REPORT NO. DOT-TSC-OST-77-62

PROPAGATION OF DISTURBANCES IN TRAFFIC FLOW

Pierre Dersin
S.B. Gershwin
Michael Athans

Massachusetts Institute of Technology
Cambridge MA 02139

SEPTEMBER 1977

FINAL REPORT

DOCUMENT IS AVAILABLE TO THE U.S. PUBLIC THROUGH THE NATIONAL TECHNICAL INFORMATION SERVICE, SPRINGFIELD, VIRGINIA 22161

Prepared for

U.S. DEPARTMENT OF TRANSPORTATION
OFFICE OF THE SECRETARY
Office of the Assistant Secretary for Systems Development and Technology
Office of Systems Engineering
Washington DC 20590

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
DOT-TSC-OST-77-62		

4. Title and Subtitle	5. Report Date / 6. Performing Organization Code
PROPAGATION OF DISTURBANCES IN TRAFFIC FLOW	5. Report Date: September 1977 6. Performing Organization Code:
7. Author(s): Pierre Dersin, S.B. Gershwin, and Michael Athans	8. Performing Organization Report No.: DOT-TSC-OST-76-50
9. Performing Organization Name and Address: Massachusetts Institute of Technology* Electronic Systems Laboratory Department of Elec. Eng. and Computer Sci. Cambridge MA 02139	10. Work Unit No. (TRAIS): OS750/R8526 11. Contract or Grant No.: DOT-TSC-849
12. Sponsoring Agency Name and Address: U.S. Department of Transportation Office of the Secretary Off. of the Asst. Sec. for Sys. Dev. & Tech. Office of Systems Engineering Washington DC 20590	13. Type of Report and Period Covered: Final Report June 1975-Oct. 1976 14. Sponsoring Agency Code:

15. Supplementary Notes

*Under contract to: U.S. Department of Transportation
Transportation Systems Center
Kendall Square
Cambridge MA 02142

16. Abstract

The system-optimized static traffic-assignment problem in a freeway corridor network is the problem of choosing a distribution of vehicles in the network to minimize average travel time.

It is of interest to know how sensitive the optimal steady-state traffic distribution is to external changes including accidents and variations in incoming traffic.

Such a sensitivity analysis is performed via dynamic programming. The propagation of external perturbations is studied by numerical implementation of the dynamic programming equations.

When the network displays a certain regularity and satisfies certain conditions, we prove, using modern control theory and graph theory, that the effects of imposed perturbations which contribute no change in total flow decrease exponentially as distance from the incident site increases. We also characterize the impact of perturbations with nonzero total flow. The results confirm numerical experience and provide bounds for the effects as functions of distance.

17. Key Words	18. Distribution Statement
Traffic, Stability, Disturbance, Propagation, Assignment, Accident, Exponential Decay	DOCUMENT IS AVAILABLE TO THE U.S. PUBLIC THROUGH THE NATIONAL TECHNICAL INFORMATION SERVICE, SPRINGFIELD, VIRGINIA 22161

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	294	

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

PREFACE

This research was conducted at the Transportation group of M.I.T. Electronic Systems Laboratory and sponsored by the Department of Transportation under contract DOT/TSC/849 and a graduate fellowship of the Belgian American Educational Foundation.

We are grateful to Dr. Diarmuid O'Mathuna and Ms. Ann Muzyka of TSC, and Dr. Robert Ravera of DOT for their encouragement, support and criticism.

We wish to extend our warmest thanks to Ms. Margaret Flaherty, Ms. Susanna Natti, Ms. Myra Sbarounis, Ms. Susan Garland and Ms. Barbara Peacock. Under the skilled supervision of the latter, they have ensured a beautiful typing after a long struggle with equations and handwriting. Ms. Susanna Natti should be particularly thanked, for having taken quite more than her share of the burden.

Mr. Arthur Giordani and Mr. Norman Darling, of the drafting department, deserve a special mention for their fine craftsmanship.

METRIC CONVERSION FACTORS

Approximate Conversions to Metric Measures

Symbol	When You Know	Multiply by	To Find	Symbol
		LENGTH		
in	inches	2.5	centimeters	cm
ft	feet	30	centimeters	cm
yd	yards	0.9	meters	m
mi	miles	1.6	kilometers	km
		AREA		
in^2	square inches	6.5	square centimeters	cm^2
ft^2	square feet	0.09	square meters	m^2
yd^2	square yards	0.8	square meters	m^2
mi^2	square miles	2.6	square kilometers	km^2
	acres	0.4	hectares	ha
		MASS (weight)		
oz	ounces	28	grams	g
lb	pounds	0.45	kilograms	kg
	short tons (2000 lb)	0.9	tonnes	t
		VOLUME		
tsp	teaspoons	5	milliliters	ml
Tbsp	tablespoons	15	milliliters	ml
fl oz	fluid ounces	30	milliliters	ml
c	cups	0.24	liters	l
pt	pints	0.47	liters	l
qt	quarts	0.95	liters	l
gal	gallons	3.8	liters	l
ft^3	cubic feet	0.03	cubic meters	m^3
yd^3	cubic yards	0.76	cubic meters	m^3
		TEMPERATURE (exact)		
°F	Fahrenheit temperature	5/9 (after subtracting 32)	Celsius temperature	°C

Approximate Conversions from Metric Measures

Symbol	When You Know	Multiply by	To Find	Symbol
		LENGTH		
mm	millimeters	0.04	inches	in
cm	centimeters	0.4	inches	in
m	meters	3.3	feet	ft
m	meters	1.1	yards	yd
km	kilometers	0.6	miles	mi
		AREA		
cm^2	square centimeters	0.16	square inches	in^2
m^2	square meters	1.2	square yards	yd^2
km^2	square kilometers	0.4	square miles	mi^2
ha	hectares (10,000 m^2)	2.5	acres	
		MASS (weight)		
g	grams	0.035	ounces	oz
kg	kilograms	2.2	pounds	lb
t	tonnes (1000 kg)	1.1	short tons	
		VOLUME		
ml	milliliters	0.03	fluid ounces	fl oz
l	liters	2.1	pints	pt
l	liters	1.06	quarts	qt
l	liters	0.26	gallons	gal
m^3	cubic meters	35	cubic feet	ft^3
m^3	cubic meters	1.3	cubic yards	yd^3
		TEMPERATURE (exact)		
°C	Celsius temperature	9/5 (then add 32)	Fahrenheit temperature	°F

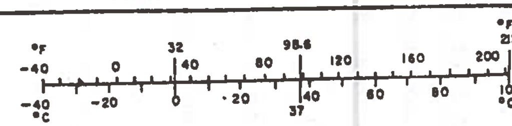

TABLE OF CONTENTS

SECTION PAGE

ILLUSTRATIONS

TABLES

1. INTRODUCTION

1.1 PROBLEM STATEMENT

1.1.1 Physical motivation

This work represents part of an effort dealing with the dynamic stochastic control of freeway corridor systems. The purpose of the research program is to derive and analyze methods of optimization, estimation, and control which may be applicable in achieving the most efficient use possible of a given freeway-corridor network system.

A freeway corridor network system is defined as a set of roads whose primary purpose is to carry large volumes of automobile traffic between a central business district (CBD) of a city and the neighboring residential areas. It consists of one or more parallel freeways and one or more parallel signalized arterials which carry the bulk of the traffic, as well as other roads (such as freeway entrance ramps) which connect these highways [1], [2]. See Fig. 1.1.1 for a graphic illustration of a simple freeway corridor. Our research objective is to achieve an improved efficiency in the use of the existing roadway structure, rather than modifying that structure. Hence, real time computer control is necessary to improve traffic flow as a whole.

The problem of real-time computer control of a freeway corridor can be looked upon as a large scale stochastic dynamic optimal control problem. Such problems cannot be solved exactly, and some approximate solution must be sought. Our approach has been to separate the estimation, incident detection, and control activities, and then to further subdivide control into static and dynamic strategies.

Static network optimization seeks to find a traffic flow configuration in the network, which is optimal in some sense, assuming all system parameters are constant in time. While the network is never actually at steady state since cars are always moving, certain parameters, such as average flows, link capacities, and others, are constant or nearly constant so that static optimization methods may be applied.

Static network optimization is closely related to the problem of traffic assignment, in which the distribution of vehicles in a roadway network is

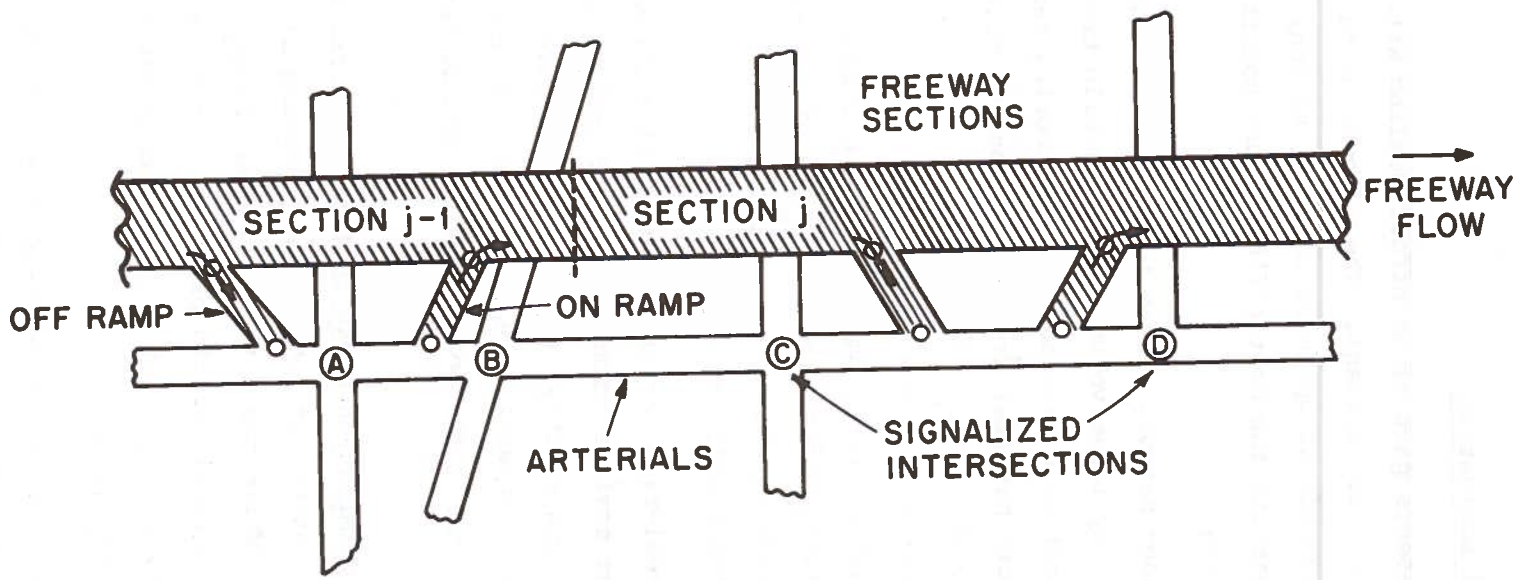

Figure 1.1.1 ILLUSTRATIVE FREEWAY CORRIDOR NETWORK

sought as the solution of a mathematical programming problem. In a system optimized assignment, a central network operator selects vehicles' routes to minimize average travel time, energy consumption, or other criterion. In user-optimization, drivers themselves choose routes to minimize their own travel time. In either case, traffic assignment produces an allocation of vehicles to roadway links, and fractions of traffic changing direction at each decision point in the network (such as ramps, and intersections). In this report, we consider system-optimized assignments, in which there is a cost function that is minimized. In particular, we minimize the average time vehicles are in the network, by finding the optimal flow of vehicles on each and every network link, given a set of demands.

When an incident occurs, or when there is a significant change in parameters that are usually constant (such as a capacity or an average arriving flow) at a node, the static assignment must be recalculated. This is expensive and time consuming (in terms of computer time), if the entire network must be re-optimized. In this report, a method for analyzing the network so as to reduce this expense and to save time is presented. An approximate solution to the assignment problem is obtained. This approximation can be used itself to generate the new assignment, or it can be used to predict the portion of the network that will be significantly affected by the incident or parameter change. The nonlinear optimization algorithm could then be used to find the new assignment within the affected region.

The method of approach is based upon perturbation analysis; it is performed by solving an optimization problem with the same constraints, but a simpler cost function than the original problem. This simpler cost function (which is quadratic) is approximately the same as the true cost function when the change is small. Because it is quadratic, the suboptimal reassignment problem is treated analytically.

By means of dynamic programming and other techniques well adapted to the problem structure, the perturbation analysis of the network can be reduced to a problem involving the multiplication of a small number of matrices and vectors. These numerical operations are very fast, and can easily be implemented on-line using minicomputers.

To illustrate the use of perturbation analysis, consider the network in Figure 1.1.2. This network consists of a freeway (top roadway), and alternate route such as an arterial road (bottom), and several entrance ramps. (This is simplified for the purpose of description.) The dashed lines show how the system is divided into subnetworks. The flows in the freeway are labeled $x_1(k)$, $k = \ldots -1,0,1,\ldots$; the flows in the arterial are labeled $x_2(k)$, $k = \ldots -1,0,1,\ldots$; and the flows in the entrance ramps are $\phi(k)$, $k = \ldots,-1,0,1,\ldots$.

Assume that before an incident, the optimal flows are given in Table 1.1.1. These flows could have been produced by the last run of a nonlinear programming algorithm. Notice that we have assumed that the arriving flow to the alternate route is too large, and hence that most of it is shifted to the freeway through the entrance ramps.

Assume that an incident takes place, and that the capacity of the freeway at subnetwork 0 has been reduced. Suppose that it is now optimal to limit the freeway flow to 4500 veh/hr at that point, and to carry 2000 veh/hr on the arterial. This requires reoptimization, and we may assume that the new optimal distribution is given in Table 1.1.2. The differences between Table 1.1.1 and Table 1.1.2 are given in Table 1.1.3.

The theory or perturbational analysis tells us that Table 1.1.3 (and thus Table 1.1.2) could have been found approximately in a simpler, cheaper way than by nonlinear programming. Instead, a set of 2x2 matrices, D, is calculated in advance. For this example, assume all the D matrices are given by

$$D(k) = \begin{pmatrix} .90 & .19 \\ .10 & .81 \end{pmatrix}$$

Then for $k = 0,1,2,\ldots$, we have

$$\delta x(k+1) = D(k) \quad \delta x(k)$$

and for $k = 0,-1,-2,\ldots$, we have

$$\delta x(k-1) = D(k) \quad \delta x(k)$$

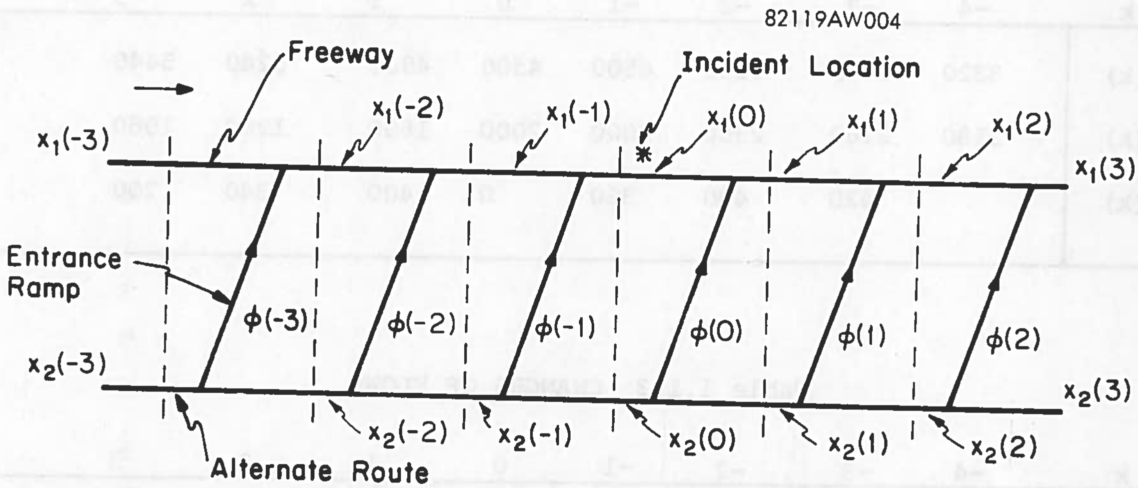

Figure 1.1.2 APPLICATION TO REROUTING TRAFFIC AROUND AN INCIDENT

Table 1.1.1 FLOWS BEFORE INCIDENT

k	-4	-3	-2	-1	0	1	2	3
$x_1(k)$	3500	4000	4500	5000	5200	5400	5600	5700
$x_2(k)$	3000	2500	2000	1500	1300	1100	900	800
$\phi(k)$		500	500	500	200	200	200	100

Table 1.1.2 FLOWS AFTER INCIDENT

k	-4	-3	-2	-1	0	1	2	3
$x_1(k)$	3320	3740	4140	4500	4500	4900	5240	5440
$x_2(k)$	3180	2760	2360	2000	2000	1600	1260	1060
$\phi(k)$		420	400	360	0	400	340	200

Table 1.1.3 CHANGES OF FLOWS

k	-4	-3	-2	-1	0	1	2	3
$\delta x_1(k)$	-180	-260	-360	-500	-700	-500	-360	-260
$\delta x_2(k)$	+180	+260	+360	+500	+700	+500	+360	+260
$\delta\phi(k)$		-80	-100	-140	-200	+200	+140	+100

where $\delta x(k)$ is given in Table 1.1.3. Both of the recurrence relations are initialized by $\delta x(0)^T = (-700, 700)$.

Before the incident, the entrance ramps are used to shift traffic from the alternate route to the freeway because the alternate is crowded and the freeway is empty. The incident limits this by lowering the capacity of the freeway at the incident site (k=0). Only entrance ramps are available, so that the only control is ramp metering. Thus, the only way to manage the traffic is to require less flow on the ramps upstream and more flow downstream. This is why $\delta\phi(k)$ is negative, for $k \leq 0$, and positive for $k > 0$.

Also, it is in the system's interest to get as much traffic on the freeway as soon as possible, as long as there is less than 4500 veh/hr at $k = 0$. As a result, most of the change from Table 1.1.1 to Table 1.1.3 must take place in the neighborhood of $k = 0$.

This example illustrates several features we will explore in greater depth in the body of this report.

a. Once the matrices $D(k)$ are calculated, simple numerical operations are required to implement ramp metering.

b. Both upstream and downstream behavior are calculated in the same way. In general, the matrices $D(k)$ vary with k, and are given by different values upstream and downstream.

c. The perturbations ($\delta x(k)$ and $\delta\phi(k)$) from the optimal flow diminish as $|k|$ increases, i.e. as we move from the incident site in both the upstream and downstream directions.

One reason for studying the effects of perturbations on traffic assignments is thus to save computer expense and time. Another is to verify the meaningfulness of the static formulation. Because traffic conditions change with time, the traffic situation may vary substantially between two successive static optimizations. If the freeway corridor network is long enough, a car may enter at one end just as an optimization is performed, and not leave from the other until after the next optimization is calculated.

If, in the meantime, the traffic conditions at the entrance nodes have changed, the optimal static assignment will no longer be the same. Therefore, a static optimization is sensible only as far as the static optimal traffic assignments are insensitive to changes in incoming traffic.

Indeed, let us suppose that a small change in the incoming traffic leads to important changes in traffic assignment downstream. Then the optimization routine will perform those changes accordingly. However, consider a car which enters the network at the time of the previous optimization and still is within the network although very far from the entrances. Such a car is not affected by the traffic which appears at the entrances at the same instant. Therefore, the previous optimal assignment is still optimal for it, but we have hypothesized that the optimization routine will change the assignment drastically. Thus, applying static optimization to a problem whose optimal solution is highly sensitive to the initial data (the entering traffic) leads to absurd policies. We show, in the case considered in this report, that the static assignment is insensitive to the arriving traffic distribution. That is, if a small change takes place at a point in the network, its effect diminishes with distance from that specific point.

To recapitulate: we have discussed the main reasons to conduct a sensitivity analysis of optimal static assignment within a freeway-corridor network system. They are:

a. to achieve great economies in computer effort by substituting a local optimization for a global one when an abrupt event drives the system suddenly away from optimality, and

b. to establish the validity of static optimization.

1.1.2 Assumptions

Several assumptions have been made to formulate the problem dealing with traffic perturbation analysis. This has enabled us to make stronger quantitative statements than otherwise. In the following, we discuss the meaning and effect that the assumptions have. It is important to observe that the resulting mathematical formulation remains physically meaningful.

First, we consider a freeway-corridor network in which all flow is in one direction (i.e., toward the city or toward the suburbs), and in which no entering or exiting takes place except at the ends of the network. This simplifies the analysis and notation by eliminating the need for distinguishing between different origins and destinations. In this case, we derive formulas for the perturbations arising from a specified change in flow at a specified location in the network.

Second, we assume that the traffic perturbations are small. This allows us to approximate the travel time cost function (which can be quite general and is often represented as a fourth order polynomial) by a quadratic form. Any cost functions which have continuous second partial derivatives may be approximated in this way if flow perturbations are small. An example appears in section 5.

Actual perturbations may not be small. Therefore, our prediction of the rate at which perturbations diminish may be in error, but not our assertion that they do tend to vanish. This is because we approximate the cost function, but not the conservation of flow constraints. In practice, this may call for a conservative use of the results we present here. If by these methods it is calculated that a two mile section of the network should be re-optimized, it may be prudent to re-optimize a four mile section. In any case, the entire network need not be re-optimized.

Third, we do not explicitly include capacity constraints in our problem foundation. However, this is not a serious restriction because we are minimizing travel time. Travel time functions which are derived from queueing theory have the following property. As flow approaches a certain value, travel time increases sharply. This "penalty" effect tends to act as a capacity constraint.

Links which are heavily loaded, i.e., in which this effect is important, have cost functions with very large first and second derivatives. This is important because these derivatives enter the quadratic approximation discussed above. Consequently, capacity constraints are implicitly taken into account here.

Fourth, we analyze in great detail the special case in which the perturbation does not change the total flow in the network. This corresponds to a traffic incident where some drivers avoid the roadway on which the accident has taken place. The total traffic passing any point along the network is unchanged, although the distribution among the different parallel roads and ramps is altered.

The analysis of this case simplifies the more general case in which total flow is changed, as a result of a surge in demand at a specific point. Instead of diminishing with distance, the resulting perturbations approach constant traffic distributions far from the site where the initial perturbation is imposed.

Fifth, we study separately downstream and upstream perturbations. In the downstream case, a perturbation is imposed at an intermediate point in the network and we calculate its effect downstream, until it reaches the exit. In the upstream case, we calculate its effect upstream until the entrances are reached. In both cases, the perturbation is specified at the intermediate point, but in the upstream case, the flow is specified at the entrances as well. Thus, the downstream case reduces to an initial value problem while the upstream case is a two-point boundary value problem. However, if the network is long enough, and the imposed perturbations do not change the total flow, the entrance boundary condition in the upstream case is not important. As a result, upstream perturbations are calculated as if they were downstream perturbations in a network whose direction of travel is reversed. For this reason, we first study the downstream problem and then the upstream problem.

Sixth, we are able to obtain our strongest results in the stationary case, where there is a great deal of spatial homogeneity. This assumption allows us to calculate, under certain assumptions, the rate at which perturbations diminish. (In a recent work [42], we have proved that perturbations diminish even in the non-stationary case. However, bounds on the rate of convergence remain unavailable.)

Further, we have found a graphical test for stationary networks which indicates whether the network has a certain property which is necessary in order for perturbations to decrease.

In real traffic systems, perturbations are often not confined to a small region. Before reaching steady state, they may grow to cover most of the corridor. In the context of this report, we would conclude that the network does not have the capacity to carry the traffic, either because the total capacity of freeways and arterials is insufficient, or because the access between the roads is unsatisfactory. At any rate, the steady-state assumptions are violated, and static optimization techniques, including the perturbation analysis presented in this report, are not applicable.

1.2 CONTRIBUTION AND MAIN RESULTS OF THIS REPORT

To conduct the sensitivity perturbation analysis, we have to use a great deal of control theory. It seems that some theoretical results which we have derived not only are extremely useful for our original traffic-oriented problem, but may also be of a wider theoretical interest. Therefore, we have divided this section into two parts: in 1.2.1, we summarize our contribution to the traffic problem presented in 1.1, and in 1.2.2, we survey our contributions to the theory of optimal control.

1.2.1 Contribution to Traffic Engineering

We show that a general freeway-corridor network system can be studied by splitting it into subnetworks, and also that is reasonable to approximate the average travel-time cost function of [2] to the second-order perturbation term about an optimal solution. This procedure allows one to reduce the perturbation-sensitivity analysis to a quadratic optimization problem with linear constraints which model flow conservation. Moreover, the quadratic cost function can be split into a sum of terms, each of which corresponds to one subnetwork. The minimization of that quadratic cost function subject to the linear constraints is performed by discrete dynamic programming, using the label of the subnetwork as the stage parameter. One can thus determine the optimal downstream or upstream perturbations from the dynamic programming equations. This is much more economical than recomputing the new optimal assignments, through the use of a nonlinear programming algorithm.

