, or reduced, circuit $C_\ell^1(n) = \phi$, or else if it can be written as a linear combination of the saturated circuits of matrix $P_B$, that is, of flows in $F_{BA}$.

Proof:

From (4.3.10) and (4.3.11)

$$z(F_{\bar{B}}) = -\pi(S^1)P_{\bar{B}}$$

and using Corollary 4.2.2

$$z_\ell(n) = -\pi(S^1)C_\ell^1(n) \qquad \text{all } f_\ell(n) \in F_{\bar{B}} \qquad (4.3.13)$$

On the other hand from Theorem 4.2.7 and the fact that $QQ^{-1} = Q^{-1}Q = I$, we have

$$\pi(S^1)[-c^1|P_B] = (1,0,\ldots 0)$$

so

$$\pi(S^1)P_B = 0 \tag{4.3.14}$$

where $\pi(S^1) < 0$ and rank $P_B = q - 1$.

But from (4.3.13)

$$z_\ell(n) = 0 \Rightarrow \pi(S^1)C_\ell^1(n) = 0 \tag{4.3.15}$$

Clearly $C_\ell^1(n) = \phi$ satisfies (4.3.15). Suppose $C_\ell^1(n) \neq \phi$; then from (4.3.14) and (4.3.15)

$$\pi(S_1)\ [P_B \mid C_\ell^1(n)] = 0 \tag{4.3.16}$$

where $[P_B|C_\ell^1(n)]$ is a qxq square matrix. But since $\pi(S^1) \neq 0$, in order for (4.3.16) to be true we need

$$\det[P_B|C_\ell^1(n)] = 0 \tag{4.3.17}$$

But since rank $P_B = q-1$, (4.3.17) will be true if

$$C_\ell^1(n) \in \text{Space spanned by columns of } P_B \tag{4.3.18}$$

This proves the "only if" statement.

To prove in the other direction observe that if $C_\ell^1(n)$ is a linear combination of the columns of $P_B$, we can write

$$C_\ell^1(n) = P_B\mu \tag{4.3.19}$$

where $\mu$ is a non-null q-1 vector. Multiplying on the left by $\pi(S^1)$ and using (4.3.14)

$$\pi(S^1)C_\ell^1(n) = \pi(S^1)P_B\mu = 0$$

so, from (4.3.13), $z_{\ell}(n) = 0$ .

Q.E.D.

Finally applying Lemmas 4.3.1-3 to Theorem 4.2.9 we have

Lemma 4.3.6

If $C(F_{BA})$ is the set of circuits for variables in $F_{BA}$ than

$$\ell \notin C(F_{BA}) \Rightarrow \bar{P}_B(\underline{r}\ell) = 0 \Rightarrow f_{\ell} = f_{\ell T} \Rightarrow S_{\ell} = \alpha^1 c_{\ell} - f_{\ell T}, \text{all } \ell \in \bar{S}^1$$

$$n \in \bar{N}^1 \Rightarrow f_{BT}(n) = b_{IT}(n)$$

Furthermore for $arc(i,\cdot) \in T(n)$

$$arc(i,\cdot) \notin C(F_{BA}) \cap T(n) \Rightarrow M_{Bn}(\underline{r}i) = 0 \Rightarrow f_{BT}(i,n) = b_{IT}(i,n)$$

□

From the analysis done in this and previous sections we have seen that all data in the updated constraint matrix and optimal basis of (LP1) can be easily obtained knowing the saturated links and the alternate paths. They are obtained via graph theoretic means, except for the inversion of a matrix Q of $|S^1| \times |S^1|$ dimension. Since except for this matrix the remaining matrices are graph-related $(0,\pm 1)$ matrices, all the computations involving these matrices consist of just additions and subtractions. Furthermore, we have also seen how to exploit the structure of the graph to easily detect most zero entries of any of the matrices, and we also saw how to compute isolated columns, rows and elements as we need them, without having to haul the whole information tableau throughout the process. Now, before moving to next sections, in which we will make a closer analysis of the saturation

matrix Q, we introduce some additional concepts that will contribute to further simplification of previous computations.

Definition 4.3.1

We will say that commodity m is a subcommodity of commodity n if arc(m,n) is the only arc terminating in node n.

Note that whenever there exists a subcommodity m having more than one alternate path for traffic (i,m) and, furthermore, $r_i(n) > 0$ then the optimal solution of (LP1), or even of the global problem, will never be unique because we have infinite ways to apportion commodities (i,m) and (i,n) among the allowable paths without modifying the saturation level of any link. To avoid this nonuniqueness we can eliminate commodity n, and therefore all variables $f_\ell(n)$, as a valid commodity and adjoin r(n) to r(m). This rule will, in addition, reduce the dimension of some of the program matrices.

But this is not the only way of eliminating elements of F.

Lemma 4.3.7

Let $\Lambda$ be a circuit, not necessarily directed, and let $\Lambda(\bar{\ell})$ be the chain resulting from taking link $\ell$ out of the circuit; then if

$$\Lambda(\bar{\ell}_1) \subset T(m)$$
$$\Lambda(\bar{\ell}_2) \subset T(n) ,$$

where $\ell_1$ and $\ell_2$ can be the same or different links, we have that the columns corresponding to $f_{\ell 1}(m)$ and $f_{\ell 2}(n)$ will be identical (or opposite if $\ell 1$ and $\ell 2$ are in different directions) in (4.2.8) and therefore only one of them can be made basic.

Proof

From Theorems 4.2.2 and 4.2.3 and the fact that the columns of $F_{BA}$ have to be linearly independent.

Q.E.D.

Lemma 4.3.8

If $\Lambda$, from Lemma 4.3.10, is a cycle then neither $f_{\ell 1}(m)$ nor $f_{\ell 2}(n)$ can be basics.

Proof

In this case, all the elements of the columns are 0 or +1 and therefore can never generate a negative cost. Q.E.D.

We shall see in Chapter V and Appendix III examples of how these two lemmas provide some additional ways to reduce the dimensionality of the problem.

## 4.4 Properties of Saturation Matrix Q

As we have seen in section 4.2, Q is a square matrix with structure

$$Q = \left[\begin{array}{c|c} -c_1 & \\ \vdots & P_B \\ -c_q & \end{array}\right] = [-c^1 \mid P_B]$$

where

$$P_B(i,j) = \begin{cases} 1 \\ 0 \\ -1 \end{cases} \quad \text{for all } i,j \qquad (4.4.1)$$

$c_\ell > 0$ all $\ell$

Definition 4.4.1

We will say that a matrix with the structure of (4.4.1) is a valid Q-matrix if there exists a network and requirement matrix such that an

optimal solution of (LP1) for this network gives a Q-matrix that is identical to the given matrix.

To fully characterize a valid Q-matrix we have to study its inverse $Q^{-1}$. From Theorem 4.2.7 $Q^{-1}$ will have the structure

$$Q^{-1} = \left[\begin{array}{c} \pi^1 \\ \hline \Omega \end{array}\right] \tag{4.4.2}$$

where $\Omega$ is a $(q-1)xq$ full rank matrix and where, for simplicity of notation,

$$\pi^1 \triangleq \pi(S^1)$$

Lemma 4.4.1

A necessary condition for a square matrix Q, with the structure in (4.4.1), to be the valid saturation matrix of an optimal stable point is, beside being non-singular, that each column of Q must have at least a 1 and a -1 as components.

Proof:

From

$$Q^{-1}Q = I \tag{4.4.3}$$

we have

$$\pi^1 P_B(\underline{c}\ell n) = \sum_{a \in S^1} \pi_a P_B(a,\ell n) = 0 \tag{4.4.4}$$

Since at a stable point $\pi_a < 0$, all $a \in S^1$, (4.4.4) cannot be true if all $P_B(a,\ell n)$, $a \in S^1$, have the same sign. On the other hand $P_B(\underline{c}\ell n)$ must be nonzero or Q would be singular.

Q.E.D.

Note that from (4.4.3) we also have

$$-\pi^1 c^1 = -\sum_{\ell \in S^1} \pi_\ell c_\ell = 1$$

which is another way of proving (3.4.1).

Corollary 4.4.1

The reduced circuits appearing in $P_B$ can neither be cycles nor empty circuits.

□

Suppose that we have a matrix Q that, in addition to satisfying the necessary condition of Lemma 4.4.1, has negative elements in the first row of $Q^{-1}$. Then, each row of matrix Q can be associated with a link and each column with a reduced circuit. From these links and circuits we are able to construct a network. The synthesis procedure introduced in Appendix II will produce a network having an optimal solution with a saturation matrix identical to the given matrix Q. Thus, we conclude

Lemma 4.4.2

A matrix Q, that satisfies Lemma 4.4.1, is a valid saturation matrix if and only if all the elements of the first row of $Q^{-1}$ are negative

Theorem 4.4.1

A necessary and sufficient condition for a matrix with the structure of (4.4.1) to be a valid Q-matrix is

$$\text{sgn}\{\det[P_B(\underline{r\ell})]\} = -\text{sgn}\{\det[P_B(\underline{r(\ell+1)})]\} \qquad (4.4.5)$$

$$\det[P_B(\underline{r\ell})] \neq 0 , \qquad \ell \in S^1$$

where "sgn", means signum and $P_B(\overline{r\ell})$ is matrix $P_B$ without row $\ell$.

Proof:

We have to prove that (4.4.5) is equivalent to Lemma 4.4.1. Assume first that $Q^{-1}$ exists, then from (4.4.1), (4.4.2) and applying Cramer rule.

$$\pi_\ell = \frac{\det[P_B(\overline{r\ell})}{\det Q}(-1)^{\ell+1} \qquad \ell \in S^1 \tag{4.4.6}$$

These elements will be negative iff det $Q \neq 0$ and, whenever

$$\det Q > 0 \quad \text{then} \begin{cases} \det[P_B(\overline{r\ell})] > 0, & \text{all } \ell = \dot{2} \\ \det[P_B(\overline{r\ell})] < 0, & \text{all } \ell \neq \dot{2} \end{cases} \tag{4.4.7}$$

and whenever

$$\det Q < 0 \quad \text{then} \begin{cases} \det[P_B(\overline{r\ell})] < 0, & \text{all } \ell = \dot{2} \\ \det[P_B(\overline{r\ell})] > 0 & \text{all } \ell \neq \dot{2} \end{cases} \tag{4.4.8}$$

where $\ell = \dot{2}$ means $\ell$ even. Furthermore using a minors expansion

$$\begin{aligned} \det Q &= \sum_{\ell=1}^{q} (-c_\ell) \det[P_B(\overline{r\ell})](-1)^{\ell+1} = \sum_{\ell=1}^{q} c_\ell \det[P_B(\overline{r\ell})](-1)^{\ell} \\ &= \sum_{\ell=\dot{2}} c_\ell \det[P_B(\overline{r\ell})] - \sum_{\ell\neq\dot{2}} c_\ell \det[P_B(\overline{r\ell})] \end{aligned} \tag{4.4.9}$$

Suppose

$$\begin{aligned} &\det[P_B(\overline{r\ell})] > 0 \qquad \text{all } \ell \neq \dot{2} \\ &\det[P_B(\overline{r\ell})] < 0 \qquad \text{all } \ell = \dot{2} \end{aligned} \tag{4.4.10}$$

then from (4.4.9) $\det Q < 0$ and (4.4.8) is satisfied. Suppose, on the other hand,

$$\det[P_B(\overline{r\ell})] < 0 \qquad \text{all } \ell \neq \dot{2}$$
$$\det[P_B(\overline{r\ell})] > 0 \qquad \text{all } \ell = \dot{2} \tag{4.4.11}$$

then from (4.4.9) det Q > 0 and (4.4.7) is satisfied. Thus, no matter whether (4.4.10) or (4.4.11) are true, det Q ≠ 0 and (4.4.7) - (4.4.8) are satisfied.

Q.E.D.

An immediate consequence of (4.4.6), (4.4.9) and (3.4.8) is

Lemma 4.4.3

Assuming the links of $S^1$ ordered from 1 to $|S^1|$, the dual variables and link weights are related to the determinant of matrix Q and submatrices of $P_B$ by the following relations

$$\frac{\pi_\ell}{\pi_a} = \frac{\theta_\ell}{\theta_a} = \frac{\det[P_B(\overline{r\ell})]}{\det[P_B(\overline{ra})]}(-1)^{\ell-a} \qquad \ell, a \in S^1 \tag{4.4.12}$$

$$\det Q = D^1 \frac{\det[P_B(\overline{r\ell})]}{\theta_\ell}(-1)^\ell \qquad \ell \in S^1$$

where $D^1$ is the total equivalent capacity of $S^1$, defined in (3.4.8).

□

Theorem 4.4.2

If $Q^{-1}(\underline{r}i)$ is the i-th row of $Q^{-1}$

$$Q^{-1}(\underline{r}i)c^1 = \sum_{\ell=1}^{q} Q^{-1}(i,\ell)c_\ell = 0 \tag{4.4.13}$$

$$Q^{-1}(\underline{r}i)Q(\underline{c}j) = Q^{-1}(\underline{r}i)c^1(j) = 0 \qquad \text{all } i \neq j \tag{4.4.14}$$

where $C^1(j)$ is the reduced circuit of variable j.

$$Q^{-1}(\underline{r}i)C^1(i) = 1 \tag{4.4.15}$$

Proof

Obvious from $Q^{-1}Q = I$

Corollary L4.4.2

No row of $Q^{-1}$ except the first can have all its elements of the same sign. Furthermore if $c_\ell = c_a$, all a, $\ell \in S^1$ the sum of the elements of any row of $Q^{-1}$ must be equal to zero.

□

## 4.5 The Capacity Normalized Network

We saw in section 3.4 that, given $S^1$ and the set of allowable paths from an optimal solution of (LP1), we are able to work, for most purposes, with a capacity normalized matrix. As we shall see in Section 4.6, this can be useful in simplifying computations and reducing round-off errors. The aim of this section is to relate the dual variables and Q-matrices for the general and normalized network. It is clear that, except for the capacity vector, the remaining matrices of (4.2.8) and (4.2.15) will be identical in both networks. Furthermore, from the discussion of Section 3.4, it is also easy to see that the link weights will be identical in both networks. Denoting by tilde "^' the variables corresponding to the normalized network we can write, see (3.4.8),

$$\begin{aligned} D^1 &= \sum_{\ell \in S^1} \theta_\ell c_\ell \\ \hat{D}^1 &= \sum_{\ell \in S^1} \theta_\ell \end{aligned} \tag{4.5.1}$$

Moreover, from (3.4.1),

$$\sum_{\ell \in S^1} \pi_\ell c_\ell = -1$$

$$\sum_{\ell \in S^1} \hat{\pi}_\ell = -1 \qquad (4.5.2)$$

and also

$$\pi_\ell = \frac{\theta_\ell}{D^1}$$

$$\hat{\pi}_\ell = \frac{\theta_\ell}{\hat{D}^1} \qquad \text{all } \ell \in S^1 \qquad (4.5.3)$$

Since most sums of this section will be over the set $S^1$, we will omit its specific citation unless we feel it necessary to avoid confusion. First we present some relations between the actual and normalized values of the dual variables that will be used later.

Lemma 4.5.1

i) $\pi_\ell D^1 = \hat{\pi}_\ell \hat{D}^1 \qquad \text{all } \ell \in S^1 \qquad (4.5.4)$

ii) $\pi_\ell = -\left(\sum_\ell \pi_\ell\right)\hat{\pi}_\ell \triangleq -\varepsilon\, \hat{\pi}_\ell$

$$\hat{\pi}_\ell = -\left(\sum_\ell \hat{\pi}_\ell c_\ell\right)\pi_\ell \triangleq -\hat{\varepsilon}\, \pi_\ell \qquad (4.5.5)$$

iii) $\left(\sum_\ell \pi_\ell\right)\left(\sum_\ell \hat{\pi}_\ell c_\ell\right) \triangleq \varepsilon\, \hat{\varepsilon} = 1 \qquad (4.5.6)$

Proof:

i) Obvious from (4.5.3)

ii) From (4.5.4), (4.5.1) and (4.5.3)

$$\pi_\ell = \frac{\hat{D}^1}{D^1}\,\hat{\pi}_\ell = \frac{\Sigma\theta_\ell}{D^1}\,\hat{\pi}_\ell = -\left(\sum\pi_\ell\right)\hat{\pi}_\ell \ .$$

Similarly for the other relation

iii) Obvious from (4.5.5)

Q.E.D.

The structures of $Q$ and $\hat{Q}$ are

$$Q = \left[\begin{array}{c|c} -c_1 & \\ \vdots & P_B \\ -c_q & \end{array}\right] \quad ; \qquad \hat{Q} = \left[\begin{array}{c|c} -1 & \\ \vdots & P_B \\ -1 & \end{array}\right] \tag{4.5.7}$$

These two matrices are related by the following theorem

Theorem 4.5.1

The inverse of the saturation matrices of the general and normalized problem satisfy

$$Q^{-1} = \frac{1}{\hat{\varepsilon}}\ \hat{Q}^{-1}\,\hat{\psi} \tag{4.5.8}$$

$$\hat{Q}^{-1} = \frac{1}{\varepsilon}\,Q^{-1}\,\psi$$

where

$$\psi\hat{\psi} = \hat{\psi}\psi = I$$

$$\hat{\psi} = \hat{\varepsilon}\ I + \hat{\psi}_0 \tag{4.5.9}$$

$$\psi = \varepsilon\ I + \psi_0$$

and where

$$\hat{\psi}_0 = \begin{bmatrix} 1-c_1 \\ \cdot \\ \cdot \\ \cdot \\ 1-c_q \end{bmatrix} [\hat{\pi}_1,\ldots,\hat{\pi}_q] \triangleq (\underline{1}-c^1)\hat{\pi}^1$$

$$\psi_0 = \begin{bmatrix} c_1^{-1} \\ \cdot \\ \cdot \\ \cdot \\ c_q^{-1} \end{bmatrix} [\pi_1,\ldots,\pi_q] \triangleq (c^1-\underline{1})\pi^1 \qquad (4.5.10)$$

$\varepsilon$ and $\hat{\varepsilon}$ are given in (4.5.5).

Proof:

Note that $Q$ from (4.5.7) can be written

$$Q = \left[\begin{array}{c|c} -1 & \\ \cdot & \\ \cdot & P_B \\ \cdot & \\ -1 & \end{array}\right] + \left[\begin{array}{c|c} 1-c_1 & \\ \cdot & \\ \cdot & 0 \\ \cdot & \\ 1-c_q & \end{array}\right] = \hat{Q} + \begin{bmatrix} 1-c_1 \\ \cdot \\ \cdot \\ \cdot \\ 1-c_q \end{bmatrix} [1,0,\ldots,0] \qquad (4.5.11)$$

Applying to (4.5.11) Theorem AI.3 of Appendix I yields

$$Q^{-1} = \hat{Q}^{-1} + \frac{\hat{Q}^{-1}\hat{\psi}_0}{\sum \hat{\pi}_\ell c_\ell} = \frac{\hat{Q}^{-1}}{\hat{\varepsilon}} [\hat{\varepsilon} I + \hat{\psi}_0]$$

Similarly

$$\hat{Q} = Q + \begin{bmatrix} c_1^{-1} \\ \cdot \\ \cdot \\ \cdot \\ c_q^{-1} \end{bmatrix} [1,0,\ldots,0] \qquad (4.5.12)$$

and applying again Theorem AI.3 yields

$$\hat{Q}^{-1} = Q^{-1} + \frac{Q^{-1}\psi_0}{\sum \pi_\ell} = \frac{Q^{-1}}{\varepsilon} [\varepsilon I + \psi_0]$$

These are the expressions of (4.5.8). To prove $\psi\hat{\psi} = \hat{\psi}\psi = I$ we have just to multiply the two equations of (4.5.8) and use (4.5.6).

Q.E.D.

Thus we see that both inverses are related through simple, structured, relations. Furthermore, these relations allow us to go in either direction (from $Q^{-1}$ to $\hat{Q}^{-1}$ or vice versa) without inverting any matrix. These facts can be very helpful in simplying computation.

Since $Q^{-1}$ and $\hat{Q}^{-1}$ have the structure below

$$Q^{-1} = \left[\frac{\pi^1}{\Omega}\right] \quad ; \quad \hat{Q}^{-1} = \left[\frac{\hat{\pi}^1}{\hat{\Omega}}\right] \tag{4.5.13}$$

we should be able to obtain from Theorem 4.5.2 the results of Theorem 4.5.1. In fact, it is a very simple exercise to verify that this is true.

From previous results it is also very simple to prove

Lemma 4.5.2

$$\det \hat{\psi} = -\hat{\varepsilon}^{(q-1)}$$

$$\det \psi = -\varepsilon^{(q-1)}$$

Lemma 4.5.3

$$\det Q = [\det (Q^{-1})]^{-1} = -\hat{\varepsilon} \det \hat{Q}$$

$$\det \hat{Q} = [\det (\hat{Q}^{-1})]^{-1} = -\varepsilon \det Q$$

Lemma 4.5.4

$$\frac{\det Q}{\det \hat{Q}} = \frac{D^1}{\hat{D}^1}$$

□

Now, that the relations between the general and normalized problem are known, we will be able to work with a normalized network without loss of generality. The differences between these two problems are reduced even further if we work with link weights instead of link dual variables, since they are the same in both problems. Furthermore, we can write a theorem, entirely analogous to Theorem 4.5.2, where the dual variables are replaced by weights. For this purpose define

$$\begin{aligned} \theta_\theta^{-1} &\triangleq D^1 Q^{-1} \\ \hat{Q}_\theta^{-1} &\triangleq \hat{D}^1 \hat{Q}^{-1} \\ \psi_\theta &\triangleq D^1 \psi \\ \hat{\psi}_\theta &\triangleq \hat{D}^1 \hat{\psi} \end{aligned} \tag{4.5.14}$$

where clearly, from (4.5.9), (4.5.10) and Theorem 4.5.1,

$$\begin{aligned} \psi_\theta &= D^1 \varepsilon I + D^1 \psi_0 = -\hat{D}^1 I + \psi_{0\theta} \\ \hat{\psi}_\theta &= -D^1 I + \hat{\psi}_{0\theta} \end{aligned} \tag{4.5.15}$$

and

$$\psi_\theta = (c^1 - \underline{1})\theta^1 = -\hat{\psi}_{0\theta} \tag{4.5.16}$$

so

$$\psi_\theta + \hat{\psi}_\theta = -(D^1 + \hat{D}^1) I \triangleq -DI$$

Furthermore from $\hat{\psi}\psi = \psi\hat{\psi} = I$ we have

$$\hat{\psi}_\theta \psi_\theta = \psi_\theta \hat{\psi}_\theta = D^1 \hat{D}^1 I$$

Also, replacing (4.5.14) in (4.5.8) and using Theorem 4.5.1 yields

$$Q_\theta^{-1} = \frac{D^1}{\varepsilon} \hat{Q}^{-1}\hat{\psi} = -\hat{D}^1\hat{Q}^{-1}\hat{\psi} = -\hat{Q}^{-1}\hat{\psi}_\theta = -\hat{Q}_\theta^{-1}\hat{\psi}$$

$$\hat{Q}_\theta^{-1} = -Q^{-1}\psi_\theta = -Q_\theta^{-1}\psi$$

Using Lemmas 4.5.1-3 we summarize these results as

Theorem 4.5.2

The inverse of the weighted saturation matrices, defined in (4.5.14), of the general and normalized problem satisfy

$$\begin{aligned} \hat{Q}_\theta^{-1} &= -\hat{Q}^{-1}\hat{\psi}_\theta = -\hat{Q}_\theta^{-1}\hat{\psi} \\ \hat{Q}_\theta^{-1} &= -Q^{-1}\psi_\theta = -Q_\theta^{-1}\psi \end{aligned} \qquad (4.5.17)$$

where $\psi_\theta$ and $\hat{\psi}_\theta$ are integer matrices that satisfy

$$\begin{aligned} \psi_\theta + \hat{\psi}_\theta &= -(D^1+\hat{D}^1)I \\ \psi_\theta\hat{\psi}_\theta &= \hat{\psi}_\theta\psi_\theta = D^1\hat{D}^1 I \end{aligned} \qquad (4.5.18)$$

and

$$\begin{aligned} \det \hat{\psi}_\theta &= \hat{D}^1(D^1)^{q-1} \\ \det \psi_\theta &= D^1(\hat{D}^1)^{q-1} \end{aligned} \qquad (4.5.19)$$

Lemma 4.5.5

The elements of $Q^{-1}$, $Q_\theta^{-1}$, $\hat{Q}^{-1}$ and $\hat{Q}_\theta^{-1}$ are related as follows

$$Q^{-1}(i,j) = \hat{Q}^{-1}(i,j) - \frac{\pi_j}{\hat{\varepsilon}} \sum_k \hat{Q}^{-1}(i,k)c_k$$

$$\hat{Q}^{-1}(i,j) = Q^{-1}(i,j) - \frac{\pi_j}{\varepsilon}\sum_k Q^{-1}(i,k)$$

$$Q_\theta^{-1}(i,j) = -\frac{D^1}{\hat{D}^1}\hat{Q}_\theta^{-1}(i,j) + \frac{\hat{\theta}_j}{\hat{D}^1}\sum_k \hat{Q}_\theta^{-1}(i,j)c_k \qquad (4.5.20)$$

$$\hat{Q}_\theta^{-1}(i,j) = -\frac{\hat{D}^1}{D^1}Q_\theta^{-1}(i,j) + \frac{\theta_j}{D^1}\sum_k Q_\theta^{-1}(i,k)$$

**Proof:**

Since all of these results are obtained in a similar way we will only prove the first equation. From (4.5.8), the ith row of $Q^{-1}$ is

$$Q^{-1}(\underline{r}i) = \frac{1}{\hat{\varepsilon}}Q^{-1}(\underline{r}i)\hat{\psi}$$

and the (i,j)-th element

$$Q^{-1}(i,j) = \frac{1}{\hat{\varepsilon}}\hat{Q}^{-1}(\underline{r}i)\hat{\psi}(\underline{c}j) = \frac{1}{\hat{\varepsilon}}\hat{Q}^{-1}(\underline{r}i)\begin{bmatrix}(1-c_1)\hat{\pi}_j\\ \vdots\\ (1-c_j)\hat{\pi}_j + \hat{\varepsilon}\\ \vdots\\ (1-c_q)\hat{\pi}_j\end{bmatrix}$$

$$= \frac{1}{\hat{\varepsilon}}\hat{\pi}_j\sum_k \hat{Q}^{-1}(i,k)(1-c_k) + \hat{Q}^{-1}(i,j)$$

But, since from (4.4.17), $\sum_k \hat{Q}^{-1}(i,k) = 0$ we have

$$Q^{-1}(i,j) = \hat{Q}^{-1}(i,j) - \frac{\hat{\pi}_j}{\hat{\varepsilon}}\sum_k \hat{Q}^{-1}(i,k)c_k$$

Q.E.D.

## 4.6 The Inversion of the Saturation Matrix

The special structure of the saturation matrix suggests that many properties of this matrix could be used to simplify the task of inverting Q. With no intention of being exhaustive, we present here some inversion methods that make use of simplifying properties

and techniques. A simplification that can always be made, and that among other simplications could reduce round-off errors, is to perform the inverse operation on the normalized matrix $\hat{Q}$ and use, then, the results of section 4.5.

Method I (Gauss-Jordan)

This method consists on the application of Theorem AI.1 of Appendix I, to invert Q. As an example we consider the inversion of

$$Q = \begin{array}{c} \begin{matrix} \alpha^1 & & & F_{BA} \end{matrix} \\ \begin{bmatrix} -1 & -1 & 0 & 1 \\ -1 & 1 & -1 & 1 \\ -1 & 0 & 0 & -1 \\ -1 & 1 & 1 & 1 \end{bmatrix} \end{array} \begin{matrix} \ell_1 \\ \ell_2 \\ \ell_3 \\ \ell_4 \end{matrix}$$

$$[Q|I] = \left[\begin{array}{cccc:cccc} -1 & -1 & 0 & 1 & 1 & 0 & 0 & 0 \\ -1 & 1 & -1 & 1 & 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & -1 & 0 & 0 & 1 & 0 \\ -1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{cccc:cccc} -2 & -1 & 0 & 0 & 1 & 0 & 1 & 0 \\ -2 & 1 & -1 & 0 & 0 & 1 & 1 & 0 \\ -1 & 0 & 0 & -1 & 0 & 0 & 1 & 0 \\ -2 & 1 & 1 & 0 & 0 & 0 & 1 & 1 \end{array}\right]$$

$$\sim \left[\begin{array}{cccc:cccc} -2 & -1 & 0 & 0 & 1 & 0 & 1 & 0 \\ -4 & 0 & -1 & 0 & 1 & 1 & 2 & 0 \\ -1 & 0 & 0 & -1 & 0 & 0 & 1 & 0 \\ -8 & 0 & 0 & 0 & 2 & 1 & 4 & 1 \end{array}\right] \sim \left[\begin{array}{cccc:cccc} 8 & 0 & 0 & 0 & -2 & -1 & -4 & -1 \\ 0 & -4 & 0 & 0 & 2 & -1 & 0 & -1 \\ 0 & 0 & -2 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & -8 & -2 & -1 & 4 & -1 \end{array}\right]$$

$$\sim \left[\begin{array}{cccc:cccc} 8 & 0 & 0 & 0 & -2 & -1 & -4 & -1 \\ 0 & 8 & 0 & 0 & -4 & 2 & 0 & 2 \\ 0 & 0 & 8 & 0 & 0 & -4 & 0 & 4 \\ 0 & 0 & 0 & 8 & 2 & 1 & -4 & 1 \end{array}\right] = [8I|Q_\theta^{-1}] = 8[I|Q^{-1}]$$

where obviously $\hat{D}^1 = D^1 = 8$, and where many of the properties of Sections 4.4 and 4.5 can be seen inspecting $Q_\theta^{-1}$ and $Q^{-1}$. It is clear that due to the structure of Q this method can be very attractive in many cases.

Aside from the possible simplification that this procedure may introduce in the computation of $Q^{-1}$, this method can be also useful to provide some additional insight into the relations among Q, $Q^{-1}$ and the circuits of the network. One obvious immediate consequence is the complete equivalence that exists between the canonic equations and the Q-matrix, see also Section 3.6. We can easily see that finding the canonic ratio in Section 3.6 is entirely equivalent to performing row operations on [Q I] to obtain a row where only the first one of the first $|S^1|$ elements is nonzero. Similarly the process of finding the equivalent capacity of a flow variable is completely equivalent to the process of producing by row operations a row of zeroes except for the element corresponding to that variable. In fact, as we will see in Chapter VI when we study the dynamic of the first saturation problem, there exists a correspondence between the sensitivity factors of section 3.6 and the elements of $Q^{-1}$. Now we move to derive some of previously mentioned relations.

Theorem 4.6.1

If

$$\sum_{\ell \in S^1} \theta_\ell P_B(\underline{r}\ell) = 0 \qquad (4.6.1)$$

and the $\{\theta_\ell\}$ are positive integers with a greatest common divisor of 1 for all $\ell$, then $\theta_\ell$ is the weight of link $\ell$, for all $\ell \in S^1$.

Proof:

From Theorem AI.1 and the fact that the first row of $[D^1 I | Q_\theta^{-1}]$ is $[D^1 0 \ldots 0 \,|-\theta]$, that is, it has zeroes in all positions corresponding to matrix $P_B$.

Q.E.D.

Corollary 4.6.1

If all reduced circuits have the same number of links in the positive and negative directions then $\theta_\ell = 1$, all $\ell \in S^1$, because in that case

$$\sum_{\ell \in S^1} P_B(\underline{r}\ell) = 0$$

□

It follows that to obtain the link weights we have just to solve system (4.6.1). Since rank $P_B = |S^1| - 1$ this system of equations will have infinite solutions, as expected from Section 3.4. We will consider the smallest integer solution. Note that from the structure of $P_B$, system (4.6.1) will, in general, be very simple to solve. In many cases $P_B$ will have a triangular or quasi-triangular structure. For the previous example, for instance, we can reorder rows and columns of $P_B$ to give the system

$$\begin{bmatrix} 1 & 1 & 1 & -1 \\ 1 & 1 & -1 & 0 \\ 1 & -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \theta_{\ell 4} \\ \theta_{\ell 2} \\ \theta_{\ell 1} \\ \theta_{\ell 3} \end{bmatrix} = 0$$

and by inspection, conclude

$$\theta_{\ell 4} = \theta_{\ell 2}$$

$$\theta_{\ell 1} = \theta_{\ell 4} + \theta_{\ell 2} = 2\theta_{\ell 2}$$

$$\theta_{\ell 3} = \theta_{\ell 4} + \theta_{\ell 2} + \theta_{\ell 1} = 4\theta_{\ell 2}$$

This immediately gives the weights $\theta = [2 \quad 1 \quad 4 \quad 1]$. Once vector $\theta$ is known we can reduce by one the dimension of the system to be solved to obtain the remaining rows.

The advantage of the row transformation method is that, aside from the fact that $P_B$ or $P_B(\underline{\overline{ci}})$ could be written many times in quasi-triangular form, it is very easy to detect when two weights, or two variables of $Q^{-1}(\underline{ri})$, are equal in value, and thus we can reduce immediately the dimension of the system of equations to be solved. For this matter consider that two variables will be equal if

- they belong to a reduced circuit that only contains these two links or
- there exists a circuit that differs from another in $P_B$ only in the fact that two links have different directions.

Method II

Let $\nabla$ denote matrix $Q^T Q$, where here T means transpose. Using a well known result of matrix analysis we can write

$$Q^{-1} = (Q^T Q)^{-1} Q^T = \nabla^{-1} Q^T \qquad (4.6.4)$$

Therefore, if $\nabla$ has an easily invertible structure we can have some advantage by using (4.6.4). In the first place $\nabla$ is symmetric, and thus

so is $\nabla^{-1}$; therefore only half of the elements have to be computed. Furthermore, since the elements of $\nabla$ are the scalar product of two columns of Q, we have

Lemma 4.6.1

For a capacity normalized network

$\nabla(1,1) = |S^1| = q$

$\nabla(i,i)$ = # of links on the reduced circuit for the i-th variable (all $i \neq 1$)

$\nabla(1,i) = \nabla(i,1)$ = difference between the number of links on the negative and positive direction of the reduced circuit of the i-th variable (all $i \neq 1$)

$\nabla(i,j) = \nabla(j,i)$ = difference between the number of links that are common with the same sign and common with different sign on the i-th and j-th reduced circuits

Examples

i)

$$Q = \begin{bmatrix} -1 & 1 & 1 & -1 \\ -1 & 1 & -1 & 1 \\ -1 & -1 & 1 & 1 \\ -1 & -1 & -1 & -1 \end{bmatrix} ; \nabla = 4I ; \quad \nabla^{-1} = \frac{1}{4} I ; \quad Q^{-1} = \frac{1}{4} Q^T$$

ii)

$$Q = \begin{bmatrix} -1 & 1 & 1 & 0 \\ -1 & 1 & -1 & 0 \\ -1 & -1 & 1 & 1 \\ -1 & -1 & -1 & -1 \end{bmatrix} ; \nabla = \begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 4 & 2 \\ 0 & 0 & 2 & 2 \end{bmatrix} ; \nabla^{-1} = \frac{1}{4} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & -2 \\ 0 & 0 & -2 & 4 \end{bmatrix}$$

iii)

$$Q = \begin{bmatrix} -1 & 1 & 1 & 1 & 1 \\ -1 & 1 & -1 & -1 & -1 \\ -1 & -1 & 1 & -1 & -1 \\ -1 & -1 & -1 & 1 & -1 \\ -1 & -1 & -1 & -1 & 1 \end{bmatrix} ; \quad V = \begin{bmatrix} 5 & 1 & 1 & 1 & 1 \\ 1 & 5 & 1 & 1 & 1 \\ 1 & 1 & 5 & 1 & 1 \\ 1 & 1 & 1 & 5 & 1 \\ 1 & 1 & 1 & 1 & 5 \end{bmatrix} ;$$

$$V^{-1} = \frac{1}{45} \begin{bmatrix} 8 & 1 & 1 & 1 & 1 \\ 1 & 8 & 1 & 1 & 1 \\ 1 & 1 & 8 & 1 & 1 \\ 1 & 1 & 1 & 8 & 1 \\ 1 & 1 & 1 & 1 & 8 \end{bmatrix}$$

Method III

Suppose that matrix Q has some zero entires, i.e., suppose that not all staturated links are included in all reduced circuits, and that we can reorder the rows and columns of Q in such a way that we isolate a submatrix of zeroes (not necessarily including all the zeroes of the matrix Q). Consider, also, that this submatrix of zeroes is large enough to allow the existance of the following matrix, obtained from Q by moving the zero submatrix to the top right-hand corner and the column corresponding to variable $\alpha$ to a position immediately to the left of this matrix, as shown below

$$Q_\alpha = \left[\begin{array}{c|c|c} K & -c^K & 0 \\ \hline D & -c^G & \tilde{G} \end{array}\right] \triangleq \left[\begin{array}{c|c|c} K & -c^K & 0 \\ \hline D & & G \end{array}\right] \qquad (4.6.5)$$

Here K and G are kxk and gxg (k+g = q) square matrices, respectively, and $c^K$ and $c^G$ are the capacity vectors for links corresponding to the top and bottom partitions of $Q_\alpha$, respectively.

Consider the inverse $Q_\alpha^{-1}$ and partition it in similar blocks

$$Q_\alpha^{-1} = \left[\begin{array}{c|c} X & Y \\ \hline U & W \end{array}\right] \tag{4.6.6}$$

Note, from the structure of the matrix G, that we can consider it to be the Q-matrix of a subnetwork that could be obtained from the original network by deleting all links not involved with the requirements using the links of G. This new Q-matrix will have its own internal set of equivalent dual variables and link weights. Let $\pi^G$ denote the vector of inner dual variables for this subnetwork, and let G and K denote not only the submatrices of (4.6.5) but also the sets of saturated links in G and K; then we have

Theorem 4.6.2

The matrices defined in (4.6.1) and (4.6.2) are related as

$$\begin{aligned} X &= \frac{1}{1+\mu} K^{-1} \Theta \\ W &= \frac{1}{1+\mu} G^{-1} \Xi \\ Y &= h\, W(\underline{r}1) \\ U &= -WD_K \end{aligned} \tag{4.6.7}$$

where

$$\Xi = (1+\mu)I - v\pi^G$$

$$\Theta = (1+\mu)I - c^K\pi^G D_K \qquad (4.6.8)$$

$$v = \begin{bmatrix} v_1 \\ \cdot \\ \cdot \\ \cdot \\ v_g \end{bmatrix} = D_K c^K \quad ; \quad h = \begin{bmatrix} h_1 \\ \cdot \\ \cdot \\ \cdot \\ h_K \end{bmatrix} = K^{-1}c^K$$

$$D_K = DK^{-1} \qquad (4.6.9)$$

Here $\mu$ is the scalar

$$\mu = \pi^G v = \sum_{\ell \in G} \pi^G_\ell v_\ell$$

and $W(\underline{r}1)$ is the first row of matrix W and $\pi^G = G^{-1}(\underline{r}1)$.

Proof:

From Theorem AI.2 we have

$$X = (K - CG^{-1}D)^{-1}$$

$$W = (G - DK^{-1}C)^{-1} \qquad (4.6.10)$$

$$Y = -K^{-1}CW$$

$$U = -G^{-1}DX$$

where C is the upper right hand corner matrix of $\Omega_\alpha$; i.e.

$$C = [-c^K | 0] \ .$$

Manipulating

$$DK^{-1}C = DK^{-1}[-c^K|0] = [-DK^{-1}c^K|0] = [-v|0] = -v(10\ldots0)$$

and thus

$$W = [G + v(1\,0\,\ldots 0)]^{-1} \tag{4.6.11}$$

Applying Theorem AI.3 to (4.6.11)

$$W = G^{-1} - \frac{G^{-1}v\pi^G}{1+\pi^G v} = G^{-1}\left(I + \frac{v\pi^G}{1+\mu}\right) = \frac{1}{1+\mu}\,G^{-1}\Xi$$

Similarly, denoting $\tau \triangleq \pi^G D$, we have

$$CG^{-1}D = \left[-c^K \middle| 0\right]\left[\begin{array}{c}\pi^G \\ \hline \Omega^G\end{array}\right]D = -c^K\pi^G D = -c^K\tau$$

where $\tau$ is clearly a row vector. Hence

$$X = [K + c^K\tau]^{-1} \tag{4.6.12}$$

and applying Theorem AI.3 to (4.6.12) yields

$$X = K^{-1} - \frac{K^{-1}c^K\tau K^{-1}}{1+\tau K^{-1}c^K} = K^{-1}\left[I - \frac{c^K\pi^G D_K}{1+\mu}\right] = \frac{1}{1+\mu}\,K^{-1}\Xi$$

Finally for the last two expression of (4.6.10)

$$Y = K^{-1}[c^K|0]W = [K^{-1}c^K|0]W = hW(\underline{r}1)$$

$$U = -G^{-1}DK^{-1} + \frac{G^{-1}DK^{-1}c^K\pi^G DK^{-1}}{1+\mu} = -G^{-1}\left(I - \frac{v\pi^G}{1+\mu}\right)DK^{-1}$$

$$= -WD_K$$

Q.E.D.

The significant aspect of this theorem is its similarity with Theorem 4.5.2, in particular the similarity between $\hat{\psi}$ and $\Xi$, from which we can interprete W in terms of G and its subnetwork. For this purpose

note that, from Theorem 4.5.2, we can write

$$Q^{-1} = \frac{Q^{-1}}{\hat{\pi}c^1}\hat{\psi} = \frac{\hat{Q}^{-1}}{\hat{\pi}c^1}[\hat{\pi}c^1 I - (c^1 - \hat{c}^1)\pi] \tag{4.6.13}$$

where $\hat{c}^1 = (1,1,\ldots,1)^T$ is the capacity vector for $\hat{Q}$. Thus $\Delta c^1 = c^1 - \hat{c}^1$ gives the change in link capacity incurred when moving from $\hat{Q}$ to $Q$.

Similarly from (4.6.7) - (4.6.9)

$$W = \frac{G^{-1}}{1+\pi^G v}\Xi = \frac{G^{-1}}{1+\pi^G v}[(1+\pi^G v)I - v\pi^G] \tag{4.6.14}$$

But from the properties of dual variables, (3.4.1),

$$\pi^G c^G = -1 \tag{4.6.15}$$

so that

$$1 + \pi^G v = \pi^G v - \pi^G c^G = \pi^G(v - c^G) \tag{4.6.16}$$

Defining

$$c^W = c^G - v \tag{4.6.17}$$

and substituting in (4.6.14) we have

$$W = \frac{G^{-1}}{\pi^G c^W}[\pi^G c^W I - (c^W - c^G)\pi^G] \tag{4.6.18}$$

Furthermore, from (4.2.31) and (4.6.5)

$$\pi^1 \triangleq \pi(S^1) = Q^{-1}(\underline{r}1) = Q_\alpha^{-1}(\underline{r}(k+1)) = [U(\underline{r}1) \mid W(\underline{r}1)] \tag{4.6.19}$$

Comparing (4.6.18) with (4.6.13) and using (4.6.19) and the results of Section 4.5, we conclude:

Lemma 4.6.2

The weights of links of G, generated by considering subnetwork G alone, are the same as the weights of these links obtained by considering the entire network.

A related result is

Lemma 4.6.3

v is a vector of negative numbers, obtained independently of G, whose absolute value has to be added to vector $c^G$ to provide subnetwork G with a correct value of the equivalent capacity $\alpha^1$ that corresponds to the entire network.

Proof:

Although this Lemma also follows from the same results, we find it useful to prove it explicitly.

The value of the dual variables that corresponds to links of G when this subnetwork is considered in isolation is

$$\pi_\ell^G = - \frac{\theta_\ell^G}{D_G^1} \qquad \text{all } \ell \in G \tag{4.6.20}$$

where

$$D_G^1 = \sum_{\ell \in G} \theta_\ell^G c_\ell = \theta^G c^G \tag{4.6.21}$$

The value of the dual variables of these same links when they are considered immersed in the whole network are, considering Lemma 4.6.2,

$$\pi_\ell = - \frac{\theta_\ell^G}{D^1} \qquad \text{all } \ell \in G \tag{4.6.22}$$

where now

$$D^1 = \theta^G c^G + \theta^K c^K \qquad (4.6.23)$$

and $\theta^K$ is the set of weights of links of K when the problem is solved on the entire network. But from (4.6.18) we also have

$$D^I = \theta^G c^W = \theta^G (c^G - v) = \theta^G c^G - \theta^G v \ . \qquad (4.6.24)$$

Thus from (4.6.24) and (4.6.23)

$$\theta^K c^K = -\theta^G v \ , \qquad (4.6.25)$$

Since $\theta^K$, $c^K$ and $\theta^G$ are greater than zero, we conclude

$$v < 0 \qquad (4.6.26)$$

Furthermore (4.6.24) also proved the second part of the Lemma

Q.E.D.

From these results we can write

Theorem 4.6.3

If, by means of row and column permutation, matrix Q can be written as in (4.6.5), we can conclude:

i) There is a hierarchical division of the requirements of $R^1$ into two sets: the set using links of G alone, and its complementary set in $R^1$.

ii) The net effect of the upper level network on subnetwork G, is that of modifying the capacity vector of G from $c^G$ to $c^G - v$, where v is a negative vector given in (4.6.9).

iii) Subnetwork G, that in addition to its internal traffic will receive some traffic from the outside network K,

can be treated independently of the rest of the network by associating the external traffic coming into it with an increment in some corresponding internal requirements. In this way a reoptimization involving only subnetwork G can be initiated by means of a gxg saturation matrix. The actual Q-matrix of the whole network can then be updated using the results of Theorem 4.6.3.

iv) If matrix G has also some zero entries, and can be partitioned as in (4.6.5), the same procedure can be applied to it. In this way, we have $R^1$, $S^1$ and the network itself, divided into hierarchies, and we can reoptimize first within each hierarchy, and then transmit the necessary information outwards.

□

Although we will go back to these results in Chapter VI, when we study the reoptimization process, we will write now, for the sake of completeness, the relation between the dual variables of Q and G.

Lemma 4.6.4

The dual variables of the whole network are related to the inner pseudo dual variables of subnetwork G, with capacity vector v, by the following relations:

$$\pi_\ell = \frac{\pi_\ell^G}{1+\mu} \quad , \qquad \text{all } \ell \in G$$

$$\pi_\ell = - \sum_{a \in G} \pi_a D_K(a,\ell) \qquad \text{all } \ell \in K \tag{4.6.27}$$

where $D_K(a,\ell)$ is the $(a,\ell)$-th element of matrix $D_K$.

Proof:

From (4.2.31) and (4.6.5)

$$\pi^1 \triangleq \pi(s^1) = Q^{-1}(\underline{r}1) = Q_\alpha^{-1}(\underline{r}(k+1)) = [U(\underline{r}1) \mid W(\underline{r}1)]$$

But from Theorem 4.6.2

$$W(\underline{r}1) = \frac{\pi^G}{1+\mu}\,\Xi$$

$$U(\underline{r}1) = -W(\underline{r}1)D_K\,,$$

so the $\ell$-th components of these vectors are

$$W(1,\ell) = \frac{\pi^G}{1+\mu}\,\Xi(\underline{c}\ell) = \frac{1}{1+\mu}\left[\pi_\ell^G - \pi_\ell^G\sum_i \pi_i^G v_i + \pi_\ell^G\sum_i \pi_i^G v_i\right]$$

$$= \frac{\pi_\ell^G}{1+\mu}$$

$$U(1,\ell) = -\sum_{a\in G} W(1,\ell)D_K(1,\ell) = -\sum_{a\in G} \pi_a D_K(a,\ell)\,.$$

Q.E.D.

D is obviously the matrix that relates K to G. That is, it has the information of how the links of G appear in the saturated circuits of commodities originated outside subnetwork G. Clearly if D=0, then $D_K = 0$, $v = 0$ and there is no sharing between the two subnetworks; this case will correspond to an unstable situation.

The problem with this third method is that we need a saturation matrix that can be written in the form (4.6.5). If all, or most, saturated links appear in all, or most, saturated circuits then it is not possible to write Q in the form of (4.6.5). We anticipate that many practical cases, however, will exhibit the structure of (4.6.5), so that

the techniques of Method III will be applicable. We will present some example networks in Appendix III.

# CHAPTER V

## THE SUCCESSIVE SATURATION PROBLEM

### 5.1 Introduction

In this chapter we study the problem of lower saturation levels when the solution to (LP1) is not unique. Since the ideas supporting a lower level optimization were presented in section 2.4, we will not repeat them here. On the other hand we will present, in this introduction, some facts and results concerning nonuniqueness and the space of solutions of (LP1) that will be used in later sections.

We have already said that the solution to (LP1) will not, in general, be unique. We proceed now to characterize a non-unique solution in terms of the various subsets in which links and flows can be partitioned, given an optimal solution to (LP1). But first we introduce some definitions, many of which have already been established but which are repeated here for convenience. Since we will be working with various levels simultaneously we will use a superscript to characterize these levels. Given an optimal basic solution to (LP1), where $B^1$ denotes the optimal basis [*], we define:

- for the links, the sets

$$S_0^1 = \{\ell | s_\ell = 0, \ \ell \in L\}; \quad S_0^1 \cap B^1 = \phi,$$

$$S^1 = \{\ell | s_\ell = 0, \pi_\ell < 0, \ell \in L\}; \quad S^1 \subset S_0^1, \tag{5.1.1}$$

$$\bar{S}_0^1 = \{\ell | s_\ell > 0, \ell \in L\}; \quad \bar{S}_0^1 \subset B^1,$$

---

[*] $B^1$ stands not only for the basic matrix but also for the set of elements in the basis.

$$\bar{S}^1 = L - S^1 = \bar{S}_0^1 \cup (S_0^1 - S^1), \quad \text{and}$$

- for the flows the sets

$$\begin{aligned}
F_B^1 &= \{f_\ell(n) \,|\, f_\ell(n) > 0,\ f_\ell(n) \in F\}\ ; \quad F_B^1 \subset B^1, \\
F_{BT}^1 &= \{f_\ell(n) \,|\, \ell \in T(n),\ f_\ell(n) \in F_B^1\}; \quad F_{BT}^1 \subset F_B^1\ , \\
F_{BA}^1 &= F_B^1 - F_{BT}^1\ ; \quad F_{BA}^1 \subset F_B^1\ , \\
F_{\bar{B}}^1 &= \{f_\ell(n) \,|\, f_\ell(n) = 0,\ f_\ell(n) \in F\} = F - F_B^1\ ; \quad F_{\bar{B}}^1 \cap B^1 = \phi, \\
F_+^1 &= \{f_\ell(n) \,|\, z_\ell(n) > 0,\ f_\ell(n) \in F_{\bar{B}}^1\}; \quad F_+^1 \subset F_{\bar{B}}^1\ , \\
F_0^1 &= \{f_\ell(n) \,|\, z_\ell(n) = 0,\ f_\ell(n) \in F_{\bar{B}}^1\}; \quad F_0^1 \subset F_{\bar{B}}^1\ ,
\end{aligned} \tag{5.1.2}$$

where

$$F_{\bar{B}}^1 = F_+^1 \cup F_0^1$$

In Fig. 5.1.1 we represent a tableau that corresponds to an optimal solution partitioned according to previous definitions. As we have seen in Chapter IV we can represent the transition, when finding an optimum solution, from Step 1, (4.2.8), to Step 2, (4.2.21), as a transition from Tableau I to Tableau II of Fig. 5.1.1 below. We have also included the RHS and partial cost vectors in these tableaus for convenience and the sake of completeness.

Note that Tableaus I and II only correspond to the upper parts of (4.2.8) and (4.2.21) respectively. Since the spanning trees are invariants in the process, the lower part is of lesser interest for us, at this moment, than the upper part.

| $S^1$ | $\bar{S}^1$ | $\alpha^1$ | $F_{BA}$ | $F_0^1$ | $F_+^1$ | RHS |
|---|---|---|---|---|---|---|
| I | | $-c^1$ | $P_B^1$ | $P_{B0}^1$ | $P_B^1$ | $-r_T^1$ |
| | I | $-\bar{c}^1$ | $\bar{P}_B^1$ | $\bar{P}_{B0}^1$ | $P_{B+}^1$ | $-\bar{r}_T^1$ |
| | | 1 | | | | |

Tableau I

| $S^1$ | $\bar{S}^1$ | $\alpha^1$ | $F_{BA}$ | $\bar{F}_0^1$ | $F_+^1$ | RHS |
|---|---|---|---|---|---|---|
| $(Q^1)^{-1}$ | | I | | $P_0^1$ | $P_+^1$ | $b_0^1$ |
| $G^1$ | I | | | $\bar{P}_0^1$ | $\bar{P}_+^1$ | $\bar{b}_0^1$ |
| $-\pi^1(S^1)$ | | | | | $z^1(F_+^1)$ | $-\alpha_0^1$ |

Tableau II

Fig. 5.1.1

To infer, from these tableaus, the uniqueness or not of problem (LP1) we make use of Theorem 3.2.2 to write the following, rather obvious, result:

Lemma 5.1.1

The solution to (LP1) is unique iff $F_0^1$ and $S_0^1 - S^1$ are empty sets.

□

According to Definition 3.5.1 if $S_0^1 = S^1$ then the optimal basic solution is minimal. From the comments made in Section 3.5 it is apparent that we are interested in minimal solutions, since a nonminimal solution cannot be optimal for lower saturation levels. If the present solution is not minimal, but the problem is not unique and there exists an optimal basic solution that is minimal, we will shift to a minimal solution. Later on this chapter we will see how to deal with problems where none of the optimal basic solutions is minimal. We, thus again, make the assumption that we are dealing with a minimal solution.

We will now prove a result complementary to Lemma 5.1.1.

Lemma 5.1.2

If the solution to (LP1) is not unique then any possible variable entering the base must enter at a finite level. In consequence, every zero-cost column must have at least one positive component.

Proof:

From linear programming it is known ([29], page 167) that any zero cost variable can be made basic and enter at a finite level if it has a positive component in its column. Moreover based on physical grounds,

we already know that, for this problem, none of the variables can become unbounded, therefore each column with zero cost must have at least a positive component.

Q.E.D.

We can now extend another linear programming result, for bounded problems, to (LP1) and write

Lemma 5.1.3 ([29], pag 99, [28], pag 22)

If the solution to (LP1) is not unique then there are an infinite number of solutions, where any convex combination of two or more optimal basic solutions is also an optimal solution. Thus the solution space is a convex polyhedron whose vertices, or extreme points, are the optimal basic solutions of (LP1).

□

The purpose of the lower level problems is to successively reduce the size of the polyhedron until it is converted into a single point, or into the smallest possible polyhedron, in case the solution to the successive problem is not unique. Suppose this single point, or "smallest" polyhedron, contains one of the vertices of the original polyhedron, defined in Lemma 5.1.3. Then, the task of finding an optimal solution, for the successive problem, would be rather simple because it would consist of finding the appropriate optimal basic solution of (LP1), and there already exist procedures to solve this problem (see [29]). In general, though, the solution will be at the interior of the convex polyhedron, and the satement of problem (LP1) will not be enough to find the point. How to state the new problem, once an optimal solution to (LP1) is known,

and how to exploit the information provided by the solution of (LP1) as much as possible, are the subjects of this chapter.

## 5.2 The Second Saturation Level

The purpose of the second saturation level is to shrink the solution space of (LP1), whenever this solution space is more than a single point, to a subspace containing solutions that will also minimize the maximum saturation level of links not in $S^1$. From (2.4.1) and (2.2.6) this problem can be formulated as

$$\begin{aligned} &\min \ \alpha^2 \\ &\text{s.t.} \\ &\sum_n f_\ell(n) \leq \begin{cases} \alpha_0^1 c_\ell, & \ell \in S^1 \\ \alpha^2 c_\ell, & \text{otherwise} \end{cases} \qquad (5.2.1) \\ &E(\underline{rn})\underline{f}(n) = r(n), \quad \text{all } n \in N \\ &\underline{f} \geq 0 \\ &\alpha^2 \geq 0 \end{aligned}$$

where $\alpha_0^1$ is the optimal value of $\alpha^1$ from (LP1).

Note, however, that formulated in this way the only use that is made of the information contained in an optimal solution of (LP1) is the value of $\alpha_0^1$ and the composition of set $S^1$. Thus, to solve (5.2.1), many of the computations already performed during the solution of (LP1), will have to be repeated. This circumstance can be avoided if we formulate the second problem in a slightly different manner. For this purpose suppose that $S^1$ denotes the convex polyhedron of optimal solutions of (LP1); then (5.2.1) is equivalent to

$$\min \alpha^2$$

$$\sum_n f_\ell(n) + s_\ell = \begin{cases} \alpha_0^1 c_\ell \ , \ \ell \in s^1 \\ \alpha^2 c_\ell \ , \ \text{otherwise} \end{cases} \qquad (5.2.2)$$

$$(\alpha_0^1, s^T, f^T) \in S^1$$

It seems reasonably clear that, knowing an optimal solution to (LP1), Tableau II of Fig. 5.1.1, completed with its corresponding lower part, can be a starting point to solve (5.2.2). To proceed with the optimization from this tableau we first have to add the column of variables $\alpha^2$. Since, from (5.2.1), this column has zeroes in all positions corresponding to links of $S^1$, and all pivots needed to arrive to Tableau II take place in rows of $s^1$, it is obvious that $\alpha^2$ does not need to be updated to enter the tableau. Once $\alpha^2$ is inserted in the tableau the first pivot operation must take place at $\alpha^2$, since it is evident that $\alpha^2$ must be made basic and, on the other hand, the tableau was optimal before $\alpha^2$ was entered. Nevertheless, since $\alpha^2$ is a column whose nonzero elements are equal to -1 and whose cost is equal to 1, making $\alpha^2$ basic makes the system lose its primal feasibility. Although for optimization purposes this is not necessarily an undesirable property, it turns out that it is easier to study the problem if we avoid this situation. This problem can be easily overcome if instead of $\alpha^2$ we use a new variable $\beta^2$ defined as the slack between $\alpha_0^1$ and $\alpha^2$, that is

$$\beta^2 = \alpha_0^1 - \alpha^2 \qquad (5.2.3)$$

In this case the new objective function is

$$\max \beta^2 = \min -\beta^2 \qquad (5.2.4)$$

and the column $\beta^2$ is the opposite of the $\alpha^2$ column. It is worthwhile to observe that $\beta^2$, or $\alpha^2$, is the variable that will allow the system to leave the boundary of $S^1$ and move to interior points.

The new pivots of this optimization method will take place in this expanded tableau. Note, however, that the condition of not leaving $S^1$ prevents the system from making pivots within $S^1$ or $F_+^1$, since sone of the variables in these sets can be made basic. The variables in $S^1$ and $F_+^1$, become, in fact, null variables after the first problem and, see Section 3.2, they can be deleted from the tableau and adjoined to the constraints through the equalities $s(S^1) = 0$, $F_+^1 = 0$.

To continue our investigation, it will be convenient to write the objective function of (5.2.2.) first as

$$\min \alpha_0^1 - \beta^2 \tag{5.2.5}$$

and then, since $\alpha^1$ cannot be smaller than $\alpha_0^1$, as

$$\min \alpha^1 - \beta^2 \tag{5.2.6}$$

$$\alpha^1 \leq \alpha_0^1$$

In view of thse facts the second saturation problem, in matrix form, can be written as

$$\min \alpha^1 - \beta^2$$

$$\hat{A}^2 x = \hat{b}^2 \qquad (5.2.7)$$

$$x \geq 0$$

$$s(S^1) = 0,\ F_+^1 = 0$$

where

$$x^T = (\delta^1,\ \beta^2,\ \alpha^1,\ s^T(\bar{S}^1),\ (f^1)^T) \qquad (5.2.8)$$

$$\hat{A}^2 = \left[\begin{array}{c|c|c} 1 & 0 & 1\ 0\cdots0\cdots\cdots0 \\ \hline 0 & 0 & \\ \hline 0 & \bar{c}^1 & A^{11} \\ \hline 0 & 0 & \end{array}\right] \begin{array}{l} \uparrow \bar{I} \\ \downarrow \\ \uparrow \\ L \\ \downarrow \\ \uparrow \\ N(N-1) \\ \downarrow \end{array} \quad ; \quad b^2 = \left[\begin{array}{c} \alpha_0^1 \\ \hline b \end{array}\right] \qquad (5.2.9)$$

and where $s^T(\bar{S}^1)$ are the slacks for links in $\bar{S}^1$, $f^1$ is the flow vector without the elements of $F_+^1$, $A^{11}$ is matrix $A^1$ of (3.1.3) without the columns corresponding to links of $S^1$ or flows of $F_1^+$, $\delta^1$ is the slack between $\alpha^1$ and $\alpha_0^1$ (i.e., $\alpha^1 + \delta^1 = \alpha_0^1$), and $\bar{c}^1$ and b are defined in (4.2.3) and (3.1.4) respectively.

To solve this problem one can start, as we said before, with a modified version of Tableau II. This new tableau is obtained from the previous one by deleting the columns corresponding to the variables in $s(S^1)$ and $F_+^1$, adding the updated columns related to variables $\delta^1$ and $\beta^2$ and adding the updated row corresponding to the constraint $\alpha^1 + \delta^1 = \alpha_0^1$.

The two new columns will, of course, enter the tableau without any updating since they have zero coefficients for the variables in $s(S^1)$. With respect to the new row note that its updating will be given by the structure of the cost row of Tableau II. Since this row has zero coefficients in all places corresponding to variables $\alpha^1$, $s(\bar{S}^1)$ and $f^1$, the updated row will have zero coefficients everywhere except for variable $\delta^1$ that will have a unit coefficient. Thus $\delta^1$ actually becomes a zero value basic variable.

From the new tableau the optimization will proceed by first making $\beta^2$ basic, remember that it has a negative partial cost, and then entering into the basis the elements of $F_0^1$ as they are needed to successively reduce the value of the objective function $\alpha^2$. Observe that none of this basis changes will affect the structure of the row corresponding to $\alpha^1 + \delta^1 = \alpha_0^1$, that will remain as $\delta^1 = 0$. Therefore we can conclude

Lemma 5.2.1

The constraint $\alpha^1 \leq \alpha_0^1$ is redundant in (5.2.7). For every optimal solution of (5.2.7) the optimal values of $s(L)$ and F give an optimal solution for (LP1).

□

An immediate consequence of this lemma is that whenever the set of nonzero flows, related to a basic solution of (5.2.7), do not correspond with the nonzero flows of a basic solution of (LP1), as will be the case in general unless $S^2$ is composed of a single link, none of the optimal solutions of (5.2.7) will lie on the vertices of the convex polyhedron $S^1$. Thus they will be convex combinations of basic optimal

solutions of (LP1), as we predicted earlier. In this general case the solution of (5.2.7) will generate a new set of links $S^2$ that will saturate below $\alpha_0^1$, unless (LP1) is unstable, and a new set of alternate paths $F_{BA}^2$ that will balance the saturation level of the links of $S^2$.

Another consequence of Lemma 5.2.1 is that the constraints of (5.2.7) can be written as

$$\begin{aligned} A^2 x &= b \\ x &\geq 0 \\ s(S^1) &= 0, \quad F_+^1 = 0 \end{aligned} \tag{5.2.10}$$

where

$$x^T = (\beta^2, \alpha^1, s^T(\bar{S}^1), (f^1)^T) \tag{5.2.11}$$

$$A^2 = \left[\begin{array}{c|c} \begin{matrix} 0 \\ \hline \bar{c}^{-1} \\ \hline 0 \end{matrix} & A^{11} \end{array}\right] \begin{matrix} \bar{\dagger} \\ L \\ \downarrow \\ \bar{\dagger} \\ N(N-1) \\ \downarrow \end{matrix} \tag{5.2.12}$$

and b was defined in (3.1.4).

In order to avoid the need of explicitely writing the equalities for null variables, at the second and lower level problems, we now introduce the following:

Notation Assumption. Whenever a defined link slack or flow variable does not appear in the statement of a given level problem the missing variable will be considered a null variable for that problem.

Summaring these results and assumptions we can write

Theorem 5.2.1

The second saturation problem can be stated as

(LP2):

$$\begin{aligned} \min\ & \alpha^1 - \beta^2 \\ & A^2 x = b \\ & x \geq 0 \end{aligned} \tag{5.2.13}$$

where x and $A^2$ are defined in (5.2.11) and (5.2.12), respectively, and b was defined in (3.1.4).

□

The dual problem of (LP2) will be

(DP2):

$$\begin{aligned} \max\ & \lambda^2 b \\ & \lambda^2 A^2 \leq \gamma^2 \end{aligned} \tag{5.2.14}$$

where

$$\lambda^2 = (\pi^2, \sigma^2) = (\pi^2, \sigma^2(1), \ldots, \sigma^2(N))$$

is an $L + N(N-1)$ vector of dual variables and $\gamma^2 = (1\text{-}1,1,0,\ldots 0)$ is an $NL + 1 - |S^1| - |F_+^1| + 1$ row vector.

We could now begin a study completely parallel to that carried out in Chapter III, and define and describe the same type of concepts and results that were studied for (LP1). In particular we could define a linear region, a stable point, the set of canonic equations, etc. Since most results will be totally analogous we omit them here and refer instead to Chapter III whenever needed. Having this in mind we devote the remaining part of this section, as well as section 5.3 and 5.4, to the study of the main characteristics of problem (LP2) and (DP2) and their analogies and differences from homologous properties of (LP1) and (DP1). Although totally equivalent to Definition 3.2.2 it is worthwhile to write explicitly

Definition 5.2.1

For a stable optimal solution of (LP2), link $\ell$ belongs to the

second saturation set $S^2$ iff $\pi_\ell < 0$ and $\ell \in \bar{S}^1$.

The main difference between Definitions 5.2.1 and 3.2.2 is that in 5.2.1 we have to add the condition $\ell \in \bar{S}^1$. The reason, as we will see in Theorem 5.2.2 below, is that applying the complementary slackness condition to (LP2) and (DP2) we cannot guarantee that $\pi_\ell > 0$ for all $\ell \in S^1$.

We now state, without proof since it is entirely analogous to the proof of Chapter III, a theorem that is, for this problem, the equivalent of Theorem 3.4.1 for the first saturation problem.

Theorem 5.2.2

Let x and $\lambda^2$ be feasible solutions for (LP2) and (DP2) respectively. A necessary and sufficient condition that they both be optimal solutions is that for all $\ell \in L$ and all $n \in N$

i) For all $\ell \in \bar{S}^1$

$$\pi^2_\ell < 0 \Rightarrow s_\ell = 0$$

$$\pi^2_\ell = 0 \Leftarrow s_\ell > 0 \qquad (5.2.15)$$

For all $\ell \in S^1$

$\pi^2_\ell$ unrestricted in sign

ii) $$\sum_{\ell \in S^1} \pi^2_\ell c_\ell = 0$$

$$\sum_{\ell \in S^2} \pi^2_\ell c_\ell = -1 \qquad (5.2.16)$$

iii) $$f_\ell(n) > 0 \Rightarrow -\pi^2_\ell = \sigma^2_{o(\ell)}(n) - \sigma^2_{t(\ell)}(n) \qquad (5.2.17)$$

□

From (5.2.15) and (5.2.16) we have

Corollary 5.2.2

Either $\pi^2_\ell = 0$ for all $\ell \in S^1$, of else there is at least one link of $S^1$ for which $\pi^2_\ell > 0$.

□

Although the solution to (LP2) will simultaneously minimize the saturation levels both of arcs in $S^1$ and $S^2$, the fact that components of the dual variables $\pi^2$ may be positive prevents a direct assignment of distances to the arcs of $S^1$ and $S^2$. As we will see later in section 5.4, when we talk about distance assignments, it will be convenient if we can find a transformation that, applied to vector $\pi^2$, will give an image vector $\pi^{2+}$ such that $\pi^{2+} \leq 0$. It turns out that we can easily obtain such a correspondence, by eliminating the artifact that makes some $\pi_\ell$ positive. From (4.2.28) it is very simple to see that the result $z_\ell = -\pi_\ell$, where $z_\ell$ is the partial cost of link $\ell$, generalizes to this problem. Since at the optimal solution all costs must be positive or zero, the only reason why, for some $\ell \in S^1$, the cost may be negative is because the links (slacks) of $S^1$ are treated as artificial, and not real, variables. Suppose we decide to incorporate the links of $S^1$ and the elements of $F^1_+$ as active variables and keep the constraint $\alpha^1 \leq \alpha^1_0$; then we can formulate the second problem as

$(LP2^+)$:

$$\begin{aligned} \min\ & \alpha^1 - \beta^2 \\ & A^{2+}x = b^{2+} \\ & x \geq 0 \end{aligned} \qquad (5.2.18)$$

where

$$x^T = (\delta^1, \beta^2, \alpha^1, s^T, (f)^T)$$

and

$$A^{2+} = \left[\begin{array}{c|c|ccc} 1 & 0 & 1 & 0 \ldots . & 0 \\ \hline & 0 & & & \\ 0 & \frac{-1}{c} & & A^1 & \\ & 0 & & & \end{array}\right] ; \qquad b^{2+} = b^2 \tag{5.2.19}$$

The matrix $A^1$ is given in (3.1.3) and the remaining quantities are defined in (5.2.9).

The dual of $(LP2^+)$ is

$(DP2^+)$

$$\begin{aligned} &\max \eta^{2+}\alpha_0^1 + \lambda^{2+}b \\ &(\eta^{2+}, \lambda^{2+})A^{2+} \leq \gamma^{2+} \end{aligned} \tag{5.2.20}$$

where

$$\lambda^{2+} = (\pi^{2+}, \sigma^{2+})$$

and

$$\gamma^{2+} = (0,-1,1,0,\ldots 0)$$

Similarly to Theorem 5.2.2 we can apply the complementary slackness condition to obtain

Theorem 5.2.3

Let x and $(\eta^{2+}, \lambda^{2+})$ be feasible solutions for $(LP2^+)$ and $(DP2^+)$ respectively. A necessary and sufficient condition that they both be

optimal solutions is that for all $\ell \in L$ and all $n \in N$

i) For all $\ell \in \bar{S}^1$

$$\pi_\ell^{2+} < 0 \Rightarrow s_\ell = 0 \qquad (5.2.21)$$

$$\pi_\ell^{2+} = 0 \Leftarrow s_\ell > 0 \ ,$$

for all $\ell \in S^1$

$$\pi_\ell^{2+} \leq 0$$

ii) $$\sum_{\ell \in S^1} \pi_\ell^{2+} c_\ell = \eta^{2+} \leq 0 \qquad (5.2.22)$$

$$\sum_{\ell \in S^2} \pi_\ell^{2+} c_\ell = -1$$

iii) $$f_\ell(n) > 0 \Rightarrow -\pi_\ell^{2+} = \sigma^{2+}_{o(\ell)}(n) - \sigma^{2+}_{t(\ell)}(n) \ . \qquad (5.2.23)$$

□

We can see, through this theorem, that $\pi^{2+}$ has, in fact, the desired properties. The obtaining of the relation between $\pi^2$ and $\pi^{2+}$ will be left for section 5.4

Because of the similarity between problems (LP2) and ($LP2^+$), and considering the additional fact that the solution of ($DP2^+$) can be immediately associated with a valid distance assignment, we could ask the question: why do we not consider problem ($LP2^+$) only and delete problem (LP2)? There are two main reasons that can be used to support our decision of working with (LP2) instead of ($LP2^+$): a) ($DP2^+$) has, as we will see later, a non-unique solution, thus making that, in addition

to other possible inconveniences, we could not give a sensitivity interpretation to the variables of ($DP2^+$); and b), ($LP2^+$) is a problem of higher dimension than problem (LP2). Therefore ($LP2^+$) can be better considered as an auxiliar problem whose only purpose is to provide with a transformation that will generate a valid set of link distance assignments from the dual variables of (LP2); once the transformation is obtained problem ($LP2^+$) is no longer needed.

## 5.3 Matrix Analysis of the Second Saturation Problem

In this section we want to extend the results of Chapter IV to the second saturation problem. Since, according to the formulation of the second saturation problem, an optimal solution to this problem includes a solution to (LP1) as a necessary starting point and by product, we will distinguish the matrices that properly characterize one problem or the other by the superscript "1" or "2". As in Section 4.2, suppose that we know from an optimal solution of (LP2) the set of saturated links and basic flows. That is, we know

$$S^2 = \{\ell | s_\ell = 0,\ \pi^2_\ell < 0,\ \ell \in S^1\}$$
$$F^2_B = \{f_\ell(n) | f_\ell(n) > 0,\ \text{all } \ell \in L,\ \text{all } n \in N\} \qquad (5.3.1)$$

We also assume that the solution is minimal, and that we know $S^1$. Furthermore, from Chapter IV and Sections 5.1 and 5.2, we also have a partition of F into sets $F^2_B$ and $F - F^2_B$. In particular, Tableaus I and II of Fig. 5.1.1 give this partition except for the spanning tree set $F_{BT}$, which has been omitted, and for set $F^2_{BA}$ that is not specified. Observe first

that the set $F_{ET}^2$ obtained from $F_B^2$ need not be the same as set $F_{BT}^1$ obtained from $F_B^1$, because in the case of the first saturation level we did not have constraints for commodities outside $R^1$. Nevertheless it should be clear that replacing $F_{BT}^1$ by $F_{BT}^2$ in (4.2.21) would generate a totally equivalent basic optimal solution for (LP1). Thus, we assume that $F_{BT}^1 \equiv F_{BT}^2$ and we omit the superscript to avoid confusion.

We saw in sections 5.1 and 5.2 that (4.2.21) can be taken as a starting point for the solution of (LP2). If we already know the set $F_{BA}^2$, the transformation that has to be applied to (4.2.21) (extended with column $\beta^2$) to reach the optimal tableau for (LP2) will be the transformation necessary to make $\beta^2$ and $F_{BA}^2$ basics. First we introduce some definitions

$$\begin{aligned} S^{12} &= S^1 \cup S^2 \\ \bar{S}^{12} &= L - S^{12} \\ |S^2| &= q^2 \ . \end{aligned} \tag{5.3.2}$$

The remaining variables and sets will have a definition entirely equivalent, except for the superscript, to other variables and sets already defined, and will not be repeated here.

If we assume the components of vector x of (5.2.13) ordered as

$$x^T = (S^1, S^2, \bar{S}^{12}, \alpha^1, F_{BA}^1, \beta^2, F_{BA}^2, F_{BT}, F_B^2) \tag{5.3.3}$$

we can write the modified version of (4.2.21), before $\beta^2$ and $F_{BA}^2$ are made basic, as in (5.3.4) below:

$$A^{12} = \begin{array}{c|c|c|c|c|c|c|c|c} & S^1 & S^2 & \bar{S}^{12} & \alpha^1 & F^1_{BA} & \beta^2 & F^2_{BA} & F_{BT} & F^2_{\bar{B}} \\ \hline & (Q^1)^{-1} & & & & I & 0 & P^{21}_B & & P^{21} \\ \hline & \bar{G}^{12} & I & & & & \bar{c}^2 & P^2_B & & P^2_{\bar{B}} \\ \hline & \bar{G}^{10} & & I & & & \bar{c}^2 & \bar{P}^2_3 & & \bar{P}^2_{\bar{B}} \\ \hline & G^1 & & & & & 0 & M^2_B & I & M^2_{\bar{B}} \end{array} \qquad (5.3.4)$$

Considering that

$$F^1_{\bar{B}} = F^2_{BA} \cup F^2_{\bar{B}} \qquad (5.3.5)$$

the definition of all matrices of (5.3.4) as a function of the matrices of (4.2.21) is straightforward and, therefore is not undertaken here, except for the matrices under $F^2_{BA}$. Since the elements of $F^2_{BA}$ must have a zero cost at the first level we have, from Theorem 4.3.1, that the elements of $F^2_{BA}$, either, have a zero reduced circuit, or else, their reduced circuits can be written as a linear combination of the reduced circuits of $F^2_{BA}$. In the first case $P^{21}_B = 0$ and the interpretation of the columns of $F^2_{BA}$ is completely analogous to the interpretation of the columns of $F^1_{BA}$, since $F^2_{BA}$ has not been affected by the pivoting in $Q^1$.

In the second case $P_B^{21} \neq 0$; to understand the implication of this fact we can isolate the columns corresponding to $\alpha^1$, $F_{BA}^1$, $\beta^2$ and $F_{BA}^2$, as shown below in (5.3.6).

| $\alpha^1$ | $F_{BA}^1$ | $\beta^2$ | $F_{BA}^2$ |
|---|---|---|---|
| $-c^1$ | $P_B^1$ | 0 | $P_{\bar{B}}^{21}$ |
| $-c^2$ | $\bar{P}_B^{12}$ | $c^2$ | $P_{\bar{B}}^2$ |
| $-\bar{c}^2$ | $\bar{P}_B^{10}$ | $\bar{c}^2$ | $\bar{P}_{\bar{B}}^2$ |

a)

| $\alpha^1$ | $F_{BA}^1$ | $\beta^2$ | $F_{BA}^2$ |
|---|---|---|---|
| I | | 0 | $P_B^{21}$ |
| 0 | | $c^2$ | $P_B^2$ |
| | | $\bar{c}^2$ | $\bar{P}_B^2$ |

b)

(5.3.6)

where $\bar{c}^2 = c(\bar{S}^{12})$.

Since the elements of $F_{BA}^2$ are zero-cost at the first level, the first row of $P_B^{21}$ is a row of zeroes. Thus with respect to the elements of $F_{BA}^2$ the only change from a) to b) of (5.3.6) is a change in the way of representing the circuits of $F_{BA}^2$. In a) these circuits are represented in terms of their link components, in b) these circuits are represented in terms of the circuits of $F_{BA}^1$ and the links of $\bar{S}^{-1}$. It is very easy to see, from (5.3.6), that

$$P_B^{21}(\underline{r}1) = 0$$

$$P_B^{21}(\overline{\underline{r}1}) = [P_B^1(\overline{\underline{r}1})]^{-1} P_{\overline{B}}^{21}(\overline{\underline{r}1})$$

$$P_B^2 = P_{\overline{B}}^2 - P_B^{12} P_B^{21}(\overline{\underline{r}1}) \qquad (5.3.7)$$

$$\overline{P}_B^2 = P_{\overline{B}}^2 - \overline{P}_B^{12} P_B^{21}(\overline{\underline{r}1})$$

Because of the transformations (5.3.7) we have no guarantee that $P_B^2$ will be, in general, a 0, $\pm$ 1 matrix. Nevertheless we can still define a saturation matrix for the second level as

$$Q^2 \triangleq \left[\begin{array}{c|c} -c^1 & P_B^2 \end{array}\right] \qquad (5.3.8)$$

where $Q^2$ will still satisfy many of the saturation matrix properties, in particular

$$(Q^2)^{-1} = \left[\begin{array}{c} \pi^2(s^2) \\ \hline \Omega^2 \end{array}\right] \qquad (5.3.9)$$

We must, of course, take note that the matrix we have to use to make $\beta^2$ and $F_{BA}^2$ basics is not $Q^2$ but $Q_+^2$ where

$$Q_+^2 = \left[\begin{array}{c|c} c^2 & P_B^2 \end{array}\right] \qquad (5.3.10)$$

But if $(Q_+^2)^{-1}$ is the inverse of this matrix we can easily write, from (5.3.8) - (5.3.10)

$$(Q_+^2)^{-1} = \left[\begin{array}{c} -\pi^2(s^2) \\ \hline \Omega^2 \end{array}\right] \qquad (5.3.11)$$

One important observation that we can make whenever $P_B^{21} \neq 0$ is that it is no longer obvious how to partition the set $F_{BA}^1 \cup F_{BA}^2$ into subsets $F_{BA}^1$ and $F_{BA}^2$. In fact, it will be more accurate to interpret $\alpha^1$, $\beta^2$ and $F_{BA}^1 \cup F_{BA}^2$ as forming a larger saturation matrix with the links of $s^1 \cup s^2$. In this case the global or composite saturation matrix will be $Q^{12}$ below

$$Q^{12} = \begin{array}{c|c|c} & \alpha^1 & \beta^2 \qquad F_{BA}^1 \cup F_{BA}^2 \\ \hline -c^1 & 0 & \\ \hline -c^2 & c^2 & P_B^{12} \end{array} \quad ; \quad (Q^{12})^{-1} = \begin{array}{c|c} \pi^1(s^1) & 0 \\ \hline \pi^1(s^1)-\pi^2(s^1) & -\pi^2(s^2) \\ \hline \multicolumn{2}{c}{\Omega^{12}} \end{array} \qquad (5.3.12)$$

and we will now prove

Lemma 5.3.1

$(Q^{12})^{-1}$ has the structure given in (5.3.12) where $\Omega^{12}$ is a full rank $|s^1| + |s^2| - 2$ by $|s^1| + |s^2|$ matrix.

Proof:

Since the fact that $\Omega^{12}$ has to be full rank and have those dimensions is obvious we only have to prove that the first two rows

of $(Q^{12})^{-1}$ are as given in (5.3.12).

It is convenient to reorder the columns of $Q^{12}$ in the same way as the columns of (5.3.6) and to consider the process of moving from $Q^{12}$ to $(Q^{12})^{-1}$ as broken into two steps. First pivoting at $Q^1$, and then at $Q^2_+$, we have that the first row of the future inverse matrix after the first step is $[\pi^1(s^1)\,|\,0]$. Thus, since the entries of this row corresponding to links of $s^2$ are zero, the second step does not affect the structure of this row. Furthermore since $a^1$ is the first column of both (5.3.6) and (5.3.13) $[\pi^1(s^1)\,|\,0]$ will be the first row of $(Q^{12})^{-1}$.

With respect to the second row of $(Q^{12})^{-1}$, note that if we call to the first and second rows of $(Q^{12})^{-1}$ w1 and w2, respectively, then, from (4.2.27) applied to (LP2), we can write

$$\pi^2(s^1, s^2) = (1,-1,0,\ldots,0)\ (Q^{12})^{-1} = \\ = w1 - w2 \qquad (5.3.13)$$

so

$$w2 = [\pi^1(s^1) - \pi^2(s^1), -\pi^2(s^2)]$$

Q.E.D.

In general not all columns of $P^{21}_E$ of (5.3.6) will be nonzero. Furthermore, it may happen that if only a few elements of $F^2_{BA}$ use links of $s^1$ we could find a new partition of set $F^1_{BA} \cup F^2_{BA} \cup F^2_0$ into $F^1_{BA}$, $F^2_{BA}$ and $F^2_0$ having a $P^{21}_E = 0$. To see that consider that $F^2_0$ is non-empty, i.e., (LP2) has a nonunique solution, and that $F^2_0$ has elements whose circuits do not use links of $s^1$; then it may happen that some of these elements of $F^2_0$ can be moved to $F^2_{BA}$ to replace the elements of this set that where

using links of $S^1$. To complete the new partition operation we have to check if some of the elements that have been replaced from $F^2_{BA}$ must be nonzero and, thus, should be included in $F^1_{BA}$. In this case we have to be able to move as many elements from $F^1_{BA}$ to $F^2_0$ as we move from $F^2_{BA}$ to $F^1_{BA}$. If this reordering operation is not feasible, then it is not possible to obtain a partition with $P^{21}_{\overline{B}} = 0$. Note that $P^{21}_{\overline{B}} = 0$ is an interesting property to be satisfied, both for computation and hierarchical purposes, since in this case there is a clear division of allowable paths with respect to which level fixes their values. We should therefore try to perform such a reordering whenever possible.

To obtain the updated constraint matrix for the optimal solution (5.3.1), we have to pivot at matrix $Q^2_+$ in (5.3.4). As a result of this operation we would obtain the matrix given in (5.3.14), where the various submatrices can be easily obtained from the matrices of (5.3.4). Note that whenever $P^{21}_{\overline{B}} = 0$ neither the first column of submatrices of (5.3.4), nor the first row, will be altered by the transformation.

$$A^2_0 = \begin{array}{|c|c|c|c|c|c|}
\multicolumn{1}{c}{S^1} & \multicolumn{1}{c}{S^2} & \multicolumn{1}{c}{\overline{S}^{12}} & \multicolumn{1}{c}{\alpha^1 F^1_{BA}} & \multicolumn{1}{c}{\beta^2 F^2_{BA}} & \multicolumn{1}{c}{F_{BT}} \quad F^2_{\overline{B}} \\
\hline
G^{11} & G^{21} & & I & & \quad P^{12} \\
\hline
G^{12} & (Q^2_+)^{-1} & & & I & \quad P^2 \\
\hline
\overline{G}^{12} & \overline{G}^2 & I & & & \quad \overline{P}^2 \\
\hline
\overline{G}^{\overline{12}} & G^2 & & & & I \quad M^2 \\
\hline
\end{array} \qquad (5.3.14)$$

Note from (5.3.12) that

$$G^{11}(\underline{r}1) = \pi^1(S^1) \; ; \quad G^{21}(\underline{r}1) = 0 \qquad (5.3.15)$$

$$G^{12}(\underline{r}1) = \pi^1(S^1) - \pi^2(S^1)$$

and except for the ordering of rows the block of four matrices of the left upper corner correspond to $(Q^{12})^{-1}$ of (5.3.12).

Similarly we can see that the inverse matrix of the basis for the optimal solution of (LP2) defined by (5.3.1) is

$(B^2)^{-1} =$

| $S^1$ | $S^2$ | $\bar{S}^{12}$ | $R$ |
|---|---|---|---|
| $G^{11}$ | $G^{21}$ | | $V^1$ |
| $G^{12}$ | $(Q_+^2)^{-1}$ | | $V^2$ |
| $\bar{G}^{12}$ | $\bar{G}^2$ | $I$ | $\bar{V}^2$ |
| $\bar{G}^{\bar{12}}$ | $G^2$ | | $U^2$ |

(5.3.16)

where again all submatrices can be easily obtained from (4.2.25).

It is clear that we could now expand the analysis of sections 4.2 to cover these new transformations. In particular we could partition the V, U, M, P, and some of the G, matrices by commodities and perform a micro-analysis of their characteristics. How and when they will change

in the transformation depends on whether or not a given commodity belongs to $R^1$ or $R^2$, and whether or not a commodity of $R^1$ uses only links of $S^1$ or, on the other hand, if it also flows through links of $S^2$, etc. Although important, this type of analysis is not so relevant for us here, since it can be considered as a rather straightforward generalization of the study done in sections 4.2 and 4.3.

The new RHS and partial cost vector can also be obtained in a similar way. From this last vector we can also partition $F_{\bar{B}}^2$ in subsets completely analogous to those in which $F_{\bar{B}}^1$ was partitioned, i.e., $F_0^2$, $F_+^2$ etc., see (5.1.2) and (5.1.3). As an example of all these calculations we will obtain now the new value of the routing variables $F_{BA}^1$ and $F_{BA}^2$, and also the relation of $\pi^2(S^1)$ with $\pi^2(S^2)$, $\pi^1(S^1)$ and the graph matrices of the circuits of $F_{BA}^1$ and $F_{BA}^2$.