By using dynamic programming, we show numerically that a change in incoming flow distribution which does not affect the total incoming flow gives rise to perturbations whose magnitudes decrease very rapidly away from the point where the traffic perturbation has happened. Such a perturbation can result from a traffic incident. This method can be applied to any succession of subnetworks, all totally different: both the number of entrances and the number of links may vary from one to the next.

We then further consider the case when all subnetworks are identical; i.e., consist of the same pattern of entrances and exits connected by links, and contribute identical terms in the cost function. In this special case, we have been able to prove rigorously the following statements about the downstream and upstream perturbations, under some technical assumptions.

a. If the network is long enough, the downstream flow perturbations (caused by a change in incoming traffic which does not affect the total flow) decrease in magnitude as an exponential function of the distance. We derive a formula that can be used to obtain bounds for those flow perturbation magnitudes as a function of the distance. This enables one to determine exactly the neighborhood of that incident point over which the traffic assignment must be re-optimized. Obtaining these bounds involves computing only the eigenvalues and eigenvectors of a small matrix. As an illustration, the flow-perturbation magnitudes will typically be decreased by a factor of 10^{-10} after 30 stages in a network with 2 parallel roadways.

b. An initial perturbation which alters the total flow converges exponentially fast to a limiting distribution constant. This limiting distribution is also easily calculated.

c. The same results hold for upstream flow perturbations, provided that the initial perturbation occurs far enough from the entrance to the network. To obtain the bounds for the upstream perturbations, the same procedure is followed as for the downstream perturbations except for interchanging two matrices in all calculations. It amounts to solving the downstream perturbations problem for the same network but where the direction of traffic is reversed.

d. The optimal value of the quadratic perturbation cost is a quadratic function of the incoming traffic perturbation. It increases with the number of subnetworks in such a way that the cost of adding one subnetwork goes to an easily calculated constant.

e. The rate at which perturbations diminish depends on two things: the topological structure of the subnetworks and the quadratic part of the cost function. In general, when the structure is altered by adding more links, perturbations diminish faster. In fact, if insufficient links exist, perturbations may never diminish. Similarly, when links that carry flow away from sites of congestion are made cheaper, perturbations decrease faster. We can summarize this in the following rule of thumb: The greater the access between parallel roadways (lateral access), the faster perturbations tend to vanish.

Remark 1. When we say "far enough from the entrances or from the exits," we mean "far enough for some converging quantities to have practically reached their limit" and, according to our numerical experience, 10 or even 5 subnetworks are quite enough.

Remark 2. All the above results rest mainly on the property that a certain dynamical system is controllable. To help the reader check this property, we have formulated a geometric as well as an algebraic criterion based on the topological structure of the subnetwork. Roughly speaking, the required property holds if there are enough links connecting the entrances with the exits of a typical subnetwork.

1.2.2 Contribution to Optimal Control Theory

We have shown that the equations describing the conservation of flow in a network can be transformed into a discrete dynamical system in standard form by defining new states and controls. This we call a reduced system. We have characterized the controllability of such a system directly from the network topological structure by associating a directed graph to it in a natural manner. This graph is of the same type as those which are used to describe the possible transitions of a Markov chain. The controllability of the reduced systems can be related to properties concerning the final classes of any Markov chain whose possible transitions are described by that graph. In a wide class of networks, the uniqueness of the final class is equivalent to controllability of the reduced systems. In all of them, the aperiodicity of a unique final class guarantess controllability. Thus, we have related a notion of importance in modern control theory; namely, controllability, with an important criterion in the classification of Markov chains, in the case of systems arising from flow conservation in networks. Markov chain theory appears naturally in this context although our problem is strictly a deterministic one.

We have also characterized the spectrum of the propagation matrix over infinitely many stages. That is, we have studied the limit of the matrix derivative of the optimal state at one stage with respect to the state at the previous stage when the number of stages goes to infinity. When a reduced system is controllable (and under two auxiliary assumptions, probably not necessary), this matrix has the number 1 as simple eigenvalue and its other eigenvalues have a magnitude less than 1. These properties are exactly those enjoyed by the transition probability matrix of a Markov chain which has only one final class, when that class is aperiodic. We have thus been led to conjecture that the propagation matrix over infinitely many steps is a stochastic matrix. This fact we have observed numerically for the propagation matrices in any number of steps, but we have not proven rigorously.

In the same spirit, we characterize both the asymptotic marginal cost per unit of flow and the fixed point of the infinite-step propagation matrix by a minimal property. We next conjecture a simple relation giving the propagation matrix over infinitely many steps as an analytical function of the one-step propagation matrix. We have observed those conjectures to hold in the numerical examples.

1.3 LITERATURE SURVEY

This report is a part of a research project whose goal is the efficient use of road networks. The overall approach as well as the coordination of all parts within the whole is explained in [1].

The global static optimization problem is described in [2]. It uses as data: traffic volumes, roadway conditions, and possibly occurrences of accidents. These parameters are estimated by an extended Kalman filter [4] using information from roadway sensors [5]. A dynamic control algorithm [3] compares the estimations of the actual traffic variables with the desired values given by the static algorithm and is intended to diminish the deviations. The present work is a sensitivity analysis of the static optimization problem.

The problem has been formulated from the point of view of "system-optimization" rather than "user-optimization." This is the terminology used by Payne and Thompson [8] and Dafermos and Sparrow [9], to distinguish between the two principles proposed by Wardrop in 1952[6]. Our static optimization algorithm makes the average travel time within the network a minimum. Therefore, some vehicles will possibly spend more time than others. Both principles are discussed in references [6], [7], [8], [9], [10], [11], and [12].

The behavioral model consists of three main assumptions. First, it assumes the existence of a precise deterministic relation (the fundamental diagram of traffic) between flow and density on all freeway sections (Fig. 1.3.1). The flow is the number of vehicles per hour, and the density, the number of vehicles per mile. This relation is discussed in [15], and has been demonstrated experimentally ([16], [17]).

Second, the model of signalized intersections does not take cycle time into account, as is done in [14], and [18], or synchronization-offset effects [14]. However, it shares the most important features with those more refined models.

Third, an M/M/1 queuing model has been adopted at freeway entrance ramps and signalized intersections; more refined models are found in the literature [18] and [19], but not in the context of freeway networks.

As early as 1956, Bellman [29] has pointed out the applicability of dynamic programming to transportation problems. General concepts and applications of modern optimal control, in particular dynamic programming, can be found in [26] and [27]. Our cost function is quadratic, and the dynamical system, expressing the conservation of flow, is linear. A transformation of the state and control variables casts our problem into a format near to the classical linear-quadratic problem [25] and [28]. Under the assumption that all subnetworks are identical, the system becomes stationary. We can then apply the results on the asymptotic behavior of the optimal cost for the linear-quadratic problem as given by Dorato and Levis [21], and also a direct analytical expression for the solution of an algebraic Riccati equation derived by Vaughan [22]. This, combined with the state-costate approach [23], gives a basic relation which is the crux of our proof that perturbations tend to

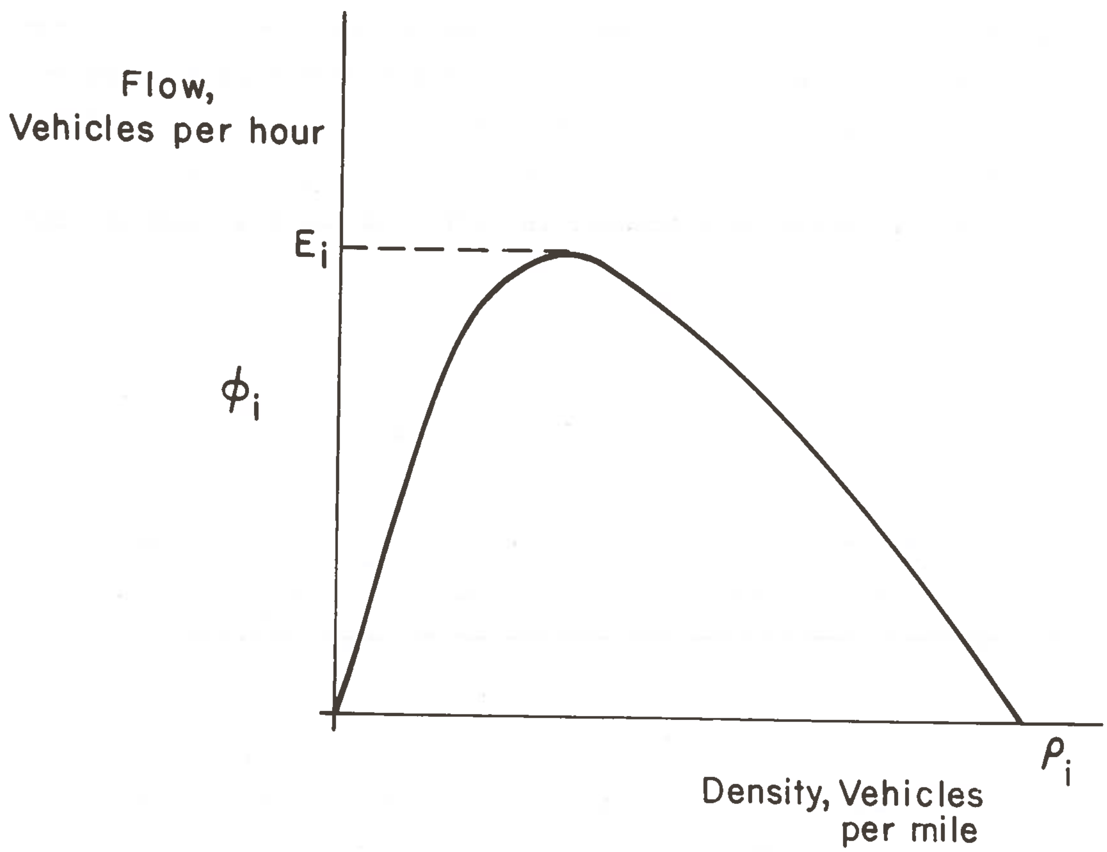

Figure 1.3.1 FUNDAMENTAL DIAGRAM OF TRAFFIC

diminish away from the incident location. In this context, the spectral theorems on powers of matrices [31] prove very useful.

Those asymptotic results rely mainly on controllability properties. To make the meaning of that notion precise in the context of networks, we have used graph theory [35], [36], and [37]; in particular, the intrinsic classification of directed graphs associated with Markov chains [32] and [33].

Thus, the tools we have been using come, on the one hand, from modern control theory: in particular, dynamic programming and the linear-quadratic problem; and on the other hand, from graph and Markov chain theory. This is sensible because we are doing dynamic programming, with quadratic cost, and over directed networks.

1.4 OUTLINE OF THE REPORT

Section 2 is the cornerstone of this work, and contains all the results about downstream perturbations. In Section 2.2, our approach is related to the static network optimization problem. We introduce the notation, describe the constraints, and set up the cost function. Next, we solve the quadratic minimization problem by dynamic programming, and derive the chain of equations which leads to the calculation of optimal downstream flow perturbations.

Section 2.3 is devoted to establishing the main results about downstream perturbations in a succession of identical subnetworks, which we call a <u>stationary network</u>. However, some constructions performed there apply to more general types of networks as well. Also, the restriction is not as stringent as it may appear at first sight, since it applies to networks such as in Fig. 2.3.9-3. We show that, by a suitable transformation, one can replace our conservation of flow constraints by a linear dynamic control system. We give conditions ensuring the controllability of that new system because we later show that controllability is necessary for perturbations to diminish.

Next, we extend slightly some classical results of linear-quadratic optimal control theory to the case when the cost function is quadratic, but not homogeneously quadratic. In that manner, we can apply this extended linear-quadratic theory to our problem, and prove the numerically observed properties, and derive a crucial relation which yields bounds on the perturbations. To do so, we have to make another transformation of the system, and to combine the state-costate approach with an explicit solution of the algebraic Riccati equation; also, to use a standard theorem on powers of matrices.

In Section 3, we are concerned with upstream perturbations. We show that the same steps as in Section 2 can be repeated, and indicate the adjustments to be made so as to derive the corresponding results.

In Section 4, we characterize steady-state constants of the network which emerged in Sections 2 and 3 by a variational property.

Section 5 methodically describes how a general network can be divided into subnetworks, and how the cost function of [2] can be expanded to the second order so as to make dynamic-programming applicable. The method is used on an example which is presented in [2], and values for the perturbations are compared with those obtained in [2] when dealing with the complete cost function.

Section 6 is devoted to certain conjectures. We motivate those conjectures, formulate them, and present numerical as well as intuitive evidence in their favor.

In Section 7, we list and comment briefly on a representative variety of examples numerically studied and show the corresponding computer outputs. On that material, the proved results as well as the conjectures can be checked.

Section 8 presents our conclusions.

Several technical theorems or proofs are grouped in appendices. In appendix A, graph-theoretical terminology and results are collected. (We have done so to make it self-contained, and because we needed an intrinsic graph formulation of a classification often found in the literature of Markov chains.)

In appendix B, the system transformation described in Section 2.3 is more carefully examined for a special category of subnetworks. Also, the controllability condition studied in Section 2.3, as well as the result from linear-quadratic optimal control theory, are extended to those subnetworks.

Appendix C is a short theorem on convergence of recursive linear difference equations, used in Section 2.3.

Appendix D is a proof of a theorem on convergence of the powers of some matrices, usually found in the literature for stochastic matrices, but valid for a more general class and used here.

In appendix E, a theorem is proved on the equivalence between the general linear-quadratic optimal control problem and a class of such problems with "diagonal" cost function. This theorem is stated and used in Section 2.3.

Appendix F is a presentation and very brief discussion of the computer program we have used.

2. DOWNSTREAM PERTURBATIONS IN A FREEWAY CORRIDOR NETWORK WITH QUADRATIC COST

2.1 INTRODUCTION

In this section, we establish all the results concerning downstream perturbations caused in a freeway network by a small change in incoming traffic at the entrance. In section 2.2, we contrast the present work with the static optimization approach presented in [2] and introduce notation. Next, we discuss the structure of a general network. It is split up into subnetworks; no assumptions are made as to the individual structures of the various subnetworks. We then set up the linear constraints which arise from flow conservation and the quadratic cost which results from a Taylor series expansion of the cost function used in [2]. Finally, we solve the quadratic minimization problem by dynamic programming, and we derive and summarize the key equations which yield the optimal solution and can be implemented on a digital computer.

In section 2.3, we prove rigorously the nature of the behavior apparent from the numerical implementation of the equations of section 2.2. We do so for the case when all the subnetworks have the same structure (i.e., topology), and contribute identical terms in the cost function. Even in that special case, several technical obstacles have to be resolved, using a great deal of linear-quadratic optimal control theory, graph theory, and matrix analysis. The central result is the calculation of bounds for the downstream perturbations as functions of the distance away from the entrance, counted in number of subnetworks.

2.2 PROBLEM STATEMENT AND SOLUTION BY DYNAMIC PROGRAMMING

2.2.1 Summary of the Static Optimization Problem

We briefly review the way the static optimization problem has been formulated and solved in [2], and the conjectures to which it has led.

The static traffic assignment is intended to minimize the total vehicle-hours expended on the network per hour; namely,

$$J = \sum_{\text{all } i} \phi_i \tau_i , \tag{2.1}$$

where ϕ_i is the flow along link i; i.e., the number of vehicles per hour flowing through link i, and τ_i the average time a vehicle spends on link i. The cost J given by Eq. (2.1) is therefore expressed in units of vehicles, or vehicle-hours per hour (veh-hr/hr).

It is shown in [2] that minimizing J is equivalent to minimizing the average time which vehicles spend on the network. The average time τ_i is composed of two parts: the time to traverse at the average velocity, and the time spent waiting in queues, present only if i is an entrance ramp or leads to a traffic light.

Using the "fundamental diagram" [15] which relates the flow ϕ_i along link i and the density ρ_i (the number of vehicles per mile along link i), and assuming an M/M/1 queuing model, one can show that the cost J is given by

$$J = \sum_{\text{all } i} \ell_i \rho_i(\phi_i) + \sum_{i \notin A} \frac{\phi_1^2}{E_i(E_i - \phi_i)} \tag{2.2}$$

where ℓ_i is the length of link i, in miles; A is the class of freeway links (as opposed to entrance ramps or signalized arterials); E_i is the effective capacity of link i (see [2]); $\rho_i(\ell_i)$ is the inverse relation of the fundamental diagram obtained by taking $\phi_i \leq E_i$.

The flows must be non-negative and must not exceed the effective capacities; i.e.,

$$0 \leq \phi_i \leq E_i . \tag{2.3}$$

The expression for E_i differs according to whether i is an entrance ramp or a signalized arterial.

If i is an entrance ramp,

$$E_i = \phi_{i\,max}(1 - \phi_j/\phi_{j\,max}) , \tag{2.4}$$

where $\phi_{i\,max}$ and $\phi_{j\,max}$ are the maximum capacities of links i and j, respectively;

and link j is the portion of freeway which ramp i impinges upon.

If i is a signalized arterial, then

$$E_i = g_i \phi_{i\ max}, \tag{2.5}$$

where g_i is the green split; i.e., the proportion of time the light is green for traffic coming along link i.

$$0 \leq g_i \leq \alpha, \tag{2.6}$$

for some number $\alpha \leq 1$ which is part of the data.

Problem Statement

The static problem is that of minimizing J, given by (2.2), by choosing both the flows ϕ_i and the green splits g_i, subject to the constraints (2.3) and (2.6) and to the flow conservation constraints:

$$\sum_{\text{incoming}} \phi_i = \sum_{\text{outgoing}} \phi_i . \tag{2.7}$$

There is one such constraint per node; the left-hand side sum in (2.7) is over all links whose flow enters the node, and the right-hand side sum is over all links whose flow leaves it.

Discussion

An accelerated gradient projection algorithm is used in [2] to solve that static traffic-assignment problem on various examples of networks, and to study the effect of incidents and congestions. In that minimizing problem, there is an exogeneous parameter: the flow entering the network. To see how the optimal traffic assignments ϕ_i^* and g_i^* vary as a result of a change in entering flow, one can solve the problem with each set of new data, using the gradient-projection algorithm. That is what has been done in [2]; and, based on that numerical experimentation, the author has conjectured the following. A modification in entering traffic which does not affect the total flow gives rise to perturbations in the optimal assignments which decrease in magnitude downstream from the entrances. However, reworking the whole problem each time the exogeneous parameter is changed (from one optimization time to the next, or because of an incident) is very costly and cannot lead to general statements. Therefore, it is of interest to derive a way to deal directly with perturbations; i.e., with changes $\delta\phi_i$ and δg_i in optimal assignments resulting from changes in the external parameter. This is what we do here.

2.2.2 Structure of the Network

Throughout this work, we shall avail ourselves of a simple, but important, property. A freeway corridor network can be split into a succession of subnetworks; the exits of subnetwork k are the entrances to subnetwork (k+1). This can always be achieved, possibly by adding fictitious links which do not contribute to the cost. The procedure is explained, and an example is treated, in section 5. This decomposition of the network is schematically illustrated in Fig. 2.2.2, with the notation which we introduce below.

The number of subnetworks the total network consists of is N-1. Each box in Fig. 2.2.2 represents a subnetwork. Subnetwork k consists of n_k entrances, n_{k+1} exits and m_k links connecting them. Nothing is assumed as yet as to m_k, n_k, n_{k+1}, for different values of k (except that they are positive integers). In other words, the number of entrances as well as of links may vary from one subnetwork to the next. The flows (in vehicles/hour) along the m_k links of subnetwork k constitute an m_k-dimensional vector that we denote by $\underline{\phi}(k)$. Its i^{th} component $\phi_i(k)$ is the flow along link i of subnetwork k.

The flows entering subnetwork k constitute an n_k-dimensional vector denoted by $\underline{x}(k)$. Its i^{th} component $x_i(k)$ is the flow (in vehicles per hour) entering subnetwork k through entrance i or, equivalently, leaving subnetwork (k-1) through exit i.

It is possible to express the cost function (2.2) as a sum of (N-1) terms, each of them depending only on the vector of flows along the links of a single subnetwork. Thus,

$$J(\underline{\phi}(1), \ldots, \underline{\phi}(N-1)) = \sum_{k=1}^{N-1} J_k(\underline{\phi}(k)). \tag{2.8}$$

A suitable partitioning into subnetworks leads to this decomposition of the cost function. In this analysis, we temporarily neglect the green splits: we show in section 5 how they can be included without modifying the basic framework.

Figure 2.2.2 SPLITTING OF A NETWORK INTO SUBNETWORKS

2.2.3 Linear Constraints

The constraints are of two types: (a) positivity and capacity constraints, and (b) flow-conservation constraints. We concentrate here on the flow-conservation constraints (2.7) because the former ones (2.3) and (2.6) will not play an important role in our sensitivity analysis (see section 2.2.4). Once the decomposition of the freeway corridor in subnetworks has been performed, the conservation of flow is expressed by the following linear relations among $\underline{x}(k)$, $\underline{\phi}(k)$, and $\underline{x}(k+1)$, for each k.

$$\underline{x}(k) = \underline{Y}(k)\ \underline{\phi}(k), \tag{2.9}$$

$$\underline{x}(k+1) = \underline{Z}(k)\ \underline{\phi}(k), \tag{2.10}$$

where $\underline{Y}(k)$ and $\underline{Z}(k)$, the entrance and exit node-link incidence matrices, respectively, are defined as follows.

Definition 2.1:

The matrix $\underline{Y}(k)$ has n_k rows and m_k columns. Its (i,j) entry is equal to 1 if, and only if, the j^{th} link originates from the i^{th} entrance; and is equal to 0 otherwise. Thus, each column of $\underline{Y}(k)$ has exactly one entry equal to 1, and all other entries equal to 0, since column j of $\underline{Y}(k)$ corresponds to link j, and link j originates from exactly one entrance.

Definition 2.2:

The matrix $\underline{Z}(k)$ has n_{k+1} rows and m_k columns. The (i,j) entry of $\underline{Z}(k)$ is equal to 1 if, and only if, link j leads to exit i. Each column of $\underline{Z}(k)$ has exactly one entry equal to 1, and all other entries equal to 0, since column j of $\underline{Z}(k)$ corresponds to link j, and link j leads to exactly one exit.

An example of a subnetwork is given in Fig. 2.2.3. The links are labeled, and the entrances and exits are labeled with numbers between parentheses. For the example of Fig. 2.2.3, $n_k = n_{k+1} = 3$, $m_k = 5$,

$$\underline{Y}(k) = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}; \quad \underline{Z}(k) = \begin{bmatrix} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

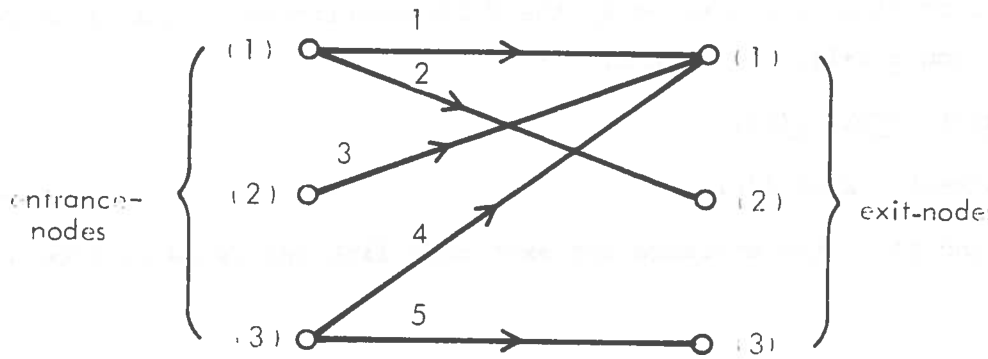

Figure 2.2.3 EXAMPLE OF SUBNETWORK

For the conservation of flow to hold, $\underline{Y}(k)$ and $\underline{Z}(k)$ have to be such that, for every k,

$$\sum_{i=1}^{n_k} x_i(k) = F, \tag{2.11}$$

where, by definition,

$$F \triangleq \sum_{i=1}^{n_1} x_i(1), \tag{2.12}$$

or

$$\underline{\nu}^T_{n_k} \underline{x}(k) = \underline{\nu}^T_{n_1} \underline{x}(1) \triangleq F, \tag{2.13}$$

for every k. Note that in Eq. (2.13), for any integer J, $\nu^T_J = (1,\ldots 1)$ denotes a row vector of dimension J with all components equal to 1. In fact, it follows from the previous remarks on the columns of the matrices $\underline{Y}(k)$ and $\underline{Z}(k)$ that

$$\underline{\nu}^T_{n_k} \underline{Y}(k) = \underline{\nu}^T_{n_{k+1}} \underline{Z}(k) = \underline{\nu}_{m_k}, \tag{2.14}$$

so that the constraints (2.13) are implied by (2.9) and (2.10).

2.2.4 Quadratic expansion of cost function

The problem treated in [2] is what we call here problem P.

Definition of Problem P

$$\text{minimize } J(\underline{\phi}) = \sum_{k=1}^{N-1} J_k(\underline{\phi}(k)), \tag{2.15}$$

subject to: $\underline{x}(k) = \underline{Y}(k)\,\underline{\phi}(k)$

$\underline{x}(k+1) = \underline{Z}(k)\,\underline{\phi}(k)$ $\quad k = 1,\ldots N-1,$ (2.16)

$\underline{\phi}(k) \geq 0$ (*),

and to the initial condition

(*) For a vector $\underline{v}$, $\underline{v} \geq 0$ means that each component of $\underline{v}$ is non negative.

$$\underline{x}(1) = \underline{\xi}. \tag{2.17}$$

Suppose the solution $\underline{\phi}^*(k)$, $k = 1,\ldots N-1$ has been found, applying the algorithm used in [2].

Definition of Problem P'

Now we are interested in problem P', exactly the same as problem P except that the condition (2.17) has been replaced by

$$x(1) = \underline{\xi} + \underline{\delta\xi}, \tag{2.18}$$

where $\underline{\delta\xi}$ is small, compared to $\underline{\xi}$.

Comparison of Problems P and P'

The nominal solution $\underline{\phi}^*$ satisfies all the constraints (2.3), (2.6), (2.7). Let us call $\hat{\underline{\phi}}$ the solution of problem P', and set

$$\delta\underline{\phi}^*(k) = \hat{\underline{\phi}}(k) - \underline{\phi}^*(k). \tag{2.19}$$

We assume that the initial perturbation $\underline{\delta\xi}$ is small enough, so that the positivity and capacity constraints are binding for $\hat{\underline{\phi}}$ if and only if they are for $\underline{\phi}^*$. Therefore, we delete those links on which these contraints are binding, and ignore those constraints on the remaining links. If $\phi_i^*(k) = 0$, we assume that in the perturbed solution, $\hat{\phi}_i(k) = 0$ will hold, so that $\delta\phi_i^*(k) = 0$. One way to impose that $\delta\phi_i(k) = 0$ is to delete link i of subnetwork k since we know that it will be effectively absent in the optimal solution. Deleting links amounts to deleting columns in the $\underline{Y}(k)$ and $\underline{Z}(k)$ matrices, and consequently reducing the dimension of $\underline{\phi}(k)$.

Since the flow conservation constraints are linear and are satisfied both by $\underline{\phi}^*$ and $\hat{\underline{\phi}}$, with initial condition $\underline{\xi}$ and $\underline{\xi} + \underline{\delta\xi}$ respectively, they are satisfied by $\underline{\delta\phi}^*$ with initial condition $\underline{\delta\xi}$.

$$\left.\begin{aligned} \delta\underline{x}(k) &= \underline{Y}(k)\ \delta\underline{\phi}(k) \\ \delta\underline{x}(k+1) &= \underline{Z}(k)\ \delta\underline{\phi}(k) \end{aligned}\right\} k = 1,\ldots N-1, \tag{2.20}$$

$$\delta\underline{x}(1) = \underline{\delta\xi}, \tag{2.21}$$

where $$\delta\underline{x}(k) = \hat{\underline{x}}(k) - \underline{x}^*(k), \tag{2.22}$$

and $\hat{\underline{x}}(k)$, $\underline{x}^*(k)$ are the optimal state-vectors in problems P' and P, respectively.

Now, let

$$\delta J(\delta\underline{\phi}) = J(\underline{\phi}^* + \delta\underline{\phi}) - J(\underline{\phi}^*), \tag{2.23}$$

for any $\delta\underline{\phi}^T = (\delta\underline{\phi}^T(1), \delta\underline{\phi}^T(2), \ldots, \delta\underline{\phi}^T(N-1))$.

Let us define:

$$J_1 = \min_{\underline{\phi}} J(\underline{\phi}), \quad \text{subject to: (2.16), and to initial condition: (2.17)}, \tag{2.24}$$

$$J_2 = \min_{\delta\underline{\phi}} J(\underline{\phi}^* + \delta\underline{\phi}), \quad \text{subject to: (2.20), and to initial condition: (2.21)}, \tag{2.25}$$

$$J_3 = \min_{\delta\underline{\phi}} \delta J(\delta\underline{\phi}), \quad \text{subject to: (2.20), and to initial condition: (2.21)}. \tag{2.26}$$

Then,

$$J_2 = J(\underline{\phi}^*) + J_3. \tag{2.27}$$

Therefore, problem P' is equivalent to:

$$\min \quad \delta J(\delta\underline{\phi}),$$

$$\text{subject to: } \left.\begin{aligned} \delta\underline{x}(k) &= \underline{Y}(k)\ \delta\underline{\phi}(k) \\ \delta\underline{x}(k+1) &= \underline{Z}(k)\ \delta\underline{\phi}(k) \end{aligned}\right\} k = 1,\ldots N-1, \tag{2.28}$$

and to the initial condition: $\delta\underline{x}(1) = \underline{\delta\xi}$.

Now we use a second order Taylor expansion for $\delta J(\delta\underline{\phi})$.

$$\delta J(\delta\underline{\phi}) = J(\underline{\phi}^* + \delta\underline{\phi}) - J(\underline{\phi}^*) = \sum_{k=1}^{N-1} \{J_k(\underline{\phi}^*(k) + \delta\underline{\phi}(k)) - J_k(\underline{\phi}^*(k))\}$$

$$\simeq \sum_{k=1}^{N-1} \{\tfrac{1}{2}\ \delta\underline{\phi}^T(k)\ \underline{L}(k)\ \delta\underline{\phi}(k) + \underline{h}^T(k)\ \delta\underline{\phi}(k)\}, \tag{2.29}$$

where

$$\left.\begin{aligned} \underline{L}(k) &= \frac{\partial^2 J_k}{\partial \phi(k)^2} (\underline{\phi}^*(k)) \\ \underline{h}(k) &= \frac{\partial J_k}{\partial \phi(k)} (\phi^*(\underline{k})) \end{aligned}\right\}, \tag{2.30}$$

that is, $\underline{L}(k)$ is the hessian matrix of J_k, and $\underline{h}(k)$ is the gradient of J_k, both evaluated at the optimal solution $\underline{\phi}^*(k)$ of the unperturbed problem P.

2.2.5 Solution by Dynamic Programming

In the previous section, we have shown that the sensitivity analysis is reduced to the following quadratic problem:

$$\text{minimize} \quad \delta J(\delta\underline{\phi}) = \sum_{k=1}^{N-1} \left(\frac{1}{2}\, \delta\underline{\phi}^T(k)\ \underline{L}(k)\ \delta\underline{\phi}(k) + \underline{h}^T(k)\ \delta\underline{\phi}(k)\right)$$

$$\equiv \sum_{k=1}^{N-1} \delta J_k(\delta\underline{\phi}(k)),$$

$$\text{subject to:} \quad \left.\begin{aligned} \delta\underline{x}(k) &= \underline{Y}(k)\ \delta\underline{\phi}(k) \\ \delta\underline{x}(k+1) &= \underline{Z}(k)\ \delta\underline{\phi}(k) \end{aligned}\right\} \quad k = 1, \ldots N-1,$$

$$\text{and to the initial condition:} \quad \delta\underline{x}(1) = \underline{\delta\xi},$$

where $\underline{\delta\xi}$ is an exogeneous parameter.

Notation. Now this is merely a mathematical problem, and the rest of this section will be devoted to its analysis. It is irrelevant how we denote the symbols, provided that we retain their precise meaning. Therefore, we shall write $\underline{\phi}(k)$, $\underline{x}(k)$, J_k, $\underline{\xi}$ instead of $\underline{\delta\phi}(k)$, $\underline{\delta x}(k)$, δJ_k, $\underline{\delta\xi}$ for notational convenience. It has to be borne in mind, however, that we deal with perturbations as it is precisely stated in 2.2.4. Let us rewrite the problem in that notation.

Mathematical Problem Statement

Problem P_1:

$$\min_{\underline{\phi}} J(\underline{\phi}) = \sum_{k=1}^{N-1} \left(\frac{1}{2} \underline{\phi}^T(k)\ \underline{L}(k)\ \underline{\phi}(k) + \underline{h}^T(k)\ \underline{\phi}(k)\right) \equiv \sum_{k=1}^{N-1} J_k(\underline{\phi}(k)), \tag{2.31}$$

$$\text{subject to: } \left.\begin{aligned} \underline{x}(k) &= \underline{Y}(k)\ \underline{\phi}(k) \\ \underline{x}(k+1) &= \underline{Z}(k)\ \underline{\phi}(k) \end{aligned}\right\} k = 1, \ldots N-1, \tag{2.32}$$

$$\text{and to the initial condition: } \underline{x}(1) = \underline{\xi}. \tag{2.33}$$

To solve this problem by dynamic programming, we define the partial optimization problems and the value functions V_k for $k = 1, \ldots N$, by:

$$V_k(\underline{\eta}) = \min_{\underline{\phi}} \sum_{i=k}^{N-1} J_i(\underline{\phi}(i)), \tag{2.34}$$

$$\text{subject to: } \left.\begin{aligned} \underline{x}(i) &= Y(i)\ \underline{\phi}(i) \\ \underline{x}(i+1) &= \underline{Z}(i)\ \underline{\phi}(i) \end{aligned}\right\} i = k, k+1, \ldots, N-1,$$

$$\text{and to the initial condition: } \underline{x}(k) = \underline{\eta}.$$

In particular,

$$V_N(\cdot) \equiv 0. \tag{2.35}$$

The minimum of the cost function (2.31) under the constraints (2.32) and (2.33) is equal to $V_1(\underline{\xi})$.

The sequence of value functions V_k is recursively determined backward from (2.35) by the Bellman functional equations:

$$V_k(\underline{x}(k)) = \min_{\underline{\phi}(k)} [J_k(\underline{\phi}(k)) + V_{k+1}(\underline{Z}(k)\ \underline{\phi}(k))], \tag{2.36}$$

$$\text{subject to: } \underline{Y}(k)\ \underline{\phi}(k) = \underline{x}(k). \tag{2.37}$$

Let us emphasize that, at stage k, the only constraint is (2.37) since $\underline{x}(k)$ is specified. The constraint $\underline{x}(k+1) = \underline{Z}(k)\ \underline{\phi}(k)$ in the minimization problem is taken into account by substituting $\underline{Z}(k)\ \underline{\phi}(k)$ for $\underline{x}(k+1)$ as the argument of

V_{k+1} in the right-hand side of (2.36). The results concerning the solution of our problem through the recursive determination of the sequence $V_k(\cdot)$ by solving Bellman's equations are summed up in Theorem 2.1.

Theorem 2.1

Assume that the matrices $\underline{L}(k)$ in the cost function (2.31) are all positive definite. Then,

a. The minimization problem (2.31) is well defined and has a unique minimizing point $\underline{\phi}^{*T} = (\underline{\phi}^{*T}(1), \underline{\phi}^{*T}(2), \ldots \underline{\phi}^{*T}(N-1))$. The minimum value is $V_1(\underline{\xi})$, where $\underline{\xi}$ is the exogeneous parameter in (2.33) and V_1 is defined by (2.34).

b. The sequence of value functions (2.34) is given by:

$$V_k(\underline{x}) = \frac{1}{2}\,\underline{x}^T\,\underline{C}(k)\,\underline{x} + \underline{b}(k)^T\,\underline{x} + a(k), \tag{2.38}$$

where

1) $\underline{C}(k)$ are positive definite $(n_k \times n_k)$ matrices (for $k < N$), and are recursively obtained from the relations

$$\underline{C}_N = 0,$$
$$\underline{C}^{-1}(k) = Y(k)M^{-1}(k)Y(k)^T \tag{2.39}$$

2) $\underline{b}(k)$ are $(n_k \times 1)$ vectors, recursively obtained from:

$$\underline{b}_N = 0,$$
$$\underline{b}(k) = \underline{C}(k)\,\underline{Y}(k)\,\underline{M}^{-1}(k)\,(\underline{h}(k) + \underline{Z}^T(k)\,\underline{b}(k+1)), \tag{2.40}$$

where

$$\underline{M}(k) \triangleq \underline{L}(k) + \underline{Z}^T(k)\,\underline{C}(k+1)\,\underline{Z}(k). \tag{2.41}$$

3) $a(k)$ are scalars, recursively obtained from:

$$a_N = 0,$$
$$a(k) = a(k+1) + \frac{1}{2}(\underline{h}(k) + \underline{Z}^T(k)\,\underline{b}(k+1))^T(\underline{M}^{-1}(k)\,\underline{Y}^T(k)\,\underline{C}(k)\cdot$$
$$\cdot\underline{Y}(k)\,\underline{M}^{-1}(k) - \underline{M}^{-1}(k))(\underline{h}(k) + \underline{Z}^T(k)\,\underline{b}(k+1)). \tag{2.42}$$

c. The unique minimizing vector is given by the feedback law:

$$\underline{\phi}^*(k) = \underline{M}^{-1}(k)\underline{Y}^T(k)\underline{C}(k)\underline{x}(k) + (\underline{M}^{-1}(k)\underline{Y}^T(k)\underline{C}(k)\underline{Y}(k)\underline{M}^{-1}(k)-\underline{M}^{-1}(k))\cdot$$
$$\cdot(\underline{h}(k) + \underline{Z}^T(k)\ \underline{b}(k+1)). \tag{2.43}$$

Proof. To determine the value functions defined by (2.33) we solve Bellman's functional equations:

$$V_k(\underline{\eta}) = \min_{\underline{\phi}} \ [J_k \ (\underline{\phi}) + V_{k+1}(\underline{Z}\ \underline{\phi})], \tag{2.44}$$

$$\text{subject to: } \underline{Y}(k)\ \underline{\phi} = \underline{\eta}.$$

The sequence of functions V_k given recursively by (2.44) from (2.35) is unique provided that, at each stage, the minimization problem is well defined. We shall now show that it is the case and that the value functions are quadratic:

$$V_k(\underline{x}) = \frac{1}{2}\underline{x}^T\ \underline{C}(k)\underline{x} + \underline{b}(k)^T\underline{x} + a(k), \tag{2.45}$$

with positive definite matrices $\underline{C}(k)$ $(k < N)$, and determine the recurrent relations satisfied by $\underline{C}(k)$, $\underline{b}(k)$, $a(k)$.

Using expression (2.31) for $J(k)$, and adjoining a Lagrange multiplier $\underline{\lambda}(k)$ of dimension n_k to the constraint (2.37), we obtain

$$V_k(\underline{x}(k)) = \min_{\underline{\phi}(k)}\left\{\frac{1}{2}\underline{\phi}^T(k)\ \underline{L}(k)\ \underline{\phi}(k) + \underline{h}^T(k)\underline{\phi}(k) + \underline{\lambda}^T(k)(x(k) - Y(k)\underline{\phi}(k))\right.$$
$$\left. + V_{k+1}(\underline{Z}(k)\underline{\phi}(k))\right\}, \tag{2.46}$$

$$\text{subject to: } \underline{Y}(k)\ \underline{\phi}(k) = \underline{x}(k),$$
$$\text{with initial condition: } V_N(\cdot) \equiv 0.$$

The initial condition shows that (2.45) is satisfied for $k = N$, with $\underline{C}(N) = 0$, $\underline{b}(N) = 0$, and $z(N) = 0$. It remains to prove (2.45) for general k by induction. Thus, assuming that (2.45) holds for V_{k+1}, with $\underline{C}(k+1)$ positive definite if $k+1 < N$ and $\underline{C}(k+1) = 0$ if $k + 1 = N$, we substitute (2.45) for V_{k+1} in (2.46) and obtain:

$$V_k(\underline{x}(k)) = \min_{\underline{\phi}(k)} \left\{\frac{1}{2}\ \underline{\phi}^T(k)L(k)\underline{\phi}(k) + \underline{h}^T(k)\ \underline{\phi}(k) + \underline{\lambda}^T(k)(\underline{x}(k)-Y(k)\underline{\phi}(k))\right.$$
$$\left. + \frac{1}{2}\underline{\phi}^T(k)\underline{Z}^T(k)\underline{C}(k+1)\underline{Z}(k)\underline{\phi}(k) + \underline{b}^T(k+1)\underline{Z}(k)\underline{\phi}(k) + a(k+1)\right\}, \tag{2.47}$$

subject to: $\underline{Y}(k)\underline{\phi}(k) = \underline{x}(k)$.

The minimization is performed by equating to zero the derivative of the right-hand side of (2.46) with respect to $\underline{\phi}(k)$, to obtain

$$\underline{L}(k)\underline{\phi}(k) + \underline{h}(k) - \underline{Y}^T(k)\underline{\lambda}(k) + \underline{Z}^T(k)\underline{b}(k+1) + \underline{Z}^T(k)\underline{C}(k+1)\underline{Z}(k)\underline{\phi}(k) = 0. \tag{2.48}$$

Let

$$\underline{M}(k) \overset{\Delta}{=} \underline{L}(k) + \underline{Z}^T(k)\underline{C}(k+1)\underline{Z}(k). \tag{2.49}$$

Since $\underline{L}(k)$ is positive definite by assumption and $\underline{Z}^T(k)\underline{C}(k+1)\underline{Z}(k)$ is positive semi-definite (because $\underline{C}(k+1)$ is, by induction hypothesis), $\underline{M}(k)$ is positive definite and therefore invertible.

Accordingly,

$$\underline{\phi}^*(k) = -\underline{M}^{-1}(k)(\underline{h}(k) + \underline{Z}^T(k)\underline{b}(k+1)) + \underline{M}^{-1}(k)\underline{Y}^T(k)\underline{\lambda}(k). \tag{2.50}$$

Now the constraint (2.37) implies that

$$\underline{x}^*(k) = -\underline{Y}(k)\underline{M}^{-1}(k)(\underline{h}(k) + \underline{Z}^T(k)\underline{b}(k+1) + \underline{Y}(k)\underline{M}^{-1}(k)\underline{Y}^T(k)\underline{\lambda}(k). \tag{2.51}$$

Equation (2.51) may be solved for $\underline{\lambda}(k)$ because $\underline{Y}(k)\underline{M}^{-1}(k)\underline{Y}^T(k)$ is invertible. Let us prove this statement.

Each column of $\underline{Y}(k)$ contains exactly one 1 and zeros for other entries. It is possible to find n_k different columns with the 1 on a different row in each because there are n_k different links originating each from a different entrance. Thus,

$$\sum_i x_i \; Y_{ij}(k) = 0 \text{ for } j = 1,\ldots m_k, \tag{2.52}$$

implies $\underline{x} = 0$; or, $\underline{x}^T\underline{Y}(k) = 0$ implies $\underline{x} = 0$.

On the other hand, $\underline{Y}(k)\underline{M}^{-1}(k)\underline{Y}^T(k)$ is positive semi-definite. Since $\underline{M}(k)$, and therefore $\underline{M}^{-1}(k)$, is positive definite. However, $\underline{x}^T\underline{Y}(k)\underline{M}^{-1}(k)\underline{Y}^T(k)\underline{x} = 0$ implies $\underline{x}^T\underline{Y}(k) = 0$, and therefore $\underline{x} = 0$ from (2.52), so that $\underline{Y}(k)\underline{M}^{-1}(k)\underline{Y}^T(k)$ is positive definite, hence invertible.

Define

$$\underline{N}_k \triangleq (\underline{Y}(k)\underline{M}^{-1}(k)\underline{Y}^T(k))^{-1}. \tag{2.53}$$

The matrix $\underline{N}_k$ is positive definite, so that equation (2.51) can be transformed into

$$\underline{\lambda}(k) = N(k)[\underline{x}(k) + \underline{Y}(k)\underline{M}^{-1}(k)(\underline{h}(k) + \underline{Z}^T(k)\underline{b}(k+1)], \tag{2.54}$$

and so that equation (2.50) becomes:

$$\underline{\phi}(k) = -\underline{M}^{-1}(k)(\underline{h}(k) + \underline{Z}^T(k)b(k+1)) + \underline{M}^{-1}(k)\underline{Y}^T(k)\underline{N}(k)[\underline{x}(k) + \underline{Y}(k)\underline{M}^{-1}(k)(\underline{h}(k) + \underline{Z}^T(k)\underline{b}(k+1))]. \tag{2.55}$$

The matrix of second derivatives of the right-hand side of (2.46) with respect to $\underline{\phi}(k)$ is equal to $\underline{L}(k) + \underline{Z}^T(k)\underline{C}(k+1)\underline{Z}(k) = \underline{M}(k)$ and is positive definite; therefore, $\underline{\phi}(k)$ given by equation (2.55) is the unique minimizing vector. Substituting its value from equation (2.55) into equation (2.47) gives the minimum value $V_k(\underline{x}(k))$, and one can show that it is quadratic in $\underline{x}(k)$, and obtain $\underline{C}(k)$, $\underline{b}(k)$, $z(k)$ in terms of $\underline{C}(k+1)$, $\underline{b}(k+1)$, $a(k+1)$. In particular,

$$\begin{aligned}\underline{C}(k) &= \underline{N}(k)\underline{Y}(k)\underline{M}^{-1}(k)\underline{L}(k)\underline{M}^{-1}(k)\underline{Y}^T(k)\underline{N}(k) \\ &\quad + \underline{N}(k)\underline{Y}(k)\underline{M}^{-1}(k)\underline{Z}^T(k)\underline{C}(k+1)\underline{Z}(k)\underline{M}^{-1}(k)\underline{Y}^T(k)\underline{N}(k) \\ &= \underline{N}(k)\underline{Y}(k)\underline{M}^{-1}(k)(\underline{L}(k) + \underline{Z}^T(k)\underline{C}(k+1)\underline{Z}(k))\underline{M}^{-1}(k)\underline{Y}^T(k)\underline{N}(k) \\ &= \underline{N}(k)\underline{Y}(k)\underline{M}^{-1}(k)\underline{Y}^T(k)\underline{N}(k) = \underline{N}(k)\underline{N}^{-1}(k)\underline{N}(k) = \underline{N}(k).\end{aligned}$$

Therefore, $\underline{C}(k)$ is positive definite since the induction hypothesis implies the positive definiteness of $\underline{N}(k)$. Equivalently, $\underline{C}^{-1}(k) = \underline{Y}(k)(\underline{L}(k) + \underline{Z}^T(k)\underline{C}(k+1)\underline{Z}(k))^{-1}\underline{Y}^T(k)$ which is equation (2.39). In like manner, equations (2.40) and (2.41) are easily derived. Equation (2.55) is equivalent to into (2.43). Q.E.D.

Remark: Calculation of perturbations

Having solved the quadratic minimization problem, we can now find the sequence of optimal states, $\{x^*(k)\}$, recurrently, from (2.43) and the constraint $\underline{x}(k+1) = \underline{Z}(k)\underline{\phi}(k)$. Thus,

$$\underline{x}^*(k+1) = \underline{Z}(k)\ \underline{\phi}^*(k),$$

or

$$\underline{x}^*(k+1) = \underline{Z}(k)\underline{M}^{-1}(k)\underline{Y}^T(k)\underline{C}(k)\underline{x}^*(k)$$
$$+ \underline{Z}(k)(\underline{M}^{-1}(k)\underline{Y}^T(k)\underline{C}(k)\underline{Y}(k)\underline{M}^{-1}(k)-\underline{M}^{-1}(k)(\underline{h}(k) + \underline{Z}^T(k)\underline{b}(k+1)). \quad (2.56)$$

Those results are valid for any quadratic minimization problem with a cost function given by Eq. (2.31) and constraints (2.32) and (2.33), provided that the assumptions of Theorem 2.1 are fulfilled. However, in our particular problem, where $\underline{x}(k)$ represents the perturbations at stage k caused by an initial perturbation $\underline{\xi}$, the coefficients $\underline{L}(k)$ and $\underline{h}(k)$ are not arbitrary; rather, they are given by equation (2.30).

If $\underline{\xi} = 0$, the downstream perturbations will certainly be equal to zero since the full problem has been optimized (Section 2.2.4) for $\underline{\xi} = 0$. Likewise, a zero perturbation at stage k induces a zero perturbation at stage (k+1). Therefore, our particular $\underline{L}(k)$ and $\underline{h}(k)$ are such that the constant coefficient in (2.56) vanishes, and equation (2.56) becomes:

$$\underline{x}^*(k+1) = \underline{Z}(k)\underline{M}^{-1}(k)\underline{Y}^T(k)\underline{C}(k)\underline{x}^*(k). \quad (2.57)$$

We sum up below the key equations which lead to the calculation of downstream perturbations.