Lemma 5.3.2

The optimal value of variables $\beta^2$, $\alpha^2$, $f_\ell(n) \in F_{BA}^1 \cup F_{BA}^2$, corresponding to the optimal solution (5.3.1) are

$$\begin{aligned} \beta_0^2 &= \alpha_0^1 - \alpha_0^2 \\ \alpha_0^2 &= -\pi^2(S^2) f^1(S^2) \\ F_{BA}^2 &= -\Omega^2 f^1(S^2) \\ {}^2(F_{BA}^1) &= F_{BA}^1 - P_B^{21}(\overline{\underline{r}^1}) F_{BA}^2 \end{aligned} \qquad (5.3.17)$$

where $f^1(S^2)$ is the vector of aggregate flows on links of $S^1$ as obtained from the (LP1) optimal solution (see Theorem 4.2.9), $\Omega^2$ is

given in (5.3.9), $P_B^{21}(\underline{r}1)$ is the matrix that describe how the elements of $F_{BA}^2$ use links of $S^1$ and can be obtained from (5.3.6), and $^2(F_{BA}^1)$ is the new set of values for the variables of $F_{BA}^1$.

Proof

Consider the RHS vector corresponding to the optimal solution of (LP1). Adjoining this RHS vector to matrix $A^{12}$ of (5.3.4), and involving it in the pivot operation that transforms (5.3.4) into (5.3.14) we will have, for the upper subvectors of the transformed RHS, using (5.3.10), (4.2.36), (4.2.39)

$$\begin{bmatrix} B_0^2 \\ F_{BA}^2 \end{bmatrix} = (Q_+^2)^{-1} s(S^2) = \begin{bmatrix} -\pi^2(S^2) \\ \Omega^2 \end{bmatrix} [\alpha_0^1 c^2 - f^1(S^2)]$$

$$= \begin{bmatrix} -\alpha_0^1 \pi^2(S^2) c^2 \\ \alpha_0^1 \Omega^2 c^2 \end{bmatrix} + \begin{bmatrix} \pi^2(S^2) f^1(S^2) \\ -\Omega^2 f^1(S^2) \end{bmatrix}$$

Moreover, using (5.2.16) and (4.4.17) applied to $Q^2$

$$\begin{bmatrix} \beta_0^2 \\ F_{BA}^2 \end{bmatrix} = \begin{bmatrix} \alpha_0^1 \\ 0 \end{bmatrix} + \begin{bmatrix} \pi^2(S^2) f^1(S^2) \\ -\Omega^2 f^1(S^2) \end{bmatrix}$$

The definition of $\beta^2$ completes the proof of the first three equations. The fourth equation is an immediate consequence of the pivot operation on subvector $F_{BA}^1$.

Q.E.D.

Obverve that $P_B^{21}(\overline{\underline{r}1}) = 0$ implies $^2(F_{BA}^1) = F_{BA}^1$, as expected. Clearly the same values as in (5.3.17) would have been obtained if

instead of using a two step operation, first pivoting at $Q^1$ and then at $Q^2_+$, we were to make a single pivot at $Q^{12}$, defined in (5.3.12), starting with the RHS vector corresponding to Tableau I of Fig. 5.1.1. For this single pivoting operation the values of (5.3.17) could have been written as

$$\alpha_0^1 = -\pi^1(S^1) f_T(S^1)$$

$$\beta_0^2 = \alpha_0^1 - \alpha_0^2 = -[\pi^1(S^1) - \pi^2(S^1)] f_T(S^1) + \pi^2(S^2) f_T(S^2) \qquad (5.3.18)$$

$$F_{BA}^{12} = -\Omega^{12}[f_T(S^1),\ f_T(S^2)]$$

where $F_{BA}^{12}$ is a new vector defined to include both the variables of $F_{BA}^1$ and $F_{BA}^2$, and represent not only the set $F_{BA}^1 \cup F_{BA}^2$, but also their optimal values at the (LP2) optimal solution. Observe that to obtain $F_{BA}^{12}$ we do not need the whole $(Q^{12})^{-1}$ but only the bottom part $\Omega^{12}$.

Now we will compute $\pi^2(S^1)$ as a function of $\pi^2(S^2)$ and the graph matrices corresponding to the circuits of $F_{BA}^1$ and $F_{BA}^2$. For this purpose consider the row of partial costs attached to matrix (5.3.4) and apply to that row the pivot transformation that takes $A^{12}$ of (5.3.4) into $A_0^2$ of (5.3.14). Doing so, and realizing that $\pi^2(S^1)$ and $\pi^2(S^2)$ are the partial costs corresponding to links in $S^1$ and $S^2$, respectively, we can write

$$-\pi^{21} = -\pi^{11} - (-10\ldots0)\ (Q_+^2)^{-1}\overline{G}^{12}$$

$$= -\pi^{11} - \pi^{22}\overline{G}^{12}$$

where we used (5.3.11) and we called, for simplicity,

$$\pi^{ij} \triangleq \pi^i(S^j) \qquad (5.3.19)$$

so

$$\pi^{21} = \pi^{11} + \pi^{22}\bar{G}^{12} \tag{5.3.20}$$

But from (5.3.4) and the tableuas of Fig. 5.1.1

$$\bar{G}^{12} = -\bar{\Phi}^{12}(Q^1)^{-1} \tag{5.3.21}$$

where

$$\bar{\Phi}^{12} = [-c^2 | \bar{P}_B^{12}] \tag{5.3.22}$$

is the submatrix of matrix $\bar{\Phi}^{-1}$ of (4.2.24) corresponding to the links of $S^2$, and $\bar{P}_B^{12}$ is the collection of reduced circuits from the point of view of $S^2$ for the variables of $F_{BA}^1$. Partitioning, also, $(Q^1)^{-1}$ as in (4.4.2) and substituting in (5.3.21)

$$\bar{G}^{12} = -[-c^2 | \bar{P}_B^{12}] \begin{bmatrix} \pi^{11} \\ \hline \Omega^1 \end{bmatrix}$$

$$= \begin{bmatrix} c_1^2 \pi^{11} \\ \vdots \\ c_{q2}^2 \pi^{11} \end{bmatrix} - \bar{P}_B^{12} \Omega^1 \tag{5.3.23}$$

From (5.3.20) and (5.3.23) the $\ell$-th element of $\pi^{21}$ will be

$$\pi_{\ell}^{21} = \pi_{\ell}^{11} + \pi^{22}\bar{G}^{12}(\underline{c}\ell)$$

$$= \pi_{\ell}^{11} + \pi^{22}[\pi_{\ell}^{11} c^2 - \bar{P}_B^{12}\Omega^1(\underline{c}\ell)]$$

$$= \pi_{\ell}^{11}[1 + \pi^{22}c^2] - \pi^{22}\bar{P}_B^{12}\Omega^1(\underline{c}\ell)$$

Since, from (5.2.16), $\pi^{22} c^2 = -1$

$$\pi_{\ell}^{21} = -\pi^{22} \bar{P}_B^{12} \Omega^1(\underline{c}\ell) \tag{5.3.24}$$

But from the definition of $\bar{P}_B^{12}$ we can define

$$\zeta_a^{21}(m) = -\sum_{\ell \in S^2} \pi_{\ell}^{22} \bar{P}_B^{12}(\underline{c}\, am) \quad \text{all } f_a(m) \in F_{BA}^1 \tag{5.3.25}$$

to be the weight of the reduced circuit $\bar{P}_B^{12}(\underline{c}\, am)$ of variable $f_a(m) \in F_{BA}^1$. Thus manipulating (5.3.24) and substituting (5.3.25)

$$\pi_{\ell}^{21} = \sum_{am \in F_{BA}^1} \Omega^1(am, \ell)\ \zeta_a^{21}(m) \qquad (*)$$

$$= \sum_{am \in F_{BA}^1} (Q^1)^{-1}(am, \ell)\, \zeta_a^{21}(m)$$

The interpretation of these results is summarized in

Lemma 5.3.3

Given $F_{BA}^1$, from an optimal solution of (LP1), and the set $S^2$ we can obtain the components of vector $\pi^2(S^1) \triangleq \pi^{21}$ as follows

$$\pi_{\ell}^{21} = \sum_{am \in F_{BA}^1} (Q^1)^{-1}\ (am, \ell)\, \zeta_a^{21}(m) = \zeta^{21} \Omega^1(\underline{c}\ell) \tag{5.3.26}$$

where $(Q^1)^{-1}(am, \ell)$ is the $(am, \ell)$-th component of $(Q^1)^{-1}$ and $\zeta_a^{21}(m)$, given by (5.3.25), is the reduced circuit weight with respect to $S^2$ for variables $f_a(m) \in F_{BA}^1$. $\zeta^{21}$ is a $|F_{BA}^1|$ row vector.

□

As a check on the result of (5.3.26) we can compute

(*) $am \in F_{BA}^1$ is a shorthand notation for: all am such that $f_a(m) \in F_{BA}^1$.

$$\sum_{\ell \in S^1} \pi_\ell^{21} c_\ell = \sum_{\ell \in S^1} c_\ell \sum_{am \in F_{BA}^1} (Q^1)^{-1}(am,\ell)\zeta_a^{21}(m)$$

$$= \sum_{am \in F_{BA}^1} \zeta_a^{21}(m) \sum_{\ell \in S^1} (Q^1)^{-1}(am,\ell) c_\ell = 0$$

from (4.4.17), which agrees with (5.2.16).

It is obvious that if $P_B^{12} \neq 0$ then any feasible set $F_{BA}^1$ will give the same answer (5.5.17), since no matter how we order the columns of $Q^{12}$ we must obtain the same vectors of dual variables (we are assuming, of course, that the point is stable.)

An immediate consequence of (5.3.24) is

Lemma 5.3.4

If $\bar{P}_B^{12} = 0$, that is, if commodities of $R^1$ do not use links of $S^2$, then $\pi^{21} = 0$.

□

Observe that from Lemma 5.4.10 we can only infer that $R^1$ does not interact with $S^2$. From this result we could be tempted to conclude that, whenever $P_B^{12} = 0$, if $\alpha_0^2$ becomes larger than $\alpha_0^1$ and (LP2) gets transformed in (LP1) and vice versa, the new $R^1$ will also be independent of the new $S^2$. This conjecture is not true, in general, as can be seen with the help of the sketch of Fig. 5.3.1.

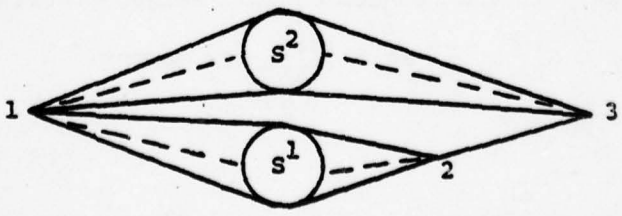



Fig. 5.3.1

In this figure $R^1 = r_1(2)$ does not use links of $S^2$, and therefore $\bar{P}_B^{12} = 0$. On the other hand if $\alpha_0^2$ becomes larger than $\alpha_0^1$ commodity (1,3) can use the paths via node 2 and thus for the new (LP1) - (LP2) pair, $\bar{P}_B^{12} \neq 0$.

Following the same procedure used to prove Lemma 5.3.3 we could prove

Lemma 5.3.5

The commodity dual variables of an optimal solution of (LP2) are given by

$$\sigma_m^2(n) = - \sum_{\ell \in T^1(m,n)} \pi_\ell^{21} - \sum_{\ell \in T^2(m,n)} \pi_\ell^{22} \qquad (5.3.27)$$

Furthermore for any variable $f_\ell(n) \in F_B^2$

$$z_\ell^2(n) = \sigma_{t(\ell)}^2(n) - \sigma_{o(\ell)}^2(n) - \pi_\ell^2$$

analogously to (4.2.35).

□

The conclusion that we infer from this matrix analysis and matrix calculations is twofold: First, as in the case of the first saturation problem, knowing the saturated links and the set of allowable paths we can readily compute the updated constraint and inverse basis matrices. Second, we can obtain expressions that directly give any particular elements of these matrices and/or RHS and partial cost vectors. Furthermore these expressions can take advantage of some properties that can save part of the computation that we would have incurred by applying pivoting operations alone. Consider,

for instance, the proof of Lemma 5.3.2 where by using (5.3.17) directly we can save the computation of $\pi^2(S^2)c^2 = 1$ and $\Omega^2 c^2 = 0$. Of course, a change of basis performed by pivoting operations, or equivalent matrix manipulations, is, in principle, blind to all properties that can be used to save computations. It turns out that in our problem we have many of these computation-saving properties that can be used very efficiently.

We will now show some of these matrix calculations applied to the three node example of Section 4.2.

Example 5.3.1

From Example 4.2.3 we have

$$F^1_{BA} = \{f_c(1), f_b(3)\}$$

$$F_{BT} = \{f_b(1), f_g(1), f_a(2), f_c(2), f_e(3), f_d(3)\}$$

$$F^1_+ = \{f_d(1), f_g(2), f_a(3)\}$$

$$F^1_0 = \{f_e(2)\}$$

Since $F^1_0 \neq \phi$ the solution to the (LP1) problem for this example is not unique and, therefore, we can attempt the solution of a second saturation problem. If $s_a$ is the smallest slack variable among the links of $\bar{S}^1$, then we can reduce the saturation level of link a by using the alternate path {e,c} provided by variable $f_e(2) \in F^1_0$. This new path will, obviously, increase the saturation level of both links e and c. If we assume that the level on link c is higher than the level on link e, then, $S^2 =$ {a,c} and the tableaus (5.3.4) and (5.3.14) of the corresponding (LP2)

problem are those given below. For simplicity we only write, for the second tableau, the most important vectors and matrices. Note that all crossed out matrices are nonzero ones whose value need not be calculated to obtain the optimal solution. Empty spaces correspond to zeros.

Observe also that if $s_a$, corresponding to (LP1), is smaller than $s_c$ or $s_e$, then, after making $\beta^2$ basic at e or c (depending on whether $s_e < s_c$ or vice versa) $f_e(2)$ will appear with a positive cost; activating $f_e(2)$ will clearly cause $\alpha_c$ and $\alpha_e$ to increase. In this case the set $S^2$ will have a single element, e or c, and the second saturation problem will not only not increase the number of alternate paths, but will also reduce the set $F_0^2$ to $\phi$.

For the case where $S^2 = \{a,c\}$ the set of allowable paths is given in Fig. E.5.3.1. The numerical values correspond to the requirement vector given in Example 4.2.3.

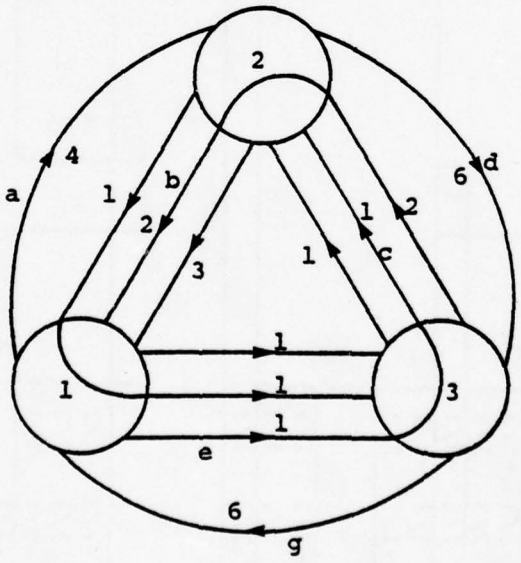
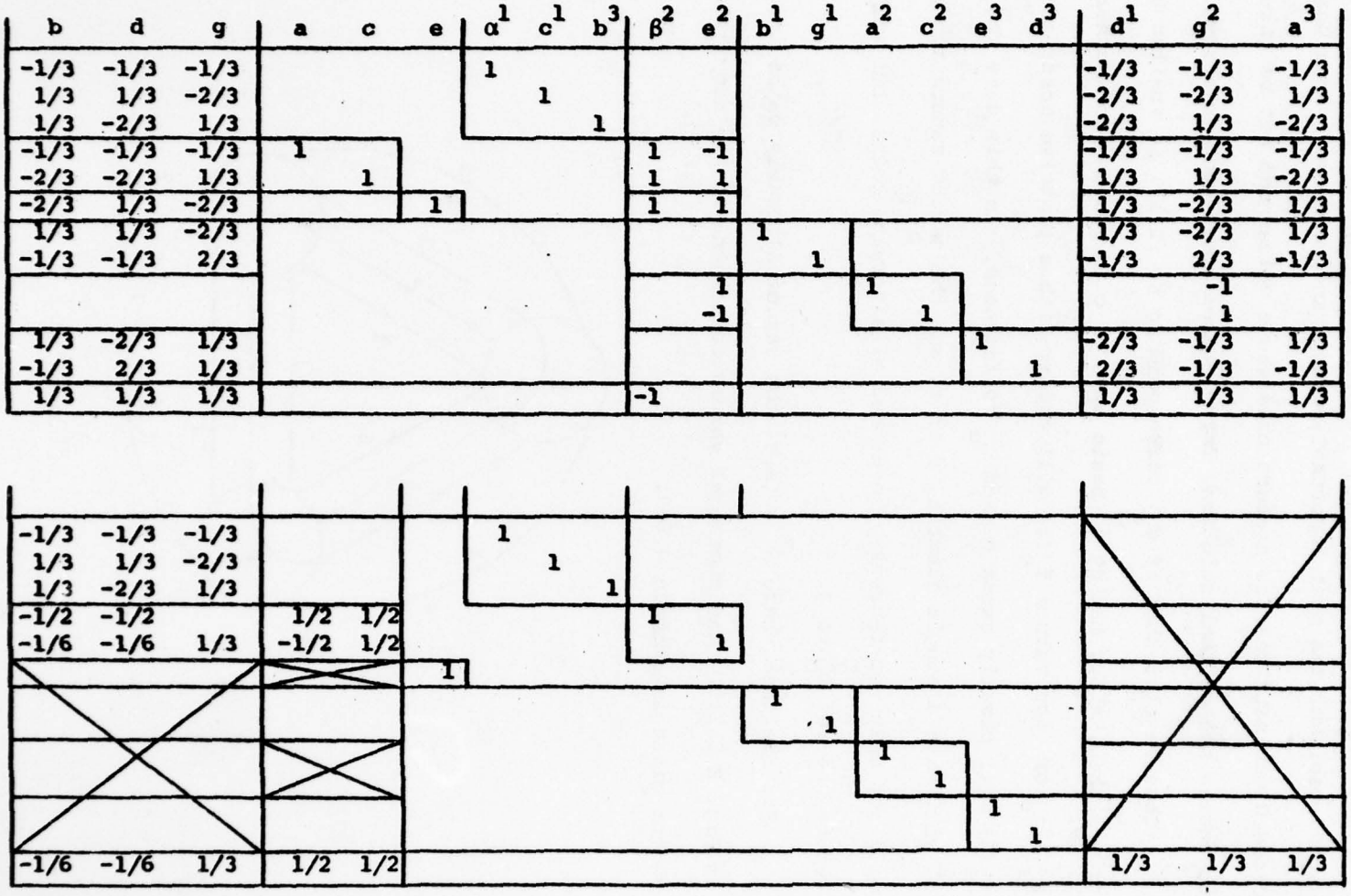



Fig. E.5.3.1


b d g | a c e | $\alpha^1$ $c^1$ $b^3$ | $\beta^2$ $e^2$ | $b^1$ $g^1$ $a^2$ $c^2$ $e^3$ $d^3$ | $d^1$ $g^2$ $a^3$
-1/3 -1/3 -1/3 | 1 | 1/3 1/3 -2/3 | 1 | 1/3 -2/3 1/3 | 1 | -1/3 -1/3 -1/3 | 1 -1 | -1/3 -1/3 -1/3
-2/3 -2/3 1/3 | 1 | 1 | 1 | -2/3 -2/3 1/3 | 1 1 | 1/3 1/3 -2/3
-2/3 1/3 -2/3 | 1 | 1 | 1 | -2/3 1/3 -2/3 | 1 1 | 1/3 -2/3 1/3
1/3 1/3 -2/3 | 1 | 1/3 -2/3 1/3
-1/3 -1/3 2/3 | 1 | -1/3 2/3 -1/3
1 | -1 | 1 | -1
-1 | 1
1/3 -2/3 1/3 | 1 | -2/3 -1/3 1/3
-1/3 2/3 1/3 | 1 | 2/3 -1/3 -1/3
1/3 1/3 1/3 | -1 | 1/3 1/3 1/3


-1/3 -1/3 -1/3 | 1 | 1/3 1/3 -2/3 | 1 | 1/3 -2/3 1/3 | 1
-1/2 -1/2 | 1/2 1/2 | 1 | 1/3 1/3 1/3
-1/6 -1/6 1/3 | -1/2 1/2 | 1 | 1
1 | 1
1 | 1
1 | 1
1
-1/6 -1/6 1/3 | 1/2 1/2 | 1/3 1/3 1/3

The optimal dual variables are

$$\pi_b^2 = \pi_d^2 = \frac{1}{6}, \quad \pi_g^2 = -\frac{1}{3}$$

$$\pi_a^2 = \pi_c^2 = -\frac{1}{2}, \quad \pi_e^2 = 0$$

$$\sigma_2^2(1) = \sigma_2^2(3) = -\frac{1}{6}, \quad \sigma_1^2(2) = \sigma_3^2(2) = \frac{1}{2}$$

$$\sigma_3^2(1) = \frac{1}{3}, \quad \sigma_1^2(3) = 0$$

## 5.4 Dual Variables Properties of the Second Level Problem. Distance Assignment

From Theorems 5.2.2 and 3.4.1 we can see that the dual variables in $\pi^2(S^2)$ will have the same properties as the dual variables in $\pi^1(S^1)$. Since these last variables have been extensively studied in Chapters III and IV we will not repeat that analysis for $\pi^2(S^2)$ here. On the other hand $\pi^2(S^1)$ is a set of dual variables with different characteristics, that appear in problem (LP2) for the first time. We have already seen in Section 5.3 the relation of $\pi^2(S^1)$ to $\pi^2(S^2)$, $\pi^1(S^1)$ and the circuit matrices of $F^1_{BA}$ and $F^2_{BA}$. Here we will show other properties of $\pi^2(S^1)$ in relation to the allowable paths of commodities in $R^1$ and $R^2$, and we will also give a distance assignment for (LP2).

Since part iii) of Theorems 5.2.2 and 3.4.1 are identical we can immediately generalize Lemmas 3.4.1 and 3.4.2 of (LP1) to a solution of (LP2). The new results are

### Lemma 5.4.1

For commodities not in $R^1$

$$\sigma_m^2(n) = -\sum_{\ell \in S^2 \cap P^2(m,n)} \pi_\ell^2 \geq 0, \quad \text{all } (m,n) \in \bar{R}^1, \tag{5.4.1}$$

$$\sigma_m^2(n) = 0 \iff s^2 \cap P^2(m,n) = \phi \quad \text{all } (m,n) \in \bar{R}^1, \tag{5.4.2}$$

where $P^2(m,n)$ is any allowable path for commodity $(m,n)$ given by an optimal solution of (LP2).

□

From this Lemma we can immediately write the following formal definition of $R^2$,

Definition 5.4.1

$$(m,n) \in R^2 \iff \sigma_m^2(n) \neq 0, \quad (m,n) \in \bar{R}^1 .$$

□

Lemma 5.4.2

For commodities of $R^1$

$$\sigma_m^2(n) = - \sum_{\ell \in P^2(m,n)} \pi_\ell^2 =$$

$$= - \sum_{\ell \in P^{21}(m,n)} \pi_\ell^2 - \sum_{\ell \in P^{22}(m,n)} \pi_\ell^2 , \tag{5.4.3}$$

where

$$P^{2i}(m,n) \triangleq s^i \cap P^2(m,n) . \tag{5.4.4.}$$

□

An immediate consequence of this Lemma, Corollary 5.2.2 and Lemma 3.1.1 is

Lemma 5.4.3

$\sigma_m^2(n)$ can be interpreted as the sensitivity factor of $\alpha_0^2$ to changes in commodity (m,n), that is

$$\Delta\alpha_0^2 = \sigma_m^2(n)\ \Delta r_m(n). \qquad (5.4.5)$$

$\sigma_m^2(n)$ will be negative for some (m,n) $\in R^1$, unless $\pi^2(S^1) = 0$. For such commodities a positive increment $\Delta r_m(n)$ will make $\alpha_0^2$ decrease

□

Lemma 5.4.4

If

$$P^{22}(m,n) = \phi, \quad \text{for all } P^2(m,n), \text{ all } (m,n) \in R^1, \qquad (5.4.6)$$

then

$$\pi^2(S^1) = 0 .$$

Proof:

From (5.4.3) and $P^{22}(m,n) = \phi$ we have

$$\sigma_m^2(n) = - \sum_{\ell \in P^{21}(m,n)} \pi_\ell^2 \quad .$$

But, from (5.4.5), $\sigma_m^2(n)$ must be equal to zero also because $P^{22}(m,n) = \phi$, all (m,n) $\in R^1$, means that commodities of $R^1$ do not use links of $S^2$ and therefore cannot affect the value of $\alpha_0^2$. Thus

$$0 = - \sum_{\ell \in P^{21}(m,n)} \pi_\ell^2 \quad , \qquad \text{all } P^2(m,n), \text{ all } (m,n) \in R^1$$

and therefore

$$\pi_{\ell}^{2} = 0, \quad \text{all } \ell \in S^{1} .$$

Q.E.D.

These results show for the first time, as compared to the results of Chapter III, a case where some $\sigma$ variable is negative, and, thus, where some requirements may cause a reduction of a saturation level by means of a positive increment in value. In Example 5.3.1, for instance, $\sigma_2^2(1) = -\frac{1}{6}$ thus $\Delta\alpha_0^2 = -\frac{1}{6}\Delta r_2(1)$. From Fig. E.5.3.1 we see that an increment in commodity (2,1) will reduce the amount of commodity (3,1) flowing on links b and c, thus, reducing $\alpha_c$ and $\alpha_0^2$.

If we consider how problem (LP2$^{+}$), instead of problem (LP2), the generalization of Lemmas 5.4.1-2 to this other problem is straightforward.

Lemma 5.4.5

For commodities in $\bar{R}^{1}$ the results are identical to those of Lemma 5.4.1 with the substitution of $\sigma_m^{2+}(n)$ for $\sigma_m^2(n)$ and $\pi_{\ell}^{2+}$ for $\pi_{\ell}^2$.

Lemma 5.4.6

For commodities of $R^{1}$

$$\sigma_m^{2+}(n) = -\sum_{\ell \in P^{21}(m,n)} \pi_{\ell}^{2+} - \sum_{\ell \in P^{22}(m,n)} \pi_{\ell}^{2+}$$

$$= -\pi^{2+}(S^{1})P^{21}(m,n) - \pi^{2+}(S^{2})P^{22}(m,n) , \qquad (5.4.7)$$

where now

$$\sigma_m^{2+}(n) > 0 , \quad \text{all } (m,n) \in R^1. \tag{5.4.8}$$

□

Now we will generalize the results of Section 3.4 to the second level problem. From Lemmas 3.4.2 and 5.4.1 the generalization of Theorem 3.4.2 to (LP2) will be

Theorem 5.4.1

If $-\pi_\ell^2$ is the length of link $\ell$, $\ell \in \bar{S}^1$, and we assume that the links of $S^1$ have been removed (or assigned an arbitrarily large distance), then all allowable paths for commodity $(m,n) \in \bar{R}^1$, as given by any optimal solution of (LP2), are those with total minimum distance. This minimum distance is equal to the optimal value of dual variable $\sigma_m^2(n)$. The total distance for commodities in $\bar{R}^1 - R^2$ is zero, and therefore all paths for these commodities are through links of zero distance.

□

We would similarly generalize Lemmas 3.4.4-5 and Theorem 3.4.2, and obtain a set of integer values that will satisfy conditions i) and iii) of Theorem 5.2.2. This set of integer values is a set of weights for the links of $\bar{S}^1$ that can replace the dual variables of Theorem 5.4.1 to describe the set of allowable paths as indicated in this theorem.

As we said in Section 5.2 the reason why the dual variables of (LP2) cannot be used to obtain a distance assignment that will describe the allowable paths for commodities of both $R^1$ and $R^2$, is because (LP2)

does not consider the links of $S^1$ and the elements of $F_+^1$ as active variables. Thus it may happen that in most cases (see Corollary 5.2.2) the solution of (DP2) will give negative partial costs for elements of $S^1$ and $F_+^1$, that is, negative distances for elements that were supposed to have a positive distance. This difficulty can be overcome by the ($LP2^+$) formulation of the problem, since this formulation, see (5.2.18) includes $S^1$ and $F_+^1$ as active members of the set of primal variables. Thus an optimal solution of ($LP2^+$) must be such that neither the elements of $S^1$ nor the elements of $F_+^1$ will have a negative cost. For the case of the elements of $S^1$ this means that none of the links of $S^1$ will have a positive dual variable, i.e., negative distance. For the elements of $F_+^1$ the condition means that these elements cannot have a negative cost and that they will remain at a zero value. We will come again to this point later on in this section.

Due to the importance of the optimal dual variables of ($LP2^+$), we will spend some time now studying problems ($LP2^+$) and ($DP2^+$) and the relation of variables $\pi^{2+}$ and $\sigma^{2+}$ with variables $\pi^2$ and $\sigma^2$. If $(\pi^1, \sigma^1)$, $(\pi^2, \sigma^2)$ and $(\eta^{2+}, \pi^{2+}, \sigma^{2+})$ are optimal solutions of problems (DP1), (DP2) and ($DP2^+$), respectively, then they must satisfy Theorems 3.1.1, 5.2.2 and 5.2.3 (i.e., the complementary slackness conditions) respectively. Condition iii) of these theorems can also be written, according to Lemmas 3.4.3, 5.4.1-2 and 5.4.5-6, as

$$\begin{aligned}
&\sigma_m^1(n) = -\pi^1 p^1(m,n), && \text{all } p^1(m,n),\ (m,n) \in R,\\
&\sigma_m^2(n) = -\pi^2 p^2(m,n), && \text{all } p^2(m,n),\ (m,n) \in R, \qquad (5.4.9)\\
&\sigma_m^{2+}(n) = -\pi^{2+} p^{2+}(m,n), && \text{all } p^{2+}(m,n),\ (m,n) \in R,
\end{aligned}$$

where now $P^i(m,n)$, which represents <u>any</u> allowable path obtained from the solution of problem (DPi), is an L-component vector with 1's in the positions corresponding to the links of the allowable path and 0's elsewhere. Since the optimal solution space $S^2$ is a subset of the space $S^1$, and furthermore [from the definition of problems (LP2) and ($LP2^+$) (see Section 5.2)], the set of allowable paths of (DP2) and ($DP2^+$) are the same, we can write

$$\left.\begin{aligned} \sigma_m^1(n) &= -\pi^1 P^2(m,n) \\ \sigma_m^2(n) &= -\pi^2 P^2(m,n) \\ \sigma_m^{2+}(n) &= -\pi^{2+} P^2(m,n) \end{aligned}\right\} \quad \text{all } P^2(m,n),\ (m,n) \in R, \qquad (5.4.10)$$

Thus, all $(\pi^1, \sigma^1)$, $(\pi^2, \sigma^2)$ and $(\pi^{2+}, \sigma^{2+})$ satisfy the relations, corresponding to the variables $f_\ell(n)$ that generate the paths $P^2(m,n)$, of the third condition of Theorems 3.1.1, 5.2.2 and 5.2.3. One of the reasons why $(\pi^1, \sigma^1)$ or $(\pi^2, \sigma^2)$ are <u>not</u> solutions of ($DP2^+$), however, is because neither of them satisfy the other two conditions of Theorem 5.2.3; in particular $\pi^2$ is not, in general, a non-positive vector, and $\pi^1$ is not nonzero for the links of $S^2$. Another reason why $(\pi^2, \sigma^2)$ is not an optimal solution of ($DP2^+$) is because as mentioned before, $z^2(f_\ell(n))$, some $f_\ell(n) \in F_+^1$, can be negative, therefore forcing some $f_\ell(n) \in F_+^1$ to be basic, thus violating some of the relations included in iii) of Theorem 5.2.3.

Consider now $P^2(m,n)$ partitioned into the subvectors intersecting with $S^1$, $S^2$ and $\bar{S}^{12} \triangleq L - S^1 \cup S^2$, respectively, and let

$$p^{2i}(m,n) \triangleq p^2(m,n) \cap s^k, \quad i=1,2$$

$$\pi^{ji} \triangleq \pi^j(s^i), \quad i=1,2, \quad j=1,2,2+ \quad . \tag{5.4.11}$$

Then, we can write

$$\sigma_m^1(n) = \begin{cases} -\pi^{11}p^{21}(m,n) , & \text{all } (m,n) \in R^1 \\ 0 & \text{otherwise} \end{cases}$$

$$\sigma_m^2(n) = \begin{cases} -\pi^{21}p^{21}(m,n) - \pi^{22}p^{22}(m,n), & \text{all } (m,n) \in R^1 \\ -\pi^{22}p^{22}(m,n) , & \text{all } (m,n) \in R^2 \\ 0 & \text{otherwise,} \end{cases} \tag{5.4.12}$$

$$\sigma_m^{2+}(n) = \begin{cases} -\pi^{21+}p^{21}(m,n) - \pi^{22+}p^{22}(m,n), & \text{all } (m,n) \in R^1 \\ -\pi^{22+}p^{22}(m,n), & \text{all } (m,n) \in R^2 \\ 0 & \text{otherwise.} \end{cases}$$

It is very easy to see, from the part of equations (5.4.12) corresponding to $(m,n) \in R^2$, and from the part of conditions i) and ii) of Theorems 5.2.2-3 corresponding to the links of $s^2$, that

$$\pi^{22+} = \pi^{22} \quad . \tag{5.4.13}$$

With respect to the links and commodities in $s^1$ and $R^1$, respectively, note that if $(\pi^2, \sigma^2)$ satisfy (5.4.12), and if for any real number $\eta$

we define

$$\pi^{21\eta} = \pi^{21} + \eta\pi^{11},$$

$$\pi^{22\eta} = \pi^{22}, \tag{5.4.14}$$

$$\pi_{\ell}^{2\eta} = 0, \quad \text{all } \ell \in \bar{S}^{12},$$

$$\sigma_m^{2\eta}(n) = -\pi^{2\eta}p^2(m,n),$$

then the vectors $(\pi^{2\eta}, \sigma^{2\eta})$ also satisfy (5.4.12). In particular note that

$$\begin{aligned}\sigma_m^{2\eta}(n) &= -\pi^{21\eta}p^{21}(m,n) - \pi^{22\eta}p^{22}(m,n) = \\ &= -(\pi^{21} + \eta\pi^{11})p^{21}(m,n) - \pi^{22}p^2(m,n) = \\ &= \sigma_m^2(n) + \eta\sigma_m^1(n), \qquad \text{all } (m,n) \in R^1,\end{aligned} \tag{5.4.15}$$

and note also that

$$\sum_{\ell \in S^1} \pi_{\ell}^{2\eta}c_{\ell} = \sum_{\ell \in S^1} \pi_{\ell}^{2}c_{\ell} + \eta \sum_{\ell \in S^1} \pi_{\ell}^{1}c_{\ell} = -\eta . \tag{5.4.16}$$

Comparing (5.4.16) and ii) of Theorem 5.2.3 we can conclude from that theorem, that if $\eta$ is a nonnegative real number such that

$$\pi^{21\eta} \leq 0$$

and

$$z^{2\eta}(f_\ell(n)) = -\pi^{2\eta}C_\ell(n) \geq 0, \quad \text{all } f_\ell(n) \in F_+^1 , \tag{5.4.17}$$

where $C_\ell(n)$ is the circuit corresponding to variable $f_\ell(n)$, then $(-\eta, \pi^{2\eta}, \sigma^{2\eta})$ is an optimal solution of (DP2$^+$). Furthmore suppose that $(-\eta, \pi^{2\eta}, \sigma^{2\eta})$ is an optimal solution of (DP2$^+$); then, it is very easy to see that, for the same reasons, $(-\bar{\eta}, \pi^{2\bar{\eta}}, \sigma^{2\bar{\eta}})$ [where $\bar{\eta} = k\eta$, k a positive real number] is another optimal solution of (DP2$^+$). The minimum value of k for which $(-\bar{\eta}, \pi^{2\bar{\eta}}, \sigma^{2\bar{\eta}})$ is an optimal solution is fixed by conditions (5.4.17). If $\bar{\eta}^{2+}$ is the value of $-\bar{\eta}$ corresponding to the minimum value of k, then we can write, from (5.4.14) and (5.4.17)

$$\bar{\eta}^{2+} = \min \left\{ \min_{\ell \in S^1} \frac{\pi_\ell^2}{\pi_\ell^1}, \quad \min_{f_\ell(n) \in F_+^1} \frac{z^2(f_\ell(n))}{z^1(f_\ell(n))} \right\} \leq 0 \tag{5.4.18}$$

where, as before,

$$z^i(f_\ell(n)) = -\pi^i C_\ell(n) , \quad i = 1,2 .$$

The corresponding dual variables for this value of the scalar $-\eta$ will be

$$\begin{aligned}
&\bar{\pi}^{21+} = \pi^{21} - \bar{\eta}^{2+}\pi^{11} \leq 0 , \\
&\bar{\pi}^{22+} = \pi^{22} < 0 , \\
&\bar{\pi}_\ell^{2+} = 0 , \quad \text{all } \ell \in \bar{S}^{12},
\end{aligned} \tag{5.4.19}$$

$$\begin{aligned}
&\bar{\sigma}_m^{2+}(n) = \sigma_m^2(n) - \bar{\eta}^{2+}\sigma_m^1(n), && \text{all } (m,n) \in R^1, \\
&\bar{\sigma}_m^{2+}(n) = \sigma_m^2(n), && \text{all } (m,n) \in R^2, \\
&\bar{\sigma}_m^{2+}(n) = 0 && \text{otherwise} ,
\end{aligned} \tag{5.4.20}$$

and the costs for the variables of $F_+^1$

$$\bar{z}^{2+}(f_\ell(n)) = -\bar{\pi}^{2+}C_\ell(n) \geq 0, \quad \text{all } f_\ell(n) \in F_+^1 . \tag{5.4.21}$$

These results can be summarized in the following theorem,

Theorem 5.4.2

Problem ($DP2^+$) defined in (5.2.20) has infinite solutions. These infinite solutions are defined by expressions (5.4.18) - (5.4.20) replacing $\bar{\eta}^{2+}$ by $\eta^{2+} = k\bar{\eta}^{2+}$, where k is a scalar real variable $k \leq 1$.

□

Thus Theorem 5.4.2 says that ($DP2^+$) has an infinite number of solutions, each one of which corresponds to a point of the positive ray of the undimensional real space whose origin is the point $\bar{\eta}^{2+}$ defined in (5.4.18). The non-uniqueness of problem ($DP2^+$) can also be inferred from the analogous result to Theorem 3.2.3 obtained for ($LP2^+$). For this matter note that ($LP2^+$) is, by definition, non-stable since any positive increment in some $r_m(n)$, $(m,n) \in R^1$, will make the problem unfeasible, thus ($DP2^+$) must have infinite solutions.

Now that we have obtained the solutions of ($DP2^+$) as a simple function of the solutions of (LP1) and (LP2), we are ready to obtain a distance assignment to describe the set of allowable paths for the second saturation problem by applying the shortest route criterion. But before doing that we will briefly study an apparent anomaly that the solution to ($DP2^+$) might present some times. Observe, from (5.4.21), that

$\bar{z}^{2+}(F_+^1)$ is not a positive but rather a non-negative vector; thus $\bar{z}^{2+}(F_+^1)$ can have some zero components. This event will occur whenever

$$\bar{\eta}^{2+} = \frac{z^2(f_\ell(n))}{z^1(f_\ell(n))}, \quad \text{for some } f_\ell(n) \in F_+^1 . \tag{5.4.22}$$

If we assume that $\bar{z}^{2+}(f_\ell(n)) = 0$ then the alternate path introduced by $f_\ell(n)$ is a shortest route, and therefore an allowable, path for commodity $(o(\ell), n)$. Thus it might appear that the solution to $(DP2^+)$ is introducing allowable paths that were not allowed at the first level. Suppose that instead of $\bar{\eta}^{2+}$ we consider the solution corresponding to $(1+\varepsilon)\bar{\eta}^{2+}$ where $\varepsilon$ is an arbitrarily small positive real number then we will have

$$\begin{aligned} z^{2+}(f_\ell(n)) &= -\pi^{2+}c_\ell(n) = \\ &= -\bar{\pi}^{2+}c_\ell(n) - \varepsilon\pi^{11}c_\ell^1(n) \\ &= \varepsilon\, z^1(f_\ell(n)) > 0 , \end{aligned} \tag{5.4.23}$$

since by hypothesis $f_\ell(n) \in F_+^1$. Therefore for any other point of the ray that generates the optimal solutions, different from its origin, this path will not be allowed. What is happening here is that, in spite of the fact that variable $f_\ell(n)$ has a zero cost and thus can be made basic and generate an alternative optimal solution, the problem is degenerate and if we make $f_\ell(n)$ basic it will enter the basis at a zero level. The degeneracy of the optimal solution of $(LP2^+)$ is due to the first constraint of (5.2.18) since $\delta^1 = 0$ at the optimal point. Observe also that the process followed to obtain an optimal solution of $(LP2^+)$

from an optimal solution of (LP2) can be interpreted as the process of increasing the lengths of the links of $S^1$, with respect to the lengths given by the solution of (LP2), until this length is large enough not to allow any commodity not in $R^1$ to use links of $S^1$. Of course, this length increasing process has to be made so that all allowable paths introduced by (LP1) and (LP2) remain allowable, thus obtaining the expressions given in (5.4.18)-(5.4.20). These expressions correspond to the smallest length for the links of $S^1$ with the desired properties. It can happen though, as we saw before, that in some cases this "smallest" assignment is not enough yet and we have to move to an interior point of the ray described by Theorem 5.4.2.

We can summarize previous results as follows

Theorem 5.4.3

The set $\{-\pi_\ell^{2+}\}$, where $\pi^{2+}$ is any optimal link dual variable vector obtained accordingly to the result of Theorem 5.4.2 for some $k > 1$, is a valid distance assignment to characterize all allowable paths under the shortest route criterion for $S^2 (S^2 \subset S^1)$, the optimal solution space of (LP2). The commodity dual varibles, $\{\sigma_m^{2+}(n)\}$, will be the set of total minimum distances for the various commodities under the $\{-\pi_\ell^{2+}\}$ assignment.

If $\bar{z}^{2+}(F_+^1)$, defined in (5.4.21), is strictly positive, i.e., if (5.4.22) is not satisfied for any $f_\ell(n) \in F_+^1$, then $\{-\bar{\pi}_\ell^{2+}\}$, obtained from Theorem 5.4.2 by making k=1, is also a valid assignment.

□

Theorem 5.4.3 is the result that we were looking for in order to have a valid distance assignment for the allowable paths of the second saturation

problem. Using now similar arguments to those used in Section 3.4 we can write a lemma analogous to Lemma 3.4.4:

Lemma 5.4.7

Conditions i) and iii) of Theorem 5.2.3 define, for a given optimal value $\eta^{2+}$, a ray on the $(\eta,\lambda^+)$ space. The optimal solution of (DP2$^+$), corresponding to the given value $\eta^{2+}$, is the point $(\eta^{2+}, \pi^+, \sigma^+)$ of that ray which satisfies

$$\sum_{\ell \in S^1} \pi_\ell^{2+} c_\ell = \eta^{2+}$$

$$\sum_{\ell \in S^2} \pi_\ell^{2+} c_\ell = -1 \ .$$

□

Furthermore, using again similar arguments to those of Section 3.4 we can also write

$$\pi_\ell^{2+} = - \frac{\theta_\ell^{2+}}{\Delta^{2+}} \ , \qquad (5.4.24)$$

$$\sigma_m^{2+}(n) = \frac{d_m^{2+}(n)}{\Delta^{2+}} \ ,$$

$$\Delta^{2+} = \sum_{\ell \in S^2} \theta_\ell^{2+} c_\ell \ ,$$

$$\eta^{2+} = \frac{1}{\Delta^{2+}} \sum_{S^1} \theta_\ell^{2+} c_\ell \ ,$$

where $\theta_{\ell}^{2+}$ and $d_m^{2+}(n)$ are non-negative integers. We could also prove

Lemma 5.4.8

$s^1$, $s^2$, $\eta^{2+}$ and the set of allowable paths define, for a given network, a ray that can be characterized by an integer vector $(-\theta^{2+}, d^{2+})$, defined by (5.4.24). The smallest such integer vector is obtained when $\Delta^{2+}$ = minimum common multiple of the denominators of the $\pi$'s.