<u>Theorem 2.2</u>

In the quadratic approximation to the static optimization problem stated in Sections 2.2(1-2-3), the downstream perturbations $\underline{x}^*(1)$, $\underline{x}^*(2)$,...$\underline{x}^*(k)$,... are given recurrently from the initial perturbation $\underline{\xi}$, by:

$$\underline{x}^*(1) = \underline{\xi}, \quad (2.58)$$

$$\underline{x}^*(k+1) = \underline{D}(k)\ \underline{x}^*(k), \quad k = 1,...N-1, \quad (2.59)$$

where

$$\underline{D}(k) = \underline{Z}(k)(\underline{L}(k) + \underline{Z}^T(k)\underline{C}(k+1)\underline{Z}(k))^{-1}\underline{Y}^T(k)\underline{C}(k), \quad (2.60)$$

and the $\underline{C}(k)$ are found recurrently (backward in k) from:

$$\underline{C}_N = 0, \tag{2.61}$$

$$\underline{C}^{-1}(k) = \underline{Y}(k)(\underline{L}(k) + \underline{Z}^T(k)\,\underline{C}(k+1)\,\underline{Z}(k))^{-1}\underline{Y}^T(k) \tag{2.62}$$

$$k = N-1,\ N-2,\ \ldots, 2, 1.$$

Proof. Theorem 2.2 is an immediate corollary of Theorem 2.1, and the remark following it.

Remarks . a. We call $\underline{D}(k)$ the propagation matrix at stage k. Its study is essential for our purposes which are to prove that, for an initial perturbation $\underline{\xi}$ (not too large in magnitude, as compared with the optimal initial steady-state flow allocation [2], so that the quadratic approximation in 2.2.4 is valid), not affecting the total flow, the sequence $|\delta x_i(k)|$ goes to zero as k goes to infinity, for every $i = 1, \ldots n_k$. Also, we want to find bounds for $\max_i |\delta x_i(k)|$ as functions of k. In Section 2.3, we shall concentrate on the case of a "stationary network"; i.e., a network whose subnetworks are identical. Some results for more general networks appear in [42].

b. The only assumption of Theorem 2.1 (and therefore of Theorem 2.2) is that the $\underline{L}(k)$ be positive definite matrices. Given the definition of $\underline{L}(k)$ by (2.30), this is equivalent to requiring that the cost function of Section (2.2.1) (equation 2.2) be locally strictly convex in the neighborhood of the optimal solution to the unperturbed problem. As explained in Section 2.2.4 (and, with further details, in Section 5), we delete the links corresponding to binding constraints for the quadratic sensitivity analysis. In the global optimization problem solved by a numerical algorithm, there are always some nonbinding constraints. Consequently, the solution point lies in the interior of some manifold of reduced dimension. Therefore, the nonlinear cost function (used in [2]) must be locally convex around this point in that manifold.

2.3 ASYMPTOTIC ANALYSIS IN THE STATIONARY CASE

2.3.1 Introduction

In this section, we concentrate on the case when all subnetworks are identical and contribute to identical costs. We call this the stationary case.

We summarize, in 2.3.2, the main features observed when applying the equations of 2.2 to calculate perturbations in a variety of stationary networks. We go through several steps to explain and prove those features and to derive quantitative bounds. In subsection 2.3.3, we perform a transformation to describe the flow-conservation constraints by a linear system in standard form, and express the cost function in the new variables. In 2.3.4, we investigate the controllability of the new system because this property determines the system's asymptotic behavior. We present a geometric and an algebraic criterion.

In subsection 2.3.5, we extend slightly the classical results on the linear-quadratic optimal control problem to the case when the cost function contains linear terms, and apply that extension to the study of the asymptotic behavior of the cost function. In subsection 2.3.6, we give an expression for the propagation matrices, and show in particular that the asymptotic propagation matrix can be found by solving an algebraic Riccati equation and a linear equation.

Subsection 2.3.7 is devoted to the sensitivity analysis itself. We show there that, if a reduced system is controllable and two minor technical conditions hold, then the magnitudes of zero-flow perturbations decrease geometrically with distance at a rate which depends on the eigenvalues of a certain matrix. We give a direct method to compute those eigenvalues. Also, we show how to find the stationary distribution to which perturbations with non-zero total flow converge exponentially. Finally, we discuss the importance of the assumptions made and illustrate the theory by applying it to a special class of examples in subsection 2.3.8. In subsection 2.3.9, we show how a class of networks, which are apparently not stationary, can be studied as stationary networks. We call them quasi-stationary networks. The entire analysis of sections 2.3.3 to 2.3.7 is applicable to them.

2.3.2 Numerical Experiments

The equations of section 2.2 have been applied to various examples of stationary networks; i.e., networks where m_k, n_k, $\underline{L}(k)$, $\underline{Y}(k)$, and $\underline{Z}(k)$ are the same for all k. The complete set of numerical examples is presented in section 7. In particular, we shall often refer to the subnetwork of Fig. 7.1

as the "standard two-dimensional example." This example appears in Bellman's 1956 work [29]. Also, several subnetworks with three entrances and three exits (therefore giving rise to a three-dimensional state vector $\underline{x}(k)$) have been studied with various structures and various costs.

We briefly state the main features which have been observed in all our numerical experience. The numerical examples are discussed in greater detail in section 7.

Main observed features

A perturbation $\underline{x}(1)$ which leaves the total flow unchanged gives rise to a sequence $\underline{x}(k)$ of perturbations which decrease very fast in magnitude downstream.

If $\sum_{j=1}^{n} x_j(1) = 0$, we observe: $|x_i(k)| \to 0$ for $k \to \infty$, for all i. Typically, for $k = 30$, $|x_i(30)| \sim 10^{-10}|x_i(1)|$. Assume that an external cause, like an accident, modifies the distribution of the incoming traffic in subnetwork 1 among the various entrances. If the traffic assignments are changed so as to remain optimal, the resulting change in the downstream distribution of traffic will become negligible not very far from the accident. Therefore, if such an external modification happens, it will be sufficient to recalculate the assignments in the immediate neighborhood of the accident. The main purpose of section (2.3) is to prove this in the general stationary case and, in addition, to derive a quantitative estimate of that neighborhood given a tolerance. That is, given $\varepsilon > 0$, determine k_0 such that $|x_i(k)| < \varepsilon$ for $i = 1,\ldots n_k$, for all $k \geq k_0$. This amounts to establishing a bound for $|x_i(k)|$ given $\underline{x}(1)$.

For the sequence $\{\underline{C}(k)\}$, the interesting asymptotic behavior occurs uptream; i.e., toward the entrance, when $(N-k) \to \infty$, since the sequence is specified downstream by $\underline{C}_N = 0$. Using the computer program (see section 7 and appendix F), it is observed that, as $(N-k) \to \infty$, $C_{ij}(k) \to \infty$. This simply means that the cost to travel an infinite distance is infinite.

More informative is the property that the matrix $\underline{\Delta C}(k)$ defined by

$$\underline{\Delta C}(k) \triangleq \underline{C}(k) - \underline{C}(k+1), \qquad (2.63)$$

goes to a constant as (N-k) → ∞. Moreover, this constant matrix has all its entries equal.

$$\underline{\Delta C}(k) \to \underline{\Delta C}.$$

$$\underline{\Delta C} = \alpha \, \underline{\nu}_n \underline{\nu}_n^T = \alpha \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \cdots & \cdots & \cdots & \cdots \\ \vdots & \vdots & & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix}, \qquad (2.64)$$

where α is a positive number.

(Recall that $\underline{\nu}_n \in R^n$; $\underline{\nu}_n^T = (1, \ldots 1)$; and $n_k = n$ for all k since the network is stationary.)

For instance, in the standard two-dimensional example, with

$$\underline{L} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}, \qquad (2.65)$$

α was observed equal to 1/3.

Thus, the limiting behavior of $\underline{C}(k)$ as (N-k) → ∞ is linear in k:

$$\underline{C}(k) \sim \overline{\underline{C}} + (N-k)\alpha \underline{\nu}_n \underline{\nu}_n^T, \qquad (2.66)$$

with $\overline{C}$ positive definite and finite.

<u>Qualification</u>: In the disconnected subnetwork of Fig. 7.7, where the top part is the standard two-dimensional example, what is observed is:

$$\underline{\Delta C}(k) \to \underline{\Delta C} = \begin{bmatrix} \alpha & \alpha & 0 \\ \alpha & \alpha & 0 \\ 0 & 0 & \beta \end{bmatrix},$$

where α is the same as in the standard example of Fig. 7.1, and $\beta = L_5$ is the entry of the cost matrix $\underline{L}$ corresponding to link 5. This is due to the impossibility, in this case, of describing the flow conservation by a controllable system because of the disconnectedness of the subnetwork (see 2.3.4).

The propagation matrices $\underline{D}(k)$ are observed to have positive entries with all entries of each column adding up to 1. By definition, $\underline{D}^T(k)$ <u>is observed to be a stochastic matrix</u>. Also, it has been found that the sequence $\{\underline{D}^T(k)\}$ converged to a constant stochastic matrix $\underline{D}^T$ as $(N-k) \to \infty$. For instance, in the standard two-dimensional example with $\underline{L}$ given by (2.65), one finds:

$$\underline{D} = \begin{bmatrix} 2-\sqrt{2} & \sqrt{2}-1 \\ \sqrt{2}-1 & 2-\sqrt{2} \end{bmatrix}.$$

2.3.3 System Transformation. Corresponding Expression for Cost Function.

Because we want to use results which are found in the literature (e.g.,[27] for dynamical systems in standard form, and ours is in implicit form, we have to perform a transformation of the states and the controls, so that the new system will be in standard form. We describe this transformation in this section, and express the cost function in the new variables. We then illustrate it on the standard two-dimensional example.

<u>Definition 2.3</u>: <u>A dynamical system is described in implicit form if the state equations are of the type</u>:

$$\underline{F}(k)(\underline{x}(k), \underline{x}(k+1), \underline{u}(k)) = 0, \quad k = 0,1,2,..., \qquad (2.67)$$

where $\underline{F}(k)$ is a function of the states $\underline{x}(k)$ and $\underline{x}(k+1)$ at times k and k+1, respectively, and of the control $\underline{u}(k)$ at time k.

<u>Definition 2.4</u>: A dynamical system is said to be in <u>standard form</u> if the state equations are of the type:

$$\underline{x}(k+1) = \underline{f}(k)(\underline{x}(k), \underline{u}(k)), \qquad (2.68)$$

with some function f(k).

<u>Definition 2.5</u>: <u>A dynamical system in standard form is linear if the function</u> $\underline{f}(k)$ <u>of (2.68) is a linear function of the state and the control at time k</u>:

$$\underline{f}(k)(\underline{x}(k), \underline{u}(k)) = \underline{A}(k)\underline{x}(k) + \underline{B}(k)\underline{u}(k). \qquad (2.69)$$

Example. The state equations (2.9) and (2.10), expressing flow conservation , are in implicit form, with

$$\underline{F}_k(\underline{x}(k), \underline{x}(k+1), \underline{\phi}(k) = \begin{bmatrix} \underline{x}(k) - \underline{Y}(k)\underline{\phi}(k) \\ \underline{x}(k+1) - \underline{Z}(k)\underline{\phi}(k) \end{bmatrix}. \qquad (2.70)$$

We define below a new state $\underline{z}(k)$ and a new control $\underline{v}(k)$ instead of the former state $\underline{x}(k)$ and the former control $\underline{\phi}(k)$, so that the system, expressed in $\underline{z}(k)$ and $\underline{v}(k)$, will be in standard form and linear. Also, we express the cost functions of (2.31) in terms of $\underline{z}(k)$ and $\underline{v}(k)$.

This transformation is valid for stationary networks.

Before describing the transformation, we have to classify the subnetworks in three categories according to their structure.

Structural classification of networks

By definition, in the stationary case, the number of entrances is equal to the number of exits. We may label the n entrances in any manner we like, but we have to label the exits in the same manner, so that entrance i corresponds to exit i; otherwise the i^{th} component of $\underline{x}(k)$ will not have the same physical meaning as the i^{th} component of $\underline{x}(k+1)$. Therefore, we can speak of the entrance-exit pairs bearing the same label, or of the exit corresponding to some entrance.

Definition 2.6: We shall call a subnetwork where at least one entrance is connected by a link to the corresponding exit a subnetwork of class 1.

Definition 2.7: A subnetwork of class 2 is a subnetwork where no entrance is is connected by a link to the corresponding exit, and where there is at least one entrance from which more than one link originates.

Definition 2.8: A subnetwork of class 3 is a subnetwork where no entrance is connected by a link to the corresponding link, and where only one link originates from each entrance.

Examples. The subnetworks of Examples 7.1, 7.2, 7.3, 7.5, 7.7 are of class 1. Those of examples 7.4 and 7.6 are of class 2. That of Fig. 2.3.3 is of class 3.

Remark: Subnetworks of class 3 are really unimportant for our study because the route followed by a car in the entire network is completely determined from the time it enters the network. Indeed, there is no choice at any node, and therefore the traffic allocation problem is trivially solved since the flow along each link is completely determined by the entering traffic distribution $\underline{\xi}$.

Accordingly, we shall completely exclude subnetworks of class 3 in the sequel. The proofs will only be given in the main course of the work for subnetworks of class 1. The proofs concerning subnetworks of class 2 as well as technical proofs for class 1 are to be found in Appendix B. Subnetworks of class 1 are by far the most important because a connection between an entrance and an exit bearing the same label represents a through roadway. In our context of main roads and side roads connecting them, those through roadways will most often be present.

We state all the results concerning state and control transformation, and the corresponding expression for the cost function, in the following theorem.

Theorem 2.3

1) For a subnetwork of class 1:

If

a) The n entrance-exit pairs are so labeled that entrance n is connected by a link to exit n.

b) The m links are so labeled that the n first links originate from n different entrances, and link n is the one which connects entrance n to exit n.

c) The new state $\underline{z}(k)$ is an (n-1)-dimensional vector defined by

$$\underline{x}(k) = \begin{pmatrix} \underline{z}(k) \\ x_n(k) \end{pmatrix} . \qquad (2.71)$$

d) The new control $\underline{v}(k)$ is an (m-n)-dimensional vector defined by

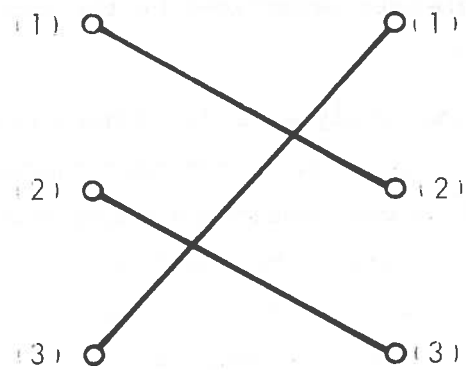

Figure 2.3.3 EXAMPLE OF SUBNETWORK OF CLASS 3

$$\underline{\phi}(k) = \begin{pmatrix} \underline{u}(k) \\ \underline{v}(k) \end{pmatrix}, \tag{2.72}$$

where $\underline{u}(k)$ is an n-dimensional vector.

Then

a. The flow conservation constraints are expressed by a dynamical system in $\underline{z}(k)$ that is in standard form, linear and stationary:

$$\underline{z}(k+1) = \underline{A}\,\underline{z}(k) + \underline{B}\,\underline{v}(k). \tag{2.73}$$

b. The cost function is quadratic but not homogeneously quadratic in the states and the controls. It is given by:

$$J = \sum_{k=1}^{N-1} J(k)(\underline{z}(k), \underline{v}(k)), \tag{2.74}$$

with

$$J(k)(\underline{z}(k), \underline{v}(k)) = \frac{1}{2}\,[\underline{z}^T(k), \underline{v}^T(k)] \begin{bmatrix} \underline{Q} & \underline{M}^T \\ \underline{M} & \underline{R} \end{bmatrix} \begin{bmatrix} \underline{z}(k) \\ \underline{v}(k) \end{bmatrix} + \underline{p}^T \underline{z}(k)$$

$$+ \underline{q}^T \underline{v}(k) + r. \tag{2.75}$$

2) For a subnetwork of class 2:

There exists a possible labeling of the entrance-exit pairs and an (n-1)-dimensional vector $\underline{d}$ such that, if

$$z'(k) = z(k) + \underline{d}, \tag{2.76}$$

where $\underline{z}(k)$ is defined by (2.71),

Then

a. The flow-conservation constraints are expressed by the dynamical system (2.73) in $\underline{z}'(k)$ instead of $\underline{z}(k)$, with the same coefficients $\underline{A}$ and $\underline{B}$ as in class 1.

b. The cost function is expressed in terms of $\underline{z}'(k)$ and $\underline{v}(k)$ by (2.75) with the same coefficients $\underline{Q}$, $\underline{M}$, $\underline{R}$ as in class 1 and with coefficients $\underline{p}'$, $\underline{q}'$, r' different from $\underline{p}$, $\underline{q}$, r of class 1.

Remarks

1) For a subnetwork of class 1, the construction described in part 2 of the theorem is also possible; that is: one has the choice between describing flow conservation by the state $\underline{z}(k)$ given by (2.71) or by the state $\underline{z}'(k)$ defined by (2.76).

These descriptions, however, will often correspond to two different labelings of the entrance-exit pairs. We shall refer to them as "subsystem in $\underline{z}$" and "subsystem in $\underline{z}'$."

For a subnetwork of class 2, only the parametrization of part 2 of the theorem is possible (see appendix B).

2) The parametrization of part 1 is applicable to a nonstationary network, provided that the number of entrances is the same for each subnetwork (n_k is constant with k), and all subnetworks are of class 1. Exactly the same construction can be performed. The system (2.73) becomes nonstationary, and the coefficients $\underline{Q}$, $\underline{M}$, $\underline{R}$, $\underline{p}$, $\underline{q}$, r depend on k because $\underline{Y}(k)$, $\underline{Z}(k)$, $\underline{L}(k)$, $\underline{h}(k)$ depend on k.

3) Conditions (a) and (b) of the theorem (part 1) are not restrictions on the structure of the subnetwork: they are just indications as to how the nodes and links have to be labeled to define the subsystem. It is possible to label the entrance-exit pairs so that (a) be satisfied because the subnetwork is of class 1. It is possible to label the links so that (b) be satisfied because at least one link originates from each entrance.

Proof of theorem 2.3

We give here only the proof of part 1 (class 1 links). For part 2 (class 2 links), the reader is referred to appendix B.

Proof of part 1(a)

The coefficients $\underline{A}$, $\underline{B}$, $\underline{Q}$, $\underline{M}$, $\underline{R}$, $\underline{p}$ $\underline{q}$, $\underline{r}$ are obtained by successive partitioning of the incidence matrices $\underline{Y}$ and $\underline{Z}$ and the cost matrix and vector $\underline{L}$ and $\underline{h}$.

Specifically, one goes through the following steps:

1) $\underline{Y}$ is partitioned into

$$\underline{Y} = (\underline{I}_n, \underline{\tilde{Y}}), \tag{2.77}$$

where $\underline{I}_n$ is the (nxn) identity matrix and $\underline{\tilde{Y}}$ is an n x (m-n).

$\underline{Z}$ is partitioned into

$$\underline{Z} = (\underline{Z}_1, \underline{Z}_2), \tag{2.78}$$

where $\underline{Z}_1$ is nxn and $\underline{Z}_2$ is n x (m-n).

2) $\underline{Z}_1$ is further partitioned into

$$\underline{Z}_1 = \begin{pmatrix} \underline{\tilde{Z}}_1 & (\underline{Z}_1)_{1n} \\ (\underline{Z}_1)_{n1} & (Z_1)_{nn} \end{pmatrix}, \tag{2.79}$$

where $\underline{\tilde{Z}}_1$ is (n-1) x (n-1), $(\underline{Z}_1)_{1n}$ is (n-1) x 1, $(\underline{Z}_1)_{n1}$ is 1 x (n-1), and $(Z_1)_{nn}$ is scalar.

$\underline{Z}_2-\underline{Z}_1\underline{Y}$ is partitioned into

$$\underline{Z}_2-\underline{Z}_1\underline{\tilde{Y}} = \begin{pmatrix} \widetilde{\underline{Z}_2-\underline{Z}_1\underline{\tilde{Y}}} \\ (\underline{Z}_2-\underline{Z}_1\underline{\tilde{Y}})_n \end{pmatrix}, \tag{2.80}$$

where $(\widetilde{\underline{Z}_2-\underline{Z}_1\underline{\tilde{Y}}})$ is (n-1) x (m-n), and $(Z_2-Z_1\tilde{Y})_n$ is 1 x (m-n).

3) $\underline{L}$ and $\underline{h}$ are partitioned into

$$\underline{L} = \begin{pmatrix} \underline{L}_{11} & \underline{L}_{12} \\ \underline{L}_{21} & \underline{L}_{22} \end{pmatrix}, \tag{2.81}$$

and

$$\underline{h} = \begin{pmatrix} \underline{h}_1 \\ \underline{h}_2 \end{pmatrix}, \tag{2.82}$$

where $\underline{L}_{11}$ is nxn, $\underline{L}_{12}$ is n x (m-n), $\underline{L}_{21}$ is (m-n) x n, and $\underline{L}_{22}$ is (m-n) x (m-n); $\underline{h}_1$ is nx1, and $\underline{h}_2$ is (m-n) x 1.

4) $\underline{L}_{11}$ and $\underline{h}_1$ are further partitioned into

$$\underline{L}_{11} = \begin{pmatrix} \tilde{\underline{L}}_{11} & (\underline{L}_{11})_{1n} \\ (\underline{L}_{11})_{n1} & (L_{11})_{nn} \end{pmatrix}, \tag{2.83}$$

$$\underline{h}_1 = \begin{pmatrix} \tilde{\underline{h}}_1 \\ (h_1)_n \end{pmatrix}, \tag{2.84}$$

where $\tilde{\underline{L}}_{11}$ is (n-1) x (n-1), $(\underline{L}_{11})_{1n}$ is (n-1) x 1, $(\underline{L}_{11})_{n1}$ is 1 x (n-1), and $(L_{11})_{nn}$ is scalar; $\tilde{\underline{h}}_1$ is (n-1) x 1, and (h_1) is scalar.

Setting

$$\underline{\Lambda} = \underline{L}_{12} - \underline{L}_{11}\tilde{\underline{Y}}, \tag{2.85}$$

one partitions $\underline{\Lambda}$ into

$$\underline{\Lambda} = \begin{pmatrix} \tilde{\underline{\Lambda}} \\ \underline{\Lambda}_n \end{pmatrix}, \tag{2.86}$$

where $\tilde{\underline{\Lambda}}$ is (n-1) x (m-n), and $\underline{\Lambda}_n$ is 1 x (m-n).

5) The coefficients occurring in (a and b) are given by

$$\underline{A} = \tilde{\underline{Z}}_1 - (\underline{Z}_1)_{1n}\underline{\nu}_{n-1}^T, \tag{2.87}$$

$$\underline{B} = \widetilde{\underline{Z}_2 - \underline{Z}_1\tilde{\underline{Y}}}, \tag{2.88}$$

$$\underline{Q} = \tilde{\underline{L}}_{11} - (\underline{L}_{11})_{1n}\underline{\nu}_{n-1}^T - \nu_{n-1}(L_{11})_{n1} + (\underline{L}_{11})_{nn}\underline{\nu}_{n-1}\underline{\nu}_{n-1}^T \tag{2.89}$$

$$\underline{M} = \tilde{\underline{\Lambda}}^T - \underline{\Lambda}_n^T\underline{\nu}_{n-1}^T, \tag{2.90}$$

$$\underline{R} = \tilde{\underline{Y}}^T\tilde{\underline{L}}_{11}\tilde{\underline{Y}} - \tilde{\underline{Y}}^T\underline{L}_{12} - L_{12}^T\tilde{Y} + \underline{L}_{22} \tag{2.91}$$

$$\underline{p} = \tilde{\underline{h}}_1 - \underline{\nu}_{n-1}{}^*\underline{h}_1)_n + F(\underline{L}_{11})_{1n} - F(L_{11})_{nn}\underline{\nu}_{n-1}, \tag{2.92}$$

$$q = \underline{h}_2 - \tilde{\underline{Y}}^T\underline{h}_1 + F\underline{\Lambda}_n^T, \tag{2.93}$$

$$r = \frac{F^2}{2}(L_{11})_{nn}, \tag{2.94}$$

where F is the total entering flow perturbation.

$$F \equiv \underline{\nu}_n^T\underline{x}(1). \tag{2.95}$$

Let us now establish the above equations.

From the definition of $\underline{u}(k)$ and $\underline{v}(k)$ (2.72), the partitioning of $\underline{Y}$ given by (2.77) and that of $\underline{Z}$ given by (2.78), it follows that the flow-conservation constraints

$$\underline{F}(k)(\underline{x}(k), \underline{x}(k+1), \underline{\phi}(k)) = 0, \tag{2.96}$$

where $\underline{F}(k)$ is defined by (2.70), become:

$$\left.\begin{aligned} \underline{x}(k) &= (\underline{I}_n, \underline{\tilde{Y}}) \begin{pmatrix} \underline{u}(k) \\ \underline{v}(k) \end{pmatrix}, & (2.97) \\ \underline{x}(k+1) &= (\underline{Z}_1, \underline{Z}_2) \begin{pmatrix} \underline{u}(k) \\ \underline{v}(k) \end{pmatrix}. & (2.98) \end{aligned}\right\}$$

From (2.97) it follows that

$$\underline{u}(k) = \underline{x}(k) - \underline{\tilde{Y}}\ \underline{v}(k), \tag{2.99}$$

which, combined with (2.98), yields:

$$\underline{x}(k+1) = \underline{Z}_1\ \underline{x}(k) + (\underline{Z}_2 - \underline{Z}_1\underline{\tilde{Y}})\underline{v}(k). \tag{2.100}$$

Now the definition of $\underline{z}(k)$ (2.71), together with the conservation of total flow (2.13), implies that:

$$x_n(k) = F - \underline{\nu}^T_{n-1}\ \underline{z}(k), \tag{2.101}$$

for every k.

We now take the (n-1) first components of (2.100), and express $x_n(k)$ in terms of $\underline{z}(k)$ and F through (2.101). With the further partitioning of $\underline{Z}_1$ and $\underline{Z}_2$ defined by (2.79) and (2.80), we can express the (n-1) first components of (2.100):

$$\underline{z}(k+1) = \underline{\tilde{Z}}_1\ \underline{z}(k) + (\underline{Z}_1)_{1n} x_n(k) + (\widetilde{\underline{Z}_2 - \underline{Z}_1\underline{\tilde{Y}}})\underline{v}(k), \tag{2.102}$$

or, using (2.101),

$$\underline{z}(k+1) = [\underline{\tilde{Z}}_1 - (\underline{Z}_1)_{1n}\underline{\nu}^T_{n-1}]\underline{z}(k) + (\widetilde{\underline{Z}_2 - \underline{Z}_1\underline{\tilde{Y}}})\underline{v}(k) + F(\underline{Z}_1)_{1n}, \tag{2.103}$$

which will be (2.73) (with $\underline{A}$ defined by 2.87 and $\underline{B}$ by 2.88) if the term $(\underline{Z}_1)_{1n}$

is zero.

Let us now show that indeed $(\underline{Z}_1)_{1n} = 0$. In fact, (2.79) and the definition of $\underline{Z}$ (Definition (2.2)) show that the i^{th} component of $(\underline{Z}_1)_{1n}$ is 1 if and only if u_n is the flow along a link leading to exit i. All other components of $(\underline{Z}_1)_{1n}$ are equal to zero. However, condition 1(b), together with (2.72), shows that u_n is the flow along link (n,n), so that $(\underline{Z}_1)_{1n} = 0$ and $(Z_1)_{nn} = 1$. Thus, part 1(a) of the theorem is proved.

Proof of part 1(b)

We now express the cost function (2.31) in terms of $\underline{z}(k)$ and $\underline{v}(k)$.

From the partitioning (2.81) and (2.82) of $\underline{L}$ and $\underline{h}$, and from (2.77), it follows that

$$J(k)(\underline{x}(k),\ \underline{v}(k)) = \frac{1}{2}[\underline{x}^T(k)\underline{L}_{11}\underline{x}(k) + 2\underline{x}^T(k)(\underline{L}_{12}-\underline{L}_{11}\tilde{\underline{Y}})\underline{v}(k)$$
$$+ \underline{v}^T(k)(\tilde{\underline{Y}}\ \underline{L}_{11}\tilde{\underline{Y}} - 2\tilde{\underline{Y}}^T\underline{L}_{12} + \underline{L}_{22})\underline{v}(k)] + [\underline{h}_1^T\underline{x}(k)$$
$$+ (\underline{h}_2^T-\underline{h}_1^T\tilde{\underline{Y}})\underline{v}(k)]. \tag{2.104}$$

We can now express $\underline{x}(k)$ in terms of $\underline{z}(k)$ and F through (2.101), and use the further partitioning (2.83), (2.84), and (2.86), so that:

$$\underline{x}^T(k)\underline{L}_{11}\underline{x}(k) = \underline{z}^T(k)\tilde{\underline{L}}_{11}\underline{z}(k) + 2\underline{z}^T(k)(\underline{L}_{11})_{1n}x_n(k)$$
$$+ (L_{11})_{nn}[x_n(k)]^2,$$
$$= \underline{z}^T(k)\tilde{\underline{L}}_{11}\underline{z}(k) + 2\underline{z}^T(k)(\underline{L}_{11})_{1n}(F - \underline{\nu}_{n-1}^T\underline{z}(k))$$
$$+ (F - \underline{\nu}_{n-1}^T\underline{z}(k))^2(L_{11})_{nn}, \tag{2.105}$$

and

$$2\underline{x}^T(k)\ \underline{\Lambda}\ \underline{v}(k) = 2[z^T(k)\ \tilde{\Lambda}\ v(k) + (F - \underline{z}^T(k)\underline{\nu}_{n-1})\underline{\Lambda}_n\underline{v}(k)]$$
$$= 2\underline{z}^T(k)(\tilde{\underline{\Lambda}} - \underline{\nu}_{n-1}\underline{\Lambda}_n)\underline{v}(k) + 2F\underline{\Lambda}_n\underline{v}(k), \tag{2.106}$$

and also

$$\underline{h}_1^T\underline{x}(k) = \tilde{h}_1^Tz(k) + (h_1)_n(F - \underline{\nu}_{n-1}^T\underline{z}(k))$$
$$= [\tilde{\underline{h}}_1^T - (h_1)_n\underline{\nu}_{n-1}^T]\underline{z}(k) + (h_1)_nF. \tag{2.107}$$

Replacing the corresponding terms by the right-hand sides of (2.105), (2.106), and (2.107), respectively, in (2.104) gives equation (2.75) with the coefficients defined by equations (2.87) through (2.94). Q.E.D.

Illustration: Standard two-dimensional example

Let us apply the transformation of Theorem (2.3) to the standard two-dimensional example 7.1. This is a subnetwork of class 1.

The labeling required in part 1 (a and b) of the theorem is shown in Fig. 7.1. In this case, $n = 2$ and $m = 4$.

From (2.72), $\underline{u}^T(k) = (\phi_1(k), \phi_2(k))$

and $\underline{v}^T(k) = (\phi_3(k), \phi_4(k))$.

Also, from (2.71), the state is scalar: $z(k) = x_1(k)$.

It is seen directly from Figure 7.1 that, in this example,

$$x_1(k+1) = \phi_1(k) + \phi_4(k),$$

$$x_2(k+1) = \phi_2(k) + \phi_3(k),$$

which, in the new notation, becomes (using 2.13)

$$\left.\begin{aligned} z(k+1) &= u_1(k) + v_2(k) \\ F - z(k+1) &= u_2(k) + v_1(k) \end{aligned}\right\} . \tag{2.108}$$

Also, from Fig. 7.1,

$$\begin{aligned} u_1(k) &= x_1(k) - v_1(k) = z(k) - v_1(k), \\ u_2(k) &= x_2(k) - v_2(k) = F - z(k) - v_2(k), \end{aligned} \tag{2.109}$$

so that equations (2.108) become:

$$z(k+1) = z(k) - v_1(k) + v_2(k), \tag{2.110}$$

which is a special case of (2.73), with

$$A = 1, \tag{2.111}$$

$$\underline{B} = (-1, 1), \tag{2.112}$$

The same result can be found by applying equations (2.87) and (2.88) with

$$\underline{Z}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \tag{2.113}$$

and

$$\underline{Z}_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}. \tag{2.114}$$

The cost function is given in $\underline{z}(k)$ and $\underline{v}(k)$ by 2.75.

The corresponding coefficients are:

$$Q = L_1 + L_2, \tag{2.115}$$

$$\underline{M} = \begin{pmatrix} -L_1 \\ L_2 \end{pmatrix}, \tag{2.116}$$

$$\underline{R} = \begin{pmatrix} L_1+L_3 & 0 \\ 0 & L_2+L_4 \end{pmatrix}, \tag{2.117}$$

$$p = -FL_2, \tag{2.118}$$

$$\underline{q} = \begin{pmatrix} 0 \\ -FL_2 \end{pmatrix}, \tag{2.119}$$

$$r = \frac{F^2}{2} L_2 , \tag{2.120}$$

obtained by applying equations (2.89) through (2.94) with

$\tilde{L}_{11} = L_1,$

$(L_{11})_{nn} = L_2,$

$(L_{11})_{1n} = (L_{11})_{n1} = 0,$

$\tilde{Z}_1 = 1 = (Z_1)_{nn},$

$(Z_1)_{1n} = (Z_1)_{n1} = 0.$

2.3.4 Controllability of Reduced Systems

In this section, we investigate the controllability of the reduced systems defined in section 2.3.3. Let us recall that the reason why we are interested in controllability is to be able to make the statements about the asymptotic behavior in section 2.3.5. We shall not apply the standard controllability tests to the matrices $\underline{A}$ and $\underline{B}$ of equations (2.87) and (2.88) because of the complicated sequence of transformations to obtain them. Instead, we shall go back to the definition of controllability, and take advantage of both the graph structure of the network and the way the reduced state $\underline{z}$ or $\underline{z}'$ is related to the original state $\underline{x}$. For the proofs, we concentrate on subnetworks of class 1 (and on the reduction to subsystems in $\underline{z}$, rather than in $\underline{z}'$). The proofs for subnetworks of class 2 are included in appendix B. In 2.3.4.1, we consider the meaning of controllability of a reduced system in the context of repeated subnetworks. In 2.3.4.2, we formulate a condition for controllability in the language of graph theory. It is necessary and sufficient for class 1 subnetworks, and at least sufficient for class 2. The reader is referred to appendix A and to the bibliography, for graph terminology. In 2.3.4.3, we translate the condition into an algebraic one by means of Markov chain theory.

2.3.4.1 The Meaning of Controllability in This Context

Remark The state reduction performed in 2.3.3 is artificial as far as it seemed to favor one entrance-exit pair by labeling it n and treating it differently from the others. We have to do so to substitute a system in standard form for one in implicit form, but that operation has no physical meaning. Any entrance-exit pair may have been labeled n, provided that the entrance and the exit are connected by a link. Thus, we have the choice between as many subsystems as there are entrance-exit pairs connected by links (for subnetworks of class 1). We shall see that, provided that the subnetwork does not have a certain peculiar feature, either all possible subsystems are controllable or none are. Therefore, we now interpret the controllability of a subsystem in $\underline{z}$ in terms of the full state $\underline{x}$.

Theorem 2.4

1) For subnetworks of class 1, the system (2.73) describing the evolution of the reduced state $\underline{z}$ defined by (2.71) is controllable, if and only if, for any initial state $\underline{x}(1)$, it is possible to drive the full-state $\underline{x}$ in some number of steps; say, k, to

$$\underline{x}(k) = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ F \end{pmatrix}, \tag{2.121}$$

by a feasible choice of the flow perturbations $\underline{\phi}(1)$, $\underline{\phi}(2),\ldots,\underline{\phi}(k-1)$.

2) For subnetworks of class 2, there is a set of nodes i_1, $i_2,\ldots,i_r$, defined in Appendix B. The system (2.73) describing the evolution of the reduced state $\underline{z}'$ defined by (2.76) is controllable, if and only if, for any initial state $\underline{x}(1)$, it is possible to drive the full state $\underline{x}$ in some number of steps say k, to $\underline{x}(k)$, where

$$x_j(k) = \begin{cases} \frac{F}{r} & \text{if } j \in \{i_1, i_2, \ldots i_r\} \\ 0 & \text{otherwise.} \end{cases} \tag{2.122}$$

Remark. Part 2 is also applicable to subsystems in z' describing the flows in subnetworks of class 1 (see remark after theorem 2.3).

Proof of Theorem 2.4

We prove only part 1 here. See appendix B for part 2. Controllability of the reduced system in $\underline{z}(k)$ means, by definition, that it is possible to drive the reduced state $\underline{z}$ in some number of steps; say, k, to $\underline{z}(k) = 0$, by an appropriate choice of $\underline{v}(1)$, $\underline{v}(2),\ldots\underline{v}(k-1)$. However, since $\underline{\nu}_n^T\underline{x}(k) = \underline{\nu}_n^T\underline{x}(1) \triangleq F$ (2.13), $\underline{z}(k) = 0$ is equivalent to (2.121) given the definition (2.71) of $\underline{z}(i)$. On the other hand, from the definition (2.72) of $\underline{v}(i)$, choosing the sequence $\underline{v}(i)$, $i = 1,\ldots k-1$, is equivalent to choosing the sequence $\underline{\phi}(i)$, $i = 1,\ldots(k-1)$, so as to satisfy the flow conservation constraints (2.9), (2.10). Q.E.D.

Remark. Since $\underline{z}$ or $\underline{z}'$ satisfies a linear system (2.73), once $\underline{z}(i)$ has been brought to 0 (at stage k), one can choose the controls $\underline{v}(k+1)$, $\underline{v}(k+2)$ at subsequent stages k+1, k+2,... to be 0, so that $\underline{z}(j) = 0$ for all $j \geq k$. Therefore, equations (2.121) or (2.122) hold, in the full state, for all $j \geq k$. Physically, they mean that, for subnetworks of class 1, the entire incoming flow perturbation of traffic can be driven to entrance n after k subnetworks, and made to enter all downstream subnetworks through entrance n. Likewise, for subnetworks ot class 2, it can be spread out evenly between entrances $i_1, i_2, \ldots i_r$, to subnetwork k, and made to enter all downstream subnetworks in the same manner.

2.3.4.2 Graph Formulation of Controllability

The meaning of negative flow perturbations

Since we deal with flow perturbations, and not with flows (see 2.2), some components of the vectors $\underline{\phi}(\ell)$ can be chosen to be negative. This possibility makes the reduced systems more easily controllable because the conditions of theorem 2.4 will be more easily satisfied. In addition, an important class of perturbations we consider are those which do not change the total flow passing through the network. Some components of $\underline{x}(1)$, and therefore of $\underline{\phi}$, will certainly be negative in this case. We have been describing the subnetworks by directed bipartite graphs with 2n nodes and m arcs. However, the possibility of choosing some components of $\underline{\phi}(\ell)$ to be negative allows one to disregard the arrows in those graphs; i.e., to consider them to be undirected graphs. Indeed, driving a flow perturbation $\alpha > 0$ from exit j to entrance i of subnetwork ℓ is possible by assigning $\phi_r(\ell) = -\alpha$ if r is the label of link (i,j). This is illustrated in Fig. 2.3.4-1. Thus, we drive a negative-flow perturbation $-\alpha$ from entrance i to exit j.

In Fig. 2.3.4-2, it is illustrated how the use of negative flow perturbations makes it possible to drive the reduced state to 0 in one step in that example. Even when no link exists from entrance i to entrance j, we may be able to follow some path back and forth between the entrances and exits of the subnetwork to send some flow perturbations from entrance i to exit j. With that construction in mind, we shall introduce the notion of the accessibility graph to describe the connectivity properties of a subnetwork.

Definition 2.8

Let S be a subnetwork. Its associated accessibility graph G is obtained as follows:

1) Graph G has n nodes, one for each entrance-exit pair in S.

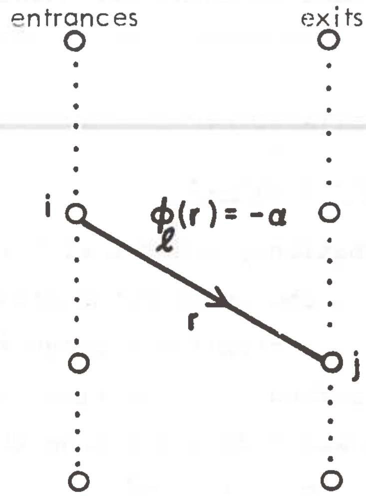

Figure 2.3.4-1   NEGATIVE-FLOW PERTURBATION

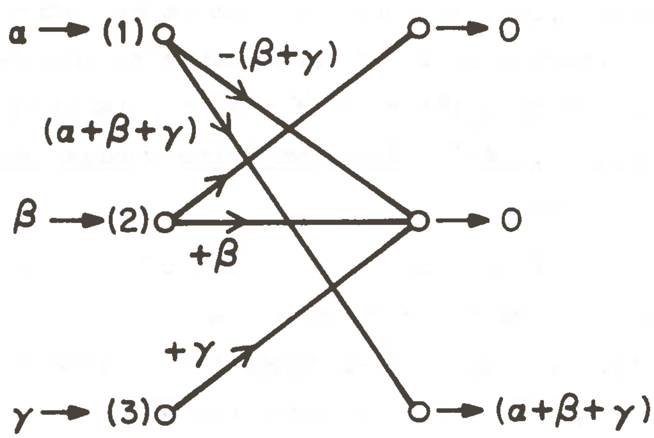

Figure 2.3.4-2   RAPID CONTROL, BY NEGATIVE-FLOW PERTURBATIONS

2) Graph G is directed and arc (i,j) exists if and only if there is a path from entrance i to exit j in S (without taking the arrows into account in S).

Remark: Consider the entire network, whose typical subnetwork is S. From the definition of G, it follows immediately that the existence of a path from entrance i in subnetwork k to exit j in subnetwork ℓ ($\ell > k$) in the network is equivalent to the existence of a path of length (ℓ-k) from node i to node j in G.

Example. The accessibility graph corresponding to the subnetwork of Fig. 2.3.4-2 is drawn in Fig. 2.3.4-3. It is complete.

Preview. The definition 2.8 of G will enable us to characterize the controllability of reduced systems in the following theorem. We consider here only those subnetworks whose accessibility graph does not contain any transient node (appendix A). Subnetworks giving rise to graphs with transient nodes are somewhat peculiar and are considered in appendix B. An example is given in Fig. 2.3.4-4, and the corresponding accessibility graph is drawn in Fig. 2.3.4-5.

Theorem 2.5.

Let S be a subnetwork. Assume that its accessibility graph G does not have any transient node. Then

1) If S is a subnetwork of class 1

A necessary and sufficient condition for any reduced system in $\underline{z}$ to be controllable is that G be strongly connected(*).

2) If S is a subnetwork of class 2

a) If the accessibility graph G is strongly connected and aperiodic(*), all reduced systems (in $\underline{z}'$) are controllable.

b) If the accessibility graph G is not strongly connected, no reduced system is controllable.

---

*See appendix A for definition.

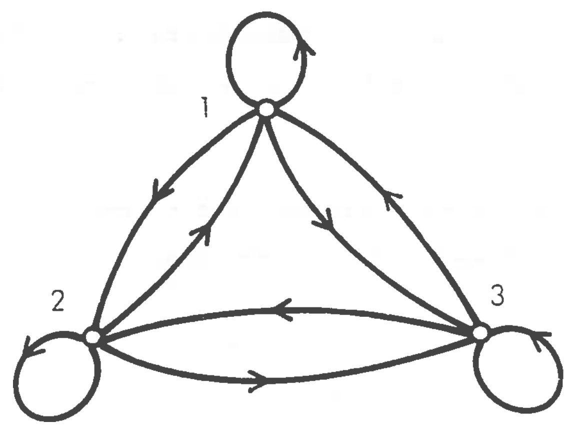

Figure 2.3.4-3   ACCESSIBILITY GRAPH OF THE SUBNETWORK OF RAPID CONTROL

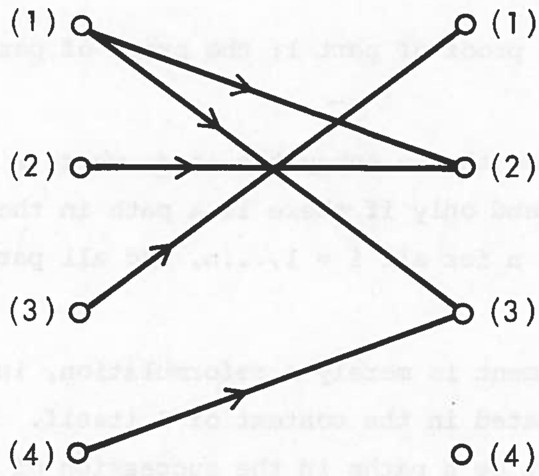

Figure 2.3.4-4 SUBNETWORK WHOSE ACCESSIBILITY GRAPH HAS TRANSIENT NODES

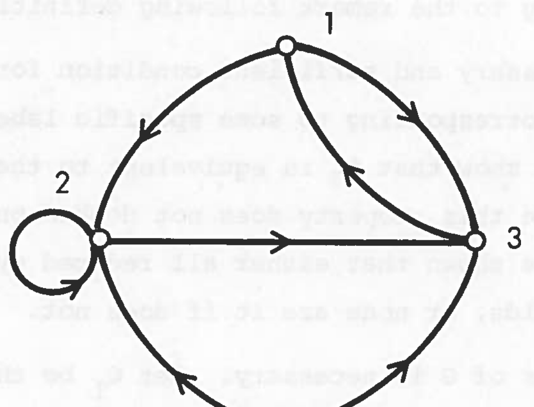

Figure 2.3.4-5 ACCESSIBILITY GRAPH OF THE SUBNETWORK OF FIGURE 2.3.4-4

Proof of Theorem 2.5.

We shall give only the proof of part 1; the proof of part 2 is in appendix B.

1) First, let us notice that a subsystem in $\underline{z}$, where $\underline{z}$ is defined by (2.71), is controllable if and only if there is a path in the accessibility graph G from node i to node n for all i = 1,...n, and all paths have the same length.

2) Indeed, that statement is merely a reformulation, in terms of G, of a condition which can be stated in the context of S itself. That condition is the following: there must be n paths in the succession of identical subnetworks, with origin 1,2,...n, respectively, and extremity n. In addition, these n paths must go through the same number of subnetworks, so that flow perturbations along those n paths will arrive at entrance n at the same stage. This condition is necessary and sufficient for being able to drive any incoming flow perturbation to the nth entrance at some stage k, and that is precisely controllability in our problem, according to theorem 2.4.(1). On the other hand, the statement of (1) is clearly an appropriate translation of this condition in terms of G, according to the remark following definition 2.8.

3) We thus have a necessary and sufficient condition for controllability of the reduced system in $\underline{z}$ corresponding to some specific labeling of the entrances and exits. Let us show that it is equivalent to the strong connectedness of graph G. Since this property does not depend on the labeling of the nodes of G, we shall have shown that either all reduced systems are controllable if the property holds, or none are it if does not.

4) Strong connectedness of G is necessary. Let C_1 be the strongly connected component of node n. If G is not strongly connected, $C_1 \neq G$. If i is a node in G and $i \notin C_1$, then there is no path from node i to node n in G. Therefore, according to (1), the subsystem corresponding to that particular choice of n is not controllable.

5) Strong connectedness of G is sufficient. If G is strongly connected, then for all nodes i = 1,2,...n, there is a path in G with origin i and extremity n. It remains to be shown that it is possible to find n paths with

the same length. Since the subnetwork S is of class 1, there is at least one entrance connected to the corresponding exit by a link. Accordingly, there is a loop at the corresponding node in G, which implies that G is aperiodic. Aperiodicity in turn implies (by elementary arithmetic reasoning, in appendix A.4) the existence of n paths of same length from nodes 1,2,...n, respectively, to node n. Q.E.D.

Remark. For subnetworks of class 2, we do not know if condition 2(a) is necessary as well as sufficient.

Examples of controllable and uncontrollable systems

Fig. 2.3.4-2 shows a subnetwork of class 1 which gives rise to controllable subsystems. The corresponding accessibility graph is drawn in Fig. 2.3.4-3. The network of Fig. 2.3.4-6, also of class 1, gives rise to uncontrollable subsystems: its accessibility graph (Fig. 2.3.4-7) has two strongly connected components.

The network of Fig. 2.3.4-8 is of class 2, and gives rise to controllable subsystems; its accessibility graph is shown in Fig. 2.3.4-9. That of Fig. 2.3.4-10 is of class 2, and gives rise to uncontrollable subsystems. Its accessibility graph (2.3.4-11) is periodic (with period 2).

Remark. In the above examples, when controllability holds, it is possible to achieve it in one step; that is, to drive the reduced state to 0 in one step. In Fig. 2.3.4-12, an example is shown of a subnetwork which gives rise to controllable subsystems since it is of class 1 and its accessibility graph, shown in Fig. 2.3.4-13, is strongly connected and without transient nodes. However, in this example, the reduced state $\underline{z}$ cannot be driven to 0 in only one step. Indeed, there is only one entrance connected to the corresponding exit by a link so that we have to label it n = 3 to define a subsystem in $\underline{z}$. It is impossible to drive the flow entering entrance 1 to entrance 3 in one step. Controllability is possible in two steps because there are paths of length two with origins 1, 2, and 3, respectively, and extremity 3.