□

From Lemmas 5.4.7-8 we can write a Theorem completely analogous to Theorem 3.4.2.

Theorem 5.4.4

To any point of the ray defined by Lemmas 3.4.4 or 3.4.5 there corresponds a different valid distance assignment to describe the allowable paths according to Theorem 3.4.1. In particular the integer numbers defined in (5.4.24) constitute a valid assignment. For this assignment the total distance for any allowable path of commodity (m,n) will be $d_m^{2+}(n)$ given in (5.4.24).

□

Observe that here we have two degrees of freedom in choosing the distance assignment, as compared to the one degree we had for the solution of the first saturation problem. The new degree of freedom introduced at the second level is due to the dual variable $\eta^{2+}$ that can take infinite optimal values. Note, from Theorem 5.4.2, that $\eta^{2+}$ depends on the value of $\pi^{11}$. Suppose that, instead of $\{-\pi_{\ell}^1\}$, we choose any other

valid distance assignment for the first level problem, say $\{d_\ell^1\}$, and that we replace $\pi^{11}$ by $-d^{11}$ in (5.4.18)-(5.4.20). Then, (5.4.18) will give the new value of $\bar{\eta}^{2+}$ corresponding to the new first level assignment and, therefore, the new starting point for the $\eta$-ray of the second level problem. Expressions (5.4.19) and (5.4.20) will give the corresponding second level assignment for the new starting point. Thus we can say that $\eta^{2+}$ is translating to the second level problem the freedom of choice we had at the first level. Hence the two degrees of freedom of the second problem can be easily interpreted by considering one of them to be a legacy from the first level problem (recall that $S^2 \subset S^1$).

We can summarize all these results in the following

Lemma 5.4.9

To any valid distance assignment of the first level problem, say $\{d_\ell^1\}$, expression (5.4.18) gives the starting point of the $\eta$-ray that generate valid distance assignments for the second level problem, by replacing $\pi^{11}$ by $-d^{11}$, the vector whose components are $-d_\ell^1$. Furthermore to any point of the $\eta$-ray there corresponds, according to Lemma 5.4.7, another ray on the $(\pi,\sigma)$-space all of whose points also generate valid distance assignments for the second level problem.

□

An almost immediate consequence of this result is that, in order to obtain an integer assignment from $\pi^{2+}$ we can use $-\theta^{11}$ instead of $\pi^{11}$ in (5.4.18)-(5.4.20) and then look for the minimum common multiple of the denominators of $(\eta^{2+}, \pi^{21})$, instead of those of the larger vector

($\pi^{21+}$, $\pi^{22+}$). Most of the advantages of working with weights, instead of dual variables, which were mentioned in Sections 3.4 and 4.5 can be easily extended to this case. In particular the independence of link weights and capacities makes many of the computations easier, since once a solution is known (i.e., saturation sets and allowable paths) most of the calculations can be made with saturation matrices with link capacities equal to one. Only in a few computations will we need the actual capacity value and for those cases we can use the transformations of Section 4.5.

Throughout these sections we have seen how the results of Chapters III and IV generalize to the second saturation problem. The remaining part of this chapter will be devoted to a generalization of results to lower saturation levels, but first we present some examples that will illustrate some of this section results for the second level.

Example 5.4.1

Consider Example 5.3.1, that is a three node case whose successive saturation problem consists of two levels. The nonzero dual variable vectors for each level are

$$\pi^{11} = (-1/3, -1/3, -1/3)$$

$$\pi^{21} = (1/6, 1/6, -1/3)$$

$$\pi^{22} = (-1/2, -1/2) ,$$

where

$$S^1 = \{b,d,g\}$$

$$S^2 = \{a,c\} \ ,$$

and furthermore

$$z^1(F_+^1) = z^1(f_d(1),\ f_g(2),\ f_a(3)) = (1/3,\ 1/3,\ 1/3)$$

$$z^2(F_+^1) = \qquad\qquad\qquad\qquad (1/3,\ 1/3,\ 1/3) \ .$$

Substituting these values in (5.4.18) we have

$$\bar{\eta}^{2+} = -\frac{1}{2} \ ,$$

and using now (5.4.19) - (5.4.21)

$$\bar{\pi}^{21+} = (1/6,\ 1/6,\ -1/3) + \frac{1}{2}(-1/3,\ -1/3,\ -1/3)$$

$$= (0,\ 0,\ -1/2) \ ,$$

$$\bar{\pi}^{22+} = (-\frac{1}{2},\ -\frac{1}{2}) \ ,$$

$$\bar{\sigma}_2^{2+}(1) = 0, \quad \bar{\sigma}_3^{2+}(1) = \frac{1}{2} \ , \quad \bar{\sigma}_1^{2+}(2) = \frac{1}{2} \ ,$$

$$\bar{\sigma}_3^{2+}(2) = \frac{1}{2}, \quad \bar{\sigma}_1^{2+}(3) = 0, \quad \bar{\sigma}_2^{2+}(3) = 0 \ ,$$

$$\bar{z}^{2+}(F_+^1) = (\frac{1}{2},\ \frac{1}{2},\ \frac{1}{2}) > 0 \ .$$

Since $\bar{z}^{2+}(F_+^1) > 0$, $\{-\bar{\pi}_\ell^{2+}\}$ is a valid distance assignment. To obtain the corresponding integer assignment we transform $\bar{\pi}^{2+}$ into an integer vector according to (5.4.24) and we obtain

$$\theta_a = \theta_c = \theta_g = 1,$$

$$\theta_b = \theta_d = \theta_e = 0 \ .$$

Looking at Fig. E-5.4.1 we can see that this distance assignment precisely defines the allowable paths of Example 5.3.1 (Fig. E.5.3.1)

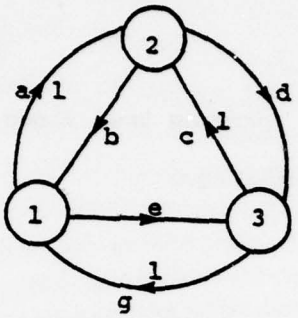
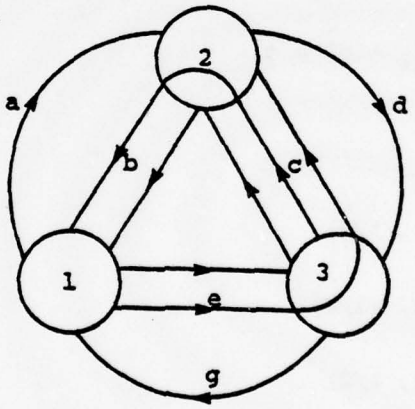



Fig. E.5.4.1

Example 5.4.2

Consider now the three node example whose set of allowable paths obtained by solving the successive saturation problem (two levels) is given by Fig. E.5.4.2a and whose link saturated sets are $S^1 = \{b,g\}$, $S^2 = \{a,c\}$.




$c_\ell = 1$, all $\ell$

Fig. E.5.4.2a

From Fig. E.5.4.2a we can easily obtain (by means of the canonic equations, for instance) the link optimal dual variables for the two levels. These dual variables are

$$\pi^{11} = (-1/2, -1/2),$$

$$\pi^{21} = (1/4, -1/4),$$

$$\pi^{22} = (-1/2, -1/2) .$$

If we choose the set of spanning tree flows to be the same as $F_{BT}$ of Example 5.3.1, then we will have

$$F^1_{BA} = \{f_c(1)\},$$

$$F^1_{+} = \{f_d(2), f_b(3)\},$$

$$F^1_{0} = \{f_d(1), f_e(2), f_a(3)\},$$

$$F^2_{BA} = \{f_e(2)\},$$

$$F^2_{+} = \{f_d(1), f_a(3)\},$$

$$F^2_{0} = F^1_{0} - F^2_{BA} - F^2_{+} = \phi,$$

and

$$z^1(F^1_{+}) = (1/2, 1/2)$$

$$z^2(F^1_{+}) = (1/4, -1/4)$$

$$z^2(F^2_{+}) = (1/2, 1/2) .$$

Using (5.4.18)-(5.4.21) with previous values we obtain

$$\bar{\eta}^{2+} = -\frac{1}{2},$$

$$\bar{\pi}^{21+} = (0, -1/2),$$

$$\bar{\pi}^{22+} = (-1/2, -1/2),$$

$$\bar{z}^{2+}(F^1_{+}) = (1/2, 0) \geq 0 .$$

Since now $\bar{z}^{2+}(f_b(3)) = 0$, where $f_b(3) \in F_+^1$, we cannot use $\{-\bar{\pi}_{\ell}^{2+}\}$ as a valid distance assignment since for this assignment $f_b(3)$ will create an allowable path that is forbidden from the solution of (LP1). To treat this anomaly we use Theorem 5.4.3 which says that we have to consider $k > 1$ in Theorem 5.4.2 to obtain an appropriate distance assignment. Suppose we choose $k = 1 + w$, where $w > 0$, then

$$\eta^{2+} = (1+w)\bar{\eta}^{2+} .$$

Using now $\eta^{2+}$ instead of $\bar{\eta}^{2+}$ in (5.4.19) and (5.4.21) we have

$$\pi^{21+} = \bar{\pi}^{21+} - \bar{\eta}^{2+}{}_w\, \pi^{11}$$

$$= (-\tfrac{w}{4}\,,\ -\tfrac{1}{2} - \tfrac{w}{4}) \ ,$$

$$\pi^{22+} = \bar{\pi}^{22+} = (-1/2, -1/2),$$

$$z^{2+}(F_+^1) = \bar{z}^{2+}(F_+^1) - \bar{\eta}^{21}{}_w\, z^1(F_+^1)$$

$$= (\tfrac{1}{2} + \tfrac{w}{4},\ \tfrac{w}{4}) \ .$$

Thus we see that $w > 0$ implies $z^{2+}(F_+^1) > 0$ and therefore $\pi^{2+}$ is a valid distance assignment. To obtain an integer assignment we first consider the denominators of the fractions that do not include w, and multiply by the minimum common multiple. The result will be

$$\tilde{\pi}^{21+} = (\tfrac{w}{2},\ 1 + \tfrac{w}{2}),$$

$$\tilde{\pi}^{22+} = (1,1) \ ,$$

and the smallest integer vector is obtained when we make w = 2, giving the distance assignment

$$\theta_b = 1,\ \theta_g = 2,\ \theta_a = \theta_c = 1,\ \theta_e = \theta_d = 0\ .$$

Using now Fig. E.5.4.2b we can easily see that this assignment characterize those and only those allowable paths given in Fig. E.5.4.2a

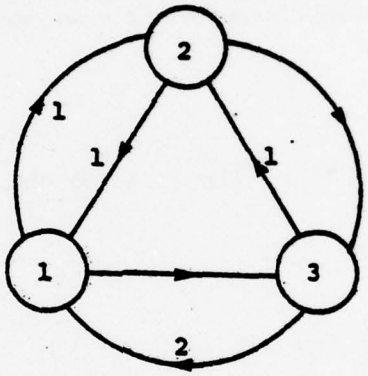



Fig. E.5.4.2b

## 5.5 The Third and Lower Saturation Problems

From previous sections it is apparent how everything studied so far generalizes and applies to any saturation level. From a conceptual point of view only two problems have to be studied, since problem 3, or any lower problem, presents no significant difference with respect to the second problem. Specifically, if we consider a net with $S^1$ and $S^2$, then these two sets are indistinguishable from the point of view of determining $S^3$. There are, nevertheless, some aspects of the third and lower problems that deserve special attention and this is the reason for this section. We are interested on the generalization of previous distance assignment methods to these lower problems; in

particular, we will consider how the degrees of freedom (or parameters) introduced by lower problems interact with those corresponding to higher problems. Furthermore we also want to introduce some general notation and make explicit some structural properties that will be used in Chapter VI when we study the dynamic of the successive problem relative to requirement changes. Some further notation and properties will be introduced in Section 5.6.

For the third saturation level we can write the corresponding problems (LP3), (DP3), ($LP3^+$) and ($DP3^+$) very easily. In particular problem (LP3) will be

$$\text{(LP3):} \qquad \min \alpha^1 - \beta^2 - \beta^3$$

$$A^3 x = b \tag{5.5.1}$$

$$x \geq 0$$

where

$$x^T = (\alpha^1, \beta^2, \beta^3, s^T(\bar{S}^{12}), (F^{12})^T) .$$

Here $\beta^3$ is the slack between $\alpha^3$ and $\alpha_0^2$, $s^T(\bar{S}^{12})$ is the vector of link slacks corresponding to the links of $L - S^1 \cup S^2$, $F^{12} \triangleq F - F_+^1 - F_+^2$, and b is the same RHS as in (LP1) and (LP2). $A^3$ is a matrix obtained from $A^2$ by dropping the columns corresponding to links of $S^2$ and the flows of $F_+^2$, and adding the column corresponding to variable $\beta^3$; this column will have zeros in all positions except in those related to the links of $\bar{S}^{12}$, whose entry will be the link capacity.

The dual of (LP3) is

(DP3):

$$\max \lambda^3 b$$
$$\lambda^3 A^3 \leq \gamma^3 , \tag{5.5.2}$$

where $\lambda^3$ has the same structure as $\lambda^1$ and $\lambda^2$, and $\gamma^3 = (1, -1, -1, 0, \ldots 0)$.

The complementary slackness conditions for the pair ( LP3) - (DP3) will be

Theorem 5.5.1

Let x and $\lambda^3$ be feasible solutions for (LP3) and (DP3) respectively. A necessary and sufficient condition that they both be optimal solutions is that for all $\ell \in L$, all $n \in N$

i)

$$\begin{aligned} \pi^3_\ell < 0 &\Rightarrow s_\ell = 0 \\ \pi^3_\ell = 0 &\Leftarrow s_\ell > 0 \end{aligned} \qquad \ell \notin s^1 \cup s^2 \tag{5.5.3}$$

$\pi^3_\ell$ unrestricted in sign, all $\ell \in s^1 \cup s^2$ ,

ii)

$$\begin{aligned} \sum_{\ell \in s^i} \pi^3_\ell c_\ell = 0 \quad , \quad i = 1,2 , \\ \sum_{\ell \in s^3} \pi^3_\ell c_\ell = -1 \quad , \end{aligned} \tag{5.5.4}$$

iii)

$$f_\ell(n) > 0 \Rightarrow -\pi^3_\ell = \sigma^3_{o(\ell)}(n) - \sigma^3_{t(\ell)}(n) . \tag{5.5.5}$$

□

Now we can define the third saturation set $S^3$ as

$$S^3 \triangleq \{\ell | \pi^3_\ell < 0, \quad \ell \in \bar{S}^{12}\} \ ,$$
$$\bar{S}^{12} \triangleq L - S^1 \cup S^2 = L - S^{12}, \tag{5.5.6}$$

and similarly the commodities in $R^3$ as

$$R^3 \triangleq \{(m,n) | \sigma^3_m(n) > 0, \ (m,n) \in \bar{R}^{12}\},$$
$$\bar{R}^{12} \triangleq R - R^1 \cup R^2 = R - R^{12}. \tag{5.5.7}$$

Using the notation introduced in (5.4.11)-(5.4.12), extended to the third level, we can write the analog of Lemmas 5.4.1-2 as follows (we will also use $\sigma^{3i}_m(n)$ instead of only $\sigma^3_m(n)$ to emphasize that $(m,n) \in R^i$).

Lemma 5.5.1

For any allowable path, $P^3(m,n)$, for commodity $(m,n)$ corresponding to an optimal solution of (LP3) we have

$$\sigma^3_m(n) = -\pi^{33} p^{33}(m,n) \geq 0, \quad (m,n) \in \bar{R}^{12} \ ,$$
$$\sigma^3_m(n) = -\pi^{32} p^{32}(m,n) - \pi^{33} p^{33}(m,n), \quad (m,n) \in R^2, \tag{5.5.8}$$
$$\sigma^3_m(n) = -\sum_{i=1}^{3} \pi^{3i} p^{3i}(m,n) \ , \quad (m,n) \in R^1,$$

where for $(m,n) \in R^{12}$ the corresponding $\sigma_m(n)$ can be negative.

□

In the same way, we can extend to this problem the remaining results of previous chapters. With respect to the pair $(LP3^+)$ - $(DP3^+)$ note that now $(LP3^+)$ will have two constraints more than (LP3), that is

$$\alpha^1 + \delta^1 = \alpha_0^1 ,$$
$$\alpha^1 - \beta^2 + \delta^2 = \alpha_0^2 , \tag{5.5.9}$$

and consequently ($DP3^+$) will have two variables, $\eta^{31+}$ and $\eta^{32+}$, more than (DP3). To each of these new variables there will correspond a relation obtained from the complementary slackness applied to ($LP3^+$) - ($DP3^+$). These relations are

$$\sum_{\ell \in S^1} \pi_\ell^{3+} c_\ell = \eta^{31+} ,$$
$$\sum_{\ell \in S^2} \pi_\ell^{3+} c_\ell = \eta^{32+} . \tag{5.5.10}$$

The remaining slackness conditions are identical to those of Theorem 5.5.1, except for the fact that now $\pi_\ell^{3+} \leq 0$, all $\ell \in L$. Thus we see that every level problem adds a new constraint, and therefore a new dual variable and a new slackness condition, to the corresponding "+" version of the problem. To obtain a distance assignment we have to use again the variables of ($DP3^+$), but following a similar procedure to that of Section 5.4 we can write the solution of ($DP3^+$) in terms of the solutions of (DP1), (DP2) and (DP3). Observe, though, that, since we already know the solution of ($DP2^+$) and also that the optimal solution set for ($DP3^+$) is a subset of the solution set of ($DP2^+$), it will be easier to relate the solution of ($DP3^+$) to the solutions of (DP1) and (DP2) via the solution of ($DP2^+$). In this respect we can write a new dual variables, following the same ideas as in (5.4.14), as follows:

$$\pi_\ell^{3\eta} = \pi_\ell^3 + \eta\pi_\ell^{2+}, \qquad \text{all } \ell \in S^1 \cup S^2 ,$$
$$\pi_\ell^{3\eta} = \pi_\ell^3 , \qquad \text{otherwise} . \tag{5.5.11}$$

Using (5.2.22) and (5.5.4) we have

$$\sum_{\ell \in S^1} \pi_\ell^{3\eta} c_\ell = \eta\, \eta^{2+},$$

$$\sum_{\ell \in S^1} \pi_\ell^{3\eta} c_\ell = -\eta\ . \tag{5.5.12}$$

Thus if we choose $\eta$ in such a way that $\pi^{3\eta} \leq 0$ and $z^{3\eta}(F_+^1 \cup F_+^2) \geq 0$, then, (5.5.11) will correspond to an optimal solution of $(DP3^+)$. It is easy to see that the minimum value of $\eta$ to have such properties will be

$$-\bar{\eta} = \min\left\{ \min_{\ell \in S^{12}} \frac{\pi_\ell^3}{\pi_\ell^{2+}},\ \min_{f_\ell(n) \in F_+^{12}} \frac{z^3(f_\ell(n))}{z^{2+}(f_\ell(n))} \right\}, \tag{5.5.13}$$

where

$$F_+^{12} = F_+^1 \cup F_+^2$$

and

$$z^i(f_\ell(n)) = -\pi^i c_\ell(n)\ , \qquad i = 3,\ 2+\ .$$

We could write a theorem completely equivalent to Theorem 5.4.2 saying that if $\bar{\eta}$ gives a solution then $k\bar{\eta}$, for any real k, $k \geq 1$, gives another solution. If, following the steps of Section 5.4 we write

$$\eta^{3+} = -k\bar{\eta}, \quad k \geq 1 \tag{5.5.14}$$

then we can write the link variables of an optimal solution of $(DP3^+)$ as

$$\pi_{\ell}^{3+} = \pi_{\ell}^{3} - \eta^{3+}\,\pi_{\ell}^{2+}, \quad \text{all } \ell \in S^{12},$$
$$\pi_{\ell}^{3+} = \pi_{\ell}^{3}, \quad \text{otherwise}, \tag{5.5.15}$$

and similarly for the $\sigma$ variables. Substituting (5.4.18) - (5.4.19) in (5.5.15) we will obtain $\pi_{\ell}^{3+}$ as a explicit function of $\pi_{\ell}^{1}$, $\pi_{\ell}^{2}$ and $\pi_{\ell}^{3}$. Also, from (5.5.14), (5.5.12) and (5.5.10) we have

$$\eta^{32+} = \eta^{3+}\,; \quad \eta^{31+} = -\eta^{3+}\eta^{2+}\,. \tag{5.5.16}$$

Using now these results we could write the lemmas and theorems corresponding to Lemmas 5.4.7-9 and Theorems 5.4.3-4 of Section 5.4. The only difference will be that in this case the optimal solutions of ($DP3^{+}$), and therefore the corresponding distance assignments, will be determined by three variables each one of them moving over a different (but related with respect to the determination of their origins) ray.

From this study it should become apparent how all these results generalize to lower problems. As a sample of what we are saying we will write below some of these generalizations, and will end the Section by presenting a three level problem example.

Lemma 5.5.2

Given the link dual variables of problem (LPk) the commodity dual variables are given by

$$\sigma_{m}^{k}(n) = -\sum_{j=i}^{k} \sum_{\ell \in P^{kj}(m,n)} \pi_{\ell}^{kj}$$
$$= -\sum_{j=i}^{k} \pi^{kj} p^{kj}(m,n) \tag{5.5.17}$$

where $P^k(m,n)$ is any allowable path for commodity (m,n) as given by the k-th problem and

$$P^{kj}(m,n) \triangleq P^k(m,n) \cap S^j \tag{5.5.18}$$

□

Observe that the summands of expressions (5.5.17) will, in general, be different depending on the allowable path $P^k(m,n)$. This is so because not all allowable paths, except for i = k, go through links of the same levels. Nevertheless the value $\sigma_m^k(n)$ is the same no matter which path is used, and is only a function of k, i and (m,n).

We can also generalize Lemmas 3.1.1 and 5.4.3 in the following way

Lemma 5.5.3

$\sigma_m^k(n)$ can be interpreted as the sensitivity factor of $\alpha_0^k$ for changes in commodity $(m,n) \in R^i$, $i \leq k$

$$\Delta\alpha_0^k = \sigma_m^k(n)\ \Delta r_m(n) \tag{5.5.19}$$

□

Another important and extremely useful result is the generalization of Lemma 5.3.3 given below.

Lemma 5.5.4

The link dual variables of problem (LPk) are given by

$$\pi^{ki} = -\zeta^{ki}\ \Omega^i\ , \qquad i < k \tag{5.5.20}$$

where

$$(\Omega^i)^{-1} = \left[\frac{\pi^{ii}}{\Omega^i}\right]\ , \qquad \text{all } i \tag{5.5.21}$$

and

$$\zeta^{ki} = \pi^{kk} \bar{P}_B^{ik}, \qquad i < k \tag{5.5.22}$$

$\bar{P}_B^{ik}$ is the graph matrix representing the part of the circuits of variables $F_{BA}^{i}$ corresponding to links of $S^{k}$. Thus

$\zeta^{ki}$ is the vector of weights of the reduced circuits, with respect to $S^{i}$, of the variables of $F_{BA}^{i}$ when the link weights are given by $\pi^{kk}$.

Proof: By induction from Lemma 5.3.3, using the same procedure used to prove that Lemma.

Q.E.D.

The appeal of this result is that we can obtain the dual variables without having to compute the G-matrices of the tableau. This fact can save an enormous amount of computation.

An equivalent way to express the result of Lemma 5.5.4 is via the global, or composite, saturation matrix. The new result will correspond to the generalization of Lemma 5.3.1, and is stated below. Since the proof of this generalized result is completely analogous to that of Lemma 5.3.1 we will not include it here.

Lemma 5.5.5

For the global saturation matrix of the successive saturation problem reordered as in (5.5.23) below, the corresponding inverse will have the structure given in (5.5.24) where $P_B^{1\ldots M}$ is the matrix of reduced circuit of the elements of $F_{BA}^{1\ldots M} \triangleq F_{BA}^{1} \cup \ldots \cup F_{BA}^{M}$ with respect

$$
\Omega^{1..M} = \begin{array}{|c|c|c|c|c|c|}
\hline
-c^1 & 0 & 0 & \cdots & 0 & \\
\hline
-c^2 & c^2 & 0 & \cdots & 0 & \\
\hline
-c^3 & c^3 & c^3 & \cdots & 0 & \\
\hline
\vdots & \vdots & \vdots & \vdots & \vdots & \\
\hline
-c^M & c^M & c^M & \cdots & c^M & \\
\hline
\end{array}
\qquad P_B^{1..M} \qquad \alpha^1 \quad \beta^2 \quad \beta^3 \quad \cdots \quad \beta^M \quad F_{BA}^1 \cup F_{BA}^2 \cup \cdots \cup F_{BA}^M \tag{5.5.23}
$$

$$
(\Omega^{1..M})^{-1} = \begin{array}{|c|c|c|c|c|}
\hline
\pi^{11} & & & \cdots & \\
\hline
\pi^{11}-\pi^{21} & -\pi^{22} & & \cdots & \\
\hline
\pi^{21}-\pi^{31} & \pi^{22}-\pi^{32} & -\pi^{33} & \cdots & \\
\hline
\vdots & \vdots & \vdots & \vdots & \vdots \\
\hline
\pi^{(M-1)1}-\pi^{M1} & \pi^{(M-1)2}-\pi^{M2} & \pi^{(M-1)3}-\pi^{M3} & \cdots & -\pi^{MM} \\
\hline
\end{array}
\qquad \Omega^{1..M} \tag{5.5.24}
$$

to the links of $S^{1..M} \triangleq S^1 \cup .. \cup S^M$ and $\Omega^{1..M}$ is a full rank matrix.

□

From (5.5.24) we see that to obtain the optimal dual variables corresponding to all levels of the successive problem we only have to compute the first M rows of matrix $(Q^{1..M})^{-1}$.

To end this section we will point out that we could easily extend the results of Section 4.4 to the global saturation matrices. Thus we could write

Definition 5.5.1

A matrix with the structure of (5.5.23) is a valid global saturation matrix for problem (LPM) iff its inverse has the structure of (5.5.24), where the elements of the first M rows have to satisfy all the properties corresponding to optimal dual variables.

It is very simple to see that Lemmas 4.4.1 and 4.4.2 generalize to this case. Thus we have that each column of $P_B^{1..M}$ must have at least a -1 and a +1 as components.

Now we will present an example that will illustrate many of these results for a three level problem.

Example 5.5.1

Consider Example 5.4.2 with the allowable paths, corresponding to a basic solution, given in Fig. E.5.4.2a but with a second saturation set $S^2 = \{a,e\}$ instead of the set $\{a,c\}$ given there. From this solution we can easily obtain the following dual variables

$$\pi^{11} = (\pi_b^1, \pi_g^1) = (-\tfrac{1}{2}, -\tfrac{1}{2})$$

$$\pi^{21} = (\pi_b^2, \pi_g^2) = (0, 0)$$

$$\pi^{22} = (\pi_a^2, \pi_e^2) = (-\tfrac{1}{2}, -\tfrac{1}{2})$$

Chossing the same $F_{BT}$ set as in Example 5.4.2 we can write

$$F_{BA}^1 = \{f_c(1)\}$$

$$F_+^1 = \{f_g(2), f_b(3)\}$$

$$F_0^1 = \{f_d(1), f_e(2), f_a(3)\}$$

$$F_{BA}^2 = \{f_e(2)\}$$

$$F_+^2 = \phi$$

$$F_0^2 = \{f_d(1), f_a(3)\} \neq \phi$$

and

$$z^1(F_+^1) = (\tfrac{1}{2}, \tfrac{1}{2})$$

$$z^2(F_+^1) = (\tfrac{1}{2}, \tfrac{1}{2})$$

Thus the solution to (LP2) is not unique and either $f_d(1)$ or $f_a(3)$ can be made basics without increasing $\alpha_0^1$ or $\alpha_0^2$. Therefore we have to solve a third saturation problem. To do that will mean reducing the saturation level of links d or c by using the alternate paths generated by $f_d(1)$ or $f_a(3)$. Note, however, from Fig. E.5.4.2a, that neither $f_d(1)$ nor $f_a(3)$ can reduce the flow of links d or c; on the contrary, they will increase

those flows if we make them (either of them or both) nonzero. Thus in this case it turns out that the solution of the third level problem corresponds to a basic solution of the second problem.

In terms of the linear programming formulation, the third problem will start with the basic solution of (LP2), adding variable $\beta^3$ and deleting the columns corresponding to $F_+^2$ (i.e., none). The first step afterwards will be to make $\beta^3$ basic, in exchange for the link of $\bar{S}^{12}$ = {c,d} with highest saturation (assume, for instance, that this link is link c). After making $\beta^3$ basic the new saturated link, link c, will have, according to (5.5.4), a nonzero dual variable

$$\pi_c^3 = -\frac{1}{c_c} = -1 .$$

To obtain the remaining dual variables we can write the composite saturation matrix [see (5.5.23)] and compute the third row of its inverse in accordance with (5.5.24). For the set $F_{BT}$ as chosen the composite matrix will be

$$Q^{123} = \begin{array}{c} \begin{matrix} \alpha^1 & \beta^2 & \beta^3 & c^1 & e^2 \end{matrix} \\ \begin{bmatrix} -1 & 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & -1 & 0 \\ -1 & 1 & 0 & 0 & -1 \\ -1 & 1 & 0 & 0 & 1 \\ -1 & 1 & 1 & 1 & 1 \end{bmatrix} \end{array} \begin{matrix} b \\ g \\ a \\ e \\ c \end{matrix}$$

where $c^1$ and $e^2$ stand for $f_c(1)$ and $f_e(2)$ respectively. Applying, for instance, the row reduction method of Section 4.6 we obtain for the third row of $(Q^{123})^{-1}$ the value

$$(Q^{123})^{-1}(\underline{r}3) = (-\tfrac{1}{2}, \tfrac{1}{2}, 0, -1, 1),$$

so that from (5.5.24) and the value of $\pi^2$ obtained before we have

$$\pi^{31} = (\tfrac{1}{2}, -\tfrac{1}{2}) ,$$

$$\pi^{32} = (-\tfrac{1}{2}, \tfrac{1}{2}),$$

$$\pi^{33} = \pi^3_c = -1 ,$$

$$\pi^3_d = 0 .$$

For these dual variables

$$z^3(f_d(1)) = -\pi^3 c_d(1) = 1,$$

$$z^3(f_a(3)) = -\pi^3 c_a(3) = 1,$$

where

$$c_d(1) = \overset{b\ \ g\ \ a\ \ e\ \ c\ \ d}{(-1,\ 1,\ 0,\ 0,\ 0,\ 1)^T},$$

$$c_a(3) = (0,\ 0,\ 1,-1,\ 0,\ 1)^T,$$

so that

$$F^3_+ = \{f_d(1),\ f_a(3)\};\quad F^3_0 = F^2_0 - F^3_+ = \phi$$

and the problem is finished. In this case [see next section] we will say that the third saturation set is a singleton.

Now we want to obtain a distance assignment. According to (5.5.15) we first need $\pi^{2+}$. Since in this case $\pi^{21} = 0$ we can immediately write

$$\pi^{21+} = -\eta^{2+}\pi^{11},$$

$$\pi^{22+} = \pi^{22},$$

and substituting in (5.5.15)

$$\pi^{31+} = \pi^{31} + \eta^{3+}\eta^{2+}\pi^{11},$$

$$\pi^{32+} = \pi^{32} - \eta^{3+}\pi^{22},$$

$$\pi^{33+} = \pi^{33} = \pi_c^3 = -1,$$

$$\pi_d^{3+} = 0.$$

If we call, for simplicity,

$$u \triangleq \eta^{3+}\eta^{2+}, \qquad v \triangleq -\eta^{3+},$$

we can write

$$\pi^{31+} = \left(\tfrac{1}{2}, -\tfrac{1}{2}\right) + u\left(-\tfrac{1}{2}, -\tfrac{1}{2}\right) = \tfrac{-1}{2}(u-1, u+1)$$

$$\pi^{32+} = \left(-\tfrac{1}{2}, \tfrac{1}{2}\right) + v\left(-\tfrac{1}{2}, -\tfrac{1}{2}\right) = \tfrac{-1}{2}(v+1, v-1),$$

and

$$z^3(F_+^1) = -\pi^{3+}C(F_+^1) = \left(\frac{u+v}{2}, \frac{u+v-2}{2}\right)$$

$$z^3(F_+^2) = \phi$$

$$z^3(F_+^3) = -\pi^{3+}C(F_+^3) = (1,1),$$

where

$$C(F_+^1) = \begin{bmatrix} C_g(2) \\ C_b(3) \end{bmatrix}^T = \begin{bmatrix} 0 & 1 & 1 & 0 & -1 & 0 \\ 1 & 0 & 0 & 1 & 0 & -1 \end{bmatrix}^T,$$

and similarly for $C(F_+^3)$.

Since we are looking for a dual vector with the property that $\pi^{3+} \leq 0$, $z^{3+}(F_+^{12}) > 0$, then any real pair $u \geq 1$, $v \geq 1$, where u and v cannot be simultaneously equal to 1, will give an appropriate solution. The reason why u and v cannot be simultaneously equal to 1 is to keep $z^3(f_b(3))$ from being equal to zero and therefore appearing as an allowable path. If we make, for instance, (u,v) = (1,3), then, we obtain the integer distance assignment

$$\theta_a = 2,\ \theta_g = \theta_e = \theta_c = 1,\ \theta_b = \theta_d = 0\ .$$

We see with the help of Fig. E.5.5.1 that the assignment describes those and only those paths of Fig. E.5.4.2a, i.e., those corresponding to the optimal solution of (LP3)

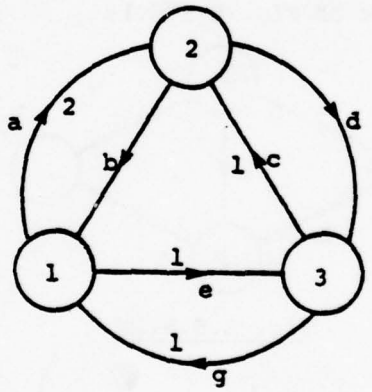



Fig. E.5.5.1

Observe that the solution of Fig. E.5.4.2b can be obtained by using the values (u,v) = (3,1).

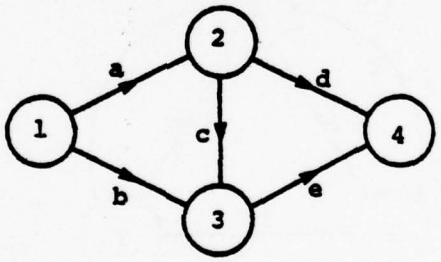
## 5.6 Other Results and Comments Relative to the Successive Saturation Problem

The purpose of this section is to study a few remaining general aspects of the successive problem. As we have seen throughout this

chapter, the action of one level on a previous level is, on the one hand, to reduce the size of the optimal space, and, on the other hand, to enlarge the number of basic alternate paths. Furthermore one level can also affect previous levels by changing the composition of sets $F^i_{BA}$, since, in general, a problem will have several choices for $F^i_{BA}$, each one of them belonging to a different basic solution. A difficulty related to this point, occurs when all basic solutions for problem i give rise to a nonminimal set $F^i_{BA}$; that is, when all basic variables have a saturated link with a dual variable equal to zero. As promised earlier, we will briefly study this situation now. We first consider an example of such a case.

Example 5.6.1

Consider the network of Fig. E.5.6.1a




$c_\ell = 1$, all $\ell$

Fig. E.5.6.1a

with the following requirement values

$r_1(2) = 1, \quad r_1(3) = 2, \quad r_1(4) = 7$

$r_2(3) = 1, \quad r_2(4) = 4$

$r_3(4) = 1,$

and for which we want to solve the successive saturation problem. According to the methodology introduced in Chapter IV we first choose a set of spanning trees, one for each commodity. Let these spanning trees be those of Fig. E.5.6.1b and, thus,

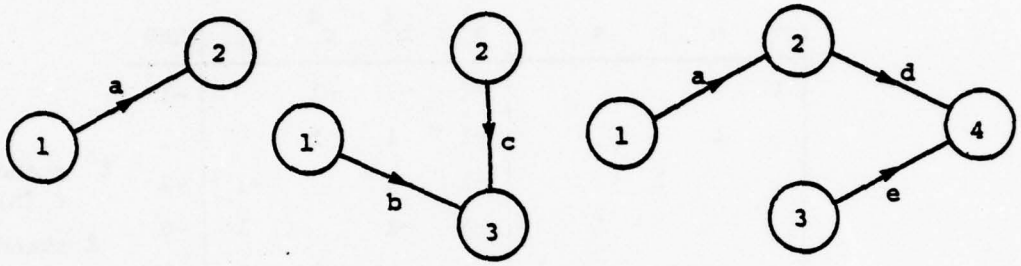



Fig. 3.5.6.1b

let

$$F_{BT} = \{f_a(2),\ f_b(3),\ f_c(3),\ f_a(4),\ f_d(4),\ f_e(4)\} \ .$$

The remaining flows, that is, those creating alternate paths, will be

$$F - F_{BT} = \{f_a(3),\ f_b(4),\ f_c(4)\} \ .$$

If only the variables of $F_{BT}$ are used to accommodate the requirements we will have the flow distribution of Fig. E.5.6.1c

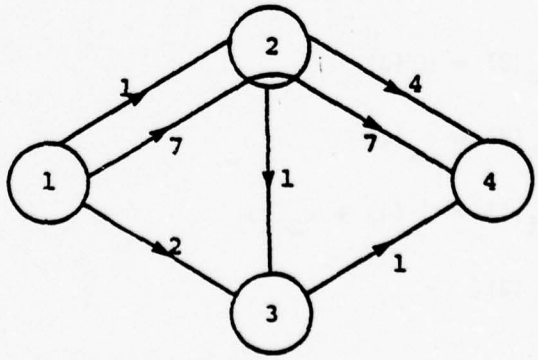



Fig. E.5.6.1c

This flow distribution corresponds to the updated constraint matrix obtained from the starting (LP1) matrix by making the elements of $F_{BT}$ basics i.e., by applying Step 1 of Section 4.1. The upper part of this updated matrix is given in Tableau E.5.6.1a,

| d | e | b | a | c | $\alpha^1$ | $b^4$ | $c^4$ | $a^3$ | RHS |
|---|---|---|---|---|---|---|---|---|---|
| 1 | | | | | -1 | -1 | -1 | | -11 |
| | 1 | | | | -1 | 1 | 1 | | -1 |
| | | 1 | | | -1 | 1 | | -1 | -2 |
| | | | 1 | | -1 | -1 | | 1 | -8 |
| | | | | 1 | -1 | | 1 | 1 | -1 |
| | | | | | 1 | | | | cost |

$\ell^n$ stands for $f_\ell(n)$

$\ell$ stands for $s_\ell$

T.E.5.6.1a

From T.E.5.6.1a it is easy to obtain an optimal solution for (LP1). Note that every time we make a change of basis, i.e., we make some element of $F-F_{BT}$ basic, we must check that the elements of $F_{BT}$ stay at a non-negative value; if not, we must change the composition of $F_{BT}$. The computation of the elements of $F_{BT}$ corresponding to a basic solution can be easily done by means of expression 4.2.39. In this particular case the computation is trivial since

$$f_a(3) + f_b(3) = r_1(3)$$

$$f_b(4) + f_a(4) = r_1(4)$$

$$f_c(4) + f_d(4) = f_a(4) + r_2(4)$$

$$f_a(2) = r_1(2),$$

etc.

An optimal solution to the (LP1) problem for this case is given in T.E.5.6.1b; since $F_{BT} > 0$, $F_{BT}$ is still a valid spanning tree set. (In fact, $F_{BT}$ will be valid throughout the example so, from now on, we will not be worried about it)

| d | e | b | a | c | $\alpha^1$ | $b^4$ | $c^4$ | $a^3$ | RHS |
|---|---|---|---|---|---|---|---|---|---|
| -1/2 | -1/2 | | | | 1 | | | | 6 |
| -1/2 | -1/2 | 1 | | | | 1 | | -1 | 4 |
| | 1 | -1 | | | | | 1 | 1 | 1 |
| -1 | -1 | 1 | 1 | | | | | | 2 |
| -1/2 | -2/3 | 1 | | 1 | | | | | 4 |
| 1/2 | 1/2 | | | | | | | | -6 = $-\alpha_0^1$ |

← $-\pi^{11}$ →

T.E. 56.1b

We see that the solution given in T.E.5.6.1b is non-minimal since $\pi_b = 0$. Furthermore the other possible basic solution, obtained by making $f_a(3)$ basic instead of $f_c(4)$, has the same property; hence we conclude that the problem is nonminimal. Suppose now that we make the slack of link b basic in T.E.5.6.1b, replacing either $f_c(4)$ or $f_a(3)$, whichever is basic. The result is given in Tableau T.E.5.6.1c, where we

| d | e | b | a | c | $\alpha^1$ | $b^4$ | $c^4$ | $a^3$ | |
|---|---|---|---|---|---|---|---|---|---|
| -1/2 | -1/2 | | | | 1 | | | | 6 |
| -1/2 | 1/2 | | | | | 1 | 1 | | 5 |
| | -1 | 1 | | | | | -1 | -1 | -1 |
| -1 | | | 1 | | | | 1 | 1 | 3 |
| -1/2 | -1/2 | | | 1 | | | 1 | 1 | 5 |
| 1/2 | 1/2 | | | | | | | | -6 |

T.E.5.6.1c

see that $s_b = -1 < 0$, but we also see that both $f_c(4)$ and $f_a(3)$, that can replace $s_b$ in the basis, have a zero cost. That means that in order to reduce the saturation level of link b below $\alpha_0^1$ we can use an alternate path that do not increase the level $\alpha_0^1$. This fact can be better seen with the help of Fig. E.5.6.1d, which represents the solution given in T.E.5.6.1c. In this figure we can use, for instance, the dotted path, which corresponds to $f_a(3)$, to reduce the flow over link b without augmenting the flow on links e or d. It is not difficult to interpret this fact

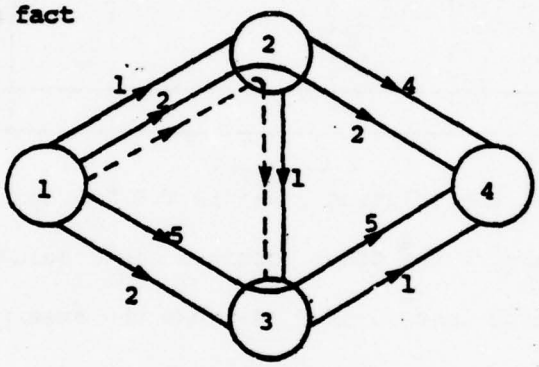



Fig. E.5.6.1d

by saying that link b should correspond to a lower saturation level. In this respect consider that we proceed to solve the second saturation problem over Tableau T.E.5.6.1c in spite $s_b < 0$. According to the results of Section 5.2 the starting tableau is given by T.E.5.6.1d (we keep also the links of $S^1 = \{d,e\}$ for the sake of completeness)

| d | e | b | a | c | $\alpha^1$ | $b^4$ | $\beta^2$ | $c^4$ | $a^3$ | |
|---|---|---|---|---|---|---|---|---|---|---|
| -1/2 | -1/2 | | | | 1 | | | | | 6 |
| -1/2 | 1/2 | | | | | 1 | | 1 | | 5 |
| | -1 | 1 | | | | | 1 | -1 | -1 | -1 |
| -1 | | | 1 | | | | 1 | 1 | -1 | 3 |
| -1/2 | -1/2 | | | 1 | | | 1 | 1 | 1 | 5 |
| 1/2 | 1/2 | | | | | | -1 | | | -6 |

T.E.5.6.1d

It is not difficult to obtain, from T.E.5.6.1d, an optimal solution to (LP2). Such a solution is given in T.E.5.6.1e and Fig. 5.6.1e. We see there that $S^2 = \{b,a\}$ and therefore that the second level problem took care of link b.

| d | e | b | a | c | $\alpha^1$ | $b^4$ | $\beta^2$ | $c^4$ | $a^3$ | |
|---|---|---|---|---|---|---|---|---|---|---|
| -1/2 | -1/2 | | | | 1 | | | | | 6 |
| | | | | | | 1 | | | -1 | 3 |
| -1/2 | -1/2 | 1/2 | 1/2 | | | | 1 | | | 1 |
| -1/2 | 1/2 | -1/2 | 1/2 | | | | | 1 | 1 | 2 |
| 1/2 | -1/2 | | -1 | 1 | | | | | | 2 |
| | | 1/2 | 1/2 | | | | | | | -5 |

T.E.5.6.1e

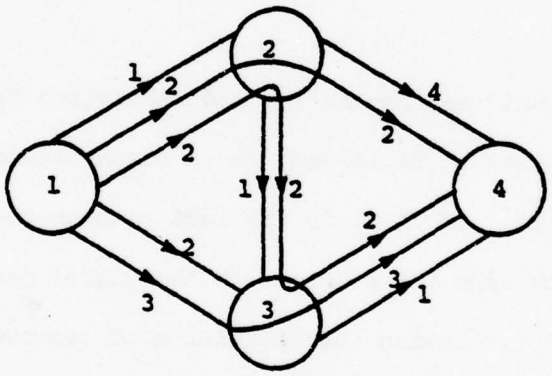



Fig. E.5.6.1e

To end this example observe that it corresponds to the case of (5.3.4) where $P_B^{21} \neq 0$, that is where the path (paths) used to balance the traffic in $S^2$ belong to a commodity of $R^1$. Note that in those cases the solution to the second saturation problem changes the value of the flow variables of $F_{BA}^1$.

□

It is not difficult to see that the procedure used in the example to overcome non-minimality can be easily generalized. This fact allows us to write

Lemma 5.6.1

Whenever we have a non-minimal i-th level problem, i.e., a problem all of whose optimal basic solutions are non-minimal, we can ignore the negativeness of the slack of the links with zero dual variable, and consider a Q matrix with only $|S^i|$ links. Continuing the optimization to lower problems, one of them will correct this negativeness. Acting in this way all previous results can be generalized to nonminimal problems.

□

As we said in Chapter II, and is obvious from the analysis, the number of problems to be solved, for a given network and requirement matrix is finite. If (LPM) is the last problem then $S^M$ must be a single point or else the solution to the global problem is not unique. In Section 4.3, following the definition of subcommodity (Definition 4.3.2), we introduced one of the causes for nonuniqueness as the possibility of having subcommodities within a network. The other possibility for nonuniqueness is given by Lemma 4.3.10. We can now see that this last nonuniqueness can also be caused by the existence of subcommodities due, not to the structure of the network itself but rather to the inaccessibility of links of $S^k$ to commodities of $R^{k+j}$, or to the particular structure of the solution. To see this fact consider, for instance, Tableau T.E.5.6.1e of Example E.5.6.1.

Since $f_a(3)$ can enter the basis only to replace $f_c(4)$ we see that, for this example, M=2 and the successive problem is not unique. The reason for the non-uniqueness lie in the fact that because of the solution we have infinite ways to apportion commodity (1,3) and, part of commodity, (1,4) via paths b or (a,c). Thus we can say that for this solution commodity (1,3) is a subcommodity of (a part of) commodity (1,4).

Summarizing we can say that the only cause of non-uniqueness is the possibility of having two or more commodities that can be arbitrarily apportioned among paths without modifying the saturation level of any link of the network. This fact can only be due to, either the existance of a network subcommodity (Definition 4.3.2), or to the existance of two flows satisfying Lemma 4.3.10. Whenever this last circumstance occurs we will say that we have a solution subcommodity. Thus we can write

Lemma 5.6.2

The successive saturation problem is non-unique iff there exists some network or solution subcommodities.

□

Whether a successive problem has a unique solution or not, in general we will have, once the last problem has been solved, some remaining links that do not belong to any $S^i$. We can consider these links as individual saturation sets with a saturation level given by the ratio between their flows and their capacities. Note,

though, that this is not the only possible way in which we can have saturation sets with only one element. Suppose we have solved the k-th level problem and found the optimal value $\alpha_o^k$. If link $\ell \in \bar{S}^{1..k}$ is the link with closest saturation level to $\alpha_0^k$ among the links of $\bar{S}^{1..k}$, then, the (k+1)-th problem will try first to reduce $\alpha_\ell$, the saturation level of link $\ell$. Suppose that all flows that could reduce the value of $\alpha_\ell$ have positive cost at some level $i \leq k$; then $\alpha_\ell$ cannot be further reduced and it will constitute a saturation level with an isolated link component. Because of this reason we will call this set a singleton. The Q-matrix of the singleton is just the negative value of the capacity of its element.

For completeness we now present some extensions of the results and bounds of Section 3.5. From Theorem 3.5.2 and using a procedure completely similar to the proof of Theorem 3.5.3 we can obtain

Theorem 5.6.1

The maximum number of alternate paths required by an optimal solution of the successive saturation problem will be $L-N-\nu_0$, where $\nu_0 = 0$ for a strongly connected network, and $\nu_0 = 1$ for a weakly connected network.

□

An immediate extention of this theorem is

Theorem 5.6.2

If M is the total number of problems solved in the hierarchy of the successive problem

$$\sum_{i=1}^{M} |S^i| \leq L - N + \nu_0 + M ,$$

where $\nu_0$ is defined in Theorem 5.6.1.

Proof:

From $|F_{BA}^i| = |S^i| - 1$ and Theorem 5.6.1.

Q.E.D.

Corollary 5.6.2a

If the bound of Theorem 5.6.2 is satisfied with equality then $M + \nu_0 < N$.

Proof:

Obvious from $\sum_{i=1}^{M} |S^i| < L$ .

Corollary 5.6.2b

The maximum number of saturated links that a single saturation level might have is $L - N + \nu_0 + 1$ and will correspond to $S^1$.

□

We will end this section and chapter with some comments on how each level modifies the partition of set F into corresponding subsets. After the first saturation problem is solved we have the first partition of F as follows

$F_{BT}^1$ = set of basic flows corresponding to the spanning trees,

$F_{BA}^1$ = set of basic flows generating alternate paths,

$F_0^1$ = set of nonbasic flows with zero partial cost

$= \{f_\ell(n) \,|\, z_\ell^{11}(n) = -\pi^{11} C_\ell^1(n) = 0$, all $\ell$, n,s.t. $f_\ell(n)$ is non basic$\}$,

$F_+^1$ = set of nonbasic flows with positive partial cost

$= \{f_\ell(n) \,|\, z_\ell^{11}(n) = -\pi^{11} C_\ell^1(n) > 0$, all $\ell$,n$\}$ .

Clearly, assuming a stable point, only $F_+^1$ will be invariant for all the optimal basic solutions of (LP1). The composition of the remaining subsets will depend on the particular optimal basic solution we choose. First we have the choice of the elements of $F_{BT}^1$ which, essentially, corresponds to a choice of coordinates. Given $F_{BT}^1$ and $\pi^{11}$ we can write

$$F_0^1 \cup F_{BA}^1 = \{f_\ell(n) \,|\, z_\ell^{11}(n) = \pi^{11} C_\ell^1(n) = 0, \quad \text{all } \ell,n\} .$$

Thus the partition of $F_0^1 \cup F_{BA}^1$ into $F_0^1$ and $F_{BA}^1$ is, somehow, arbitrary. Any subset of $|S^1| - 1$ independent elements of $F_0^1 \cup F_{BA}^1$ that could generate a valid saturation matrix will make a valid $F_{BA}^1$.

Consider now an optimal basic solution of the second saturation problem. We can similarly define subsets $F_{BT}^2$, $F_0^2$ and $F_+^2$ as follows

$F_{BT}^2$ = set of spanning tree flows corresponding to an optimal basic solution of the second level problem.

$F_0^2 = \{f_\ell(n) \,|\, z_\ell^2(n) = -\pi^2 C_\ell(n) = 0, \quad f_\ell(n) \notin F_+^1$ ,

$f_\ell(n)$ non basic$\}$

$F_+^2 = \{f_\ell(n) \,|\, z_\ell^2(n) = -\pi^2 C_\ell(n) > 0, \; f_\ell(n) \notin F_+^1\}$ .

We can define now a subset

$$F_{BA}^{12} = \{f_\ell(n) \mid f_\ell(n) \text{ basic}, f_\ell(n) \notin F_{BT}^2\}$$

= set of flows corresponding to the alternate paths of an optimal solution of (LP2),

where

$$|F_{BA}^{12}| = |S^1| + |S^2| - 2 .$$

Since the variables of $F_+^1$ are not active at (LP2) we have

$$F_{BT}^2 \cup F_{BA}^{12} \cup F_0^2 \cup F_+^2 = F - F_+^1 = F_{BT}^1 \cup F_{BA}^1 \cup F_0^1 .$$

Although (LP2) can change the structure of the spanning tree flows given by (LP1) we can immediately change coordinates in (LP1) and write the corresponding optimal solution in terms of an $F_{BT}^1 = F_{BT}^2$. This change of coordinates will affect, of course, the structure of $F_{BA}^1$ and $F_0^1$, which will exchange elements with the old $F_{BT}^1$. Therefore, without loss of generality and for simplicity of notation we will assume $F_{BT}^1 = F_{BT}^2 = F_{BT}$. In this case we have

$$F_{BA}^{12} \cup F_0^2 \cup F_+^2 = F_{BA}^1 \cup F_0^1 .$$

Note that $F_+^2$ can have elements of $F_{BA}^1$, that is, the second saturation problem could change the structure of $F_{BA}^1$. In this case $F_{BA}^1$ and $F_0^1$ will exchange elements. This situation could happen if $z_\ell^{22}(n) > 0$ for some $f_\ell(n) \in F_{BA}^1$. Thus $F_{BA}^1$ need not be a subset of $F_{BA}^{12}$. In any case, once we know $F_{BA}^{12}$ we can partition it into two subsets $F_{BA}^1$ and $F_{BA}^2$ such that $F_{BA}^1$ will generate a valid saturation matrix for (LP1) and $F_{BA}^2$ will do the same for (LP2) once the elements of $F_{BA}^1$ are made basic. Observe,

also, that once $F_{BT}$, $\pi^{11}$ and $\pi^{22}$ are known we can write, as before

$$F_0^2 \cup F_{BA}^{12} = \{f_\ell(n)\} | z_\ell^{11}(n) = z_\ell^2(n) = 0, \quad \text{all } \ell, n\} .$$

Thus $F_{BA}^{12}$ is any subset of $|S^1| + |S^2| - 2$ independent elements of $F_0^2 \cup F_{BA}^{12}$ that can form a valid global saturation matrix $Q^{12}$ as defined in (5.3.13).

This type of reasoning can be directly applied to the third and lower saturation problems, where any problem k could alter the composition of the sets of alternate paths of upper levels by properly exchanging elements of sets, say, $F_{BA}^j$ and $F_0^j$, $j < k$. If the successive saturation problem consists of M level problems we can write the following general set relationships, assuming $F_{BT}^1 = .. = F_{BT}^M = F_{BT}$,

$$\begin{aligned} F_{BA}^{12..k} \cup F_0^k \cup F_+^k &= F - (F_+^1 \cup F_+^2 \cup \ldots \cup F_+^{k-1}), \quad \text{all } 1 \leq k \leq M \\ &= F_{BA}^{12..(k-1)} \cup F_0^{k-1} , \end{aligned} \tag{5.6.1}$$

where

$$F_+^k = \{f_\ell(n) | z_\ell^k(n) > 0, \quad f_\ell(n) \notin F_+^1 \cup .. \cup F_+^{k-1}\} , \tag{5.6.2}$$

and

$$z_\ell^k(n) = -\pi^k C_\ell(n) = -\sum_{i=1}^{k} \pi^{ki} C_\ell^i(n) . \tag{5.6.3}$$

From (5.6.1) we can also write

$$\begin{aligned} F_{BA}^{12..k} \cup F_0^k &= F - (F_+^1 \cup .. \cup F_+^k) \\ &= \{f_\ell(n) | z_\ell^1(n) = \ldots = z_\ell^k(n) = 0, \quad \text{all } \ell, n\} . \end{aligned} \tag{5.6.4}$$

Note from (5.6.3), the definition of $Q^{1..k}$ given in (5.5.23) and the fact that $(Q^{1..k})^{-1} Q^{1..k} = I$, where $(Q^{1..k})^{-1}$ is given in (5.5.24), that

$$z_\ell^1(n) = \ldots = z_\ell^k(n) = 0 \iff wi[C_\ell(n) \cap S^{1..k}] = 0, \quad \text{all } 1 \leq i \leq k, \tag{5.6.5}$$

where wi is the i-th row of $(Q^{12..k})^{-1}$ and $S^{1..k} = S^1 \cup \ldots \cup S^k$. Thus (5.6.5) provides us with a way to write an alternate definition of $F_{BA}^{12..k} \cup F_0^k$.

Suppose now that we have a set $F_{BA}^{12..k}$ but we do not know a partition of this set into subsets $F_{BA}^1, F_{BA}^2, \ldots, F_{BA}^k$ and we want to find one such partition. The first condition that we have to impose on these subsets is

$$|F_{BA}^k| = |S^i| - 1, \qquad \text{all } 1 \leq i \leq M. \tag{5.6.6.}$$

To find a possible set $F_{BA}^1$ we have to look for a set of $|S^1| - 1$ elements of $F_{BA}^{12..k}$ that can form a valid first saturation matrix (see Section 4.4). This is an easy task, according to Section 4.4, and we can readily obtain a set $F_{BA}^1$. Given $F_{BA}^1$ we can then find a corresponding set $F_{BA}^2$ via one of two equivalent methods:

i) Making $F_{BA}^1 \cup F_{BA}^2$ generate a valid global saturation matrix for problem (LP2) according to Definition 5.5.1, or

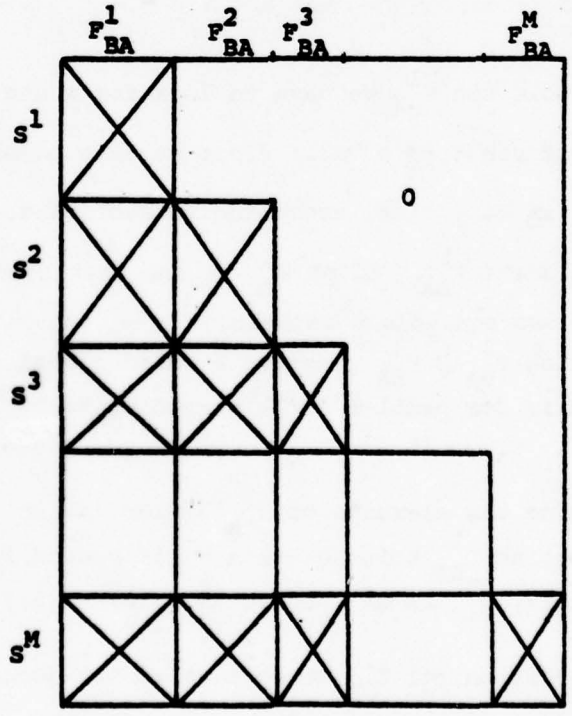
ii) Making the elements of $F_{BA}^1$ basics, as in (5.3.6), and choosing $F_{BA}^2$ to generate a valid second level saturation matrix, $Q_+^2$, as defined in (5.3.10) - (5.3.11).

A necessary condition for $F_{BA}^2$ to be a valid set according to method i) is that the circuit matrix $P_B^{12}$ must be full rank and each column

must have at least a +1 and a -1. Similarly a necessary condition, according to method ii), is that $P_B^2$ must be full rank and each column must have at least a positive and a negative component.

The procedure can be obviously generalized to lower levels. A case where the partition of $F_{BA}^{12..k}$ is immediate is when the columns of matrix $P_B^{12..M}$ can be reordered to give a lower block triangular matrix with the structure of (5.6.7) below. In this case, which for (LP2) corresponds to $P_B^{21} = 0$ (see Section 5.3), the solution of the partitioning problem is unique. In general, it can be expected that the partitioning problem will not fall either into the completely general case nor into the most simplified one. This means that we can expect to be able to reorder $P_B^{1..M}$ into a lower block triangular matrix where some of the upper diagonal blocks or subblocks will be nonzero.

$P_B^{1..M}$ =




(5.6.7)

# CHAPTER VI

# DYNAMIC ANALYSIS OF THE SUCCESSIVE SATURATION PROBLEM

## 6.1 Introduction

The object of this chapter is the study of the dynamic evolution of the successive problem optimal solution due to requirement changes. In Section 3.2 we saw that, for a stable point, there exists a certain margin of change of the requirements that will not cause a change of optimal basis. We called the linear region for a given basis the region where vector r and the optimal point can move without forcing a change of optimal basis. Although the definition of section 3.2 was introduced for the first saturation problem, it is clear from Chapter V that it generalizes to lower saturation problems. Because of this fact it seems obvious that we will have two different types of situations regarding the dynamic analysis of the successive problem: one will be the dynamic evolution of the optimal point within a linear region, and the other will be the reoptimization required when the system is forced to change linear region.

Observe that the successive problem as a whole, in spite of being a nonlinear problem, has also a linear region. This linear region will correspond to the optimal basis of the last saturation problem. If we consider the linear region for a given level as the region within which r can change without altering the optimality of the set of alternate paths $F^i_{BA}$ for that level, it is not difficult to see that the linear region for the global problem will be the intersection of the linear regions

of all existing level problems. Thus we can monitor the global problem by monitoring each problem individually; as soon as a level needs basis reoptimization, the global problem will have to change linear region. Therefore the study of the linear evolution of the optimal point will be focused on the study of the evolution of individual levels.

To enumerate the causes that can induce a change of basis consider the different type of basic variables that exist at any optimal solution. We have basic variables corresponding to the value of the flow on alternate paths, which we called routing or flow variables in Section 3.6. One also has the basic variables corresponding to the slacks between the value of two consecutive saturation levels, and finally one has the basic variables corresponding to the spanning trees. Therefore there are three different causes that can induce a change of basis, depending upon which type of variable is becoming negative. The physical reasons behind these three causes can be identified as follows:

i) A slack becoming negative means that two saturation levels are coalescing and the level that was below will surpass the level that was on top, thereby gaining the right to use the saturated links of this level.

ii) A flow variable becoming negative implies that a saturation level is breaking apart.

iii) Finally, a spanning tree flow becoming negative will mean a change in the structure of this spanning tree.

We will study all these three causes in the coming sections of this chapter.

## 6.2 Optimal Control and Information Flow Within a Linear Region

In this and the next sections we will study the necessary computations and information flow needed for tracking the optimal solution within a linear region.

Suppose that the optimal solution of the successive problem generates an M-level problem and that node m receives an increment in commodity $(m,n) \in R^k$ of value $\Delta r_m(n)$. Consider also that this value is not enough to cause a change of optimal basis. We want to study the operations that the system has to perform in order to reach the new optimal point. Assuming a central control, the reoptimization process will have the following steps:

i) Node m will send the value $\Delta r_m(n)$, specifying the commodity, to the central control

ii) The central control will compute, according to expressions (4.2.39), (5.3.29), or their analogs for lower levels, the new value of all the variables of $F_{BA}^i$, $1 \leq i \leq M$, that are affected by that change.

iii) The central control will send the information on the new values to those nodes whose flow variables have changed value.

iv) The nodes that receive the information proceed to implement the necessary changes.

Now we will study each step in more detail. Step i) presents no special difficulties. With respect to step ii) let us, first, obtain a general expression for $F_{BA}^i$.

Since the solution, of the successive problem provides a valid partitioning of alternate paths into subsets $F_{BA}^i$, $1 \leq i \leq M$, we can draw a simplified tableau including: a) the circuits corresponding to

the variables of $F^i_{BA}$, all i, b) the capacity columns corresponding to saturation levels and slacks between consecutive levels, and c) the RHS vector resulting from the completion of step 1 of Section 4.2. This tableau will have the form shown below, where we choose M = 3 for simplicity of notation:

| | $\alpha^1$ \| $F^1_{BA}$ | $\beta^2$ \| $F^2_{BA}$ | $\beta^3$ \| $F^3_{BA}$ | RHS |
|---|---|---|---|---|
| $s^1$ | $1_Q{}^1$ | $1_\phi{}^{21}$ | $1_\phi{}^{31}$ | $1_b{}^1$ |
| $s^2$ | $1_\phi{}^{12}$ | $1_{Q_+}{}^2$ | $1_\phi{}^{32}$ | $1_b{}^2$ |
| $s^3$ | $1_\phi{}^{13}$ | $1_{\phi_+}{}^{23}$ | $1_{Q_+}{}^3$ | $1_b{}^3$ |

(6.2.1)

Here, for i > j

$$1_\phi{}^{ij} = [-c^i | P_B^{ij}]; \quad 1_{\phi_+}{}^{ij} = [c^i | P_B^{ij}]; \quad 1_\phi{}^{ji} = [0 | P_B^{ji}]$$

$$1_{Q_+}{}^i = [c^i | P_B^{ii}] \tag{6.2.2}$$

$$P_B^{ij} = C(F^i_{BA}) \cap s^j ,$$

where $C(F^i_{BA})$ is the circuit matrix of the variables of $F^i_{BA}$ and, $-1_b{}^i \triangleq f^i_T \triangleq f_T(s^i)$ is the aggregate flow of the links of $s^i$ assuming all requirements flow only through links of spanning trees (see Lemma 4.2.4 and its corollary).

The left superscript 1 in all matrices means that these matrices correspond to a tableau obtained after step 1 of Section 4.2 is completed but before the variables of $\{\alpha^1, F^1_{BA}\}$ have become basics. In general,

a left superscript i will mean that the tableau corresponds to the case when step 1 has been completed and, furthermore, the variables of sets $\{\alpha^1, F^1_{BA}\} \ldots \{\alpha^{i-1}, F^{i-1}_{BA}\}$ have become basics. Thus, for instance

$$ {}^{i}Q^{i}_{+} = Q^{i}_{+} \tag{6.2.3} $$

where $Q^i_+$ represents the saturation matrix for level i, that is where $(Q^i_+)^{-1}(\underline{r}1) = \pi^i(S^i)$. Hence we have from (6.2.3) and tableau (6.2.1) that

$$ {}^{1}\Phi^{ji} = 0 \text{ all } i < j \Rightarrow Q^{j}_{+} = {}^{1}Q^{j}_{+} \tag{6.2.4} $$

In terms of computation, (6.2.4) is, of course, a very important simplifying property that we must exploit whenever it is present. In words, (6.2.4) describes the case in which the routing of upper-level commodities will impact the routing at level j, but the routing of level j does not have impact back on the routing at upper levels. Similar simplifying rules can be obtained, in the other direction, whenever some ${}^{1}\Phi^{ij} = 0$ for some, or all, $i < j$. To make these simplifying results more apparent let us study first the case where the $\Phi$-matrices of the upper diagonal part of the tableau (6.2.1) are zero.

Case I $({}^{1}\Phi^{ji} = 0$ all $i < j$, all $j$, $2 \leq j \leq M)$

If we make $\alpha^1$, $F^1_{BA}$, $\beta^i$, $F^i_{BA}$, all $2 \leq i \leq M$, basic in (6.2.1) the optimal value of these new basic variables, after some matrix manipulation and using the results of Chapters IV and V, is given by

Lemma 6.2.1

The value of the basic flows generating alternate paths for commodities of level i, under the assumption of case I, are given by

$$F^i_{BA} = -\Omega^i [f^i_T + \sum_{j=1}^{i-1} P^{ji}_B F^j_{BA}] , \quad i \geq 1 \tag{6.2.5}$$

□

Expression (6.2.5) states clearly the hierarchical ordering of the $\{F^i_{BA}\}$ for this case. If we assume now an increment $\Delta r_m(n)$, where $(m,n) \in R^k$, then clearly this increment cannot affect the variables of $F^j_{BA}$, $j < k$. The sequence of expressions that will give the correct terms for the $F^i_{BA}$ (and thus the new optimal point) can be easily obtained from (6.2.5) and Lemma 4.2.4. These results are given by the following lemma.

Lemma 6.2.2

The increment on the value of the routing variables generated by an increment $\Delta r_m(n)$, $(m,n) \in R^k$, under the assumptions of case I, is

$$\Delta F^k_{BA} = -\Omega^k \Delta f^k_T$$

$$\Delta F^i_{BA} = -\Omega^i [\Delta f^i_T + \sum_{j=1}^{i-1} P^{ji}_B \Delta F^j_{BA}], \quad \text{all } i > k$$

where the components of $\Delta f^i_T$ are

$$\Delta f^i_\ell = \begin{cases} \Delta r_m(n) , & \text{all } \ell \in T(m,n) \cap s^i \\ 0 , & \text{otherwise} \end{cases} \tag{6.2.7}$$

□

Thus to compute (6.2.6) we need, in principle(*), all inverses $(Q^i)^{-1}$,

(*) Some $\Delta f^i_T$ and $\Delta F^j_{BA}$ could be equal to zero, reducing the number of equations.

$i \geq k$, and the graph matrices and vectors $P_B^{ji}$ and $T(m,n)$. Hence, instead of working with a big inverse matrix we are able to perform all computations by means of small $q_i \times q_i$ matrices. The remaining calculations are performed by graph theoretic means.

Observe, from (6.2.6), that the cause for a $\Delta f_\ell^i(n) \neq 0$ can be two-fold: as a result of $\Delta f_T^i$ being non-zero, or as a result of the effect of previous level rearrangements expressed through the term $\Sigma P_B^{ji} \Delta F_{BA}^j$. Note that each row of $P_B^{ji}$ corresponds to a link of $S^i$; thus the product of the row corrsponding to link $\ell$ by the column vector $F_{BA}^j$ gives the net flow that the variables of the j-th problem add to the flow of link $\ell$.

For the sake of completeness we now obtain the new value of the saturation levels $\alpha_0^i$, $i \geq k$. Following an entirely analogous procedure to that used for the proofs of Lemmas 6.2.1-2 we could prove

Lemma 6.2.3

Given a solution to the successive saturation problems, the value of the various saturation levels are

$$\alpha_0^i = -\pi^{ii}[f_T^i + \sum_{j=1}^{i-1} P_B^{ji} F_{BA}^j], \quad i \geq 1 \quad . \tag{6.2.8}$$

Furthermore for a $\Delta r_m(n)$, $(m,n) \in R^k$ we have

$$\begin{aligned} \Delta\alpha_0^k &= -\pi^{kk} \Delta f_T^k \\ \Delta\alpha_0^i &= -\pi^{ii}[\Delta f_T^i + \sum_{j=1}^{i-1} P_B^{ji} \Delta F_{BA}^j], \quad i > k, \end{aligned} \tag{6.2.9}$$

where all vector and matrices are defined in Lemmas 6.2.1-2.

□

When the Φ-matrices of the upper triangular part are nonzero then Lemmas 6.2.1-2 are no longer valid for the computation of the new variables. We now study this general case.

Case II (general case)

For this case the ordering and partitioning given in tableau (6.2.1) proves to be inappropriate. We can instead use (5.5.23) and define a combined global saturation matrix as follows (we again choose M = 3 for simplicity)

$$Q^{123} = \begin{array}{c|ccc|c} & \alpha^1 & \beta^2 & \beta^3 & F^1_{BA} \cup F^2_{BA} \cup F^3_{BA} \\ \hline & -c^1 & 0 & 0 & \\ & -c^2 & c^2 & 0 & P^{123}_B \\ & -c^3 & c^3 & c^3 & \end{array} \quad ; \quad \begin{array}{c} \text{RHS} \\ \hline \\ -f^{123}_T \\ \\ \end{array} \qquad (6.2.10)$$

where $P^{123}_B$ is now the reduced circuit matrix of the variables of $F^1_{BA} \cup F^2_{BA} \cup F^3_{BA}$ with respect to the links of $S^1 \cup S^2 \cup S^3$. From Lemma 5.5.5 we have

$$(Q^{123})^{-1} = \begin{array}{|ccc|} \hline \pi^{11} & & \\ \pi^{11} - \pi^{21} & -\pi^{22} & \\ \pi^{21} - \pi^{31} & \pi^{22} - \pi^{32} & -\pi^{33} \\ \hline & \Omega^{123} & \\ & & \\ \hline \end{array} \qquad (6.2.11)$$

Thus the optimal value of the flow of the alternate paths will be given by

$$F_{BA}^{123} = -\Omega^{123} f_T^{123} , \tag{6.2.12}$$

where now $F_{BA}^{123}$ is the vector whose components are those of $F_{BA}^1 \cup F_{BA}^2 \cup F_{BA}^3$.

Similarly, for the optimal value of the saturation levels

$$\begin{aligned} \alpha_0^1 &= -\pi^1(S^1) f_T(S^1) , \\ \alpha_0^2 &= -\pi^2(S^1 \cup S^2) f_T(S^1 \cup S^2) , \\ \alpha_0^3 &= -\pi^3(S^1 \cup S^2 \cup S^3) f_T(S^1 \cup S^2 \cup S^3) . \end{aligned} \tag{6.2.13}$$

Considering now an increment in $\Delta r_m(n)$ the corresponding change in the flow variables will be given by

$$\Delta F_{BA}^{123} = -\Omega^{123} \Delta f_T^{123} ; \tag{6.2.14}$$

in particular for $f_\ell(n) \in F_{BA}^{123}$

$$\Delta f_\ell(n) = -\Omega^{123}(\underline{r}\ell n) \Delta f_T^{123} . \tag{6.2.15}$$

Thus the main difference between the general and simplified case is that whereas in the simplified case we can perform the reoptimization via a collection of small saturation matrices for the general case we have to work with a single, larger, composite matrix.

In general it might be expected that a problem will neither fall into the simplified case not into the more complicated one, where all $\Phi$-matrices are nonzero. If it happened, for instance, that matrices

${}^{1}\phi^{21}$ and ${}^{1}\phi^{31}$ of (6.2.1) are zero matrices, then we could use a composite matrix for the links of $S^2$ and $S^3$ and use the simplified method between $\Omega^1$ and this composite matrix.

Observe from (6.2.15) that since $f_T$ and $\Omega^{123}$ are constant for a given linear region we can give the information contained in $f_T$ and $\Omega^{123}$ $(\underline{r}\ell n)$ to the node $o(\ell)$ and then this node can compute $\Delta f_\ell(n)$ for a given change $\Delta f_T$. This procedure, which is also applicable to the simplified case (for this case the corresponding $\Omega(\underline{r}\ell n)$, and $f_T$ vectors will be smaller) can give rise to a decentralized control procedure within a linear region, in which the corresponding increment $\Delta r_m(n)$ is broadcast to nodes at the origin of alternate paths everytime there is a change in some requirements.

The change in the saturation levels, due to some $\Delta r_m(n)$, can be obtained from (6.2.13) or from general expressions derived from the generalization given below of Lemmas 3.1.1 and 5.4.3:

Lemma 6.2.4

For a $\Delta r_m(n) \neq 0$, $(m,n) \in R^k$, the increment in the saturation level of the various problems is given by

$$\Delta\alpha_0^i = 0 \ , \qquad i < k$$

$$\Delta\alpha_0^i = \sigma_m^{ik}(n) \ \Delta r_m(n) \ , \quad i \geq k$$

□

Finally, it deserves to be mentioned that all this information, regarding changes within a linear region due to requirement changes, could have been obtained, similarly, via the canonic equations, see

Section 3.6. In this case we will have a collection of canonic equations for each level, where some of the routing variables could appear in more than one level. It is clear that when the number of links and levels grows the canonic equation procedure becomes very inefficent, and it is better to concentrate the necessary information into the saturation matrices and $f_T$ vector.

With this we conclude the analysis of step ii) of the four step classification put forth at the beginning of this section. With respect to steps iii) and iv) we make the following observations:

a) Step iii) is essentially a problem of efficient coding, and it will not be treated here. For a very recent and excellent work in the area of coding for communication networks see [43]. We also assume that this updating traffic does not increase the saturation level of the links, which is equivalent to assuming that there exists a collection of small special channels used for this purpose.

b) Regarding step iv) we first recall the assumption made in Chapter II considering the flow and requirement variables as continuous variables. If this is not the case and the network is, for instance, a telephone network, then we could think of a suboptimal procedure that will work at a point close to the optimal solution while still maintaining the integrity of the calls. This problem will not be studied in this thesis.

c) Another problem concerning step iv) is the possibility of needing some ordering in the implementation of the necessary changes. This will be the case if some $\Delta f_\ell(n)$ is more negative than the actual value of the existing flow

from $0(\ell)$ to n. In this case, since we assumed that $\Delta r_m$ was not enough to cause a change of basis, node $o(\ell)$ cannot implement the necessary change completely until some upstream node implements its changes, so that the flow $(o(\ell), n)$ increases. The same ordering problem will appear if by implementing a computed $\Delta f_\ell(n)$ node $o(\ell)$ would exceed the capacity of link $\ell$. These situations will be, nevertheless, very easily detected and simple to solve by waiting until some upstream node acts.

We end this section with some comments about the stability of the reoptimization within a linear region. Assuming that the network is reliable (i.e., no link or node failures, good error control, etc.). The main cause of inestability will be due to the simultaneous, or almost simultaneous, appearance of several requirement changes in the network. If only one increment, $\Delta r_m(n)$, appears at a time and the separation between successive increments (we obviously consider positive or negative increments) is long enough to allow for a complete settlement of the changes caused by the first increment, then, from previous sections, we see that there is no reason for inestability.

Assuming a centralized control procedure the settling problem can be partially solved as follows: Let us assume that an increment $\Delta r_m(n)$ appears in the network while the processing of another change is taking place. Since the changes may cause opposite results we may decide either to implement the first change, or to wait until the second increment is processed and implement the results of both changes simultaneously. It seems clear that in this second case we have to fix a maximum time of response to a given requirement change. Otherwise, if too many increments come too close to each other the control may keep waiting until

the "next" change is processed before implementing any new routing thereby introducing an intolerable response delay for those changes that arrived first.

## 6.3 Change of Optimal Basis

In Section 6.1 we enumerated the causes that may induce a basis change. In this section we will study each one of them separately, with the assumption that a central control is in command of the network operation.

### 6.3.1 A Slack Between Consecutive Saturation Levels Becomes Negative (I)

This situation corresponds, obviously, to the case of a saturation level first coalescing and then surpassing the next higher saturation level. Suppose that because of some requirement changes we find

$$\beta_0^{k+1} = \alpha_0^k - \alpha_0^{k+1} < 0 \quad . \tag{6.3.1}$$

This means that level k+1 can now use links of $S^k$ to decrease its saturation level $\alpha_0^{k+1}$. Thus in terms of the successive saturation problem we have to solve a new k-th level problem, since, in general, the old (k+1)-th level solution will not be optimal for the new k-th level. To solve the new k-th level problem we need a solution of the corresponding (k-1)-th problem. If, we assume that (6.3.1) is the only reason for non-optimality, or if this non-optimality is simultaneously caused by other negative basic variables belonging to saturation levels lower than k-1, then the present solution of the (k-1)-th level problem, existing within the actual optimal basis of the successive problem, is

still optimal for the new requirement vector. Thus the corresponding existing composite saturation matrix that includes all routing variables and saturation slacks up to level k-1 is still optimal. Observe that if one does not have a partition of $F_{BA}^{1..M}$ into subsets $F_{BA}^{1},\ldots, F_{BA}^{M}$ we know, from Section 5.6, how to obtain it. Proceeding in this way and assuming, for simplicity, that $F_{BT}$, the spanning tree flows, remain invariants during the whole reoptimization process (we will study later on in this chapter how to handle changes in $F_{BT}$) the starting tableau of the new k-th level linear programming problem will be Tableau T.VI.II below, which is obtained from Tableau T.VI.I by:

- making $\alpha^1$, $\beta^2$, ..., $\beta^{k-1}$ and the elements of $F_{BA}^{1\ldots(k-1)}$ basics
- deleting the elements of $F_{+}^{k-1}$, that is, all the elements of the (k-1)-th level optimal tableau with positive cost.
- adjoining the column corresponding to the slack variable $\beta^k$.

In Tableau T.VI.I, the matrix $\underline{Q^{k-1}} \triangleq Q^{1..(k-1)}$ is defined in (5.5.23), $\underline{\phi^{k-1}}$ is given by

$$\underline{\phi^{k-1}} = \begin{array}{c|c|c|c|c} \alpha^1 & \beta^2 & \ldots & \beta^{k-1} & F_{BA}^{\underline{k-1}} \\ \hline -\underline{c}^{\underline{k-1}} & \underline{c}^{\underline{k-1}} & \ldots & \underline{c}^{\underline{k-1}} & \underline{P}_{BA}^{\underline{k-1}} \end{array} \qquad (6.3.1)$$

where

$$\underline{c}^{\underline{k-1}} \triangleq c(\underline{S}^{\underline{k-1}}) \qquad (6.3.2)$$

and $\underline{k}$ is a new symbol that denotes the sequence 1 2...k. All the P-matrices correspond to circuit matrices as defined in Chapter IV, and

| $s^{k-1}$ | $\bar{s}^{k-1}$ | $\alpha^1\beta^2\ldots\beta^{k-1}F_{BA}\underline{\quad}^{k-1}$ | $F_0^{k-1}$ | $F_+^{k-1}$ | RHS |
|---|---|---|---|---|---|
| I | | $Q^{k-1}$ | $P_{B0}^{k-1}$ | $P_{B+}^{k-1}$ | $b^{k-1}$ |
| | I | $\Phi^{k-1}$ | $\bar{P}_{B0}^{k-1}$ | $\bar{P}_{B+}^{k-1}$ | $\bar{b}^{k-1}$ |
| | | 1-1...-1 | | | |

T.VI.I

| $s^{k-1}$ | $\bar{s}^{k-1}$ | $\alpha^1\beta^2\ldots\beta^{k-1}F_{BA}\underline{\quad}^{k-1}\beta^k$ | $F_0^{k-1}$ | RHS |
|---|---|---|---|---|
| $(Q^{k-1})^{-1}$ | | I | $P_0^{k-1}$ | $b_0^{k-1}$ |
| $G^{k-1}$ | I | | $\bar{P}_0^{k-1}$ | $\bar{b}_0^{k-1}$ |
| $-\pi^{k-1}(s^{k-1})$ | | -1 | | $\beta_0^{k-1}$ |

T.VI.II

$$b^{k-1} = -f_T(S^{\underline{k-1}}) \triangleq -f_T^{\underline{k-1}} \tag{6.3.3}$$

$$\bar{b}^{k-1} = -f_T(\bar{S}^{k-1}) \triangleq -\bar{f}_T^{\underline{k-1}} ,$$

where the vector $f_T$ is defined in Lemma 4.2.2. In T.VI.II $b_0^{k-1}$ is the vector that gives the optimal value of the variables $\alpha^1, \beta^2, \ldots, \beta^{k-1}$, $F_{BA}^{\underline{k-1}}$ for this optimal solution of the (k-1)-th level problem. From T.VI.I and T.VI.II we can easily write

$$b_0^{k-1} = (Q^{\underline{k-1}})^{-1} b^{k-1} . \tag{6.3.4}$$

Using a procedure similar to that used to prove Lemma 5.3.2 we could also write, from (6.3.3) and (6.3.4),

$$\begin{aligned} \alpha_0^1 &= -\pi^{11} f_T^1 , \\ \beta_0^i &= \alpha_0^{i-1} - \alpha_0^i , \qquad 2 \leq i \leq k-1 , \\ \alpha_0^i &= \pi^i b_0^{k-1} = -\pi^i (S^{\underline{i}}) f_T^{\underline{i}} , \\ F_{BA}^{\underline{k-1}} &= -\Omega^{\underline{k-1}} f_T^{\underline{k-1}} , \end{aligned} \tag{6.3.5}$$

where $\Omega^{\underline{k-1}}$ is the bottom submatrix of matrix $(Q^{\underline{k-1}})^{-1}$ as defined in (5.5.24). Thus these variables can be easily obtained from (6.3.5). It should be pointed out, as we did in Section 5.3 with respect to the results of Lemma 5.3.2, that (6.3.5) represents a significant reduction in the amount of computation relative to (6.3.4). Note that (6.3.4) will be the standard computational procedure that we would have to use to compute $b_0^{k-1}$ with a direct application of the revised simplex method. Thus, (6.3.5) is another illustration of how the structural

properties of this problem can be used to save computation.

In the same way we could compute vector $\bar{b}_0^{k-1}$, which gives the value of the slack of the saturation level of the links of $\bar{s}^{\underline{k-1}}$, with respect to $\alpha_0^{k-1}$, corresponding to this optimal solution of the (k-1)-th level problem. For this vector we can write

$$\bar{b}_0^{k-1} = \bar{b}^{k-1} - \phi^{\underline{k-1}} \; b_0^{k-1} \quad . \tag{6.3.6}$$

But from (6.3.1), (6.3.3) and (6.3.5) the element of $\bar{b}_0^{k-1}$ corresponding to link $\ell$ will be the slack variable

$$s_\ell = -f_{T\ell} - c_\ell(-\alpha_0^1+\beta_0^2+ \ldots +\beta_0^{k-1}) - \bar{P}_{BA}^{\underline{k-1}}(\underline{r}\ell) \; F_{BA}^{\underline{k-1}} = \tag{6.3.7}$$

$$= \alpha_0^{k-1} c_\ell - f_{T\ell} - \bar{P}_{BA}^{\underline{k-1}}(\underline{r}\ell) \quad F_{BA}^{\underline{k-1}} \quad .$$

In vector form

$$s(\bar{s}^{\underline{k-1}}) = \bar{b}_0^{k-1} = \alpha_0^{k-1} \bar{c}^{\underline{k-1}} - f^{k-1}(\bar{s}^{\underline{k-1}}), \tag{6.3.8}$$

where, similarly to (4.2.39),

$$f^{k-1}(\bar{s}^{\underline{k-1}}) = f_{T\ell} + \bar{P}_{BA}^{\underline{k-1}} \quad F_{BA}^{\underline{k-1}} \tag{6.3.9}$$

is the vector of values of the aggregate flows of the links of $\bar{s}^{\underline{k-1}}$ corresponding to the optimal solution of the (k-1)-th level problem. Again we can see that (6.3.8) represents a reduction of computation with respect ot (6.3.6).

Our motivation for computation of the optimal value of the RHS for the (k-1)-th level problem is not only to show how easily and efficiently it can be obtained, but also because this RHS is needed for the next

steps of the solution of the k-th level problem, as we shall now see.