2.3.4.3 Algebraic Formulation of Controllability

In this paragraph, we wish to translate the geometrical conditions given

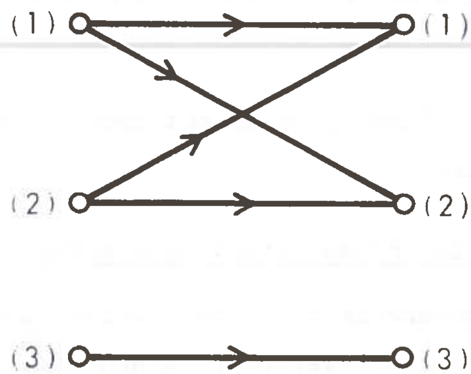

Figure 2.3.4-6 UNCONTROLLABLE SUBNETWORK OF CLASS 1

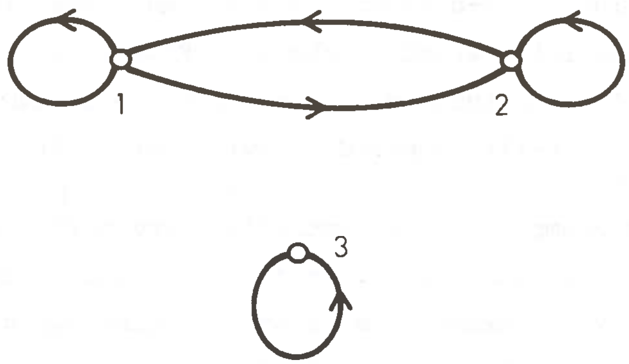

Figure 2.3.4-7 ACCESSIBILITY GRAPH OF SUBNETWORK OF CLASS 1

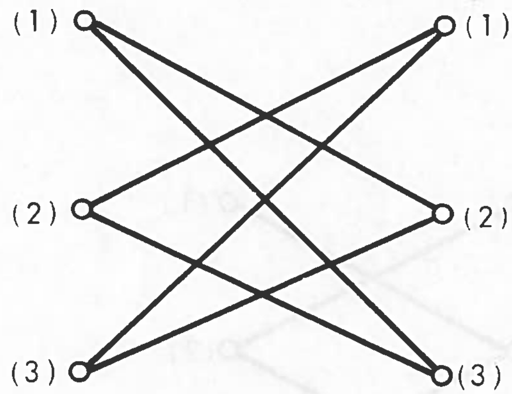

Figure 2.3.4-8   CONTROLLABLE SUBNETWORK OF CLASS 2

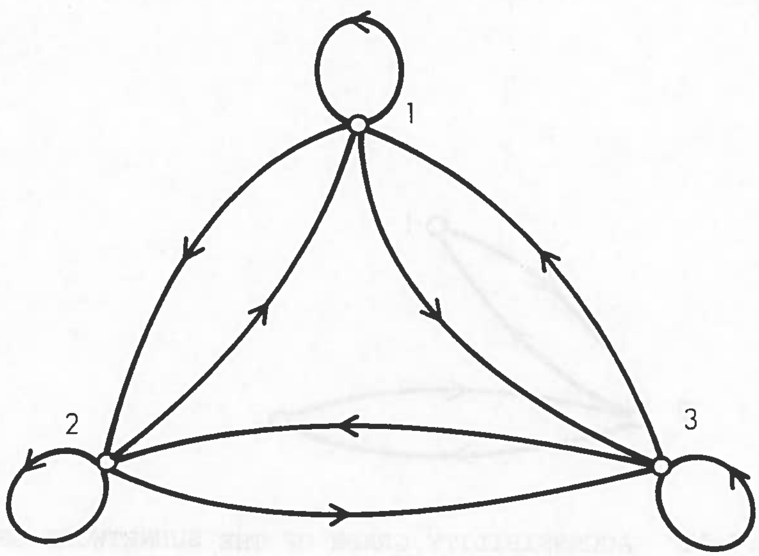

Figure 2.3.4-9   ACCESSIBILITY GRAPH OF THE SUBNETWORK OF FIGURE 2.3.4-8

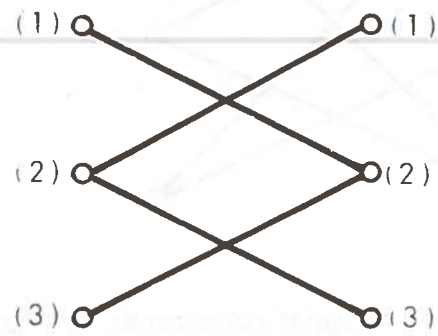

Figure 2.3.4-10 UNCONTROLLABLE SUBNETWORK OF CLASS 2

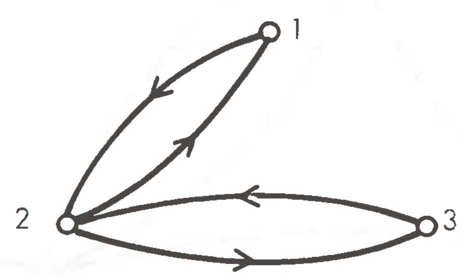

Figure 2.3.4-11 ACCESSIBILITY GRAPH OF THE SUBNETWORK OF FIGURE 2.3.4-10

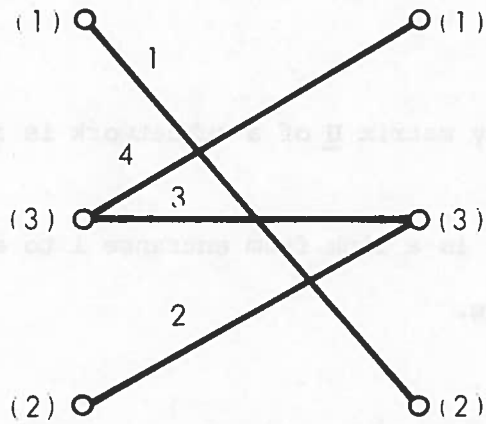

Figure 2.3.4-12 SUBNETWORK OF CLASS 1, CONTROLLABLE BUT NOT IN ONE STEP

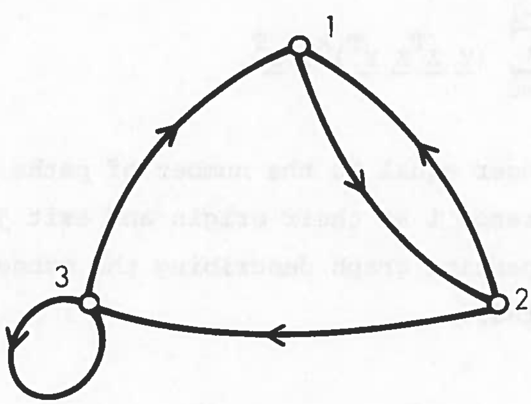

Figure 2.3.4-13 ACCESSIBILITY GRAPH OF THE SUBNETWORK OF FIGURE 2.3.4-12

in theorem 2.5 into algebraic tests. This will make easier the task of determining if a given subnetwork gives rise to controllable subsystems in the case of large subnetworks.

Definition 2.9

The entrance-exit adjacency matrix $\underline{U}$ of a subnetwork is an (nxn) matrix of zeroes and ones defined by:

$$U_{ij} = \begin{cases} 1 & \text{if there is a link from entrance } i \text{ to exit } j, \\ 0 & \text{otherwise.} \end{cases}$$

Lemma 2.6

1) If $\underline{Y}$ and $\underline{Z}$ are the entrance and exit node-arc incidence matrices respectively, defined in 2.2.3 (equations 2.9 and 2.10), then

$$\underline{U} = \underline{Y}\,\underline{Z}^T. \tag{2.123}$$

2) If $\underline{V}$ is the (nxn) matrix defined from the adjacency matrix $\underline{U}$ by:

$$\underline{V} \triangleq \sum_{k=0}^{n-1} (\underline{U}\,\underline{U}^T)^k \underline{U} = \sum_{k=0}^{n-1} (\underline{Y}\,\underline{Z}^T\underline{Z}\,\underline{Y}^T)^k \underline{Y}\,\underline{Z}^T, \tag{2.124}$$

then V_{ij} is a non-negative integer equal to the number of paths of length not greater than (2n-1) having entrance i as their origin and exit j as their extremity in the undirected bipartite graph describing the subnetwork S (i.e., in which the arrows are neglected).

Proof of Lemma 2.6

Part 1. The (i,j) entry of $\underline{Y}\,\underline{Z}^T$ is equal to $\sum_{k=1}^{m} Y_{ik}Z_{jk}$, and $Y_{ik}Z_{jk}$ is not zero (and is equal to 1) if and only if link k connects entrance i to exit j. Since, on the other hand, at most one link may connect entrance i to exit j, there will be at most one k for which $Y_{ik}Z_{jk} = 1$; therefore, $\sum_{k=1}^{m} Y_{ik}Z_{jk} =$ $= \begin{cases} 1 \text{ if } i \text{ is connected to } j \text{ by a link,} \\ 0 \text{ otherwise,} \end{cases}$ and the claim follows from the definition of $\underline{U}$.

Part 2. By induction, one can show that the (i,j) entry of $(\underline{U}\,\underline{U}^T)^k$ is the number of paths of length (2k) having entrance i as their origin and entrance j as their extremity (see appendix A). A path from an entrance to an exit is necessarily of odd length. A path of length (2k+1) from entrance i to exit j is obtained from a path of length (2k) from entrance i to some entrance ℓ by adding link (ℓ, j). Let $\underline{X} = (\underline{U}\,\underline{U}^T)^k$. Then,

$$\sum_{\ell=1}^{n} x_{i\ell}\, u_{\ell j}$$

is the number of paths of length (2k+1) from entrance i to exit j. To obtain the total number of paths from entrance i to exit j of length not greater than (2n-1), one has to sum over k, from k = 0 to k = n-1, so that (2k+1) = 1,3,...2n-1. Hence, our claim. Q.E.D.

Definition 2.10

We define the following (nxn) stochastic matrix $\underline{P}$:

$$p_{ij} = \frac{v_{ij}}{\sum_{k=1}^{n} v_{ik}} . \qquad (2.125)$$

Remark. The denominator in (2.125) is different from zero: indeed, from lemma 2.6, it is equal to the total number of paths in the bipartite graph describing S which originate from entrance i and have some exit as their extremity, and whose length is not greater than (2n-1). Since there is at least one link originating from each entrance, $\sum_{k=1}^{n} v_{ij} > 0$.

Lemma 2.8. The stochastic matrix $\underline{P}$ defined by (2.125) is naturally associated to the accessibility graph G in the sense that its (i,j) entry is non-zero if and only if arc (i,j) exists in G.

Proof of Lemma 2.8: $\underline{P}$ is a stochastic matrix by construction. By (2.125), $p_{ij} \neq 0$ if and only if $v_{ij} \neq 0$. By lemma 2.6, $v_{ij} \neq 0$ if and only if there is a path from entrance i to exit j in S with length not greater than (2n-1). By the definition 2.8 of G, this is equivalent to the existence of arc (i,j) in G:

the total number of nodes in the bipartite graph being 2n, the existence of a path implies that of a path of length not greater than (2n-1). Q.E.D.

These definitions and lemmas enable us to present the algebraic test in the following theorem.

Theorem 2.6.

Let S be a subnetwork whose accessibility graph G has no transient node. Let $\underline{P}$ be defined by (2.125) from (2.124). Define

Condition C_1: The number 1 is a simple eigenvalue of $\underline{P}$.

Condition C_2: The number 1 is a simple eigenvalue of $\underline{P}$, and $\underline{P}$ has no other eigenvalue of magnitude 1.

Then

1) For a subnetwork S of class 1

S gives rise to controllable subsystems if and only if condition C_1 holds.

2) For a subnetwork S of class 2

a) If condition C_1 does not hold, S gives rise to uncontrollable systems

b) If condition C_2 holds, S gives rise to controllable systems.

Proof of Theorem 2.6.

Taking into account lemma 2.8, it follows that the state classification of the Markov chain whose transition probabilities are the entries of $\underline{P}$ is equivalent to the two-level partition of G discussed in appendix A. Therefore, from a standard Markov chain theorem [32], [33], [34].

1) Condition $C_1 \Longleftrightarrow$ G has one single final class.

2) Condition $C_2 \Longleftrightarrow$ G has one single final class, and that final class is aperiodic.

Since, by assumption, G has no transient node, its having one single final class is equivalent to its being strongly connected, and the aperiodicity of the final class implies the aperiodicity of G. Therefore, theorem 2.6 is a mere translation of theorem 2.5. Q.E.D.

<u>Remarks</u>

1) We assume here that G has no transient nodes. If it does, some subsystems are controllable and others are not (according to whether n is transient in class 1). We discuss that point, and give an algebraic criterion to determine if some node is transient in appendix B.

2) The conditions of theorem 2.6 are valid as well for any stochastic matrix $\underline{P}$ naturally associated with G in the sense of lemma 2.8, not only for $\underline{P}$ given by (2.125). Indeed, all such matrices correspond to Markov chains with the same state classification.

<u>Examples</u>

Let us apply the above technique to some examples encountered. We obtain $\underline{V}$ through equation (2.124). For the example of Fig. 2.3.4-6,

$$\underline{V} = \begin{bmatrix} 21 & 21 & 0 \\ 21 & 21 & 0 \\ 0 & 0 & 3 \end{bmatrix},$$

so that

$$\underline{P} = \begin{bmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

whose characteristic polynomial is $s(1-s)^2$. The number 1 is an eigenvalue with multiplicity 2, so that condition C_1 is not satisfied and every subsystem is uncontrollable.

For the example of Fig. 2.3.4-10, we have

$$\underline{V} = \begin{bmatrix} 0 & 7 & 0 \\ 3 & 0 & 3 \\ 0 & 7 & 0 \end{bmatrix},$$

so that

$$\underline{P} = \begin{pmatrix} 0 & 1 & 0 \\ 1/2 & 0 & 1/2 \\ 0 & 1 & 0 \end{pmatrix},$$

with characteristic polynomial $\det(\underline{P}-s\underline{I}) = -s(1-s^2) = +s(s-1)(s+1)$. The number (-1) is an eigenvalue of $\underline{P}$, so that condition C_2 is not satisfied for this subnetwork of class 2.

In this case, theory does not allow us to conclude, because condition C_1 holds and C_2 does not. However, we have established only that C_1 is necessary and C_2 sufficient. Nevertheless, in this particular example, the subsystems are not controllable (see discussion in section 7).

2.3.5 Extension and Application of Linear-Quadratic Optimal Control Theory

The expression (2.75) for the cost function in the new variables, derived in section 2.3.3, would make our problem linear-quadratic if the coefficients $\underline{p}$ and $\underline{q}$ of the linear terms, given by (2.92) and (2.93), were equal to zero. This category of optimization problems has been studied extensively both in the continuous-time [25], [26], [27] and discrete-time cases [21]. However, linear terms are present in our cost function expressed in the new variables $\underline{z}$ and $\underline{v}$, even when the cost is homogeneously quadratic in $\underline{x}$ and $\underline{\phi}$. We state all the results concerning the extended linear-quadratic problem in the following theorem.

2.3.5.1 Extended Linear-quadratic Problem

Theorem 2.8

Let us consider the problem:

$$\text{minimize} \quad J = \sum_{i=1}^{N-1} J_i(\underline{z}(i), \underline{v}(i)), \tag{2.126}$$

over $\quad \underline{v}(1),\ldots\underline{v}(N-1)$,

$$\text{subject to:} \quad \underline{z}(i+1) = \underline{A}(i)\underline{z}(i) + \underline{B}(i)\underline{v}(i) \quad (i=1,\ldots N-1), \tag{2.127}$$

where

$$J_i(\underline{z}(i),\underline{v}(i)) = \frac{1}{2}[\underline{z}^T(i)\underline{Q}(i)\underline{z}(i)+2\underline{v}^T(i)\underline{M}(i)\underline{z}(i)+\underline{v}^T(i)\underline{R}(i)\underline{v}(i)]$$

$$+ [F\underline{p}^T(i)\underline{z}(i)+F\underline{q}^T(i)\underline{v}(i)+F^2 r(i)], \qquad (2.128)$$

$\underline{z}(i)$ is an (n_i-1)-dimensional state vector, $\underline{v}(i)$ is an (m_i-n_i)-dimensional control vector, and all the matrix and vector coefficients have the corresponding dimensions. The factor F is a nonzero scalar measured in the same units as $\underline{z}$ and $\underline{v}$, so that all coefficients: $\underline{Q}(i)$, $\underline{M}(i)$, $\underline{R}(i)$, $\underline{p}(i)$, $\underline{q}(i)$ and $r(i)$ are measured in units of cost per squared unit of $\underline{z}$.

Assume the following: Assumption A_1

$$\underline{Q}(i) \geq 0;\ \underline{R}(i) > 0;\ \underline{Q}(i) - \underline{M}^T(i)\underline{R}^{-1}(i)\underline{M}(i) \geq 0. \qquad (2.129)^{(*)}$$

Define the value functions $V(\cdot,k)$, $k = 1,\ldots N-1$ by

$$V(\underline{z}(k),k) = \min_{\underline{v}} \sum_{i=k}^{N-1} J_i(\underline{z}(i),\underline{v}(i)), \qquad (2.130)$$

subject to: $\underline{z}(i+1) = \underline{A}(i)\underline{z}(i) + \underline{B}(i)\underline{v}(i)$, $(i=k,k+1,\ldots N-1)$.

Then

1) The minimum value of J is equal to $V(\underline{z}(1),1)$.

The sequence of value functions is given by

$$V(\underline{z}(i),i) = \frac{1}{2}\underline{z}^T(i)\underline{K}(i)\underline{z}(i) + F\underline{\ell}(i)\underline{z}(i) + F^2 m(i), \qquad (2.131)$$

where $\underline{K}(i)$ is a positive semi-definite $(n_i-1)\times(n_i-1)$ matrix, $\underline{\ell}(i)$ is an (n_i-1)-dimensional vector, and $m(i)$ is a scalar all measured in cost per squared unit of $\underline{z}$.

The coefficients are recursively obtained from

$$\underline{K}(N) = 0;\ \underline{\ell}(N) = 0;\ m(N) = 0, \qquad (2.132)$$

and

*For any matrix $\underline{X}$, $\underline{X} \geq 0$ means that $\underline{X}$ is positive semi-definite, and $\underline{X} > 0$ means that $\underline{X}$ is positive definite.

$$\underline{K}(i) = \underline{Q}(i) + \underline{A}(i)^T\underline{K}(i+1)\underline{A}(i) - [\underline{B}^T(i)\underline{K}(i+1)\underline{A}(i) + \underline{M}(i)]^T\cdot$$

$$\cdot[\underline{R}(i) + \underline{B}^T(i)\underline{K}(i+1)\underline{B}(i)]^{-1}[\underline{B}^T(i)\underline{K}(i+1)\underline{A}(i) + \underline{M}(i)], \tag{2.133}$$

$$\underline{\ell}(i) = \underline{A}^T(i)\underline{\ell}(i+1) + \underline{p}(i) - [\underline{M}^T(i) + \underline{A}^T(i)\underline{K}(i+1)\underline{B}(i)]\cdot$$

$$\cdot[\underline{R}(i) + \underline{B}^T(i)\underline{K}(i+1)\underline{B}(i)]^{-1}[\underline{q}(i) + \underline{B}(i)^T\underline{\ell}(i+1)], \tag{2.134}$$

$$m(i) = m(i+1) + r(i) - \frac{1}{2}[\underline{q}(i) + \underline{B}^T(i)\underline{\ell}(i+1)]^T\cdot$$

$$\cdot[\underline{R}(i) + \underline{B}^T(i)\underline{K}(i+1)\underline{B}(i)]^{-1}[\underline{q}(i) + \underline{B}^T(i)\underline{\ell}(i+1)]. \tag{2.135}$$

2) Consider the case of <u>stationary coefficients</u> (i.e., n_i, m_i, $\underline{A}(i)$, $\underline{B}(i)$, $\underline{Q}(i)$, $\underline{M}(i)$, $\underline{R}(i)$, $\underline{p}(i)$, $\underline{q}(i)$, $r(i)$ are independent of i). Then

a) If the system (2.127) is controllable, the sequence of matrices $\underline{K}(i)$ converges to a limit matrix $\hat{\underline{K}}$, which is at least positive semi-definite, and a solution of the algebraic Riccati equation

$$\underline{K} = \underline{A}^T\underline{K}\ \underline{A} + \underline{Q} - (\underline{B}^T\underline{K}\ \underline{A} + \underline{M})^T(\underline{R} + \underline{B}^T\underline{K}\ \underline{B})^{-1}(\underline{B}^T\underline{K}\ \underline{A} + \underline{M}). \tag{2.136}$$

b) If the matrix

$$\underline{\mathcal{A}} \triangleq \underline{A}^T - (\underline{M}^T + \underline{A}^T\hat{\underline{K}}\ \underline{B})(\underline{R} + \underline{B}^T\hat{\underline{K}}\ \underline{B})^{-1}\underline{B}^T, \tag{2.137}$$

is stable(*), then the sequence $\underline{\ell}(i)$ converges to a limit $\hat{\underline{\ell}}$ given by:

$$\hat{\underline{\ell}} = (I - \underline{\mathcal{A}})^{-1}\underline{\mathcal{B}}, \tag{2.138}$$

where I is the (n-1)x(n-1) identity, and

$$\underline{\mathcal{B}} \triangleq \underline{p} - (\underline{M}^T + \underline{A}^T\hat{\underline{K}}\ \underline{B})(\underline{R} + \underline{B}^T\hat{\underline{K}}\ \underline{B})^{-1}\underline{q}. \tag{2.139}$$

c) Under the conditions of (a) and (b) (i.e., controllability and stability of $\underline{\mathcal{A}}$), the sequence of differences m(i) - m(i+1) converges to a constant $\Delta\hat{m}$ given by

$$\Delta\hat{m} = r - \frac{1}{2}(\underline{q} + F\ \underline{B}^T\hat{\underline{\ell}})^T(\underline{R} + \underline{B}^T\hat{\underline{K}}\ \underline{B})^{-1}(\underline{q} + F\ \underline{B}^T\hat{\underline{\ell}}). \tag{2.140}$$

---

*A matrix is said to be stable if all its eigenvalues have magnitudes strictly less than 1.

Proof of Theorem 2.8

Proof of Part 1. We shall apply dynamic programming. Bellman's functional equation is:

$$V(\underline{z}(i),\ i) = \min_{\underline{v}(i)} \ [J_i(\underline{z}(i), v(i)) + V(\underline{z}(i+1), (i+1))], \qquad (2.141)$$

subject to: $\underline{z}(i+1) = \underline{A}(i)\ \underline{z}(i) + \underline{B}(i)\ \underline{v}(i)$.

Let us prove (2.131) by induction. Clearly, equation (2.131) holds for $i = N$. Let us prove it for general i by construction.

Assuming equation (2.131) holds for $j = N, N-1, \ldots i+1$, with $\underline{K}(j) \geq 0$, equation (2.141) can be written:

$$V(\underline{z}(i),\ i) = \min_{\underline{v}(i)} \ [\tfrac{1}{2}\ (\underline{z}(i)^T \underline{Q}(i) \underline{z}(i) + 2\underline{v}(i)^T \underline{M}(i) \underline{z}(i)$$

$$+ \underline{v}(i)^T \underline{R}(i)\ \underline{v}(i) + F\ \underline{p}(i)^T \underline{z}(i) + F\ \underline{q}(i)^T \underline{v}(i)$$

$$+ F^2 r(i) + \tfrac{1}{2}\ \underline{z}^T(i+1)\ \underline{K}(i+1)\ \underline{z}(i+1)$$

$$+ F\ \underline{\ell}^T(i+1)\ \underline{z}(i+1) + F^2 m(i+1)], \qquad (2.142)$$

where $\underline{z}(i+1)$ is given in terms of $\underline{z}(i)$ and $\underline{v}(i)$ by (2.127).

Substituting the right-hand side of (2.127) for $\underline{z}(i+1)$ in (2.142), and equating the derivative of the quantity to be minimized to zero, we have:

$$\underline{M}(i)\ \underline{z}(i) + \underline{R}(i)\ \underline{v}(i) + F\ \underline{q}(i) + \underline{B}(i)^T \underline{K}(i+1) \underline{A}(i)\ \underline{z}(i)$$

$$+ \underline{B}(i)^T \underline{K}(i+1)\ \underline{B}(i) \underline{v}(i) + F\ \underline{B}(i)^T \underline{\ell}(i+1) = 0. \qquad (2.143)$$

The second derivative matrix of the right-hand side of (2.143) with respect to $\underline{v}(i)$ is

$$\underline{R}(i) + \underline{B}(i)^T \underline{K}(i+1)\ \underline{B}(i).$$

This matrix is positive-definite since, by assumption A_1, $\underline{R}(i) > 0$, and also $\underline{K}(i+1) \geq 0$ by the induction hypothesis. Therefore, $\underline{v}^*(i)$, the solution of (2.143), minimizes the right-hand side of (2.141), and is given by:

$$\underline{v}^*(i) = -[\underline{R}(i) + \underline{B}^T(i)\ \underline{K}(i+1)\ \underline{B}(i)]^{-1}\{[(\underline{M}(i) + \underline{B}^T(i)\underline{K}(i+1)\underline{A}(i)]z(i)$$
$$+ F\ \underline{q}(i) + F\ \underline{B}^T(i)\ \underline{\ell}(i+1)\}, \tag{2.144}$$

where the inversion indicated is justified by the positive-definiteness properties.