The first operation we have to perform in T.VI.II is to make $\beta^k$ basic. The link of $\bar{S}^{\underline{k-1}}$ that will become non-basic, i.e., saturated at the k-th level, in exchange for variable $\beta^k$ will be, according to the exit rule of the simplex method, the link $a$ for which the ratio

$$\frac{s_a}{c_a} = \alpha_0^{k-1} - \frac{f_a^{k-1}}{c_a}, \qquad a \in \bar{S}^{\underline{k-1}}, \tag{6.3.10}$$

is minimum, i.e., the link $a$ with the highest-saturation level. The consequences, with respect to the key matrices and vectors of T.VI.II, of making $\beta^k$ basic, will be the following (we reorder rows and bring the row corresponding to link a immediately below $S^{\underline{k-1}}$).

- The new saturation matrix (we will call it $\tilde{Q}^{\underline{k}}$) will have one more column (the row corresponds to link $a$ and the column to variable $\beta^k$). The corresponding new inverse will be

$$(\tilde{Q}^{\underline{k}})^{-1} = \begin{array}{|c|c|} \hline (Q^{\underline{k-1}})^{-1} & \\ & 0 \\ \hline \frac{1}{c_a} G^{\underline{k-1}}(\underline{r}_a) & \frac{1}{c_a} \\ \hline \end{array} \tag{6.3.11}$$

  where we have left the column corresponding to $\beta^k$ in the rightmost position for simplicity of notation.

- The new RHS sub-vector of optimal values of the variables in the saturation matrix (which we will call $b^k$) will be

$$b^k = \begin{bmatrix} b^{k-1} \\ \hline 0 \\ \hline s_a/c_a \end{bmatrix}. \tag{6.3.12}$$

- The new RHS sub-vector of the value of the slack variables of the links of $\bar{S}^{k-1}$ - a, call it $\tilde{\bar{b}}^{k}$, will be

$$\tilde{\bar{b}}^{k} = \bar{b}_0^{k-1}(\bar{a}) - \frac{s_a}{c_a}\,\bar{c}^{k-1}(\bar{a}) \quad , \tag{6.3.13}$$

where $v(\bar{a})$ means vector v without the component corresponding to link $a$.

- The new vector of link dual variables, call it $\tilde{\pi}^{k}$, will be

$$\tilde{\pi}^{k}(S^{k-1}, a) = \left[\pi^{k-1}(S^{k-1}) - \frac{1}{c_a} G^{k-1}(\underline{ra}), \frac{-1}{c_a}\right] \tag{6.3.14}$$

and,

- The new partial cost vector for the variables of $F_0^{k-1}$ will be

$$\tilde{z}^{k}(F_0^{k-1}) = \frac{1}{c_a}\,\bar{P}_0^{k-1}(\underline{ra}) \quad . \tag{6.3.15}$$

Let us now manipulate some of the expressions in (6.3.11) - (6.3.15) to obtain more transparent results. First we have from T.VI.I and T.VI.II that

$$G^{k-1} = -\phi^{k-1}(Q^{k-1})^{-1} \quad . \tag{6.3.16}$$

But from (5.5.24) and (6.3.1)

$$G^{k-1} = \bar{c}^{k-1}\,\pi^{k-1}(S^{k-1}) - \bar{P}_{BA}^{k-1}\,\Omega^{k-1} \quad . \tag{6.3.17}$$

The row of $G^{k-1}$ corresponding to link $a$ will then be

$$G^{k-1}(\underline{ra}) = c_a \pi^{k-1}(S^{k-1}) - \bar{P}_{BA}^{k-1}(\underline{ra})\,\Omega^{k-1} \tag{6.3.18}$$

where, from the definition of the P-matrices, $\bar{P}_{BA}^{k-1}(\underline{ra})$ is the vector of coefficients of the circuits of $F_{BA}^{k-1}$ with respect to link $a$. Therefore, $G^{k-1}$ can be obtained from the composite saturation matrix and the circuits

of $P_{BA}^{k-1}$ ; substituting (6.3.18) in (6.3.14) we have

$$\tilde{\pi}^k (S^{k-1}, a) = \frac{1}{c_a} [\bar{P}_{BA}^{k-1} (\underline{ra}) \Omega^{k-1}, \ -1] \ . \tag{6.3.19}$$

With respect to (6.3.15) note that the cost vector can also be written as

$$\tilde{z}^k (F_0^{k-1}) = -\tilde{\pi}^k (\tilde{S}^k) \ \tilde{P}_{B0}^k \ , \tag{6.3.20}$$

where

$$\tilde{S}^k = S^{k-1} \cup a \tag{6.3.21}$$

and

$$\tilde{P}_{B0}^k = \begin{bmatrix} P_{B0}^{k-1} \\ \hline \bar{P}_{B0}^{k-1} (\underline{ra}) \end{bmatrix} \ . \tag{6.3.22}$$

Therefore $\tilde{z}^k (F_0^{k-1})$ can be easily obtained from (6.3.19) and a circuit matrix, without having to compute $\bar{P}_0^{k-1}$.

To proceed with the solution of the new k-th problem we now have to look for elements of $F_0^{k-1}$ with negative partial cost, as given by (6.3.20). Assuming that $f_\ell(n) \in F_0^{k-1}$ has the largest negative cost in (6.3.20), then $f_\ell(n)$ will become basic at the next step. To find the variable that must leave the basis we have to obtain, according to the exit rule of the simplex method, the smallest positive ratio between the elements of the updated RHS(*) and the updated column corresponding

(*) Observe that from (6.3.10), (6.3.12) and (6.3.13), $\tilde{b}^k \geq 0$, $\bar{\tilde{b}}^k \geq 0$, so that RHS $\geq 0$ and we can apply the simplex method.

to $f_\ell(n)$. Since we already know the updated RHS we have to compute the updated version of the column corresponding to $f_\ell(n)$. Let $p_{\ell n}^{k-1}$ and $\bar{p}_{\ell n}^{k-1}$ denote the upper and lower parts of this column in T.VI.II, then we can write

$$\begin{aligned} p_{\ell n}^{k-1} &= (Q^{\underline{k-1}})^{-1} \quad C_{\ell}^{\underline{k-1}}(n) \\ \bar{p}_{\ell n}^{k-1} &= \bar{C}_{\ell}^{\underline{k-1}}(n) - \phi^{\underline{k-1}} \; p_{\ell n}^{k-1} \end{aligned} \tag{6.3.23}$$

where

$$C_{\ell}^{\underline{k-1}}(n) = C_\ell(n) \cap S^{\underline{k-1}}$$

$$\bar{C}_{\ell}^{\underline{k-1}}(n) = C_\ell(n) \cap \bar{S}^{k-1}$$

and $C_\ell(n)$ is, as before, the circuit generated by variable $f_\ell(n)$.

We will now compute the updating of expressions (6.3.23) after $\beta^k$ has been made basic. Since now we have one more saturated link the upper-part-vector will have one component more and the lower-part-vector one less component. If $\tilde{p}_{\ell n}^{k}$ and $\tilde{\bar{p}}_{\ell n}^{k}$ denote these new subectors we have

$$\begin{aligned} \tilde{p}_{\ell n}^{k} &= \begin{bmatrix} p_{\ell n}^{k-1} \\ \hline \frac{1}{c_a} \bar{p}_{\ell n}^{k-1}(a) \end{bmatrix} + \beta^k \\ \tilde{\bar{p}}_{\ell n}^{k} &= \bar{p}_{\ell n}^{k-1}(\bar{a}) - \frac{1}{c_a} \; \bar{p}_{\ell n}^{k-1}(a) \; \bar{c}^{\underline{k-1}}(\bar{a}) \end{aligned} \tag{6.3.24}$$

where as before, $v(\bar{a})$ represents the subvector obtained from v by deleting the component corresponding to link $a$.

Next we have to look for the smallest positive ratio between the elements of the RHS and the elements of (6.3.24). We will assume, for

simplicity, that the new point is also stable and that $\tilde{\bar{b}}^k > 0$. In this case we have to look for all the positive elements of (6.3.24). Note though that since $f_\ell(n) \in F_0^{k-1}$, the first k-1 components of $\tilde{p}^k_{\ell n}$, i.e. the components corresponding to the first k-1 rows of $(Q^{\underline{k-1}})^{-1}$, will be equal to zero, see (5.6.5), so that we do not have to compute them. Note also that since $f_\ell(n) \in F_0^{\underline{k-1}}$, $\tilde{p}^k_{\ell n}(a) = \tilde{z}^k(f_\ell(n)) < 0$, so that (obviously) the recently made basic, variable $\beta^k$, cannot become non-basic. Hence the task of looking for positive elements in (6.3.24) is confined to the components associated with the basic variables $F_{BA}^{\underline{k-1}}$ and $s(\bar{S}^{\underline{k-1}} - a)$.

Suppose we now define the following sets

$$\tilde{F}^k_{\ell n} = \{f_j(m) \mid f_j(m) \in F_{BA}^{\underline{k-1}},\ \tilde{p}^k_{\ell n}(jm) > 0\}$$
$$\tilde{L}^k_{\ell n} = \{j \mid j \in \bar{S}^{\underline{k-1}} - a,\ \tilde{\bar{p}}^k_{\ell n} > 0\} \quad . \tag{6.3.25}$$

Then we will have two cases, depending on whether the variable u that must leave the basis belongs to $\tilde{F}^k_{\ell n}$ or $\tilde{L}^k_{\ell n}$. As we said before the variable leaving the basis is obtained according to the criterion

$$\xi_{\ell n}(u) \triangleq \frac{\tilde{b}(u)}{\tilde{p}_{\ell n}(u)} =$$
$$= \min \left\{ \min_{f_j(m) \in \tilde{F}^k_{\ell n}} \frac{\tilde{b}^k(f_j(m))}{\tilde{p}^k_{\ell n}(jm)},\ \min_{j \in \tilde{L}^k_{\ell n}} \frac{\tilde{b}^k(j)}{\tilde{p}^k_{\ell n}(j)} \right\}. \tag{6.3.26}$$

We shall study these two cases separately:

Case Ia

$$u = f_\lambda(\nu) \in \tilde{F}^k_{\ell n} \ .$$

Here variable $f_\ell(n)$ will replace variable $f_\lambda(\nu)$ (*) in the basis. The consequences of this exchange of variables are studied below. First we have a change in the composition of the two important sets

$$\begin{aligned} \tilde{F}^k_{BA} &\leftarrow F^{k-1}_{BA} - f_\lambda(\nu) + f_\ell(n) \\ \tilde{F}^{k-1}_{0} &\leftarrow F^{k-1}_{0} - f_\ell(n) + f_\lambda(\nu) \end{aligned} \qquad (6.3.27)$$

where this updating symbolism, which will be used extensively below, means that the new value of the quantity on the left is obtained doing the operations, with the old values and quantities, expressed in the right.

Secondly, we have a change in the structure of the saturation matrix, where

$$\tilde{Q}^k \leftarrow \tilde{Q}^k + [\tilde{C}_\ell{}^k(n) - \tilde{C}_\lambda{}^k(\nu)] \ [0,\ldots,0,\underset{\uparrow}{1},0,\ldots 0] \qquad (6.3.28)$$
$$f_\lambda(\nu)$$

and

$$\tilde{C}_j{}^k(m) \triangleq C_j(m) \cap \tilde{S}^k \qquad (6.3.29)$$

The new inverse matrix can be immediately obtained from the old inverse by applying Theorem AI.3 of Appendix I to expression (6.3.28). Observe that, since $f_\ell(n) \in F^{k-1}_0$, the dual variables of problems higher in the

(*) Note that $\lambda$ now represents a link and should not be confused with the vector of dual variables.

hierarchy than [i.e., the rows of $(\tilde{\underline{Q}}^k)^{-1}$ corresponding to $\alpha^1, \beta^2, \ldots, \beta^{k-1}$] will not change. On the other hand from (5.5.24) it follows that the new dual variables for problem k will be

$$\tilde{\pi}^k(\tilde{\underline{s}}^k) = -(\tilde{\underline{Q}}^k)^{-1}(\underline{r}\beta^k) + \pi^{k-1}(\tilde{\underline{s}}^k), \tag{6.3.30}$$

where $(\tilde{\underline{Q}}^k)^{-1}(\underline{r}\beta^k)$ is the row of the new inverse corresponding to variale $\beta^k$. We could now use (6.3.28), (6.3.11), (6.3.14) and (6.3.24) to explicitely write $(\tilde{\underline{Q}}^k)^{-1}(\underline{r}\beta^k)$ in terms of previously known vectors. Instead we will use an equivalent and simpler form for $\tilde{\pi}^k(\tilde{\underline{s}}^k)$. This equivalent form can be obtained by applying the basis change to the cost vector. It is not difficult to see that the updating of the dual variables obtained in this way is that represented in (6.3.31)

$$\tilde{\pi}^k(\tilde{\underline{s}}^k) \leftarrow \tilde{\pi}^k(\tilde{\underline{s}}^k) + \frac{\tilde{z}^k(f_\ell(n))}{\tilde{p}^k_{\ell n}(\lambda\nu)}\,\tilde{\Omega}^k(\underline{r}\lambda\nu), \tag{6.3.31}$$

where $\tilde{\Omega}^k(\underline{r}\lambda\nu)$ is the row of $\tilde{\Omega}^k$, or $(\tilde{\underline{Q}}^k)^{-1}$ of (6.3.11), corresponding to variable $f_\lambda(\nu)$. Since, from (6.3.11)

$$\tilde{\Omega}^k(\underline{r}\lambda\nu) = [\Omega^{k-1}(\underline{r}\lambda\nu),\ 0] \tag{6.3.32}$$

the dual variable corresponding to link $a$ does not change, as should be expected.

Finally, the updating of the RHS is also easily obtained, using (6.3.26), as follows:

$$\begin{aligned}
&\tilde{b}^k(\overline{\lambda\nu}) \leftarrow \tilde{b}^k(\overline{\lambda\nu}) - \xi_{\ell n}(\lambda\nu)\tilde{p}^k_{\ell n}(\overline{\lambda\nu}),\\
&\tilde{b}^k(\lambda\nu) \leftarrow \xi_{\ell n}(\lambda\nu),\\
&\tilde{\bar{b}}^k \leftarrow \tilde{\bar{b}} - \xi_{\ell n}(\lambda\nu)\tilde{\bar{p}}^k_{\ell n},
\end{aligned} \tag{6.3.33}$$

where $\tilde{b}^k(\overline{\lambda\nu})$ means vector $\tilde{b}^k$ without the element corresponding to variable $f_\lambda(\nu)$, and $\tilde{b}^k(\lambda\nu)$ represents this element.

Once the updating is finished the reopitmization continues with the new sets and values, in the same way as before. Now we will study the case when $u \in \tilde{L}^k_{\ell n}$ .

Case Ib

$$u = \lambda \in \tilde{L}^k_{\ell n} \quad .$$

In this case $f_\ell(n)$ replaces link $\lambda$ in the basis. Thus $\lambda$ becomes a saturated link. Because of this fact, the following sets will change composition

$$\begin{aligned} \tilde{S}^k &\leftarrow \tilde{S}^k + \lambda \ , \\ \tilde{F}^k_{BA} &\leftarrow \tilde{F}^k_{BA} + f_\ell(n), \\ \tilde{F}^{k-1}_{0} &\leftarrow F^{k-1}_{0} - f_\ell(n) \quad . \end{aligned} \tag{6.3.34}$$

The new composite saturation matrix will have one more row and one more column, i.e.

$$\tilde{Q}^k \leftarrow \begin{array}{|c|c|} \hline \tilde{Q}^k & \tilde{C}^k_{\ell n} \\ \hline \lambda(\tilde{B}^k) & \\ \hline \end{array} \tag{6.3.35}$$

where $\lambda(\tilde{B}^k)$ is the vector of coefficients of link $\lambda$ for the various basic variables, i.e., $\lambda(\beta^i) = c_\lambda$, $\lambda(f_j(m)) = \pm 1,0$ depending whether

or not $\lambda \in C_j(m)$ and in which direction. The new inverse matrix can be readily updated from the old inverse by applying Theorem AI.2 of Appendix I. As in case Ia the rows corresponding to variables $\alpha^1$, $\beta^2$, ..., $\beta^{k-1}$ will not change, except for one zero component added on the right. To obtain the new dual variables we can use, again, expression (6.3.30). It is somewhat simpler, though, to obtain an updating in the same way that (6.3.31) was obtained before, i.e. by making the cost vector reflect the basis change. In this way we obtain for this case

$$\tilde{\pi}^k(\tilde{S}^k) \leftarrow [\tilde{\pi}^k(\tilde{S}^k) + \tilde{\pi}^k_\lambda \ \tilde{G}^k(\underline{r}\lambda), \ \tilde{\pi}^k_\lambda], \tag{6.3.36}$$

where

$$\tilde{\pi}^k_\lambda = \frac{\tilde{z}^k(f_\ell(n))}{\tilde{p}^k_{\ell n}(\lambda)} \tag{6.3.37}$$

and, according to (6.3.18),

$$\tilde{G}^k(\underline{r}\lambda) = c_\lambda \tilde{\pi}^k(\tilde{S}^k) - \tilde{P}^k_{BA}(\underline{r}\lambda)\tilde{\Omega}^k \ . \tag{6.3.38}$$

Finally, using (6.3.26), we can also obtain the following updating for the RHS:

$$\begin{aligned} &\tilde{b}^k(\overline{\ell n}) \leftarrow \tilde{b}^k - \xi_{\ell n}(\lambda)\ \tilde{p}^k_{\ell n} \ , \\ &\tilde{\bar{b}}^k(\bar{\lambda}) \leftarrow \tilde{\bar{b}}^k(\bar{\lambda}) - \xi_{\ell n}(\lambda)\ \tilde{\bar{p}}^k_{\ell n}(\bar{\lambda}), \\ &\tilde{b}^k(\ell n) \leftarrow \xi_{\ell n}(\lambda) \ . \end{aligned} \tag{6.3.39}$$

The reoptimization, then, continues in the same way with new Ia and Ib cases until $\tilde{z}^k(\tilde{F}_0^{k-1}) \geq 0$. At this point the problem is solved and

we define the sets corresponding to the optimal solution as follows:

$$
\begin{aligned}
S^k &\leftarrow \tilde{S}^k , \\
F_{BA}^k &\leftarrow \tilde{F}_{BA}^k , \\
F_0^k &\leftarrow \{f_j(m) \,|\, \tilde{z}^k(f_j(m)) = 0\}, \\
&\text{etc.}
\end{aligned}
\tag{6.3.40}
$$

The solution to problem k+1 is then attempted, starting with these k-th level optimal values, in the same way as for problem k.

This concludes our main study of the computation required for basis changes of type I; there are, however, some additional comments on this type of reoptimization that deserve mentioning, and we will discuss them in the next section.

### 6.3.2 Further Comments on Case I

As we said at the beginning of Section 6.3.1, Case I refers to the reoptimization induced by a saturation level, say k+1, first coalescing with and then surpassing the saturation level immediately above it. As we saw there this reoptimization involves the solution, first, of the new k-th problem and, then, of the remaining problems below k. In Section 6.3.1, the solution of the k-th problem was undertaken starting with only that information provided by the solution of the (k-1)-th level problem. Note, however, that to solve the new k-th level problem we might also be able to use the information provided by the old (k+1) problem. Of course the variables and links corresponding to the old (k+1) level solution are not optimal at the k-th level because, roughly speaking, there are

more paths available now for the commodities of the old $R^{k+1}$. But we can expect that, in most cases, the information provided by the old (k+1)-th level may still be useful at the new k-th level problem. To make use of this information, the first step to be taken in the solution of the new k-th problem will be to implement the following updatings:

$$\tilde{S}^k \leftarrow S^{k+1},$$

$$\tilde{P}^k_{BA} \leftarrow P^{k+1}_{BA},$$

$$\tilde{Q}^k \leftarrow \begin{array}{c|c|c|c} & & \beta^k & P^{k+1}_{BA} \\ \hline Q^{\underline{k-1}} & & 0 & P_{\bar{B}}^{(k+1)\,(\underline{k-1})} \\ \hline \bar{P}_{B}^{(\underline{k-1})\,(k+1)} & & C^{k+1} & P^{k+1}_{B} \end{array} \quad S^{k+1} \qquad (6.3.41)$$

where the matrix notation used in the last updating corresponds to that introduced in (5.3.6). The repercussions of the updating can be obtained following the same steps and mechanism used before, in Section 6.3.1, except that now we make several variables basic at the same time, and therefore the pivot operation is centered in an $|S^k|$-square matrix instead of a single element. To obtain $(\tilde{Q}^k)^{-1}$ we can still use Theorem AI.2 of Appendix I. Observe the great simplification of this updating operation which would result if $P_{\bar{B}}^{(k+1)\,(\underline{k-1})}$ were equal to zero. In this case the updating of $(\tilde{Q}^k)^{-1}$ would take very much the form of the

third inversion procedure introduced in Section 4.6, and thus we would have all the hierarchical implications and simplifications studied there.

To obtain the updating on the dual variables we can use procedures similar to those of Section 6.3.1, or use Lemma 5.5.4 instead. In order to exercise all different possibilities we will briefly described now how Lemma 5.5.4 is applied here. Suppose that to invert $\tilde{Q}^k$ of (6.3.41) we use Theorem AI.1 of Appendix I and we break the inversion process into two steps, first applying the transformations to bring the identity to the place where $Q^{\underline{k-1}}$ is, and secondly applying the remaining transformations to bring the identity to the whole $\tilde{Q}^k$ matrix. The first step is equivalent to making a pivot at $Q^{\underline{k-1}}$, i.e., to making $\alpha^1, \beta^2, \ldots, \beta^{k-1}$, $F_{BA}^{\underline{k-1}}$ basic, and will transform $\tilde{Q}^k$ into the matrix of (6.3.42)

$$\left[\begin{array}{c|c} I & X \\ \hline 0 & Y \end{array}\right] , \tag{6.3.42}$$

where

$$\begin{aligned} X &= (Q^{\underline{k-1}})^{-1} [0 \,|\, P_B^{\underline{(k+1)\,(k-1)}}] , \\ Y &= [c^{k+1} \,|\, P_B^{k+1}] - P_B^{\underline{(k-1)\,(k+1)}} X . \end{aligned} \tag{6.3.43}$$

Observe that, since $F_{BA}^{k+1} \subset F_0^{k-1}$, the first k-1 rows of X will be equal to zero, assuming $Q^{\underline{k-1}}$ ordered as in (5.5.23), and do not have to be computed. Furthermore from (5.3.6) and (5.3.8) we have that $\tilde{Q}^k = Y$, so

$$Y^{-1} = \left[\begin{array}{c} -\tilde{\pi}^{kk} \\ \hline \tilde{\Omega}^{kk} \end{array}\right] . \tag{6.3.44}$$

Thus the new dual variables are

$$\tilde{\pi}^k = [\tilde{\pi}^{k1}, \ldots, \tilde{\pi}^{k(k-1)}, \tilde{\pi}^{kk}, 0, \ldots, 0] \ , \qquad (6.3.45)$$

where $\tilde{\pi}^{kk}$ is given in (6.3.44) and the $\tilde{\pi}^{ki}$, $i < k$, are obtained using (5.5.20). If $P_B^{(k+1)(\underline{k-1})} = 0$ then

$$\tilde{Q}^k = [c^{k+1} | P_B^{k+1}] \ , \qquad (6.3.46)$$

and it is not difficult to see that (6.3.45) agrees with the dual variable relations obtained in Section 4.6 for inversion Method III.

Finally, we have to update the RHS. The procedure is completely analogous to that followed in case Ib of Section 6.3.1, equations (6.3.39), except that now instead of a single link $\lambda$ we have a collection of links $s^{k+1}$, and instead of a single flow variable $f_\ell(n)$ we have the set $F_{BA}^{k+1}$; therefore, instead of a scalar $\xi_{\ell n}(\lambda)$ we have a vector.

$$\xi_{F_{BA}^{k+1}}(s^{k+1}) = Y^{-1} b^{k-1}(s^{k+1}) \ , \qquad (6.3.47)$$

where $b^{k-1}(s^{k+1})$ is the vector of link slack values for the links of $s^{k+1}$ corresponding to the optimal solution of (LP(k-1)) that has been selected.

An additional comment with respect to the RHS deserves mentioning at that point. Since in the previously described process one is simultaneously making all the variables of $F_{BA}^{k+1}$ basic, without previously making basic the variables of $F_{BA}^{k}$, it may happen that some components of the updated RHS ended being negative. In this case the problem looses primal feasibility and, thus, for the remaining part of the optimization, it also looses the appropriate structure for the application

of steps Ia or Ib above. Furthermore it can also be expected that some of the previously forbidden alternate paths for the old k+1 problem will reduce the saturation level of the new k-th problem, i.e. that $\tilde{z}^k(f_\ell(n)) < 0$ for some $f_\ell(n) \in F_0^{k-1} \cap F_+^k$. Thus, we will be in a situation where the new solution is neither primal nor dual feasible. To overcome the difficulty that these unfeasibilities state for the use of both the simplex and dual simplex method we can use, what is called, a "composite algorithm", see [24], [44] and [45]. This type of algorithm is widely used in computer programs to avoid the alternate utilization of primal and dual algorithm.

One of the basic ideas of a composite algorithm consists on the parametrization of the RHS, the cost vector or both. Assume, for instance, that we add a parameter $\delta(j)$ to each $\tilde{b}^k(j) < 0$, such that $\tilde{b}^k(j) + \delta(j) > 0$, then, we regain primal feasibility and, thus, one can use the simplex method. As soon as we reach dual feasibility we make all $\delta(j)$ equal to zero. If the RHS continues to be positive then, the solution is optimal if not we can apply the dual simplex method. Note that, since there are no more flow variables with negative cost, this "dual stage" of the re-optimization will increase the value of the objective function obtained after the "primal stage" is finished. We saw in Section 6.3.1 how the basis changes and updatings take place when we use a version of the revised primal simplex method. Following a completely parallel procedure we could explain the details of the revised dual simplex method applied to this problem. Since we showed one case of the application of the dual simplex in Section 6.3.3 we will not make

the dual parallel study here.

The approach to the composite algorithm followed in this Section is based on the self-dual parametric algorithm, see [24] pag 245. A slightly different approach, where one does not have to parametrize, is suggested in [44], pag 85. In this last approach the negative basic variables are treated as artifical variables and the basis exit criterion of the simplex method is extended to cope with negative components in the RHS. The trouble with this solution is that if a link with negative slack has to be brought into the basis the pivot is made at a negative element and, therefore, this link enters the saturation set with a positive dual variable which destroys the structure of the problem.

Although we are considering, in this section, the wisdom of simultaneously entering into the basis all the variables of $F_{BA}^{k+1}$, it should be pointed out that depending on the specifics this might lead to more basis changes in the reoptimization process than the procedure explained in Section 6.3.1. One anticipates that one case where this situation can happen is when $|S^k| >> |S^{k+1}|$, since then the number of new potential alternate paths for the commodities of the old $R^{k+1}$ is much bigger than in the old (LP(k+1)) problem, and this fact can make many of the old paths, i.e., elements of $F_{BA}^{k+1}$, not "good enough" for the new k-th problem.

Finally, we want to point out an interesting case of this type I reoptimization problem. Suppose that after making the elements of $F_{BA}^{k+1}$ basics at the new k-th level problem one finds out that

$$\tilde{z}^k(f_\ell(n)) \geq 0 \qquad \text{all } f_\ell(n) \in F^k_{BA} \cup F^k_+ .$$

The implication is that all the alternate paths that previously were forbidden to the old (k+1)-th problem from the k-th problem solution were not of any help, anyhow, in reducing the saturation value of the old (k+1)-th level. In this case, for the new global optimal solution the variables and commodities of the old (LPk) and (LP(k+1)) problems will simply exchange positions in the hierarchical ordering. The remaining problems remain untouched and, thus, the only change in the optimal basis of the entire successive problem, i.e., that corresponding to the last level problem, will be that of variables $\beta^k$ and $\beta^{k+1}$. Of course, this last statement is true only if the change of requirements is such that $\beta_0^{k+1} < 0$ was the only reason for non-optimality and, therefore, none of the flow variables became negative.

### 6.3.3 A Routing Variable Becomes Negative (II)

Suppose that after making the computations, indicated in Section 6.2, to accomodate a given requirement change, the central control observes that $f_\ell(n) < 0$, for some $f_\ell(n) \in F^k_{BA}$. Then, we have a negative value in the RHS of the optimal tableau of problem (LPk) and we have to proceed to reoptimize. According to the dual simplex method, in order to find the variable that should replace $f_\ell(n)$ as a basic variable, we have to look for all the negative elements of the $f_\ell(n)$ row, i.e., row $\ell n$ in our terminology. From Chapter V we know that none of the variables of $S^i$ or $F^i_+$, i=1,...,k-1, can be made basic at the k-th problem; thus the new basic variable must

belong to either $S^k$ or $F_0^{k-1}$, see Section 5.6. We therefore have to look for the negative elements of the $\ell n$-th row of both the updated matrix corresponding to the elements of $F_0^{k-1} - F_{BA}^k$ (remember that the elements of $F_{BA}^k$ are drawn from $F_0^{k-1}$) and the submatrix of $(\underline{Q}^k)^{-1}$ corresponding to the links of $S^k$. To help in the following deliberations consider Tableau T.VI.III below, obtained from T.VI.II of Section 6.3.1 by making $\beta^k$ and the elements of $F_{BA}^k$ basics, while the links of $S^k$ become non-basic or saturated.

| $\underline{S}^k$ | $\underline{\bar{S}}^k$ | $\alpha^1\beta^2 \cdots \beta^k$ $F_{BA}\underline{k}$ | $F_0^{k-1} - F_{BA}^k$ | RHS | |
|---|---|---|---|---|---|
| $(\underline{Q}^k)^{-1}$ | | I | $P_0\underline{k}(k-1)$ | $b_0^k$ | |
| | | | | | $\leftarrow f_\ell(n) < 0$ |
| | I | | $\bar{P}_0\underline{k}(k-1)$ | $\bar{b}_0^k$ | |
| $-\pi^k(\underline{S}^k)$ | | | $z^k(F_0^{k-1} - F_{BA}^k)$ | $\beta_0^k$ | costs |

T.VI.III

From this tableau and the structure of matrix $(\underline{Q}^k)^{-1}$, see (5.5.24), we can write for the $\ell n$-th row of matrix $P_0\underline{k}(k-1)$,

$$P_0\underline{k}(k-1)(\underline{r}\ell n) = (\underline{Q}^k)^{-1}(\underline{r}\ell n)P_{\underline{B}0}\underline{k}(k-1) = \Omega^k(\underline{r}\ell n)P_{\underline{B}0}\underline{k}(k-1), \qquad (6.3.48)$$

where $P_{\underline{B}0}^{\underline{k}(k-1)}$ is the matrix of reduced circuits, for the variables of $F_0^{k-1} - F_{BA}^{k}$, with respect to the links of $S^{\underline{k}}$. The element of (6.3.48) corresponding to variable $f_j(m)$ will then be given by

$$P_{\underline{0}}^{\underline{k}(k-1)}(\ell n, jm) = \Omega^{\underline{k}}(\underline{r}\ell n) C_{\underline{j}}^{\underline{k}}(m), \qquad \text{all } f_j(m) \in F_0^{k-1} - F_{BA}^{k} . \quad (6.3.49)$$

In order to find the variable that will enter the basis we can compute, according to the dual simplex method, the ratio of the negative elements of the $\ell n$-th row and the negative value of their respective partial costs, and pick the ratio with the smallest value. From Section 5.6 we have

$$z_j^k(m) = -\pi^{\underline{k}\underline{k}} C_{\underline{j}}^{\underline{k}}(m), \quad \text{all } f_j(m) \in F_0^{k-1} - F_{BA}^{k}, \qquad (6.3.50)$$

and

$$z_j^k = -\pi_j^k \qquad \text{all } j \in S^k.$$

Thus we have to compute

$$\xi_j^k(\ell n) = \frac{\pi_j}{\Omega^{\underline{k}}(\ell n, j)}, \quad \text{all } j \in S^k \text{ s.t. } \Omega^{\underline{k}}(\ell n, j) < 0, \qquad (6.3.51)$$

$$\xi_{jm}^k(\ell n) = \frac{\pi^{\underline{k}\underline{k}} C_{\underline{j}}^{\underline{k}}(m)}{\Omega^{\underline{k}}(\underline{r}\ell n)\ C_{\underline{j}}^{\underline{k}}(m)}, \text{ all } f_j(m) \in F_0^{k-1} - F_{BA}^{k}$$

$$\text{s.t. } \Omega^{\underline{k}}(\underline{r}\ell n) C_{\underline{j}}^{\underline{k}}(m) < 0.$$

If $\{\xi_j^k(\ell n)\}$ and $\{\xi_{jm}^k(\ell n)\}$ are the set of elements defined in (6.3.51) we have to obtain

$$\xi^k(\ell n) \triangleq \min \{\xi_j^k(\ell n),\ \xi_{jm}^k(\ell n)\} \geq 0. \tag{6.3.52}$$

We now study the possible subcases that one might have depending on the set to which $\xi^k(\ell n)$ belongs.

Case IIa $\xi^k(\ell n) = \xi_a^k(\ell n)$, for some $a \in S^k$.

In this case link $a$ will become basic, i.e., unsaturated so $\tilde{S}^k = S^k - a$ and $\tilde{F}_{BA}^k = F_{BA}^k - f_\ell(n)$ (*). In terms of the saturation matrix, $Q^k$, the row corresponding to link $a$ and the column corresponding to variable $f_\ell(n)$ will be lost. The new inverse matrix can be obtained, from the old saturation matrix, by means of Theorm AI.2 of Appendix I. Using this theorem we can write

$$(\tilde{Q}^k)^{-1} = [Q^k(\underline{ra},\ \underline{c\ell n})]^{-1} = (Q^k)^{-1}(\underline{r\ell n},\ \underline{ca}) - \tag{6.3.53}$$

$$- \frac{(Q^k)^{-1}(ca-\ell n)\ (Q^k)^{-1}(r\ell n-a)}{(Q^k)^{-1}(\ell n,a)},$$

where $Q^k(\underline{ra},\underline{c\ell n})$ denotes matrix $Q^k$ without row $a$ and column $\ell n$ (similarly for $(Q^k)^{-1}(\underline{r\ell n},\ \underline{ca})$), and $(Q^k)^{-1}(r\ell n-a)$ denotes the $\ell n$-th row of $(Q^k)^{-1}$ without its a-th component (similarly for $(Q^k)^{-1}(ca-\ell n)$). From (6.3.53) we see that the matrix updating is relatively simple. In particular the new vector of dual variables will be

$$\tilde{\pi}^{kk} = \pi^{kk}(\bar{a}) - \pi_a^k \frac{\Omega^k(r\ell n-a)}{\Omega^k(\ell n,a)}, \tag{6.3.54}$$

(*) Here tilde "~" characterize the new vectors and variables.

where $\pi^{kk}_{-}(\bar{a})$ denotes vector $\pi^{kk}_{-}$ without component $\pi^{k}_{a}$.

To check if this new basis is optimal for (LPk) we have to compute the new value of the routing variables. If $b^k_0$ was the old vector value, the new one will be

$$\tilde{b}^k_0 = b^k_0(\overline{\ell n}) - \frac{b^k_0(\ell n)}{\Omega^k_{-}(\ell n, a)} (Q^k_{-})^{-1}(ca-\ell n) \ , \qquad (6.3.55)$$

where $b^k_0(\overline{\ell n})$ denotes column $b^k_0$ without the element corresponding to $f_\ell(n)$, and $b^k_0(\ell n)$ is this subtracted element. If $\tilde{b}^k_0 \geq 0$ the basis is optimal; if, on the other hand, there is some negative element in $\tilde{b}^k_0$ we have to continue the reoptimization (*). Once we reach a new optimal basis we have to check whether or not the lower levels are still optimal, and in the latter case continue with the reoptimization according to the steps presented in Sections 6.3.1-2.

Case IIb $\xi^k(\ell n) = \xi_{a\nu}(\ell n)$, for some $f_a(\nu) \in F^k_+ \subset F^{k-1}_0 - F^k_{BA}$.

In this case variable $f_a(\nu)$ becomes basic, i.e. nonzero and displaces $f_\ell(n)$ from $F^k_{BA}$. In terms of matrix $Q^k_{-}$, its $\ell n$-th column will be replaced by column $C^k_a(\nu)$, that is,

$$Q^k_{-} = Q^k_{-} + [C^k_a(\nu) - Q^k_{-}(\underline{c}\ell n)] \ [0,0,\ldots 0,\underset{\uparrow}{1},0,\ldots 0]. \qquad (6.3.56)$$
$$\ell n$$

Using Theorem AI.3 of Appendix I

$$(\tilde{Q}^k_{-})^{-1} = (Q^k_{-})^{-1} - \frac{(Q^k_{-})^{-1} u \ \Omega^k_{-}(\underline{r}\ell n)}{1 + \Omega^k_{-}(\underline{r}\ell n) u} \ , \qquad (6.3.57)$$

(*) The first k elements of $\tilde{b}^k_0$, which correspond to $\alpha^1, \beta^2, \ldots, \beta^k$, are clearly positive and do not have to be computed for this checking operation.

where

$$u = [C_a^k(v) - Q^k(\underline{c}\ell n)] \quad .$$

Again we check whether or not the new basis is optimal for (LPk) by computing, similarly to (6.3.55),

$$\tilde{b}_0^k = b_0^k(\overline{\ell n}) - \frac{b_0^k(\overline{\ell n})}{P_+^k(\ell n,\ av)} \ (Q_+^k)^{-1}(\underline{r}\ell n) \ C_a^k(v) \ , \qquad (6.3.58)$$

where $P_+^k$ is the submatrix of $P_0^{k(k-1)}$ corresponding to $F_+^k$ ; the remaining notation is explained in (6.3.55). The comments made at the end of Case IIa are also valid for this case. Note also that the updating equations (6.3.57)-(6.3.58) are again rather simple, since $C_a^k(v)$ and u are graph-related matrices. From (6.3.57) and (6.3.50) we can write for the new dual variable vector

$$\tilde{\pi}^{kk} = \pi^{kk} - \frac{z_a^k(v)}{P_+^k(\ell n, av)} \ \Omega^k(\underline{r}\ell n) \ . \qquad (6.3.59)$$

Case IIc $\xi^k(\ell n) = \xi_{av}(\ell n)$, for some $f_a(v) \in F_0^k = F_0^{k-1} - F_{BA}^k - F_+^k$ .

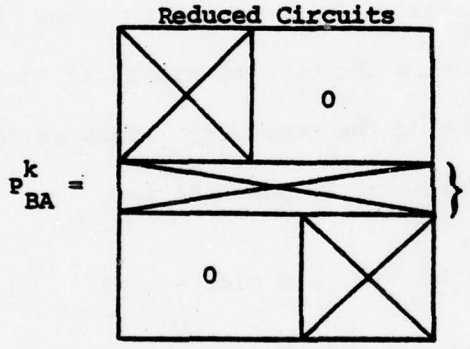
This case is completely equilvalent to IIb except that now $z_a^k(v) = 0$ and thus $\tilde{\beta}^k = \beta^k$ and $\tilde{\pi}^{kk} = \pi^{kk}$.

### 6.3.4 Further Comments on Case II

Suppose that after a basis losses optimality, reoptimization is achieved via a sequence of basis changes all of which fall under category, or type, II. If none of the links of $S^k$ has become basic, then the reoptimization consists just of a exchange of elements between sets

$F_{BA}^{k}$ and $F_{0}^{k-1} - F_{BA}^{k}$; that is, in the substitution of some alternate paths for others without modifying the structure of $S^{k}$. If, on the other hand, some link of $S^{k}$ becomes basic, then we will have the case of a saturation set either dropping a few links and decreasing in size, or else breaking apart into two or more sets.

A case where the breaking apart into two or more sets can be easily detected occurs when the link (or subset of links if there is a tie) which has to become basic is such that matrix $P_{BA}^{k}$ can be partitioned as follows




(6.3.60)

in this case [see also Theorem 4.6.2-3] sets $S_{I}^{k}$ and $S_{II}^{k}$ will belong to two different saturation levels. For cases where the saturation matrices have many zero entries, this can be a very common and easily detectable situation. Furthermore we can use Theorem 4.6.2 to easily update the new saturation matrices.

### 6.3.5 A Spanning Tree Flow Becomes Negative (III)

Since the flow on the spanning tree can be obtained by subtracting from the total commodity flow the flow taken by the alternate paths,

Section 3.6 and (4.2.42), we know that we do not need to carry the bottom, or commodity, part of the constraint matrix throughout the optimization in order to detect when a spanning tree flow variable becomes negative. Furthermore we know, from Theorem 3.5.1, that since any solution must contain the flows of a spanning tree for every destination, we must be able to replace a negative flow on the spanning tree by some other flow of some set $F^k_{BA}$ and obtain, again, a complete spanning tree for the corresponding destination. It may happen, though, that no flow variable exists among the alternate path of that commodity, which will complete a spanning tree. In this case, from Theorem 3.5.1, there is a non-basic flow that must be made basic and inserted in the spanning tree. This flow can be easily obtained by checking the cost of all flows corresponding to the links leaving the origin node of the commodity. Thus we see that if $f_\ell(n) < 0$, where $\ell \in T(n)$, we can have two posibilities.

Case IIIa. $f_a(n) \in F^k_{BA}$, some $1 \leq k \leq M$, and $o(a) = o(\ell)$.

In this case $f_a(n)$ joins the set $F_{BT}$ and $F^k_{BA}$ looses one element. In order to keep a proper saturation matrix we must also drop a link from $S^k$. To find that link consider the generalization of matrix (4.2.21) to problem (LPk). If $G^{kk}_n$ is the submatrix below $S^k$ corresponding to commodity n, then, $f_\ell(n)$ will be associated with a row of $G^{kk}_n$, say $G^{kk}_n(\underline{r}(o(\ell), n))$ (remember that there is a row of $G^k_n$ for each node that can communicate with node n.) Since $f_\ell(n)$ is leaving the basis we have to make a pivot, according to the dual simplex method, at some element of $G^{kk}_n(\underline{r}(o(\ell),n))$, where each such element corresponds, as we said before, to a link of $S^k$. Since we know, from (4.2.23), that

$$G_n^k = -\Phi_n (Q^k)^{-1},$$

where $\Phi_n$ is defined in (4.2.24), the process of finding the pivot element and making the appropriate updating is completely similar to what we did in Section 6.3.3, so we will not repeat it here. Observe that case IIIa corresponds to a case where matrix $Q^k$ looses one row and one column, so that we can apply Theorem AI.2 of Appendix I. Observe also that instead of the updating derived from that theorem we could perform an updating derived from the fact that the columns of $P_{BA}^k$, corresponding to variables of the type $f_\ell(n)$, will change because the spanning tree T(n) has changed structure. Although both updatings are equivalent it seems that the first one will involve, in general, a smaller amount of computation.

Case III b. There is no $f_a(n) \in F_{BA}^k$, $1 \leq k \leq M$, s.t. $\sigma(a) = \sigma(\ell)$

In this case a nonbasic flow has to be made basic. This corresponds, according to the generalization of matrix (4.2.21) to problem (LPk), to making a pivot at some element of the row of matrix $M_n^k$ corresponding to variable $f_\ell(n)$. All comments made to case IIIa pertain here also.

## 6.4 An Algorithm to Solve the Successive Saturation Problem

Although the most important details of a possible algorithm to solve the successive saturation problem have already been discussed in previous chapters and sections, especially Section 6.3, we will summarize here the main steps of such an algorithm.

I. Select a set of spanning trees, one for each destination. A logical starting choice will be the spanning trees with a minimum number of branches between each source-destination pair. Observe that this is a well-defined set of trees.

II. Solve problem (LPk), starting with k = 1 and continuing in increasing order of k, according to the procedure explained in Section 6.3.1.

III. After step II is completed for a given k check if either the solution is unique or else $F_0^k$ is empty. If this is the case the algorithm stops, otherwise make k = k+1 and repeat step II.

Now we will make some general comments applicable to this solution algorithm:

a) It is well known already that the candidates to become members of the set $F_{BA}^k$, at the k-th level problem, are the elements of $F_0^{k-1}$ (see Section 5.6). The decision that chooses an element of $F_0^{k-1}$ to become part or not of $F_{BA}^k$ is based on the total distance of the circuit of the candidate element with respect to the assignment corresponding to the actual set of k-th level dual variables. Several points deserve mention:

ai) For the decision test mentioned above one can use any valid distance assignment as described in Section 5.4. If we restrict our search of candidates to the elements of $F_0^{k-1}$, then the distance assignment does not have to be nonnegative and we can use the opposites of the dual variables, or any normalized integer version of them, as a valid assignment.

aii) Every time we add a variable to $F^k_{BA}$, or change its structure, we have to update the set of dual variables and/or the distance assignment.

aiii) For a given distance assignment the search for circuits with negative total distance can be undertaken via two equivalent but complementary procedures: an analytic or a synthesis procedure. The analytic procedure look for elements of $F^{k-1}_0 - F^k_{BA}$ with negative distance. The synthesis procedure look at the spanning trees and tries to find how we can close a circuit in such a way that the total distance is negative, i.e., in such a way that we traverse the links with larger weights in the negative direction. Clearly the conveniance of one method or the other will depend on how many elements the set $F^{k-1}_0$ has.

aiv) The decision to work with set $F^{k-1}_0$ or with the total set of flow variables will depend on the number of elements of $F^{k-1}_0$ relative to the number of elements of $F - \sum_{i=1}^{k-1} F^i_{BA}$. In some cases, rather than storing a very large set, it can be more convenient to apply the shortest route criterion many time over and over, since there already exists very efficient methods for implementing this criterion [11], [35], [47]).

b) If after solving problem (LPk) the corresponding basic solution is nonminimal, one should look for a miminal solution, see Sections 3.2 and 5.6. If all solutions are nonminimal we then apply Lemma 5.6.1.

c) Once a set of spanning trees has been chosen we can, in principle, forget about the commodity part of the constraint matrix and work exclusively with the upper part of the matrix, which involves only link slack and flow variables, see (4.2.21).

Since the flows on the spanning trees can be obtained by subtracting the amount of flow in the alternate paths, of a given commodity, from the total requirement of that commodity, see (3.6.3), we can easily detect when a flow on a spanning tree is negative. As soon as this fact is detected we proceed as indicated in Section 6.3.5.

d) The only differences in the solution procedure of two different level problems are in the size of the acting basis [which from problem k to problem k+1 increases by $|S^k|$], and in the number of active variables, which decreases from $F_0^{k-1}$ for problem (LPk) to $F_0^{k-1} - F_{BA}^k - F_+^k$ for problem (LP(k+1)).

In general we can say that, since the problem has a virtual working basis whose dimension is the number of saturated links, instead of the rank L+N(N-1) of the constraint matrix, the algorithm does a very good job in reducing the size of the actual working basis. Furthermore, the use of a form of the revised simplex method together with the simplifying rules that follow from the structure of the problem, should save a great deal of computation. In a large scale problem all these reductions and simplifications can be very important.

Finally, we want to point out a problem that might occur when applying the algorithm dynamically. Since every time we change basis we are changing linear region, as defined in Section 3.2, we might be faced with the case where the problem is moving through a cluster of small linear regions, thus provoking very frequent changes of basis. It seems clear that in these cases it can be better to perform sub-optiamlly than to track the optimal solution through this set of small

regions. Two mechanisms can be envisaged to reduce those too-frequent changes of basis:

I) Do not start a basis change until some variable surpass a given negative threshold.

II) Do not implement changes that lead to an optimal basis whose linear region is smaller (in surface or in max and min radius, for instance) than a preestablished value. The shape of the linear region corresponding to an optimal basis is very easy to compute and it is usually reported by commercially available computer packets ([46], pg. 106).

It should be clear that these two mechanisms can be used only if they do not lead to suboptimal solutions that are not implementable, due for instance to the violation of a capacity constraint. A detailed study of these points is left for further research.

Other areas of further research concerning this algorithm should be:

- Criteria to choose the initial solution (we already mentioned a possible one at the beginning of this section);
- Running time;
- Dynamic tracking characteristics and performance (we pointed out before some problems that may appear in the dynamic case);
- Length of code; etc.

# CHAPTER VII

# CONCLUSIONS

## 7.1 Summary of Important Results

As we said in Chapter I the main object of this thesis was to obtain new insight into the problem of dyanmic routing in data networks and, in particular, to study how one can effectively exploit the structure of the network as determined by the graph supporting the network. To achieve these goals we used a minimax objective function which divides the optimization problem into a layer hierarchy of linear programs. At each layer of this hierarchy we minimize the maximum saturation level of those links whose level has not been already fixed by a previous layer. This minimization is done under the constraints that we cannot increase the saturation level of links whose maximum saturation level was already minimized by a previous layer. Thus the optimal solution space of each layer is a subspace of the optimal solution spaces of upper layers. The solution obtained at each level specifies:

- A collection of links with common saturation level
- A collection of allowable paths that balance the saturation level of the saturated links without increasing the level of previously saturated links, and
- A collection of paths that cannot be used without increasing the level of some of the saturated links.

One important result of the thesis is that we can describe the set of allowable paths at each level by a nonnegative distance assignment to links and the shortest route criterion. This distance assignment is non-unique. In fact every level or layer introduces a new

degree of freedom in the determination of the valid assignment by means of a new ray defined in a subspace of the optimal dual space of its linear program. All rays defining valid assignments can be described in terms of the link dual variables of the various layers already optimized. Furthermore these rays can also be described by integer vectors, thus, providing an integer distance assignment to define the allowable paths. Since the dual variables are a function of the capacities of the saturated links, whenever those capacities are irrational we would obtain a set of irrational numbers as our collection of dual variables. But the integer description of the rays allow us to obtain from these dual variables an integer distance assignment to links that describes a set of allowable paths for a network independently of the (irrational) link capacities. Furthermore if $|S^i|$ is the number of saturated links at level i we can find distance assignments for level, say, k with at most $\sum_{i=1}^{k} |S^i|$ nonzero link distances. Since from Theorem 5.6.2 not all links of a network can be saturated we can describe the allowable paths of an optimal solution of the global problem by means of an integer distance assignments to links where some links will have zero distance.

Another important result of the thesis is that, given the collection of allowable paths for level k and the collection of saturated links for all levels i preceeding level k, we can obtain with a relatively small amount of effort, the updated constraint matrix and optimal basis inverse for the linear program of level k. This result is important mainly from the reoptimization point of view, when our present solution loses optimality due to requirement changes, since we do not have to store the optimal

basis inverse or constraint matrices. Furthermore, the procedure to obtain these matrices allows us to obtain any particular element, row or column with a minimal amount of computation. The amount of computation required is much less than the amount we would have needed, for computing that element or vector, via a pivoting process or ordinary linear programming matrix calculations. The reason is that we have many properties in our problem that can be used to simplify and save computation.

One of the most important properties in this sense is derived from the concept of saturation matrix. There is a saturation matrix for every level where the saturation matrix for level k is an $|S^k| \times |S^k|$ square matrix that has a column of capacities corresponding to the links of $S^k$ and whose remaining columns are $0, \pm 1$ vectors corresponding to the alternate paths for problem k. The important aspect of these saturation matrices is that they have all the relevant information concerning the solution. All remaining data concerning the current operating point can be obtained from the inverse of these saturation matrices and graph related matrices. Thus the only non-graph operation involved with the optimization process is the inversion of a collection of $|S^i| \times |S^i|$ matrices with $0, \pm 1$ elements.

In conjunction with this statement we point out that a result (closely related to the independence of the distance assignment rays with respect to capacity values) obtained in Chapter IV allows us to work for most purposes with a capacity normalized matrix, in which all capacities are equal to one. Only when we need the actual values of the dual variables or we have to work with the RHS do we have to use

the unnormalized matrices. It turns out that we can obtain the inverse of the unnormalized saturation matrices from the inverse of the normalized matrices, and vice versa, by just multiplying the corresponding matrix by another matrix easily obtained from the dual variables. Furthermore the normalized and unnormalized dual variables are also very simply related, because they belong to the same ray.

Finally, we mention that all these results can be used to simplify the reoptimization process. Reoptimization can be divided into two different types. One type is where the present optimal solution can absorb a given requirement increment without changing optimal basis. The other type is where a change of optimal basis is required. The first type of reopimization is very simple to handle, and in fact allows us to envisage a possible decentralized scheme to control the evolution of the optimal point within regions that do not require a change of optimal basis. The second type of reoptimization is more complicated, but can also be undertaken with great savings in computations thanks to previous results. In this sense the task of finding elements that have to leave or enter the basis can be associated with the task of finding the total distance of some paths of the network for a given set of link dual variables. The updating process involved with a change of basis is also very simple since only the saturation matrices involved with a given change have to be updated. All these facts lead us to propose what we expect will be an efficient algorithm to solve the sequence of linear programs for a given requirement vector.

## 7.2 Areas of Future Work

Many additional areas of research appear to be ready for further investigation.

One important point for future study concerns the efficiency of the proposed algorithm for solving the successive problem (*), and comparing it with other existing centralized routing algorithms. Another important area for future research is possible decentralized schemes that could be derived from the routing procedure used in this thesis. We have already pointed out that for regions that do not require a change of optimal basis there exists a natural and simple decentralized scheme. The main problem arises when we try to extend this decentralization to the case where a change of optimal basis is required.

Finally a third area of future research is the area of topological designs that simplify routing operations. We expect that the explicit appearance of graph related matrices in our problem formulation can contribute to the opening of this area.

---

(*) See also Section 6.4.

# APPENDIX I

## SOME MATRIX RESULTS

In this appendix we summarize most of the results and theorems of matrix analysis used in this thesis. References for proofs are [39] and [40].

AI.1 Elementary Matrix: Is the matrix which results when an elementary row (column) transformation is applied to the identity matrix. We can have the following elementary matrices:

- $H_{ij}$ - matrix I with rows i and j interchanged
- $H_i(k)$ - matrix I with k instead of 1 in the i-th diagonal position; corresponds to multiplying the i-th row by scalar k
- $H_{ij}(k)$ - matrix obtain from I by adding to the elements of the i-th row k times (k scalar) the elements of the j-th row

If previous transformations are made with columns instead of rows we have another collection of elementary matrices that will be denoted by $\overline{H}_{ij}$, $\overline{H}_i(k)$ and $\overline{H}_{ij}(k)$.

a) The effect of applying an elementary transformation to matrix A can be produced by multiplying A by an elementary matrix: on the left by H and on the right by $\overline{H}$.

b) The inverse of an elementary transformation is an elementary transformation of the same type, in particular

$$H_{ij}^{-1} = H_{ij} \qquad \bar{H}_{ij}^{-1} = \bar{H}_{ij}$$

$$H_i^{-1}(k) = H_i(1/k) \qquad \bar{H}_i^{-1}(k) = \bar{H}_i(1/k) \qquad \text{(AI-1)}$$

$$H_{ij}^{-1}(k) = H_{ij}(-k) \qquad \bar{H}_{ij}^{-1}(k) = H_{ij}(-k)$$

c) Every non-singular matrix can be expressed as a product of elementary matrices.

d) An elementary transformation does not alter the rank of a matrix

e) Two matrices A and B are called <u>equivalent</u>, $A \sim B$, (row, column equivalent) if one can be obtained from the other by a sequence of elementary transformations (row, column transformations)

$$A \sim B \Rightarrow A = H B \bar{H} \qquad \text{(AI-2)}$$

where H is a product of row elementary matrices and $\bar{H}$ a product of column elementary matrices. If H and $\bar{H}$ are obtained from elementary matrices with $k = \pm 1$, then H and $\bar{H}$ will be called <u>unitary</u> or <u>unimodular</u> elementary transformations.

f) Two matrices of the same dimension and rank are equivalents.

## AI. 2 <u>Inverstion Methods and Related Issues</u>

<u>Theorem AI.1</u>

If matrix A is reduced to an identity matrix by a sequence of row transformations alone, then $A^{-1}$ is equal to the product in reverse order of the cooresponding elementary matrices. Or, in other words, if the extended matrix (AI) is transformed into (IB) by a sequence of row transformations then $B = A^{-1}$.

Theorem AI.2

Let AB = BA = I. If we partition A and B as follows

$$A = \left[\begin{array}{c:c} A_{11} \ (m\times m) & A_{12} \ (m\times n) \\ \hdashline A_{21} \ (n\times m) & A_{22} \ (n\times n) \end{array}\right] \qquad B = \left[\begin{array}{c:c} B_{11} \ (m\times m) & B_{12} \ (m\times n) \\ \hdashline B_{21} \ (n\times m) & B_{22} \ (n\times n) \end{array}\right]$$

then, provided $A_{11}$ is nonsignular,

$$B_{11} = A_{11}^{-1} + (A_{11}^{-1}A_{12})\xi^{-1}(A_{21}A_{11}^{-1}) = A_{11}^{-1} + \frac{B_{12}\,B_{21}}{B_{22}}$$

$$B_{12} = -(A_{11}^{-1}A_{12})\xi^{-1}$$

$$B_{21} = -\xi^{-1}(A_{21}A_{11}^{-1})$$

$$B_{22} = \xi^{-1}$$

where

$$\xi = A_{22} - A_{21}(A_{11}^{-1}A_{12})$$

Theorem AI.3

If

$$A = B + uv^{T}$$

then

$$A^{-1} = B^{-1} - \lambda yz^{T},$$

where

$$y = B^{-1}u, \quad z^{T} = v^{T}B^{-1},$$

and $\lambda$ is a scalar given by

$$\lambda = \frac{1}{1+z^T u} \quad ,$$

provided A and B are nonsingular.

# APPENDIX II

## ON LINK INTEGER DISTANCE ASSIGNMENTS TO DESCRIBE AN OPTIMAL SOLUTION OF THE FIRST SATURATION PROBLEM

### AII.1 Introduction

From Section 3.4 we know that there exists an integer assignment of distance to saturated links that characterizes the allowable paths for an optimal solution of (LP1). In fact, since sets $S^1$ and $\{\pi^1_\ell\}$ are unique at a stable point, we can say, for a non-unique optimal solution of (LP1), that the distance assignment, as defined in Theorem 3.4.3, characterize the allowable paths for an optimal solution obtained as a convex combination of all basic optimal primal solutions.

We also saw in Section 3.4, Lemma 3.4.5, that there are an infinite number of assignments to links of $S^1$ that can characterize a given set of allowable paths, and we proved that all these assignments must lie on a given ray. Because of this fact we made the assumption, which we will call here the assignment assumption, that whenever we refer to an integer distance assignment we assume that this assignment corresponds to the smallest integer point of the ray. If we refer to a ray as a class of assignments then the smallest integer point of the ray will be the representative of its class. Also, since in this appendix we shall study valid sets of integers that can be distance assignments, the internal ordering of elements within $\theta$ is irrelevant.

At this point there are several questions one could naturally ask regarding integer assignments. For instance: Given a set of integers how could we check if they can correspond to an optimal solution

of some (LP1) problem? Is there any bound on the largest integer link distance that could be needed to describe an optimal solution over a network with L links?

We will answer these questions here, starting with the second one.

AII.2 On the Largest Set of Integers for a Network with L Links

We will show in this section that there exists a family of topologies that can generate arbitrarily large integer assignments, and where the value of the largest integer grows exponentially with the number of links. The procedure that we will follow is a constructive one, starting with $|S^1| = 2$ ($|S^1| = 1$ is obviously a pathological case). We will also use this procedure to show, through examples, some of the advantages that the canonic equations are able to offer in many situations. The synthesis procedure used in this section will be generalized later on in this Appendix.

For $|S^1| \triangleq q = 2$ the only existing class of assignments is $\vartheta = \{1,1\}$. A network that could provide such a solution is

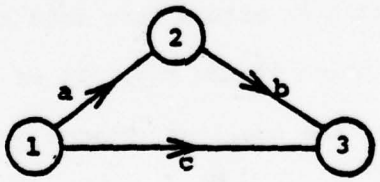



Fig. AII.2.1

with $S^1 = \{b,c\}$, $R^1 = \{r_1(3), r_2(3)\}$, the canonic equations

$$\alpha^1 = \frac{r_2(3) + y_1(3)}{c_b} = \frac{r_1(3) - y_1(3)}{c_c} \quad (*)$$

(*) Variable $y_1(3)$ gives the amount of commodity (1,3) flowing through link a.

the canonic ratio

$$\alpha^1 = \frac{r_2(3) + r_1(3)}{c_B + c_c} ,$$

dual variables

$$\pi_b = \pi_c = \frac{-1}{c_b + c_c} , \qquad \pi_a = 0$$

and distance assignment

$$\theta_b = \theta_c = -\pi_b D^1 = -\pi_c D^1 = 1$$

Numerically for $c_\ell = 1$, all $\ell$, and $r_1(2) = 1$, $r_2(3) = 2$, $r_1(3) = 4$ the solution is $\alpha^1 = 3$, $y_1(3) = 1$. In terms of a simplex solution, the nonzero flows will be

$$f_a(2) = 1,\ f_a(3) = 1,\ f_b(3) = 3,\ f_c(3) = 3$$

and the spanning trees are given in Fig. AII.2.2

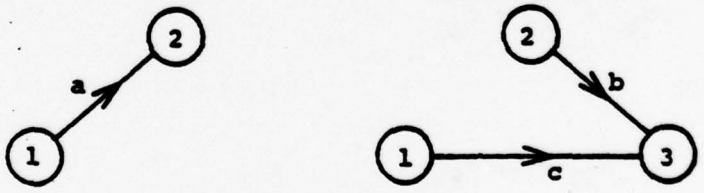



Fig. AII.2.2

Thus $f_a(3)$ is the only nonzero flow in the spanning tree and therefore the only alternate path (Note that in this case the solution to (LP1) will be unique). The routing variable $y_1(3)$ is therefore associated with $f_a(3)$. According to section 3.6 terminology this variable should have been called $y_a(3)$; note, though, that whenever we have a single alternate path for a given commodity there is no ambiguity

if we use the commodity notation as in $r_m(n)$, so we will follow this rule for simplicity.

For $|S^1| \triangleq q = 3$ is not difficult to see that we can only have two classes of assignments, either $\theta = \{1,1,1\}$ or $\theta = \{1,1,2\}$. Since we are looking for the largest assignment we take $\theta = \{1,1,2\}$; according to Lemma 3.2.1, this must correspond to a set of canonic equations of the type

$$\alpha^1 = \frac{r_m(n) - y_m(n)}{c_{\ell 1}} = \frac{y_m(n) + r_i(j) - y_i(j)}{c_{\ell 2}} = \frac{y_m(n) + y_i(j)}{c_{\ell 3}}$$

where commodity (m,n) uses two paths one going through saturated link $\ell 1$ and another going through links $\ell 2$ and $\ell 3$. Similarly commodity (i,j) has two paths, one through link $\ell_2$ and another through link $\ell_3$. The canonic ratio is

$$\alpha^1 = \frac{2r_m(n) + r_i(j)}{2c_{\ell_1} + c_{\ell_2} + c_{\ell_3}}$$

and clearly

$$\theta_{\ell 1} = 2, \quad \theta_{\ell 2} = 1, \; \theta_{\ell 3} = 1$$

A network that can provide such a solution could be the network of Fig. AII.2.3

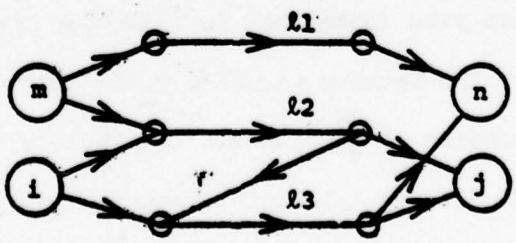



Fig. AII.2.3

which can be simplified to that of Fig. AII.2.4

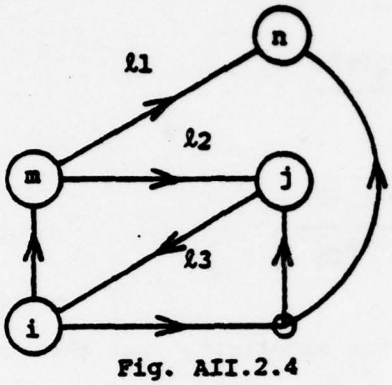



Fig. AII.2.4

Now suppose that we augment the network of Fig. AII.2.3 in such a way that we introduce a new commodity with two alternate paths: one going through links ℓ1, ℓ2 and ℓ3 and another going through a new saturated link ℓ4. Doing that we maintain a stable operating point and we have the collection of distance $\theta = \{1,1,2,4\}$ with $|S^1| = 4$. The network will be that of Fig. AII.2.5

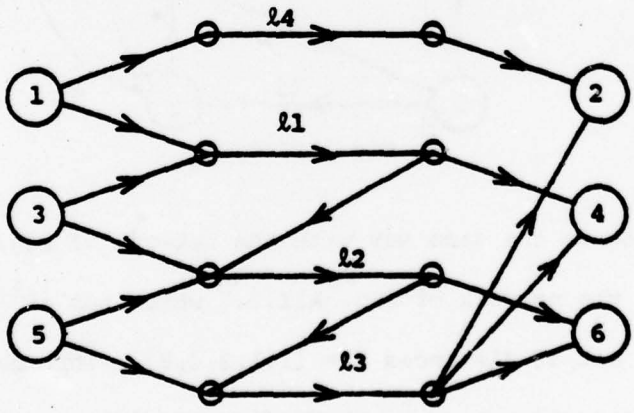



Fig. AII.2.5

and the canonic equations and canonic ratio become

$$\alpha^1 = \frac{r_1(2) - y_1(2)}{c_{\ell 4}} = \frac{y_1(2) + r_3(4) - y_3(4)}{c_{\ell 1}} = \frac{y_1(2) + y_3(4) + r_5(6) - y_5(6)}{c_{\ell 2}} =$$

$$= \frac{y_1(2) + y_3(4) + y_5(6)}{c_{\ell 3}}$$

$$\alpha^1 = \frac{4r_1(2) + 2r_3(4) + r_5(6)}{4c_\ell + 2c_{\ell 1} + c_{\ell 2} + c_{\ell 3}}$$

where we are assuming for simplicity that the remaining requirements are equal to zero.

The network of Fig. AII.2.5 can be simplified to give the network of Fig. AII.2.6

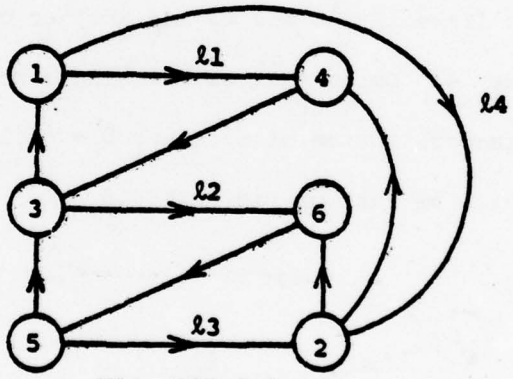



Fig. AII.2.6

Proceeding in the same way with the network of Fig. AII.2.5 we can obtain the network of Fig. AII.2.7 which has $|S^1| = 5$ and will generate the set of distances $\theta = \{1,1,2,4,8\}$. This network can be simplified to give the network of Fig. AII.2.8

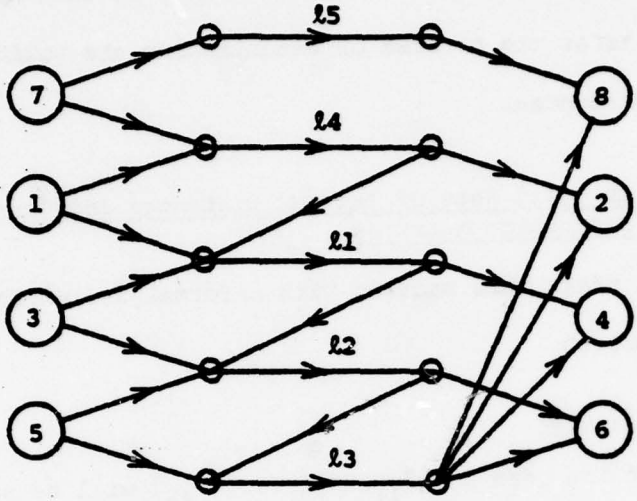



Fig. AII.2.7

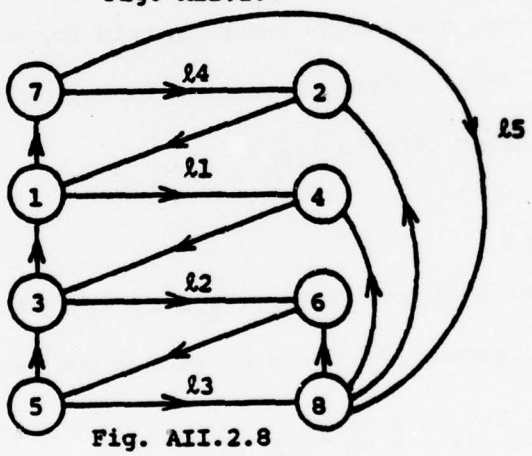



Fig. AII.2.8

Following with this family of networks we will generate the sequence of distances

$$1,1,2,4,8,\ldots,\ 2^{q-2}$$

which clearly grows exponentially with $q = |S_1|$.

Motivated by the results of this section we will present a general procedure for synthesizing (small) networks exhibiting any

valid 1st saturation set in the last section of this Appendix. Next, however, we treat the problem of establishing the validity of a set of integer as distances.

## AII.3 On Valid Sets of Integer Distances and Corresponding Saturation Matrices

We will start this section with a formal definition of valid distance assignments.

**Definition AII.3.1**

A set of integers $\theta = \{\theta_{\ell 1}, \theta_{\ell 2}, \ldots, \theta_{\ell q}\}$ will be said to be a valid distance assignment for the first saturation problem if there exist a network and a requirement matrix for which there is an optimal solution to (LP1), with q saturated links, which can be described in terms of integer weights $\{\theta_{\ell i}\}$.

**Lemma AII.3.1**

A set of integer $\theta = \{\theta_{\ell 1}, \theta_{\ell 2}, \ldots, \theta_{\ell q}\}$ will correspond to a valid distance assignment iff there exist a valid Q-matrix such that

$$\theta = -D^1 Q^{-1}(\underline{r}1) \triangleq -Q_\theta^{-1}(\underline{r}1)$$

where $D^1$ is defined in (3.4.8).

**Proof:**

Obvious from definition 4.4.1 and Theorem 4.2.7.

Q.E.D.

Since according to our previous assumptions the order of numbers within $\theta$ is irrelevant, we introduce the following definition

Definition AII.3.2

Two saturation matrices are said to be weight equivalents if they have the same dimension and give the same collection of link weights.

Lemma AII.3.2

"Equivalence" as defined by Definition AII.3.2 is a true equivalence relation. In particular the set of saturation matrices is partitioned into equivalence classes by this relation.

Proof:

Obvious from the definition of equivalence relation.

Lemma AII.3.3

If $Q$ and $\tilde{Q}$ are two saturation matrices of the same dimension and submatrices $P_B$ and $\tilde{P}_B$ are column equivalents then $Q$ and $\tilde{Q}$ are weight equivalents.

Proof:

From (AI-2) $P_B$ and $\tilde{P}_B$ will be column equivalents if

$$P_B = \tilde{P}_B \bar{H}_B \qquad \text{(AII.3.1)}$$

where $\bar{H}_B$ is a product of column elementary matrices.

From Theorem 4.6.1 we have

$$\theta \tilde{P}_B = 0$$

so multiflying (AII.3.1), on the left, by $\theta$

$$\theta P_B = \theta \tilde{P}_B \bar{H}_B = 0$$

Thus $\theta$ is also a valid assignment for $P_B$. Since according to section 4.5 the value of the capacities is inmaterial for vector $\theta$, we have proved the result.

Q.E.D.

If $\bar{H}_B$ of (AII.3.1) is only a product of elementary permutations and/or multiplication of columns by scalar (in this case the scalar must be $\pm 1$) then $Q^{-1}$ can be obtained from $\tilde{Q}^{-1}$ by means of these same elementary transformations (exchanging the role of row and column according to (AI-1)). Nevertheless the relation between $P_B$ and $\tilde{P}_B$ needs not be so simple, because Lemma AII.3.3 establishes a more general result. As an example the following three matrices are weight equivalents

$$Q = \begin{bmatrix} -1 & -1 & 0 & 0 \\ -1 & 0 & -1 & 0 \\ -1 & 1 & 1 & -1 \\ -1 & 0 & 0 & 1 \end{bmatrix} \qquad Q_\theta^{-1} = \begin{bmatrix} -1 & -1 & -1 & -1 \\ -3 & 1 & 1 & 1 \\ 1 & -3 & 1 & 1 \\ -1 & -1 & -1 & 3 \end{bmatrix}$$

$$\tilde{Q} = \begin{bmatrix} -1 & -1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ -1 & 0 & -1 & -1 \\ -1 & 1 & 0 & 1 \end{bmatrix} \qquad \tilde{Q}_\theta^{-1} = \begin{bmatrix} -1 & -1 & -1 & -1 \\ -3 & 1 & 1 & 1 \\ -1 & 3 & -1 & -1 \\ 2 & -2 & -2 & 2 \end{bmatrix}$$

$$\tilde{\tilde{Q}} = \begin{bmatrix} -1 & 1 & 1 & 0 \\ -1 & 1 & -1 & 0 \\ -1 & -1 & -1 & -1 \\ -1 & -1 & 1 & 1 \end{bmatrix} \qquad \tilde{\tilde{Q}}_\theta^{-1} = \begin{bmatrix} -1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ 2 & -2 & 0 & 0 \\ -2 & 2 & -2 & 2 \end{bmatrix}$$

where, for instance,

$$P_B = \tilde{\tilde{P}}_B \, \tilde{\tilde{H}}_B$$

and (in the notation of Appendix I)

$$\tilde{\tilde{H}}_B = \bar{H}_{12}(1)\bar{H}_1(-1/2)\bar{H}_{23}(-1)\bar{H}_{21}(1) = \begin{bmatrix} -1/2 & -1/2 & 0 \\ -1/2 & 1/2 & 0 \\ 0 & -1 & 1 \end{bmatrix}$$

Thus although Q, $\tilde{Q}$ and $\tilde{\tilde{Q}}$ have the same weights, dual variables and even some columns in common, their inverses are quite different and furthermore equal columns in the saturation matrices do not correspond with equal rows in the inverses.

Note also that, in terms of flow variables, the fact of multiplying a column of Q by -1 corresponds to an exchange of variables between $F_{BA}$ and $F_{BT}$, and therefore to a change of one link in the corresponding spanning tree. As an example consider Figs. AII.3.1a-1b below

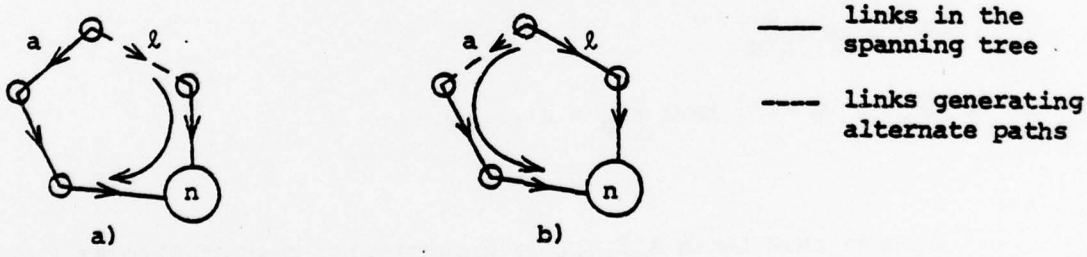



Fig. AII.3.1

If $f_\ell(n)$ was the variable corresponding to the original column of Q, multiplying that column by -1 is equivalent to replacing $f_\ell(n)$ by $f_a(n)$ and modifying the spanning tree T(n) by adding link ℓ and dropping link a.

We now give an extension of Lemma AII.3.3.

Theorem AII.3.1

A necessary and sufficient condition for Q and $\tilde{Q}$ to be weight equivalents, with $\theta$ being the weight vector, is that

$$P_B = H_B \tilde{P}_B \bar{H}_B$$

where $H_B$ and $\bar{H}_B$ are elementary matrices and where $\theta$ is an eigenvector of $H_B$.

Proof:

From e) and f) of Appendix I we can always write

$$P_B = H_B \tilde{P}_B \bar{H}_B \ .$$

Multiplying by $\theta$ on the left

$$\theta P_B = \theta H_B \tilde{P}_B \bar{H}_B \ .$$

If $\theta$ is an eigenvector of $H_B$ and $\tau$ is the corresponding eigenvalue

$$\theta P_B = \tau\theta \ \tilde{P}_B \bar{H}_B$$

so of $\theta\tilde{P}_B = 0$ we also have $\theta P_B = 0$ .

Q.E.D.

Observe that Lemma AII.3.1 is a particular case of Theorem AII.3.1, when $H_B = I$.

An easy test, from Lemma AII.3.1, to see if $\theta$ is a valid distance assignment is to check if a valid Q-matrix can be generated from $\theta$. For this purpose note that, since each column of $P_B$ represents a reduced circuit and there are q-1 reduced circuits in $P_B$, the collection of integers has to allow for the possibility of obtaining q-1 subsets of

integers with the following properties:

i) Each subset must be able to be partitioned in two non-empty blocks, such that the sum of the elements of one block has to be equal to the sum of the elements of the other block.

ii) These q-1 subsets must be independents.

Clearly each one of these subsets will correspond to a column of $P_B$. Thus to each subset we associate a 0, $\pm$ 1, q-component vector where the elements in one block will have a +1 entry, the elments of the other block a -1 entry and the remaining elements a 0 entry. Hence we will have a valid $P_B$ matrix as soon as we are able to find q-1 independent vectors obtained in this way. We can conclude then,

Lemma AII.3.4

A collection, $\theta$, of q integers can be a valid distance assignment, for the saturated links of an (LP1), iff the rank of the matrix obtained from the vectors corresponding to all possible subsets, that can be obtained from $\theta$ and satisfy property i) above, is at least q-1.

□

Example AII.3.1

Consider the collection of integers $\theta$ = (1,1,2,4). From this collection we can obtain the following subsets and blocks

{(1,1,2) - (4)} => subset 1

{(1,1) - (2)} => subset 2

{ (1) - (1)} => subset 3

If we assign each weight to a link, say $\theta = (\theta_a, \theta_b, \theta_c, \theta_d)$, then, we

can write the following three column matrix corresponding to previous subsets

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & 0 \\ -1 & 0 & 0 \end{bmatrix} \begin{matrix} a \\ b \\ c \\ d \end{matrix}$$

clearly the rank of this matrix is equal to three and thus $\theta$ is a valid distance assignment (in fact it corresponds to the network of Fig. AII.2.6).

□

The fact that, in principle, one can be forced to write all possible subsets with property i), obtained from a collection $\theta$, render the procedure inappropriate for large collections of integers. Nevertheless it can be expected that some partition techniques used in number theory, see [41], [42], could be used to simplify this task. Now we will move to the last section of this Appendix where we generalize the synthesis procedure introduced in Section AII.2, but before we will raise a question regarding the structure of a valid set $\theta$: If $\theta$ is to be a valid set of integers, is it necessary that $\theta_\ell = 1$ for at least one $\ell$? The answer is no and can be justified by means of Example 1 of Appendix III.

## AII.4 A Synthesis Procedure for the First Saturation Level

Given a valid (see section 4.4) Q-matrix it is always possible to find a capacitated network and a requirement matrix such that the solution of the corresponding first saturation problem will give a

saturation matrix indentical to the given one. Of course, the solution will not be unique by any means. A simple procedure to obtain one of these possible networks could be:

1) Take a set of isolated links corresponding to the saturated links of Q.

2) Take the reduced circuits of Q and establish the necessary connections among previous links to build these circuits. Assign the necessary commodities, source and destination, to each circuit. Note that there exists a tremendous degree of freedom at this stage since the order of connection within each path is not given by the matrix; furthermore we have many ways to assign commodities since some of them may have the same source or destination. (See also the comment at the end of this Appendix.)

3) For the saturated links take the capacities given by the Q matrix. For the remaining links the task of finding appropriate capacities is very simple, since a solution to the routing problem is known. For commodities using saturated links the allowable paths are given by the reduced circuits; for the remaining commodities the paths can be fixed arbitrarily as long as they do not use saturated links. Once the paths are known we can compute the amount of total flow through each link and give to it a capacity that, with the computed flow, will give a saturation level lower than the level of the saturated links.

Example II.4.1

Consider the following matrix Q, with $S^1 = \{a,b,c,d\}$

$$Q = \begin{bmatrix} -1 & 1 & 1 & 0 \\ -1 & 1 & -1 & 0 \\ -1 & -1 & 1 & 1 \\ -1 & -1 & -1 & -1 \end{bmatrix} \begin{matrix} a \\ b \\ c \\ d \end{matrix}$$

and $c_\ell = 1$, all $\ell \in S^1$, and where

$$4Q^{-1} = \begin{bmatrix} -1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ 2 & -2 & 0 & 0 \\ -2 & 2 & 2 & -2 \end{bmatrix}$$

Thus Q is a valid saturation matrix. Fixing an order of connection we pick the reduced circuits to be

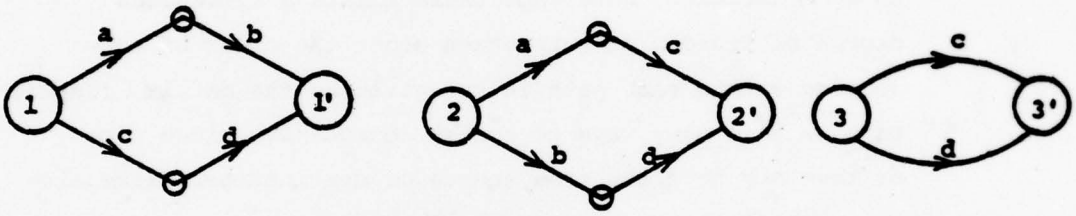



Fig. AII.4.1

Now we take the isolated links of $S^1$ and establish the appropriate connection needed for the first circuit. The result is

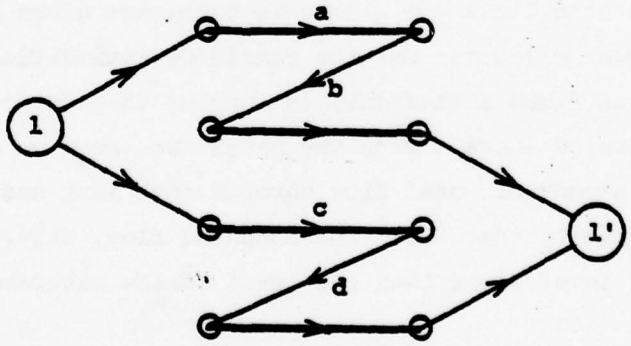



Fig. AII.4.2

We take (1,1') to be the commodity for the first circuit. Proceeding similarly with the other two circuits we would obtain the network

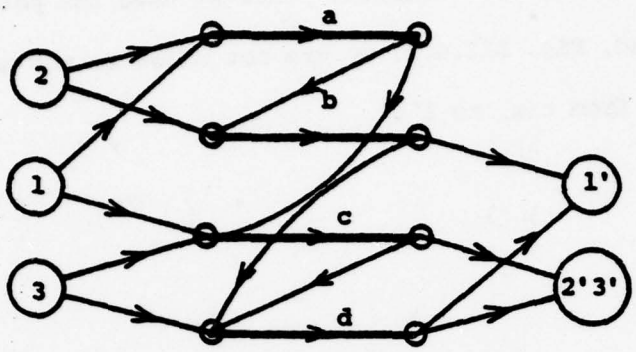



Fig. AII.4.3

Where we have chosen destinations 2' and 3' to coincide. The assignment of capacities to nonsaturated links and values to the requirement matrix is a trivial matter. It is clear that many other networks could be obtained in a similar way, all having the same set $S^1$ and link weights. In fact the network of Fig. AII.4.3 can be simplified to the network of Fig. AII.4.4.

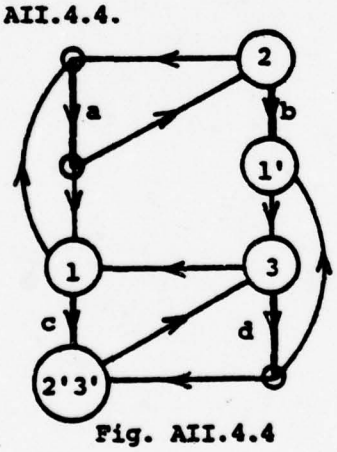



Fig. AII.4.4

Comment: One important observation has to be made with respect to the order of connecting saturated links: we have to avoid short-circuits of previously established paths; that is, we have to avoid making connections that could reduce the length of existing paths.

In previous example, for instance, once we have the paths of the first commodity fixed, Fig. AII.4.2, we are not allow to set any link, say, from t(a) to 1' or from t(c) to 1'.

# APPENDIX III

## EXAMPLES

### III.1 Introduction

In this appendix we present some examples that will illustrate many of the results of this thesis. Example 1 shows a case where none of the link weights, for a first saturation set, is equal to one; furthermore, this example is also used to illustrate some of the points made in Sections 4.4, 4.5 and 4.6. Example 2 presents a case of a first saturation matrix without zero entries and a possible network to which it can correspond. Example 3 is included to show how the simplifying properties of Sections 4.3 and 5.6, regarding subcommodities, can be used in practice; in addition a saturation matrix and distance assignment for a three level problem will be obtained. Finally Example 4 will be used to illustrate some of the dynamic reoptimization properties and rules presented in Chapter VI.

### III.2 Example 1

Consider the collection of reduced circuits of Fig. AIII.1

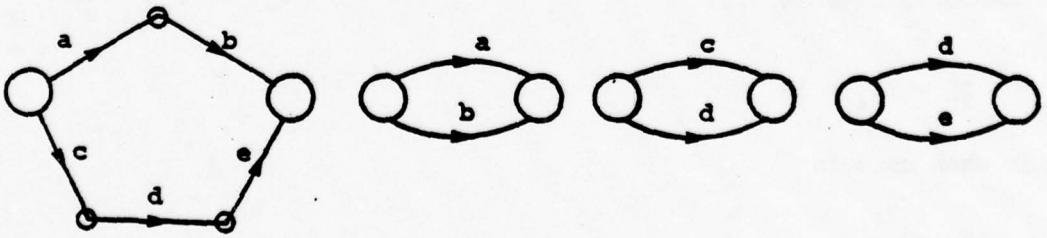



Fig. AIII.1

If we take the direction of flow of the top path on each circuit to be positive, the corresponding Q-matrix ($c_\ell = 1$, all $\ell$) will be

$$Q = \begin{bmatrix} -1 & 1 & 1 & 0 & 0 \\ -1 & 1 & -1 & 0 & 0 \\ -1 & -1 & 0 & 1 & 0 \\ -1 & -1 & 0 & -1 & 1 \\ -1 & -1 & 0 & 0 & -1 \end{bmatrix} \begin{matrix} a \\ b \\ c \\ d \\ e \end{matrix}$$

Solving for the first row of $Q^{-1}$ is equivalent to solve the system of equations (beginning with the last circuit, see Section 4.6, Method I)

$$\theta_d = \theta_e$$

$$\theta_c = \theta_d$$

$$\theta_a = \theta_b$$

$$\theta_a + \theta_b = \theta_c + \theta_d + \theta_e$$

This system implies $2\theta_a = 3\theta_c$ and the smallest set of integers that solve this equation is $\theta_a = 3$, $\theta_c = 2$. From these values we obtain the final solution

$$\theta_a = \theta_b = 3$$

$$\theta_c = \theta_d = \theta_e = 2$$

Furthermore, from (4.5.1)

$$D^1 = \Sigma\theta_\ell = 12$$

which when used in

$$\pi_\ell = \frac{\theta_\ell}{D^1}$$

will give the following dual variables values for the links of $S^1$

$$\pi_a = \pi_b = \frac{-1}{4}$$

$$\pi_c = \pi_d = \pi_\ell = \frac{-1}{6} \quad .$$

For the sake of completeness we can also obtain det Q, using (4.4.12) as follows (we take out the row with the smallest number of zeroes)

$$\det Q = \frac{\det[P_B(\overline{\underline{rd}})]}{\pi_d} (-1)^5 = 6 \det[P_B(\overline{\underline{rd}})]$$

but

$$\det[P_B(\overline{\underline{rd}})] = \begin{vmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ -1 & 0 & 0 & -1 \end{vmatrix} = - \begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix} = 2$$

so

$$\det Q = 12 \quad .$$

The fact that $\det Q = D^1$ is a coincidence and is due to the great number of zeros that appear in Q. In general we will have $\det Q > D^1$.