Substituting the expression (2.144) just found for $\underline{v}^*(i)$ in (2.142) yields the minimum $V(\underline{z}(i), i)$. One then sees that

$$V(\underline{z}(i), i) = \frac{1}{2}\underline{z}^T(i)\underline{K}(i)\underline{z}(i) + F\ \underline{\ell}(i)^T\underline{z}(i) + F^2 m(i),$$

with $\underline{K}(i)$, $\underline{\ell}(i)$, and $m(i)$ given by (2.133), (2.134), (2.135), respectively, in terms of the same coefficients at stage (i+1). Using assumption A_1 (equations 2.129), it is possible to show from (2.133) that $\underline{K}(i)$ is positive semi-definite if $\underline{K}(i+1)$ is [21].

<u>Proof of Part 2.</u>

a) In the case when $\underline{p} = 0$ and $\underline{q} = 0$, Dorato and Levis show [21] that if the system (2.127) is controllable, the sequence $\underline{K}(i)$ converges to a matrix $\hat{\underline{K}} \geq 0$ solution of (2.136). The recursive equations (2.133) for $\underline{K}(i)$ are the same as in the true linear-quadratic problem (when $\underline{p} = 0$ and $\underline{q} = 0$). Therefore, the result still holds in our problem.

b) The sequence $\underline{\ell}(i)$ is given by linear recursive equations:

$$\underline{\ell}(i) = \underline{\mathcal{A}}_i\ \underline{\ell}(i+1) + \underline{\mathcal{B}}_i,$$

with
$$\underline{\mathcal{A}}_i = \underline{A}^T - (\underline{M}^T + \underline{A}^T\underline{K}(i+1)\underline{B})(\underline{R} + \underline{B}^T\underline{K}(i+1)\underline{B})^{-1}\underline{B}^T,$$

$$\underline{\mathcal{B}}_i = \underline{p} - (\underline{M}^T + \underline{A}^T\underline{K}(i+1)\underline{B})(\underline{R} + \underline{B}^T\underline{K}(i+1)\underline{B})^{-1}\underline{q}.$$

From the convergence of $\underline{K}(i)$, it follows that $\lim_{i\to-\infty} \underline{\mathcal{A}}_i = \underline{\mathcal{A}}$ given by (2.137), and $\lim_{i\to-\infty} \underline{\mathcal{B}}_i = \underline{\mathcal{B}}$ given by (2.139). Therefore if $\underline{\mathcal{A}}$ is stable, $\underline{\ell}_i$ converges, and the limit satisfies: $\hat{\underline{\ell}} = \underline{\mathcal{A}}\ \hat{\underline{\ell}} + \underline{\mathcal{B}}$ (Appendix C). Part (c) is an immediate corollary of parts (a) and (b), given equation (2.135). Q.E.D.

2.3.5.2 Application to the Original Cost Function

The property which enables us to translate the results of the previous section (2.3.5.1) into the earlier variables $\underline{x}$ of section 2.2, is the equality of the value function expressed in terms of $\underline{x}$ as given in (2.2) to its expression in the reduced state $\underline{z}$ (or $\underline{z}'$) given by (2.131). Indeed, the change of variables performed in section 2.3.3 is merely a reformulation of the optimization problem presented in section 2.2, and does not alter the minimum value or its properties.

This allows us to identify the two expressions of the value function and to derive the asymptotic behavior of the quadratic terms of the value function expressed in $\underline{x}$ (i.e., the matrix $\underline{C}(i)$ of section 2.2). We are interested in the quadratic term, and not in the linear and constant terms $\underline{b}_i^T\underline{x}$ and a_i for the following reason.

The quantities which yield the perturbations are the propagation matrices $\underline{D}(i)$ (equations (2.59) and (2.60)). They depend only on the quadratic coefficients $\underline{L}(i)$ in the cost function (2.31) expressed in $\underline{x}$ and not on the coefficients $\underline{h}(i)$ of the linear terms. Therefore, we may set $\underline{h}(i) = 0$ for all i to obtain the sequence $\underline{D}(i)$, and this implies (through equations (2.40) and (2.42)) that $\underline{b}(i) = 0$ and $a(i) = 0$, for all i, so that the value function becomes:

$$V_i(\underline{x}) = \frac{1}{2}\,\underline{x}^T C(i)\,\underline{x}. \tag{2.145}$$

We summarize the results on the asymptotic behavior of the $\underline{C}(i)$ matrices in the following theorems.

<u>Lemma 2.9</u> Let $\underline{A}(i)$, $\underline{B}(i)$, $\underline{Q}(i)$, $\underline{M}(i)$, $\underline{R}(i)$ be defined by equations (2.87), (2.88), (2.89), (2.90), and (2.91) of section 2.3.3, respectively.

Let

$$\underline{p} = (\underline{L}_{11})_{1n} - (\underline{L}_{11})_{nn}\underline{\nu}_{n-1}, \tag{2.146}$$

$$\underline{q} = \underline{\Lambda}_n^T, \tag{2.147}$$

$$r = \frac{1}{2}\,(L_{11})_{nn}, \tag{2.148}$$

where the argument indicating the stage has been omitted for clarity. The partitions of the $\underline{L}$ matrix are defined by (2.83), and $\underline{\Lambda}$ by (2.85), (2.86). Let us partition the matrix $\underline{C}(i)$ into the submatrices

$$\underline{C}(i) = \begin{pmatrix} \underline{C}_{11}(i) & \underline{C}_{12}(i) \\ \underline{C}_{21}(i) & \underline{C}_{22}(i) \end{pmatrix}, \tag{2.149}$$

where $\underline{C}_{11}(i)$ is $(n_i-1)x(n_i-1)$, $\underline{C}_{12}(i)$ is $(n_i-1)x1$, $\underline{C}_{21}(i) = \underline{C}_{12}^T(i)$, and $C_{22}(i)$ is scalar.

The recursive equations (2.132), (2.133), (2.134), and (2.135) of Theorem 2.8 are satisfied by the sequences $\underline{K}(i)$, $\underline{\ell}(i)$, m(i), where:

$$\underline{K}(i) = \underline{C}_{11}(i) - \underline{C}_{12}(i)\underline{\nu}^T_{n_i-1} + \underline{\nu}_{n_i-1}\underline{C}_{21}(i) + C_{22}(i)\underline{\nu}_{n_i-1}\nu^T_{n_i-1}, \tag{2.150}$$

$$\underline{\ell}(i) = \underline{C}_{12}(i) - \underline{\nu}_{n_i-1}C_{22}(i), \tag{2.151}$$

$$m(i) = \frac{C_{22}(i)}{2}. \tag{2.152}$$

Proof of Lemma 2.9

Let us consider the cost function when $\underline{h}(i) = 0$ for all i. When expressed in terms of $\underline{z}$ and $\underline{v}$ (as in section 2.3.3), it is apparent that it is put in the format studied in theorem 2.8 with the coefficients $\underline{A}(i)$, $\underline{B}(i)$, $\underline{Q}(i)$, $\underline{M}(i)$, $\underline{R}(i)$, $\underline{p}(i)$, $\underline{q}(i)$, and r(i) given in this lemma.

The matrix $\underline{R}(i)$ is positive definite. Indeed, from (2.91),

$$\underline{R} = \underline{\tilde{Y}}^T\underline{L}_{11}\underline{Y} - \underline{\tilde{Y}}^T\underline{L}_{12} - L_{21}\tilde{Y} + \underline{L}_{22}.$$

Thus, let $\underline{v} \in R^{m-n}$ and define $\underline{\phi}^T = (-\underline{\tilde{Y}}\ \underline{v},\ \underline{v}) \in R^m$ as in (2.72) with $\underline{u} = -\underline{\tilde{Y}v}$. Then $\underline{v}^T\underline{R}\ \underline{v} = \underline{\phi}^T\underline{L}\ \underline{\phi} \geq 0$, and, if $\underline{v}^T\underline{R}\ \underline{v} = 0$, then $\underline{\phi} = 0$ because $\underline{L} > 0$, so that $\underline{v} = 0$. Therefore, $\underline{R} > 0$. Let us prove by induction that the value function, in the $\underline{z}$ variables, is given by (2.130) with $\underline{K}(i)$, $\underline{\ell}(i)$, and m(i) defined by (2.150), (2.151), (2.152). It is clearly true at stage i=N since $\underline{K}(N) = 0$ and $\underline{C}(N) = 0$. If it is true for stages j=N, N-1,...i+1, then $\underline{K}(i) > 0$ since to any $\underline{z} \in R^{n_i-1}$, one can associate an $\underline{x}^T = (\underline{z};\ -\underline{\nu}^T_{n-1}\underline{z}) \in R^n$, such that

$\underline{x}^T\ \underline{C}(i)\ \underline{x} = \underline{z}^T\ \underline{K}(i)\ \underline{z}$ with $\underline{K}(i)$ defined by (2.150), and we know from section 2.2 that $\underline{C}(i) > 0$.

Therefore, the matrix $\underline{R}(i) + \underline{B}^T(i)\ \underline{K}(i+1)\underline{B}(i)$ is positive definite, hence invertible, and we may apply theorem 2.8 to find that (2.131) holds and $\underline{K}(i)$, $\underline{\ell}(i)$, m(i) are derived from the same coefficients at stage (i+1) by (2.133), (2.134), (2.135).

Thus, we can identify the two expressions of the value function at stage i, using the partition (2.149) of $\underline{C}(i)$ and equation (2.101) expressing $x_n(i)$ in terms of F and $\underline{z}(i)$.

$$\begin{aligned} V_i(\underline{x}) &\triangleq \frac{1}{2}\,[\underline{z}^T(\underline{C}_{11}(i) - 2\,\underline{C}_{12}(i)\,\underline{\nu}^T_{n_i-1} + C_{22}(i)\,\underline{\nu}_{n_i-1}\underline{\nu}^T_{n_i-1})\underline{z} \\ &\quad + 2\,F(\underline{C}^T_{12}(i) - C_{22}(i)\,\underline{\nu}^T_{n_i-1})\underline{z} + C_{22}(i)F^2] \\ &\triangleq \frac{1}{2}\,\underline{z}^T\underline{K}(i)\underline{z} + F\,\underline{\ell}^T(i)\,\underline{z} + m(i)F^2, \end{aligned}$$

which yields (2.150), (2.151), and (2.152). Q.E.D.

The lemma we have just proved enables us now to characterize the asymptotic behavior of the quadratic term in the value function. This is done in the following theorem.

Theorem 2.10

Assume the network is stationary; that is: n_i, m_i, $\underline{L}_i$, $\underline{Y}_i$, $\underline{Z}_i$ do not depend on the stage parameter i.

1) If the reduced system in $\underline{z}$ is controllable, then the sequence $\underline{K}(i)$ defined by (2.133) converges to a positive semi-definite matrix $\underline{K}$ solution to the algebraic Riccati equation (2.136) (with the coefficients given as in lemma 2.9).

2) If in addition to controllability, the matrix $\underline{A}$ (of equation (2.137)) is stable, then

a) The sequence $\underline{\ell}(i)$ defined by (2.151) converges to a limit $\hat{\underline{\ell}}$, and

$$\hat{\underline{\ell}} = (I - \underline{A})^{-1}\hat{\underline{B}}, \tag{2.153}$$

where

$$\underline{\hat{B}} = \underline{p} - (\underline{M}^T + \underline{A}^T\underline{\hat{K}}\ \underline{B})(\underline{R} + \underline{B}^T\underline{\hat{K}}\ \underline{B})^{-1}\underline{q}. \tag{2.154}$$

b) The sequence m(i) defined by (2.152) is such that

$$\lim_{i\to-\infty} (m(i) - m(i+1)) = \frac{\alpha}{2}, \tag{2.155}$$

and

$$\lim_{i\to-\infty} (\underline{C}(i) - \underline{C}(i+1)) = \alpha\, \underline{\nu}_n \underline{\nu}_n^T, \tag{2.156}$$

where

$$\alpha = 2\, r - (\underline{q} + \underline{B}^T\ \underline{\hat{\ell}})^T(\underline{R} + \underline{B}^T\underline{\hat{K}}\ \underline{B})^{-1}(\underline{q} + \underline{B}^T\underline{\hat{\ell}}). \tag{2.157}$$

<u>Proof of Theorem 2.10</u>

1) According to theorem (2.18), controllability implies the convergence of $\underline{K}(i)$ to a positive semi-definite solution of the Riccati equation (2.136). According to lemma (2.9), $\underline{K}(i)$ is given by (2.150).

a) The stability of $\underline{A}$ implies the convergence of $\underline{\ell}(i)$ to $\underline{\hat{\ell}}$ given by (2.138) with $\underline{B}$ as in (2.139) (theorem 2.8).

b) Equation (2.134) then implies that $(m_i - m_{i+1})$ converges to $\alpha/2$ where α is given by (2.157). Equations (2.150), (2.151), and (2.152) can be solved for $\underline{C}(i)$ and yield:

$$C_{22}(i) = 2\, m(i),$$

$$\underline{C}_{12}(i) = \underline{\ell}(i) + 2\, m(i)\, \underline{\nu}_{n-1},$$

$$\underline{C}_{11}(i) = \underline{K}(i) + \underline{\ell}(i)\underline{\nu}_{n-1}^T + \underline{\nu}_{n-1}\underline{\ell}(i)^T + 2m(i)\underline{\nu}_{n-1}\underline{\nu}_{n-1}^T.$$

Subtracting $\underline{C}(i+1)$ from $\underline{C}(i)$ and taking into account parts 1 and (a), one obtains (2.156) since

$$C_{22}(k) - C_{22}(i+1) \to 2\,\Delta\hat{m},$$

$$\underline{C}_{12}(i) - \underline{C}_{12}(i+1) \to 2\,\Delta\hat{m}\, \underline{\nu}_{n-1},$$

$$\underline{C}_{11}(i) - \underline{C}_{11}(i+1) \to 2\,\Delta\hat{m}\, \underline{\nu}_{n-1}\nu_{n-1}^T,$$

and this implies (2.156) where $\alpha = 2\,\Delta\hat{m}$. Q.E.D.

<u>Remark.</u> The proof relies on the property that $\underline{A}$, given by (2.137), is stable. That will be shown independently in section 2.3.7, assuming only the con-

trollability of the subsystem in $\underline{z}$. We shall thus have proved theorem 2.10 under the sole controllability assumption.

Physical interpretation of equation (2.156)

Equation (2.156) is a proof of one feature of the asymptotic behavior of $\underline{C}$ which has been observed numerically (see 2.3.2). It shows that, for (N-i) large, $\underline{C}(i)$ grows linearly in the number of subnetworks (N-i):

$$\underline{C}(i) \sim \overline{\underline{C}} + (N-i)\,\alpha\,\underline{\nu}_n\underline{\nu}_n^T,$$

with some fixed finite positive definite matrix $\overline{\underline{C}}$. Thus,

$$\underline{x}^T(i)\underline{C}(i)\ \underline{x}(i) \sim \underline{x}^T(i)\overline{\underline{C}}\ \underline{x}(i) + \alpha(N-i)F^2,$$

for (N-i) large since

$$\underline{x}^T(i)\ \underline{\nu}_n\underline{\nu}_n^T\ \underline{x}(i) = F^2.$$

Therefore, as far as the quadratic part of the cost is concerned, adding one more subnetwork adds a term αF^2 to the total cost; n more subnetworks adds $n\alpha F^2$ to the cost. The rate of increase α is given by (2.157) which can be obtained by solving an algebraic Riccati equation (2.136).

Equation (2.156) implies that $\underline{x}(i)^T\underline{C}(i)\underline{x}(i)$ goes to infinity for $(N-i) \to \infty$ if $F \neq 0$. This means that the cost to travel through an infinitely long network is infinite which is to be expected. Notice, however, that if F=0, the cost of a perturbation $\underline{x}(i)$ goes to a finite limit.

We shall see in section 2.3.7 that, after a long time, flow rearranges itself, so that the same distribution is repeated. The only important parameter is then F, and one may solve a minimum cost problem on one single subnetwork under the constraint $\underline{\nu}_n^T\underline{x} = F$, to obtain α. This leads to a variational characterization of α, given in chapter 4.

2.3.5.3 Example

We illustrate the theory of this section (2.3.5) on the standard two-dimensional example already considered several times (2.3.2, 2.3.3, 2.3.4); that is, the subnetwork of Fig. (7.1). We have already calculated the coefficients $\underline{A}$, $\underline{B}$, $\underline{Q}$, $\underline{R}$, $\underline{M}$, $\underline{p}$, $\underline{q}$, r in 2.3.3.

To obtain the Riccati equation (2.136), which is here a scalar second-degree algebraic equation, we need the following expressions:

$$\underline{B}' \, K \, A = \begin{pmatrix} -K \\ K \end{pmatrix},$$

$$\underline{B}' \, K \, A + \underline{M} = \begin{pmatrix} -(K+L_1) \\ K + L_2 \end{pmatrix},$$

$$\underline{B}' \, K \, \underline{B} = \begin{pmatrix} K & -K \\ -K & K \end{pmatrix},$$

$$\underline{R} + \underline{B}' \, K \, \underline{B} = \begin{pmatrix} (L_1+L_3) + K & -K \\ -K & (L_2+L_4) + K \end{pmatrix},$$

$$(\underline{R} + \underline{B}' \, K \, \underline{B})^{-1} = [(L_1+L_2+L_3+L_4)K + (L_1+L_3)(L_2+L_4)] \begin{bmatrix} (L_2+L_4)+K & K \\ K & (L_1+L_3)+K \end{bmatrix}$$

corresponding to

$$\underline{L} = \begin{pmatrix} L_1 & 0 & 0 & 0 \\ 0 & L_2 & 0 & 0 \\ 0 & 0 & L_3 & 0 \\ 0 & 0 & 0 & L_4 \end{pmatrix},$$

and to the labeling of the links and nodes shown on Fig. (2.1).

The Riccati equation (2.136) becomes, in this case:

$$K = K + (L_1+L_2) - [-(K+L_1)(K+L_2)] \begin{bmatrix} (L_2+L_4) + K & K \\ K & (L_1+L_3) + K \end{bmatrix} \cdot$$

$$\begin{bmatrix} -(K+L_1) \\ (K+L_2) \end{bmatrix} [(L_1+L_2+L_3+L_4) + (L_1+L_3)(L_2+L_4)]^{-1}. \qquad (2.158)$$

Let us consider the numerical case when $L_1 = L_2 = 1$ and $L_3 = L_4 = 2$. Equation (2.158) then becomes:

$$K = K + 2 - \frac{1}{6K+9}\,[-(K+1)(K+1)]\begin{bmatrix} K+3 & K \\ K & K+3 \end{bmatrix}\begin{bmatrix} -(K+1) \\ (K+1) \end{bmatrix},$$

which yields $K^2 = 2$.

Since $\hat{K} \geq 0$, the solution is $\hat{K} = \sqrt{2}$. This is precisely the value found by computer implementation of equations (2.39) at stage i=20, which confirms equation (2.150) of lemma 2.9 in this particular example.

Now we can compute $\hat{\underline{\ell}}$ by (2.138). For that purpose, we need

$$\underline{A} = \underline{A}^T - (\underline{M}^T + \underline{A}^T\hat{\underline{K}}\,\underline{B})(\underline{R} + \underline{B}^T\hat{\underline{K}}\,\underline{B})^{-1}\underline{B}^T,$$

and

$$\underline{B} = \underline{p} - (\underline{M}^T + \underline{A}^T\hat{\underline{K}}\,\underline{B})(\underline{R} + \underline{B}^T\hat{\underline{K}}\,\underline{B})^{-1}\underline{q}.$$

In this particular case,

$$A = 1 - [-(K+L_1)(K+L_2)]\begin{bmatrix} (L_2+L_4) + K & +K \\ +K & (L_1+L_3) + K \end{bmatrix}\begin{bmatrix} -1 \\ 1 \end{bmatrix}\cdot$$

$$\cdot\,[(L_1+L_2+L_3+L_4)K + (L_1+L_3)(L_2+L_4)]^{-1},$$

and

$$B = -L_2 - [L_1+L_2+L_3+L_4)K + (L_1+L_3)(L_2+L_4)]^{-1}[-(K+L_1)(K+L_2)]\cdot$$

$$\cdot\begin{bmatrix} (L_2+L_4) + K & K \\ K & (L_1+L_3) + K \end{bmatrix}\begin{bmatrix} 0 \\ -L_2 \end{bmatrix},$$

In the numerical example $L_1 = L_2 = 1$ and $L_3 = L_4 = 2$, equation (2.138) becomes

$$\ell = \ell - 1 - [-(1+\sqrt{2})(1+\sqrt{2})]\begin{bmatrix} (\sqrt{2}+3) & \sqrt{2} \\ \sqrt{2} & (\sqrt{2}+3) \end{bmatrix}\begin{bmatrix} -\ell \\ \ell-1 \end{bmatrix},$$

which yields $\hat{\ell} = -\dfrac{\sqrt{2}}{2}$.

Now we can substitute the values found for $\hat{K}$ and $\hat{\ell}$ in the right-hand side of (2.157) to obtain α.

$$\alpha = 1 - \left(\frac{\sqrt{2}}{2}\right)^2 \; [1 - (1+\sqrt{2})] \begin{bmatrix} \sqrt{2}+3 & \sqrt{2} \\ \sqrt{2} & \sqrt{2}+3 \end{bmatrix} \begin{bmatrix} 1 \\ -(1+\sqrt{2} \end{bmatrix} \frac{1}{6\sqrt{2}+9} = \frac{1}{3} \, .$$

It has been indeed observed numerically (see section 7) that, in this example,

$$\underline{C}(i) - \underline{C}(i+1) \rightarrow \begin{bmatrix} 1/3 & 1/3 \\ 1/3 & 1/3 \end{bmatrix} .$$

2.3.6 Propagation Matrix and Its Asymptotic Behavior

Our main concern is the sequence of propagation matrices $\underline{D}(k)$ since it yields the perturbations recursively according to equation (2.59). Here, we relate $\underline{D}(k)$ to the coefficients of the value function expressed in the new variables $\underline{z}$ and $\underline{v}$; i.e., we express $\underline{D}(k)$ as a function of the sequence $\underline{K}(j)$ and $\underline{\ell}(j)$. The asymptotic behavior of $\underline{D}(k)$ (when $k \rightarrow -\infty$) is then derived from the results of section 2.3.5. We state the results in the following theorem.

<u>Theorem 2.11</u>

Consider the case when the dimension is stationary; i.e.: $n_k = n$ for all k.

1) The propagation matrix at stage k, $\underline{D}(k)$, is given in partitioned form by

$$\underline{D}(k) = \left[\begin{array}{c|c} \underline{\Delta}(k) + \underline{\Pi}(k)\, \underline{\nu}^T_{n_k-1} & \underline{\Pi}(k) \\ \hline \underline{\nu}^T_{n_k-1}(\underline{I}-\underline{\Delta}(k)-\underline{\Pi}(k)\, \underline{\nu}^T_{n_k-1}) & 1 - \underline{\nu}^T_{n_k-1}\underline{\Pi}(k) \end{array}\right], \tag{2.159}$$

where $\underline{\Delta}(k)$ is an $(n-1)\mathrm{x}(n-1)$ matrix, $\underline{\Pi}(k)$ is an $(n-1)$-dimensional vector, and $\underline{I}$ is the $(n-1)\mathrm{x}(n-1)$ identity.

<u>For a state reduction in $\underline{z}$,</u>

2) The matrices in $\underline{\Delta}(k)$ and $\underline{\Pi}(k)$ are given by:

$$\underline{\Delta}(k) = \underline{A}(k) - \underline{B}(k)(\underline{R}(k) + \underline{B}^T(k)\, \underline{K}(k+1)\underline{B}(k))^{-1}(\underline{M}(k) + \underline{B}^T(k)\underline{K}(k+1)\underline{A}(k)), \tag{2.160}$$

$$\underline{\Pi}(k) = -\underline{B}(k)(\underline{R}(k) + \underline{B}^T(k)\, \underline{K}(k+1)\underline{B}(k))^{-1}(\, \underline{\Lambda}^T_n(k) + \underline{B}^T(k)\underline{\ell}(k+1)), \tag{2.161}$$

where $\underline{A}(k)$, $\underline{B}(k)$, $\underline{M}(k)$, $\underline{R}(k)$ are the coefficients of the dynamical system and the cost function in the new variables as given in lemma 2.9. The matrix $\underline{K}(k)$ is obtained recursively from (2.132) and (2.133) with $\underline{\ell}(i)$ derived from (2.132) and (2.134). Also,

$$\underline{\Delta}^T(i) = \underline{\mathcal{A}}(i), \tag{2.162}$$

where $\underline{\mathcal{A}}(i)$ is the matrix occurring in (2.134); i.e.,

$$\underline{\ell}(i) = \underline{\mathcal{A}}(i)\underline{\ell}(i+1) + \underline{\mathcal{B}}(i).$$

3) If the dynamical system (2.127) is stationary and controllable, $\underline{\Delta}(k)$ goes to a limit: $\lim_{k\to-\infty} \underline{\Delta}(k) = \underline{\Delta}$, and

$$\underline{\Delta} = \underline{A} - \underline{B}(\underline{R} + \underline{B}^T\hat{\underline{K}}\,\underline{B})^{-1}(\underline{M} + \underline{B}^T\hat{\underline{K}}\,\underline{A}), \tag{2.163}$$

where $\hat{K} = \lim_{k\to-\infty} K(k)$ is a positive semi-definite matrix solution of (2.136).