Using the synthesis procedure of Appendix II we can easily obtain a network for this Q-matrix. The result is given below

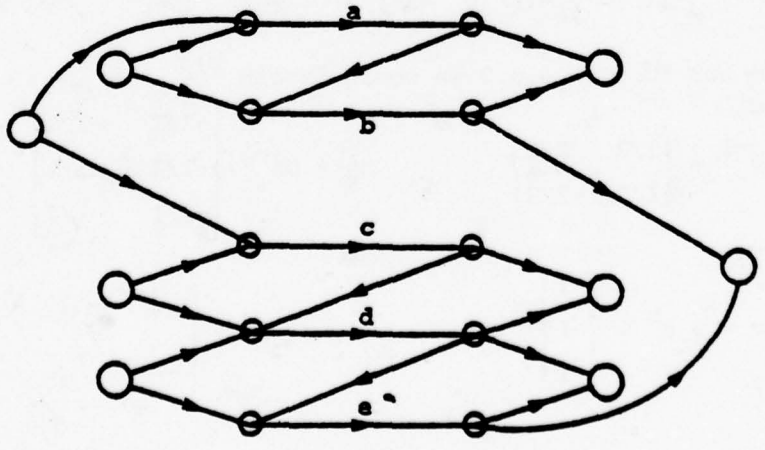



Fig. AIII.2

The network can be simplified as follows

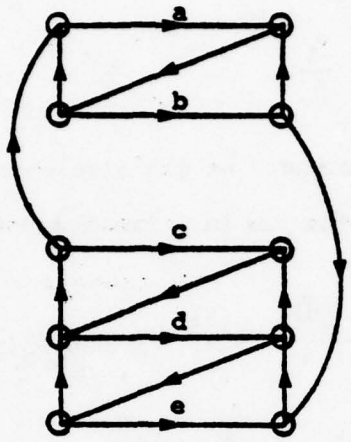



Fig. AIII.3

This scheme suggest that we have two subnetworks weakly connected to each other. This property can also be seen in terms of the Q-matrix if this matrix is written according to the procedure included in Method III of Section 4.6, that is,

$$Q_\alpha = \begin{array}{c} \\ a \\ b \\ c \\ d \\ e \end{array} \begin{array}{c} \overset{\alpha}{\overbrace{\phantom{xxxxx}}} \\ \left[\begin{array}{cc|ccc} 1 & 1 & -1 & 0 & 0 \\ 1 & -1 & -1 & 0 & 0 \\ \hline -1 & 0 & -1 & 1 & 0 \\ -1 & 0 & -1 & -1 & 1 \\ -1 & 0 & -1 & 0 & -1 \end{array}\right] \end{array}$$

Applying now Theorem 4.6.2 we would obtain

$$K^{-1} = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{bmatrix} \qquad D_K = DK^{-1} \begin{bmatrix} -1/2 & -1/2 \\ -1/2 & -1/2 \\ -1/2 & -1/2 \end{bmatrix}$$

$$v = D_K c^K = \begin{bmatrix} -1 \\ -1 \\ -1 \end{bmatrix} \qquad h = K^{-1} c^K = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$G^{-1} = \frac{1}{3}\begin{bmatrix} -1 & -1 & -1 \\ -2 & 1 & 1 \\ -1 & -1 & 2 \end{bmatrix} \qquad \text{so } \pi^G = \frac{1}{3}[-1 \quad -1 \quad -1]$$

$$\text{and } \theta^G = [1 \quad 1 \quad 1]$$

$$\mu = \pi^G v = 1$$

$$\Xi = (1+\mu)I - v\pi^G = \frac{1}{3}\begin{bmatrix} 5 & -1 & -1 \\ -1 & 5 & -1 \\ -1 & -1 & 5 \end{bmatrix}$$

$$\Theta = (1+\mu)I - c^K\pi^G D_K = \frac{1}{2}\begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix}$$

$$X = \frac{1}{1+\mu} K^{-1}\Theta = \frac{1}{4}\begin{bmatrix} 1 & 1 \\ 2 & -2 \end{bmatrix}$$

$$W = \frac{1}{1+\mu} G^{-1}\Xi = \frac{1}{6}\begin{bmatrix} -1 & -1 & -1 \\ -4 & 2 & 2 \\ -2 & -2 & 4 \end{bmatrix}$$

$$Y = hW(\underline{r}1) = \frac{1}{6}\begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

$$U = -WD_K = \frac{1}{4}\begin{bmatrix} -1 & -1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} .$$

Furthermore

$$\pi(S^1) = [U(\underline{r}1) \mid W(\underline{r}1)] = \left[-\frac{1}{4} \quad -\frac{1}{4} \quad -\frac{1}{6} \quad -\frac{1}{6} \quad -\frac{1}{6}\right]$$

which are, of course, the same values obtained before. From previous values we can easily see that all results of Method III, Section 4.6, are satisfied for this example. Note in particular the relation between

$\pi(S^1)$ and $\pi^G$ (Lemma 4.6.4).

It seems evident that Method I of inversion, would have given a better result in this case. Nevertheless the interesting aspect of Method III is how it emphasizes the hierarchical structure existing in a particular solution, and how it provides a partially decentralized way to obtain the inverse. Observe, also, that a similar hierarchy can be obtained from within subnetwork G, since we can write the saturation matrix for it as

$$G_\alpha = \left[\begin{array}{c|cc} 1 & -1 & 0 \\ \hline -1 & -1 & 1 \\ 0 & -1 & -1 \end{array}\right]$$

As a contrast to matrices with this special structure we present in the next example a case where the saturation matrix does not have zero entries and, thus, no sub-hierarchies exist within the first saturation set.

III.3 Example 2

Consider the saturation matrix

$$Q = \begin{bmatrix} -1 & 1 & 1 & -1 \\ -1 & 1 & -1 & 1 \\ -1 & -1 & 1 & 1 \\ -1 & -1 & -1 & 1 \end{bmatrix} \begin{matrix} a \\ b \\ c \\ d \end{matrix}$$

Using Method II, Section 4.6, we can easily invert this matrix to give

$$Q^{-1} = (Q^TQ)^{-1}Q^T = (4I)^{-1}Q^T = \frac{1}{4}Q^T$$

Thus, according to Appendix II, Q is a valid saturation matrix. Furthermore, using the synthesis procedure of that appendix, a possible network for this matrix is given in Fig. AIII.4.

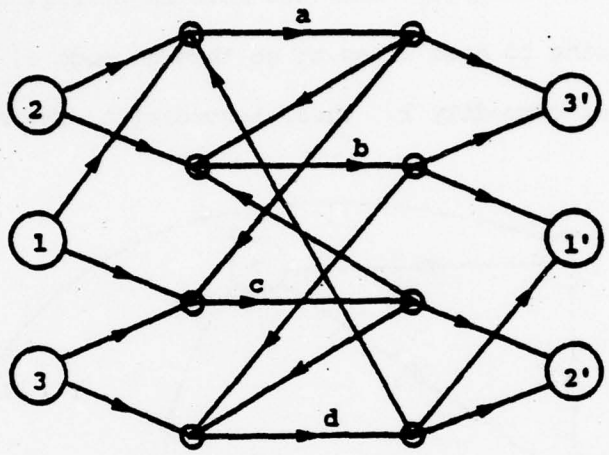



Fig. AIII.4

which can be simplified into Fig. AIII.5

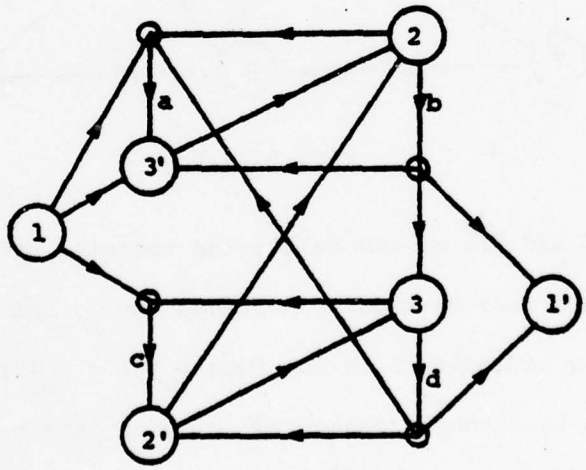



Fig. AIII.5

Next we present a three saturation level example that exhibits many simplying and hierarchical properties.

III.4 Example 3

Consider the strongly connected network of Fig. AIII.6; since all the traffic going to node 1 has to go through node 3, commodity 3 is a subcommodity of commodity 1. Thus, according to the simplifying rules

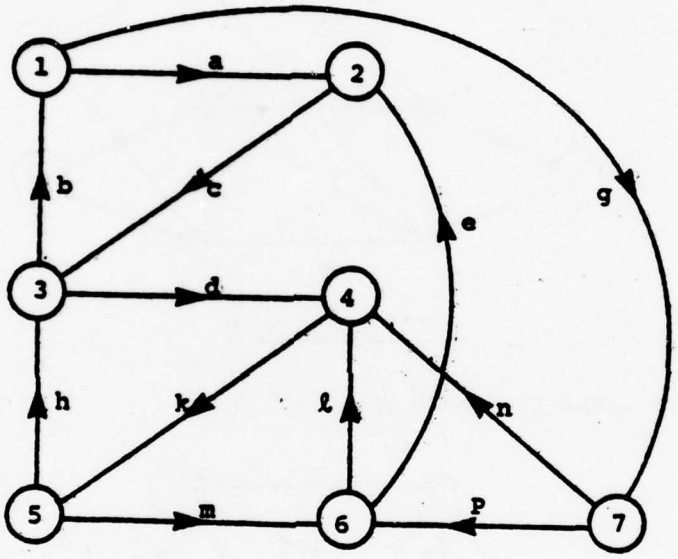



Fig. AIII.6

of Section 4.3 and 5.6 we can delete the spanning tree rooted at 1 just as if there were no traffic destined for 1, and define a new requirement for commodity 3 in the form $\tilde{r}_m(3) = r_m(3) + r_m(1)$, all m (the network is strongly connected). Since link b was in the spanning tree T(1) but is not in T(3), the amount

$$f_b(1) = \sum_m r_m(1)$$

has to be added to the current aggregate flow of link b in order to obtain the total flow of that link. Thru this re-arrangement one is able to reduce the dimension of the constraint matrix by $|T(1)| = N-1 = 6$ without destroying the identity of the problem.

Acting in the same way we can delete also the spanning trees rooted at nodes 5 and 7, giving a total reduction of dimensionality of

$$|T(1)| + |T(5)| + |T(7)| = 18$$

with respect to the initial dimension: rank $A^1 = L + N(N-1) = 12 + 42 = 54$. By means of this reduction we are, in effect, deleting rows of $A^1$.

Suppose we choose now a set of starting spanning trees, for the non-deleted commodities, as given by Fig. AIII.7. In this figure the continuous lines correspond to branches of the trees and the dotted lines to the non-tree links, i.e., those links (or flows) generating alternate paths. From this tree selection the set $F_{BT}$ will be generated by the continuous links and the complementary set $F - F_{BT}$ by the dotted links.

Since from Chapters IV and V we know that all the relevant information of a given solution can be obtained from the set $F - F_{BT}$ we will concentrate on this latter set. From Fig. AIII.7 we can easily obtain $|F - F_{BT}| = 18$; on the other hand Lemmas 4.3.7-8 provide us with some tools to simplify even further(*) the set $F - F_{BT}$. For example, note

---

(*) Note that the spanning tree deletion has already reduced the number of elements of $F - F_{BT}$ by eliminating all the flows generated by the non-tree branches of the deleted trees.

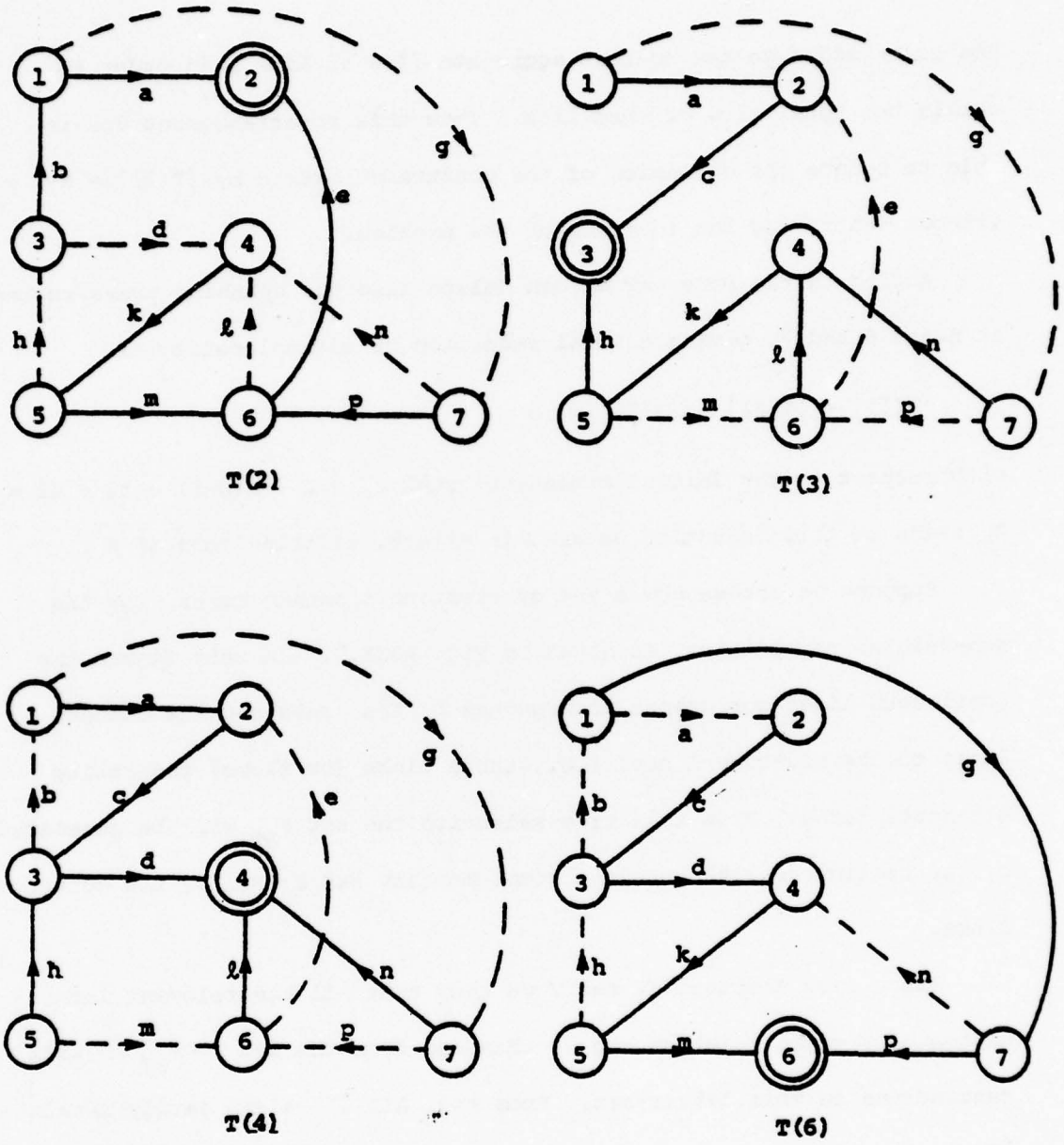



Fig. AIII.7

that $f_\ell(2)$, $f_m(3)$, $f_b(4)$ and $f_h(6)$ can never be basics because their corresponding links introduce cycles when added to the trees, so that they can be deleted from $F - F_{BT}$. Note also that $f_n(2)$ and $f_n(6)$ (or $f_p(3)$ and $f_p(4)$) generate the same circuit so one of them, say $f_n(2)$ ($f_p(3)$) can be eliminated. This last operation reduces the set $F - F_{BT}$ to the following collection of elements

$$F - F_{BT} = \{f_d(2),\ f_g(2),\ f_h(2),\ f_e(3),\ f_g(3),\ f_e(4),\ f_g(4),\ f_m(4),$$

$$f_p(4),\ f_a(6),\ f_b(6),\ f_n(6)\}$$

$$|F - F_{BT}| = 12 \ .$$

Thus through all these simplifications we have effectively reduced the number of columns of matrix $A^1$ by

$$\text{non-tree branches of } [T(1),\ T(5),\ T(7)] + 6 = 12 + 6 = 18 \ .$$

Since, as is known from Chapter IV, we can discard the spanning tree part of the constraint matrix once a set of spanning trees are known, we can conclude that one is able to solve this saturation problem working with an actual constraint matrix of (roughly)(*) dimension

$$L \times [L + (1 + |F - F_{BT}|)] = 12 \times 25 \qquad (**)$$

instead of the initial dimension, see (3.1.3), of

$$[L + N(N-1)] \times [NL + 1] = 54 \times 85 \ .$$

---

(*) Since during the solution process one might be forced to change the structure of the spanning trees, and each spanning tree could introduce different simplifications, the number of columns may change slightly during the process.

(**) The 1 stands for $\alpha^1$.

Thus the reduction of dimensionality is dramatic. But this is not all since, as we also know, one can reduce the dimension of the actual working basis to a value equal to the number of saturated links. To particularize this fact in the present example consider that we are given some vector of requirement values, for the network of Fig. AIII.6 with $c_\ell = 1$, all $\ell$, and that solving the first saturation problem we find, see Example 3.6.1,

$$S^1 = \{a, g, d, m\}$$

$$F^1_{BA} = \{f_d(2), f_m(4), f_a(6)\} .$$

The saturation matrix will then be

$$Q^1 = \begin{array}{c} \begin{matrix} \alpha^1 & d^2 & m^4 & a^6 \end{matrix} \\ \begin{bmatrix} -1 & -1 & 0 & 1 \\ -1 & 1 & -1 & 1 \\ -1 & 0 & 0 & -1 \\ -1 & 1 & 1 & 1 \end{bmatrix} \end{array} \begin{matrix} a \\ d \\ g \\ m \end{matrix}$$

where $\ell^n$ is a shorthand notation for $f_\ell(n)$. To obtain the dual variables, i.e., the first row of $Q^{-1}$, we can either apply the procedure explained in Method I, Section 4.6, or use the canonic equations, as in Example 3.6.1. In either case it is almost immediate to see that

$$\theta_g = 4, \quad \theta_a = 2, \quad \theta_d = \theta_m = 1$$

$$D^1 = \sum_\ell \theta_\ell = 8$$

$$\pi_g = -\frac{1}{2}, \quad \pi_a = -\frac{1}{4}, \quad \pi_d = \pi_m = -\frac{1}{8} ,$$

and with these values (distances) we can easily obtain

$$F_0^1 = \{f_e(3),\ f_p(4)\}$$

$$F_+^1 = \{f_g(2),\ f_h(2),\ f_g(3),\ f_g(4),\ f_e(4),\ f_b(6),\ f_n(6)\}$$

Thus all the relevant information for this first problem is contained in $Q^1$ and the spanning trees.

Since the solution to (LP1) is not unique, i.e., $F_0^1 \neq \phi$, we can attempt the solution of (LP2). Suppose that the requirement values are such that an optimal solution to (LP2) gives

$$S^2 = \{h,e\}\ ,\quad F_{BA}^2 = \{f_e(3)\}$$

Since $C_e(3) \cap S^1 = \phi$ we are in a case where the composite saturation matrix can be written in lower triangular form. As illustrated below we can therefore work with two small saturation matrices $Q^1$ and $Q^2$, whose dimensions are $|S^1|$ and $|S^2|$, respectively, instead of a larger composite matrix of dimension $|S^1| + |S^2|$.(*)

$Q^{12} =$

| $\alpha^1$ | $d^2$ | $m^4$ | $a^6$ | $\beta^2$ | $e^3$ | |
|---|---|---|---|---|---|---|
| -1 | -1 | | 1 | | | a |
| -1 | 1 | -1 | 1 | | | d |
| -1 | | | -1 | | | g |
| -1 | 1 | 1 | 1 | | | m |
| -1 | | | 1 | 1 | 1 | e |
| -1 | 1 | | 1 | 1 | -1 | h |

$=$

| | |
|---|---|
| $Q^1$ | |
| $G^{12}$ | $Q^2$ |

$(Q^{12})^{-1} =$

| | |
|---|---|
| $(Q^1)^{-1}$ | |
| $-(Q^2)^{-1}G^{12}(Q^1)^{-1}$ | $(Q^2)^{-1}$ |

(*) Empty places correspond to zeros, as usual.

The dual variables, which can be easily obtained by any of the many procedures already known, will be

$$\pi_h^2 = \pi_e^2 = -\frac{1}{2}$$

$$\pi_a^2 = -\frac{1}{4}, \quad \pi_d^2 = \frac{3}{8}, \quad \pi_g^2 = 0, \quad \pi_m^2 = -\frac{1}{8} .$$

Furthermore since

$$c_p(4) \cap (S^1 \cup S^2) = \phi$$

we have

$$F_0^2 = \{f_p(4)\}$$

$$F_+^2 = F_0^1 - F_{BA}^2 - F_0^2 = \phi .$$

Thus the solution is not unique and we shall have to solve a third saturation level problem. But first, to obtain a positive distance assignment at level two we can compute a solution to ($LP2^+$), as indicated in Section 5.4, to give

$$-\pi^{2+}(S^1 \cup S^2) = (1, 0, \tfrac{3}{2}, \tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2})$$

from which we obtain the weights

$$\theta^{2+}(S^1 \cup S^2) = (2, 0, 3, 1, 1, 1) .$$

Since for these weights

$$z^{2+}(F_+^1) > 0,$$

they represent a valid distance assignment to describe the alternate paths up to the second saturation level.

Now we move to the third level. At this level we make active (basic) the only flow variable in $F_0^2$, i.e., $f_p(4)$. Let us suppose that as a result

$$S^3 = \{n,p\}$$

Thus we have

$Q^{123} =$

| $\alpha^1$ | $d^2$ | $m^4$ | $a^6$ | $\beta^2$ | $e^3$ | $\beta^3$ | $p^4$ | |
|---|---|---|---|---|---|---|---|---|
| -1 | -1 | | -1 | | | | | a |
| -1 | 1 | -1 | 1 | | | | | d |
| -1 | | | -1 | | | | | g |
| -1 | 1 | 1 | 1 | | | | | m |
| -1 | | | 1 | 1 | 1 | | | e |
| -1 | 1 | | 1 | 1 | -1 | | | h |
| -1 | | | | 1 | | 1 | -1 | n |
| -1 | | | -1 | 1 | | 1 | 1 | p |

$=$

| | | |
|---|---|---|
| $Q^1$ | | |
| $G^{12}$ | $Q^2$ | |
| $G^{13}$ | $G^{23}$ | $Q^3$ |

which, like $Q^{12}$, also has the lower triangular property. The corresponding dual variables are

$$\pi^3(S^{123}) = (-\tfrac{1}{8}, -\tfrac{1}{16}, \tfrac{1}{4}, -\tfrac{1}{16}, 0, 0, -\tfrac{1}{2}, -\tfrac{1}{2})$$

and

$$-\pi^{3+}(S^{123}) = (\tfrac{1}{3}, 0, 0, \tfrac{1}{6}, \tfrac{1}{6}, \tfrac{1}{6}, \tfrac{1}{2}, \tfrac{1}{2})$$

so that

$$\theta^{3+} = (2, 0, 0, 1, 1, 1, 3, 3) .$$

Since, furthermore, for these weights

$$z^{3+}(F_+^1 \cup F_+^2) = z^{3+}(F_+^1) > 0$$

they represent a valid assignment for the successive saturation problem. This assignment is represented in Fig. AIII.8

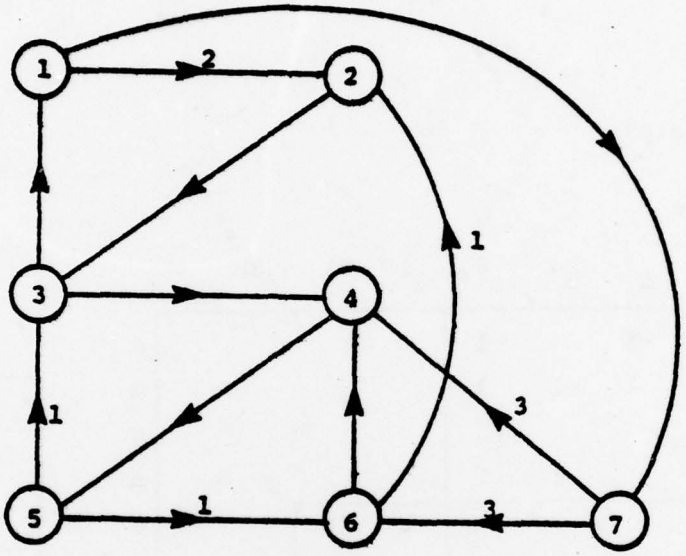



Fig. AIII.8

As we can see only half of the lines have a nonzero distance. The largest weight is 3 in contrast with the set of weights for (LP1) where the largest was 4. Thus we see that descending in the hierarchy of problems does not necessarily mean increasing the weights since new links are allowed to be nonzero from problem to problem.

This concludes the present example, which has been used to illustrate many of the simplifying properties described in the thesis. Now we move to a three node example to illustrate some dynamic behaviour of the solution in response to requirement changes.

III.5 Example 4

Here we want to illustrate some of the possible reoptimization cases explained in Chapter VI. Since the details of these reoptimizations are included in that chapter we will omit them here and consider only the evolution of the optimal solutions. Furthermore since we will be dealing with a three node example the reoptimization can be done by inspection; nevertheless the interested reader can go to Chapter VI and reconstruct the corresponding steps needed to move from one optimal solution to another.

Starting with the solution of Example 2.4.1, represented in Fig. AIII.9a, an increment $\Delta r_1^2 = 2$ will give the optimal solution of Fig. AIII.9b,

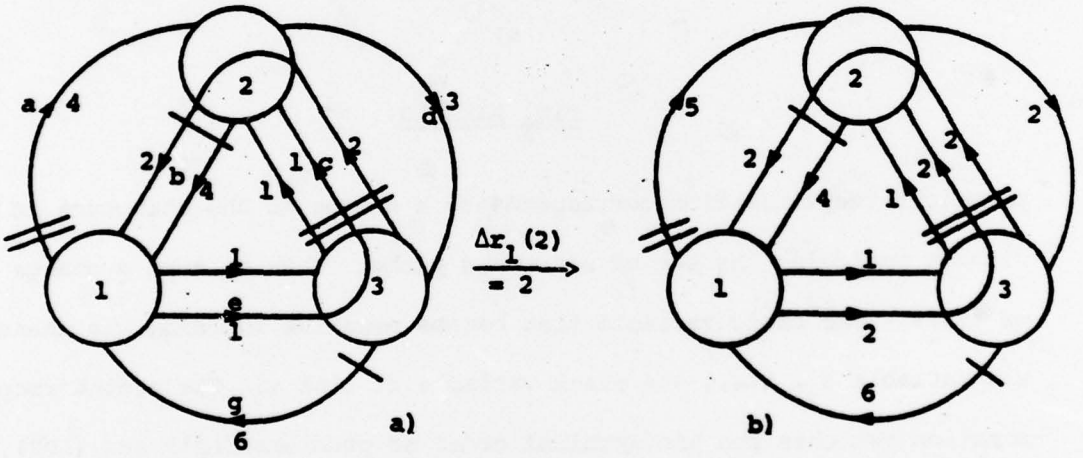



Fig. AIII.9

for which no change of basis is needed. The saturation sets still are (they are also marked in the figure)

$$S^1 = \{b,g\} \; ; \quad S^2 = \{a,c\}$$

and the new saturation levels

$\alpha_0^1 = 6$ , $\alpha_0^2 = 5$.

Now, suppose that for the solution of Fig. AIII.9b we have first an increment $\Delta r_1^3 = 3$ and then an additional increment $\Delta r_1^3 = 2$. The new solutions are given in Fig. AIII.10a and b, respectively.

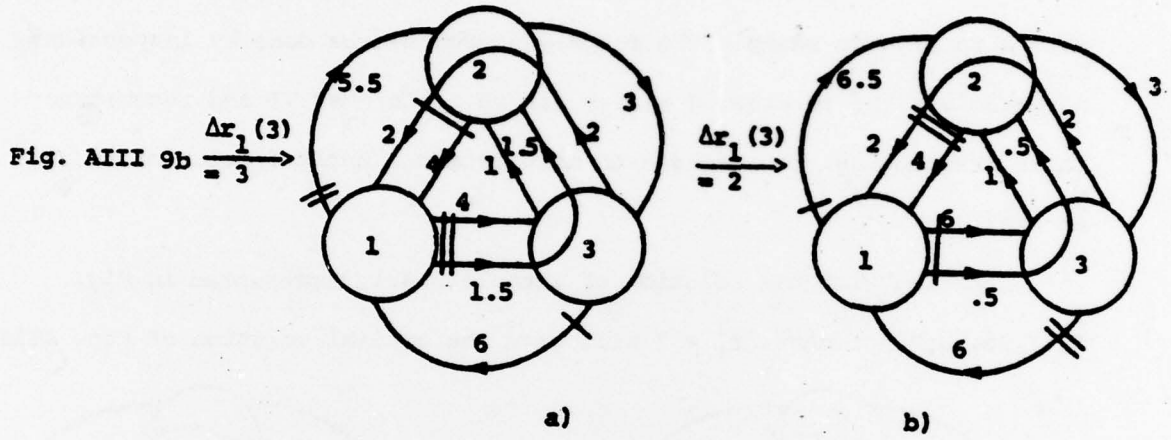



Fig. AIII.10

The first reoptimization corresponds to a change in the structure of $S^2$ without modifying the set of alternate paths. Thus we have a change of basis. The basic variable that became negative inducing the change was variable $s_e$, i.e., the slack variable of link e. The second reoptimization swtiches the hierarchical order of problems (LP1) and (LP2), thus $S^1$ becomes $S^2$ and vice versa. This change was induced by variable $\beta^2$, i.e., the slack between $\alpha_0^1$ and $\alpha_0^2$, becoming negative. The reason why the solutions switch without further modifications is twofold: on one hand, the old $S^1$ was not preventing any alternate path of the $R^2$ commodities from becoming active (basic); on the other hand the new $S^1$ does not eliminate any of the alternate paths of the old $R^1$

commodities. These facts are reflected in the composite saturation matrix by a block diagonal structure of the matrix.

$$Q^{12} = \begin{array}{|c|c|l} \hline Q^1 & 0 & s^1 \\ \hline 0 & Q^2 & s^2 \\ \hline \end{array}$$

Next consider an additional increment $\Delta r_1^3 = 3$ with respect to the requirements of the solution of Fig. AIII.10b. The new solution is given in Fig. AIII.11.

Fig. AIII.10 b $\xrightarrow{\Delta r_1(3) = 3}$

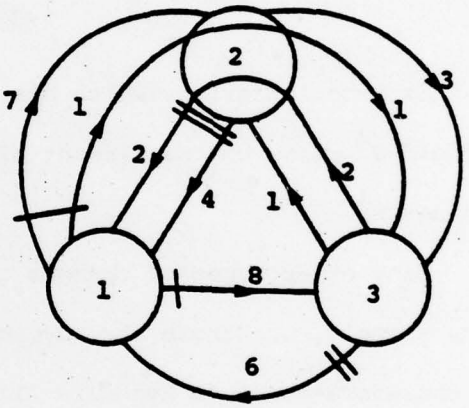



Fig. AIII.11

This last reoptimization correspond to a change in the set of basic alternate paths without changing the composition of the saturation sets. This change of basis was induced by the flow variable $f_e(2)$ becoming negative.

Finally consider that an increment of value $\Delta r_1^2 = 7$ has been added to the solution of Fig. AIII.9a, then the new solution is given in Fig. AIII.12.

Fig. AIII.9a $\xrightarrow[= 7]{\Delta r_1(2)}$

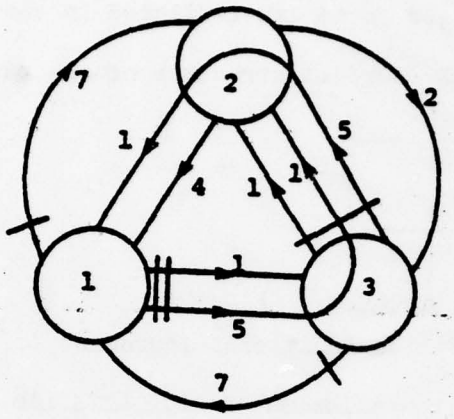



Fig. AIII.12

In this reoptimization, which has been induced by $\beta^2$ becoming negative, the old $s^2$ picks up one element of the old $s^1$ in the process of becoming the new $s^1$.

Many other types of changes could be established with this three node example, but those already shown suffice to illustrate the process. We therefore conclude Appendix III here.

## BIBLIOGRAPHY

[1] N.M. Abramson and F.F. Kuo, Eds., Computer-Communication Networks, Prentice Hall, 1973.

[2] D.W. Davies and D.L.A. Barber, Communication Networks for Computers, Wiley, 1973.

[3] P.E. Green, Jr. and R.W. Lucky, Eds., Computer Communications, IEEE Press, 1974.

[4] Computer Communications, Special Issue on IEEE Trans. on Comm. Vol. COM-25, Jan. 1977.

[5] R.G. Gallager, "Data-Communication Networks", Course 6.263, Fall 1977, MIT.

[6] L. Kleinrock, W.E. Naylor and H. Opderbeck, "A Study of Line Overhead in the Arpanet". Communications of the ACM, pp. 3-13, Vol. 19, No. 1, Jan. 1976.

[7] M. Athans, "Perspectives in Modern Control Theory", (to be published).

[8] N.R. Sandell, P. Varaiya, M. Athans and M. Safonov, "A Survey of Centralized Control Methods for Large Scale Systems", IEEE Trans. on A.C., April 1978.

[9] R.G. Gallager,"A Minimum Delay Routing Algorithm Using Distributed Computation", IEEE Trans. on Comm., Vol. COM-25, No. 1, Jan. 1977.

[10] J.M. Wozencraft, R.G. Gallager and A. Segall, "First Annual Report on Data Network Reliability", ESL internal report ESL-IR-677, M.I.T. July 1976.

[11] J.E. Defenderfer, "Comparative Analysis of Routing Algorithms for Computer Networks", MIT, ScD Thesis, March 1977, also available as ESL report N.756.

[12] M.J. Marcus, "New Approaches to the Analysis of Connecting and Sorting Networks", R.L.E. Technical Report 486, M.I.T., March 17, 1972.

[13] N. Pippenger, "The Complexity Theory of Switching Networks", R.L.E. Tech. Report 487, M.I.T., Dec. 19, 1973.

[14] D.G. Cantor and M. Gerla, "Optimal Routing in a Packet Switched Computer Network", IEEE Trans. on Computers, Vol. C-23, Oct. 1974.

[15] T.S. Stern, "A Class of Decentralized Routing Algorithms Using Relaxation", IEEE Trans. on Comm. Vol. Com-25, No. 10, Oct. 1977.

[16] F.E. Heart, R.E. Kahn, S.M. Ornstein, W.R. Crowther and D.C. Walden, "The Interface Message Processor for the ARPA Computer Network", in Conf. Rec. 1970 Spring Joint Comput. Conf. AFIPS Conf. Proc. 1970, pp. 551-66.

[17] R.G. Gallager, "Loops in Multicommodity Flows", ESL report No. 772.

[18] L. Kleinrock, Communication Nets: Stochastic Message Flow and Delay, New York: McGraw-Hill, 1964.

[19] _____, "Analytic and Simulation Methods in Computer Network Design", in Conf. Rec. Spring Joint Comput. Conf., AFIPS Conf. Proc., 1970, pp. 569-579.

[20] A. Segall, "The Modeling of Adaptive Routing in Data Communication Networks", IEEE Trans. on Comm. Vol. COM-25, No. 1, January 1977.

[21] F. Moss, Ph.D. dissertation, Massachusetts Inst. Tech., Cambridge, 1976.

[22] T.C. Hu, Integer Programming and Network Flows, Addison Wesley, 1970.

[23] H. Frank and I.T. Frisch, Communication, Transmission, and Transportation Networks, Addison-Wesley, 1971.

[24] G. Dantzig, Linear Programming and Extensions, Princeton, U.P., 1963.

[25] D.A. Pierre, Optimization Theory with Application, Wiley 1969.

[26] M.D. Mesarovic, D. Macko and Y. Takahara, Theory of Hierarchical, Multilevel Systems, Academic Press, 1970.

[27] F. Harary, Graph Theory, Addison-Wesley, 1969.

[28]. D.G. Luenberger, Introduction to Linear and Nonlinear Programming, Addison-Wesley, 1973.

[29] G. Hadley, Linear Programming, Addison-Wesley, 1962.

[30] G.B. Dantzig and R.M. Van Slyke, "Generalized Upper Bounding Techniques", J. Computer System Science, 1, 1967, pp. 213-226.

[31] S. Chen and R. Saigal, "A Primal Algorithm for Solving a Capacitated Network Flow Problem with Additional Linear Constraints", Networks, 7: 59-79, 1977.

[32] J.K. Hartman and L.S. Lasdon, "A Generalized Upper Bounding Algorithm for Multicommodity Network Flow Problems", Networks, 1.4, 1972, pp. 333-354.

[33] C.J. McCallum, Jr., "A Generalized Upper Bounding Approach to a Communications Network Planning Problem", Networks, 7, 1977, pp. 1-23.

[34] A.A. Assad, "Multicommodity Network Flows - A Survey", Oper. Res. Center M.I.T. Working Paper OR 045-75 (Dec. 1975).

[35] _____, "Multicommodity Network Flows-Computational Experience", Oper. Res. Center M.I.T. Working Paper OR 058-76 (Oct. 1976).

[36] L.S. Lasdon, "A Survey of Large Scale Mathematical Programming", in Handbook of Operations Research. Vol. I, Van Nostrand Rheinhold Co. 1978.

[37] M. Held, P. Wolfe and H.P. Crowder, "Validation of Subgradient Optimization", Math. Prog. 6,1, 62-88 (Feb., 1974).

[38] C. Berge and A. Ghouila-Houri, Programming, Games and Transportations Networks, Wiley, 1965.

[39] R. Bellman, Introduction to Matrix Analysis, McGraw Hill, 1960.

[40] F. Ayres, Matrices, Schaum Publishing Co. 1962.

[41] C.L. Liu, Introduction to Combinatorial Mathematics, McGraw Hill, 1968.

[42] G.H. Hardy and E.M. Wright, An Introduction to the Theory of Numbers, 4th Ed. Oxford Univ. Press, London, 1960.

[43] P.A. Humblet, "Source Coding for Communication Concentrators" E.S.L. Report No. 798, M.I.T., Jan. 1978.

[44] M. Simonnard, Linear Programming, Prentice-Hall 1966.

[45] W. Orchard-Hays, Advanced Linear-Programming Computer Techniques, MacGraw-Hill, 1968.

[46] S.P. Bradley, A.C. Max and T.L. Magnanti, Applied Mathematical Programming, Addison-Wesley, 1977.

[47] A. Pierce, "Bibliography on Algorithms for Shortest-Path, Shortest Spanning Tree and Related Circuit-Routing Problems (1956-1974)", Networks 5(2), 129-150 (1975).

[48] S. Seshu and M.B. Reed, Linear Graph and Electrical Networks, Addison-Wesley, 1961.

[49] M. Gerla, "The Design of Store-and-Forward (S/F) Networks for Computer Communications", Ph.D. dissertation, Dep. COmp. Sci., Univ. California, Los Angeles 1973.

A031608

## BIOGRAPHICAL NOTE

Francisco A. Ros Perán was born in Lorca (Spain) in 1950. He graduated from the Universidad Politécnica of Madrid as an Ingeniero de Telecomunicación in 1972. During his Senior year he became a Teaching Assistant in Electromagnetism and collaborated with the staff in the Department of Systems and Data Transmissions of the Escuela Técnica Superior de Ingenieros de Telecomunicación. In 1972 he became an Assistant Professor teaching undergraduate courses on System Theory and Data Transmission. In 1974 he was given a Fellowship from the ITP Foundation (Spain) for continuing education at the Massachusetts Institute of Technology where he finished the degree of Master of Science in 1976. His Master thesis was entitled "Dynamic File Assignment in a Computer Network" (also available as MIT, ESL Report No. 667, June 1976). He has published a set of class notes on Introducción a la Transmisión Digital, Departamento de Publicaciones ETSIT, Madrid 1973.

## Distribution List

| | |
|---|---|
| Defense Documentation Center<br>Cameron Station<br>Alexandria, Virginia 22314 | 12 Copies |
| Assistant Chief for Technology<br>Office of Naval Research, Code 200<br>Arlington, Virginia 22217 | 1 Copy |
| Office of Naval Research<br>Information Systems Program<br>Code 437<br>Arlington, Virginia 22217 | 2 Copies |
| Office of Naval Research<br>Branch Office, Boston<br>495 Summer Street<br>Boston, Massachusetts 02210 | 1 Copy |
| Office of Naval Research<br>Branch Office, Chicago<br>536 South Clark Street<br>Chicago, Illinois 60605 | 1 Copy |
| Office of Naval Research<br>Branch Office, Pasadena<br>1030 East Greet Street<br>Pasadena, California 91106 | 1 Copy |
| New York Area Office (ONR)<br>715 Broadway - 5th Floor<br>New York, New York 10003 | 1 Copy |
| Naval Research Laboratory<br>Technical Information Division, Code 2627<br>Washington, D.C. 20375 | 6 Copies |
| Dr. A. L. Slafkosky<br>Scientific Advisor<br>Commandant of the Marine Corps (Code RD-1)<br>Washington, D.C. 20380 | 1 Copy |

Office of Naval Research — 1 Copy
Code 455
Arlington, Virginia 22217

Office of Naval Research — 1 Copy
Code 458
Arlington, Virginia 22217

Naval Electronics Laboratory Center — 1 Copy
Advanced Software Technology Division
Code 5200
San Diego, California 92152

Mr. E. H. Gleissner — 1 Copy
Naval Ship Research & Development Center
Computation and Mathematics Department
Bethesda, Maryland 20084

Captain Grace M. Hopper — 1 Copy
NAICOM/MIS Planning Branch (OP-916D)
Office of Chief of Naval Operations
Washington, D.C. 20350

Mr. Kin B. Thompson — 1 Copy
Technical Director
Information Systems Division (OP-91T)
Office of Chief of Naval Operations
Washington, D.C. 20350

Advanced Research Projects Agency — 1 Copy
Information Processing Techniques
1400 Wilson Boulevard
Arlington, Virginia 22209

Dr. Stuart L. Brodsky — 1 Copy
Office of Naval Research
Code 432
Arlington, Virginia 22217

Captain Richard L. Martin, USN — 1 Copy
Commanding Officer
USS Francis Marion (LPA-249)
FPO New York 09501