4) If in addition, the matrix $\underline{\Delta}$ is stable, $\underline{\ell}(k)$ goes to a limit $\hat{\underline{\ell}}$ given by (2.138), and $\underline{\Pi}(k)$ goes to a limit $\underline{\Pi}$.

$$\underline{\Pi} = -\underline{B}(\underline{R} + \underline{B}^T\,\hat{\underline{K}}\,\underline{B})^{-1}(\underline{\Lambda}_n^T + \underline{B}^T\,\hat{\underline{\ell}}). \tag{2.164}$$

Consequently, $\underline{D}(k)$ goes to the corresponding limit: $\lim_{k\to-\infty} \underline{D}(k) = \underline{D}$, and

$$\underline{D} = \left[\begin{array}{c|c} \underline{\Delta} + \underline{\Pi}\,\underline{\nu}_{n-1}^T & \underline{\Pi} \\ \hline \underline{\nu}_{n-1}^T(\underline{I} - \underline{\Delta} - \underline{\Pi}\,\underline{\nu}_{n-1}^T) & 1 - \underline{\nu}_{n-1}^T\,\underline{\Pi} \end{array}\right]. \tag{2.165}$$

<u>For a state reduction in $\underline{z}'$</u>

The results in parts (2,3,4) hold, except that $\underline{\Pi}(k)$ and $\underline{\Pi}$ have to be modified: see appendix B.

Proof of Theorem 2.11

We consider here only the case of a state reduction in $\underline{z}$. Equations (2.60) and (2.62) of section 2.2 show that $\underline{D}(k)$ is independent of the linear coefficients $\underline{h}(i)$ in the cost function (2.31). Therefore, we can restrict ourselves to the case $\underline{h}(i) = 0$ for all i when studying $\underline{D}(k)$. In that case,

$$\underline{x}^*(k+1) = \underline{D}(k)\ \underline{x}^*(k) \tag{2.166}$$

(where * denotes optimality as usual) because

$$\underline{x}^*(k+1) = \underline{Z}(k)\ \underline{\phi}^*(k),$$

and the constant term in the expression (2.43) of $\underline{\phi}^*(k)$ in terms of $\underline{x}^*(k)$ vanishes (since $\underline{b}(i) = 0$ for all i, given 2.40). On the other hand,

$$\underline{z}^*(k+1) = \underline{A}(k)\underline{z}^*(k) + \underline{B}(k)\ \underline{v}^*(k), \tag{2.167}$$

where $\underline{v}^*(k)$, the optimal control at stage k in the reduced system, is given by (2.144). Therefore,

$$\underline{z}^*(k+1) = \underline{\Delta}(k)\ \underline{z}^*(k) + \underline{\Pi}(k)F, \tag{2.168}$$

where F is the total flow; i.e., the sum of the components of $\underline{x}(1)$: further, $\underline{\Delta}(k)$ is some (n-1)x(n-1) matrix, and $\underline{\Pi}(k)$ is some (n-1)-dimensional vector. Both $\underline{\Delta}(k)$ and $\underline{\Pi}(k)$ are pure numbers; i.e., without units of cost, flow, etc. We shall first transform equation (2.167), in the $\underline{z}$, F variables, into an equivalent expression in $\underline{x}$, to derive the relation (2.159) among $\underline{D}(k)$ and $\underline{\Delta}(k)$ and $\underline{\Pi}(k)$. Thereafter, we shall use equation (2.144) for $\underline{v}^*(k)$ to obtain the expressions (2.160) and (2.161) for $\underline{\Delta}(k)$ and $\underline{\Pi}(k)$, respectively.

Proof of Part 1

Replacing $\underline{z}$ and F in terms of $\underline{x}$ by means of (2.71) and (2.101) transforms (2.167) into

$$x_i^*(k+1) = \sum_{j=1}^{n-1} \Delta_{ij}(k)\ x_j^*(k) + \Pi_i(k) \sum_{j=1}^{n} x_j^*(k) \qquad (i=1,\dots,n-1),$$

$$x_n^*(k+1) = F - \sum_{i=1}^{n-1} z_i^*(k+1),$$

$$= \sum_{j=1}^{n} x_j^*(k) - \sum_{i=1}^{n-1} \sum_{j=1}^{n-1} \Delta_{ij}(k) x_j^*(k) - \left[\sum_{i=1}^{n-1} \Pi_i(k)\right] \sum_{j=1}^{n} x_j^*(k),$$

or equivalently,

$$x_i^*(k+1) = \sum_{j=1}^{n-1} [\Delta_{ij}(k) + \Pi_i(k)] x_j^*(k) + \Pi_i(k) x_n^*(k) \quad (i=1,\ldots,n-1) \tag{2.169}$$

$$x_n^*(k+1) = \sum_{j=1}^{n-1} \left[1 - \sum_{s=1}^{n-1} \Delta_{sj}(k) - \sum_{s=1}^{n-1} \Pi_s(k)\right] x_j^*(k) + \left[1 - \sum_{s=1}^{n-1} \Pi_s(k)\right] x_n^*(k). \tag{2.170}$$

Identifying the coefficients of those two equations with those of (2.166) yields:

$$D_{ij}(k) = \Delta_{ij}(k) + \Pi_i(k), \quad i=1,\ldots,(n-1); j-1,\ldots,n-1), \tag{2.171}$$

$$D_{in}(k) = \Pi_i(k), \quad i=1,\ldots,(n-1) \tag{2.172}$$

$$D_{ni}(k) = 1 - \sum_{s=1}^{n-1} \Delta_{sj}(k) - \sum_{s=1}^{n-1} \Pi_s(k), \quad (i=1,\ldots,n-1) \tag{2.173}$$

$$D_{nn}(k) = 1 - \sum_{s=1}^{n-1} \Pi_s(k), \tag{2.174}$$

which, in matrix form, becomes (2.159).

Proof of Part 2

Substituting (2.144) for $\underline{v}^*(k)$ into (2.167) yields:

$$\begin{aligned} \underline{z}^*(k+1) = [&\underline{A}(k) - \underline{B}(k)(\underline{R}(k) + \underline{B}^T(k)\underline{K}(k+1)\underline{B}(k))^{-1}(\underline{M}(k) \\ &+ \underline{B}^T(k)\underline{K}(k+1)\underline{A}(k))]\underline{z}^*(k) \\ &- \underline{B}(k)(\underline{R}(k) + \underline{B}^T(k)\ \underline{K}(k+1)\underline{B}(k))^{-1}[\underline{\Lambda}_n^T(k) + \underline{B}^T(k)\underline{\ell}(k+1)]F, \end{aligned} \tag{2.175}$$

where we have expressed $\underline{q}(k)$ from equation (2.93), with $\underline{h}(i) = 0$ for all i, to make the dependence on F explicit.

Identifying equation (2.175) with (2.168) yields $\underline{\Delta}(k)$ and $\underline{\Pi}(k)$ as given by (2.160) and (2.161), respectively. It is easy to check (2.162) by inspection of equation (2.134).

Proof of Part 3

From theorem 2.8 and lemma 2.9, we know that, if the system is controllable, $\underline{K}(i)$ converges to a limit $\hat{\underline{K}}$, a positive semi-definite solution of (2.136). Therefore, $\underline{\Delta}(i)$ goes to the corresponding limit given by (2.163).

Proof of Part 4

Because of (2.162), $\underline{A} = \underline{\Delta}^T$ where $\underline{A}$ is given by (2.137), and $\underline{A}$ is stable if and only if $\underline{\Delta}$ is. Accordingly, if $\underline{\Delta}$ is stable, theorem 2.8 implies that $\underline{\ell}(i)$ converges to $\underline{\ell}$ given by (2.138), and the sequence $\underline{\Pi}(i)$ converges to $\underline{\Pi}$ given by (2.164).

Remarks

1) The propagation matrices $\underline{D}(k)$ are homogeneous functions of degree zero in the cost matrices $\underline{L}(k)$; i.e., if all the cost matrices $\underline{L}(k)$ are multiplied by the same scalar constant, the matrices $\underline{D}(k)$ remain unchanged.

Indeed, let us notice that $\underline{M}(i)$, $\underline{Q}(i)$ and $\underline{R}(i)$ are homogeneous of degree 1 in the sequence $\underline{L}(k)$. This is easily seen from equations (2.89), (2.90) and (2.91) in the case where $\underline{h}(k) = 0$ for all k. Therefore, $\underline{K}(i)$ is homogeneous of degree 1 in the sequence $\underline{L}(k)$ according to equations (2.133), and equations (2.160) and (2.161) then show that $\underline{\Delta}(k)$ and $\underline{\Pi}(k)$ are homogeneous of degree zero; consequently, from (2.159), so is $\underline{D}(k)$.

2) The entries of every column of $\underline{D}(k)$ add up to 1. This is an immediate consequence of equation (2.159). This property is to be expected because of flow conservation. Indeed, combining the mathematical expression of flow conservation:

$$\underline{\nu}_n^T \underline{x}^*(k+1) = \underline{\nu}_n^T \underline{x}^*(k),$$

with the propagation equation (2.166), yields

$$\underline{\nu}_n^T \underline{D}(k)\underline{x}^*(k) = \underline{\nu}_n^T \underline{x}^*(k). \tag{2.176}$$

However, this equation would hold as well if the optimization were carried over stages k, k+1,...N instead of 1,2,...,N. In that case, the exogeneous parameter would be $x^*(k)$, instead of $x^*(1)$, so that

$$\underline{\nu}_n^T \underline{D}(k) = \underline{\nu}_n^T, \tag{2.177}$$

must hold.

Another way of stating this property, expressed by equation (2.177), is to say that the number 1 is an eigenvalue of $\underline{D}^T(k)$, and $\underline{\nu}_n$ a corresponding eigenvector.

Numerically, we have always observed that, in addition to that property, the entries of $\underline{D}(k)$ are positive. Therefore $\underline{D}^T(k)$ is a stochastic matrix in all numerical examples. If we could establish directly that $\underline{D}^T$ is a stochastic matrix, we would at the same time have shown that its eigenvalues lie within the closed unit circle, and, under certain conditions, within the open unit circle. Therefore, we would immediately be able to make the desired statements on the decrease of perturbations in an infinitely long network.

3) In section 2.3.7, we prove that, if the reduced system is controllable, the matrix $\underline{\Delta}$ is stable. Therefore, theorem 2.11 can be applied under the sole assumption of controllability.

Example

Let us calculate $\underline{\Delta}$ and $\underline{\Pi}$ in the numerical standard two-dimensional example considered earlier. For the example, we have:

$$\Delta = 1 - \frac{1}{6\sqrt{2}+9}\,[-1 \quad 1]\begin{bmatrix}(\sqrt{2}+3) & -\sqrt{2}\\ -\sqrt{2} & (\sqrt{2}+3)\end{bmatrix}\begin{bmatrix}-(1+\sqrt{2})\\ (1+\sqrt{2})\end{bmatrix} = 3 - 2\sqrt{2},$$

$$\Pi = \frac{-1}{6\sqrt{2}+9}\,[-1 \quad 1]\begin{bmatrix}\sqrt{2}+3 & -\sqrt{2}\\ -\sqrt{2} & \sqrt{2}+3\end{bmatrix}\begin{bmatrix}\frac{\sqrt{2}}{2}\\ -(\frac{\sqrt{2}}{2}+1)\end{bmatrix} = \sqrt{2} - 1.$$

Therefore, by (2.165),

$$\underline{D} = \begin{bmatrix} (\Delta + \Pi) & \Pi \\ 1 - (\Delta + \Pi) & 1 - \Pi \end{bmatrix} = \begin{bmatrix} (2 - \sqrt{2}) & (\sqrt{2} - 1) \\ (\sqrt{2} - 1) & (2 - \sqrt{2}) \end{bmatrix}.$$

This is indeed the matrix $\underline{D}(i)$ has been found to converge to by implementing the equations of section 2.2 on a digital computer. For instance, after 19 steps, the result displayed in the output is

$$\underline{D}(i) = \begin{bmatrix} 0.5857864376 & 0.4142135624 \\ 0.4142135624 & 0.5857864376 \end{bmatrix},$$

which is a correct approximation of $\underline{D}$ up to at least 7 decimal places.

2.3.7 Sensitivity Analysis

2.3.7.1 Physical Motivations

We now determine whether zero flow perturbations decrease, and, if so, at what rate. A simpler proof that they decrease (but with no estimate of the rate) is in [42].

The optimal flow perturbations are given recursively by equations (2.59):

$$\underline{x}^*(i+1) = \underline{D}(i)\ \underline{x}^*(i) \qquad i=1,\ldots,N.$$

We have just derived expressions for the propagation matrices $\underline{D}(i)$ and their limit $\underline{D}$ (see section 2.3.6) for $i \to -\infty$. Let us emphasize the physical meaning of that limit by a more complete notation. In a network consisting of N sub-networks, we shall denote by $\underline{D}(k, N)$ the propagation matrix at stage k. This has been previously denoted by merely $\underline{D}(k)$. Correspondingly, we write $\underline{x}^*(k, N)$ instead of $\underline{x}^*(k)$ for the perturbation at stage k in a network of length (N-1). In this notation,

$$\underline{x}^*(i+1,N) = \underline{D}(i,N)\ \underline{x}^*(i,N) \qquad i = 1,\ldots,N. \qquad (2.178)$$

While it is mathematically correct to let the parameter k go to $-\infty$ as we have done in 2.3.6, it is not notationally very satisfying because one wishes to label the subnetworks with positive integers, the most upstream being called 1 and the most downstream (N-1). However, the only physical quantity which matters for the purpose of an asymptotic behavior is the number of subnetworks which separate the k^{th} one from the exit; i.e., (N-k-1) (see Fig. 2.2.2). This is the parameter which shall go to $+\infty$.

Instead of fixing N and letting k go to $-\infty$ (which we do in 2.3.6 when writing $\underline{D} = \lim_{i \to -\infty} \underline{D}(i)$), we can as well think of k as fixed and let N, the "length" of the downstream network, go to $+\infty$. This way of viewing things shows that the physical interpretation of $\underline{D}$ is

$$\underline{D} = \lim_{N \to +\infty} \underline{D}(k, N).$$

This does not depend on k since D(k, N) depends only on N-k. Let us set $\tilde{\underline{D}}(i) = \underline{D}(k, k+i)$. Then $\underline{D} = \lim_{i \to +\infty} \tilde{\underline{D}}(i)$. Therefore, $\underline{D}$ can be viewed as the propagation matrix in an infinitely long network. Letting N go to $+\infty$ in (2.178) yields:

$$\lim_{N \to \infty} \underline{x}^*(i+1, N) = \underline{D} \lim_{N \to \infty} \underline{x}^*(i, N), \tag{2.179}$$

or, denoting $\lim_{N \to \infty} \underline{x}^*(i, N)$ by $x^*(i)$,

$$\underline{x}^*(i+1) = \underline{D}\, \underline{x}^*(i), \tag{2.180}$$

hence, by induction,

$$\underline{x}^*(i) = \underline{D}^i\, \underline{x}, \tag{2.181}$$

where $\underline{x}^*(i)$ now represents the vector of optimal flow perturbations at stage i resulting from a perturbation $\underline{x}$ at stage 1, in an infinitely long network.

Numerical experience with speed of convergence

It has been observed in all numerical experiments that the convergence

$$\underline{D}(k, N) \to D,$$

as $N \to \infty$, is very fast. This is reported in section 7.

For example, in the standard two-dimensional example with the same data as in 2.3.6, it has been found that, with N = 20, $\underline{D}(1, 20) \simeq D(2,20) \simeq \ldots \simeq \underline{D}(11, 20)$ to 14 decimal places. They are equal to $\underline{D}(14, 20)$, $\underline{D}(15, 20)$, $\underline{D}(16, 20)$ to 5 decimal places; and even to $\underline{D}(17, 20)$ and $\underline{D}(18, 20)$ to 2 decimal places. $\underline{D}(19, 20)$ is drastically different from all other matrices.

This is the generally observed behavior: $\underline{D}(N-1, N)$, the propagation matrix for one-step optimization, is very different from all the $\underline{D}(i, N)$, $N-i > 1$; i.e., the propagation matrices for optimization over more than one step. $\underline{D}(N-2, N)$ is already very close to the limit $\lim_{N\to\infty} \underline{D}(i, N) = \underline{D}$. In view of these numerical results, it is very accurate to use $\underline{D}$ instead of $\underline{D}(i, N)$ for $N-i > 10$, say; and even not at all unreasonable to do so as long as $N-i > 1$.

These considerations have to be borne in mind when we use the expression "infinitely long network;" i.e., when we let N go to infinity. It is the rapid convergence of $\underline{D}(i, N)$ which justifies using $\underline{D}$ in finite networks.

2.3.7.2 Mathematical Formulation

The mathematical developments of this section lead to a characterization of the eigenvalues of $\underline{D}$ under unrestrictive assumptions. This characterization enables one to prove the central result: <u>optimal perturbations with zero total flow decrease geometrically downstream in an infinitely long network</u>. We can even say more about general perturbations with nonzero total flow.

In the following theorem, we state what spectral properties of $\underline{D}$ we shall need, and how they will be used in proving our main point. Thereafter, we establish how these spectral properties can be derived from those of $\underline{\Delta}$. Finally, we establish the spectral properties of $\underline{\Delta}$ using controllability and two auxiliary assumptions. We can then sum up the central result, discuss the assumptions on which it rests, and illustrate it on a special class of examples.

<u>Definition 2.10</u>

The <u>algebraic multiplicity</u> of an eigenvalue is its order as a root of the characteristic polynomial. The <u>geometrical multiplicity</u> of an eigenvalue is the dimension of the vector space spanned by the corresponding eigenvectors.

<u>Theorem 2.12</u>

If

1) $\underline{D}$ is invertible.

2) The number 1 is a simple eigenvalue of $\underline{D}$.

3) All other eigenvalues of $\underline{D}$ have magnitudes strictly less than 1.

4) The eigenvalues of $\underline{D}$ all have a geometrical multiplicity equal to their algebraic multiplicity.

Then

1) There exist constant matrices $\underline{B}(1),\ldots,\underline{B}(r)$ ($r \leq n-1$ being the number of distinct eigenvalues other than 1), such that, for any positive integer k,

$$\underline{D}^k = \underline{B}(0) + \sum_{i=1}^{r} \underline{B}(i)(s_i)^k, \tag{2.182}$$

where

$$\underline{B}(0) = \underline{p}\,\underline{\nu}_n^T, \tag{2.183}$$

$s_1,\ldots,s_r$ are the r distinct eigenvalues other than 1, and $\underline{p}$ is that eigenvector corresponding to the eigenvalue 1, uniquely defined by:

$$\underline{D}\,\underline{p} = \underline{p}, \tag{2.184}$$

$$\underline{\nu}_n^T\,\underline{p} = 1. \tag{2.185}$$

2) There exists a constant M, such that, for any vector $\underline{x} \in R^n$ and any positive integer k,

$$||\underline{D}^k\underline{x} - F\,\underline{p}|| \leq M|s|^k||x||, \tag{2.186}$$ *

where

$$|s| = \max_{i=1,\ldots r} |s_i| < 1 \quad . \tag{2.187}$$

and

$$F \overset{\Delta}{=} \underline{\nu}_n^T\,\underline{x}.$$

3) We have the limiting form:

$$\lim_{k\to\infty} \underline{D}^k\underline{x} = F\,\underline{p}. \tag{2.188}$$

4) If all eigenvalues of $\underline{D}$ are simple, then

$$B_{st}(i) = V_{si}\,(V)^{-1}_{it} \tag{2.189}$$

---

* For any vector $\underline{v}$ in R^n, $||\underline{v}||$ is its euclidean norm: $||\underline{v}|| = \sum_{i=1}^{n}|v_i|^2)^{1/2}$.

and $\underline{V}$ is an (nxn) matrix of linearly independent eigenvectors of $\underline{D}$.

Remark

If the vector $\underline{x}$ is such that $\nu_n^T \underline{x} = 0$; i.e., if it describes an initial perturbation which does not alter the total incoming flow, equation (2.186) becomes: $||\underline{D}^k\underline{x}|| \leq M|s|^k||\underline{x}||$ or, according to (2.181),

$$||\underline{x}^*(k)|| \leq M|s|^k||\underline{x}^*(1)||,$$

which shows that the magnitudes of the optimal downstream flow perturbations for an infinitely long corridor network decrease exponentially to zero.

On the other hand, equation (2.186) shows that a general perturbation distributes itself among the entrances according to the distribution $\underline{p}$, which is repeated from one subnetowrk to thenext, according to (2.184).

Proof of Theorem 2.12

Proof of Part 1

See appendix D. We follow there step by step a proof which is usually found in textbooks on Markov chains, but which uses only the assumptions of this theorem.

Proof of Part 2

From equations (2.182) and (2.183),

$$||\underline{D}^k\underline{x} - \underline{p}\,\nu_n^T\,\underline{x}|| = ||\underline{D}^k\underline{x} - F\,\underline{p}|| = ||\sum_{i=1}^{r} s_i^k\underline{B}(i)\underline{x}||$$

$$\leq \sum_{i=1}^{r} |s_i|^k\;||\underline{B}(i)||\cdot||\underline{x}|| \leq r\,B\,|s|^k||x||\;,$$

where $B = \max_{i=1,\ldots r} ||\underline{B}(i)||$ and $|s| = \max_{i=1,\ldots r} |s_i|$, the matrix norm being defined, for any matrix $\underline{X}$, by $||\underline{X}|| = \max_{\text{s.t.}\;||x||=1} ||\underline{X}\;\underline{x}||$, which is equivalent to $||\underline{X}|| = \sqrt{\lambda_{max}}$, where λ_{max} is the largest eigenvalue of $\underline{X}^T\underline{X}$ (see [25]. Thus, equation (2.186) is proved with $M = rB$.

Proof of Part 3

Part 3 is a straightforward consequence of part 2 since $|s| < 1$, so that $\lim_{k\to\infty} ||\underline{D}^k\underline{x} - F\ \underline{p}|| \leq M||\underline{x}|| \lim_{k\to\infty} |s|^k = 0.$ Q.E.D.

Remark

Equation (2.189) makes it possible to compute the constant M, provided that the eigenvectors of $\underline{D}$ are known (in the case of distinct eigenvalues). Indeed, $M = (n-1) \max_{i=1,\ldots(n-1)} ||\underline{B}(i)||$, as shown in the proof of theorem 2.12, since $r = n-1$ in this case. On the other hand, the constant $|s|$ is the largest magnitude of an eigenvalue of $\underline{D}$ other than 1.

In the case of multiple eigenvalues, the $\underline{B}(i)$ can be obtained from the Jordan form.

Relation between the spectral properties of $\underline{D}$ and those of $\underline{\Delta}$

We have shown in theorem 2.11 that our main goal, i.e., obtaining bounds for perturbations, will be reached if some properties of the eigenvalues of $\underline{D}$ can be established. Now we show, in the following lemma, how those properties can be derived from those of $\underline{\Delta}$. Also, we relate the eigenvectors with eigenvalue 1 to the matrices $\underline{\Delta}$ and $\underline{\Pi}$.

Lemma 2.12

1) The following relation holds between the characteristic polynomials of $\underline{D}$ and $\underline{\Delta}$.

$$\det(\underline{D} - s\ \underline{I}_n) = (1-s)\det(\underline{\Delta} - s\ \underline{I}_{n-1}), \tag{2.190}$$

for any complex number s ($\underline{I}_n$ is the nxn identity matrix).

2) If $\underline{z}$ is an eigenvector of $\underline{\Delta}$, then $\underline{x}$, given by

$$\underline{x} = \begin{pmatrix} \underline{z} \\ \\ -\underline{\nu}_{n-1}^T\underline{z} \end{pmatrix}, \tag{2.191}$$

is an eigenvector of $\underline{D}$ with the same eigenvalue.

3) If $\underline{u}$ is an eigenvector of $\underline{D}$ with eigenvalue 1 and $\underline{\nu}_n^T\ \underline{u} = F$, then if $\underline{u}$ is

partitioned into

$$u = \begin{pmatrix} \underline{u}_1 \\ u_2 \end{pmatrix}, \tag{2.192}$$

with $\underline{u}_1 \varepsilon R^{n-1}$,

$$(\underline{I} - \underline{\Delta})\underline{u}_1 = F\,\underline{\Pi}. \tag{2.193}$$

<u>Proof of lemma 2.12</u>

<u>Proof of Part 1</u>

From equation (2.165) (section 2.3.6), it follows that

$$\det(\underline{D}-s\,\underline{I}) = \left[\begin{array}{cccc|c} \Delta_{11}-s+\Pi_1 & \Delta_{12}+\Pi_1 & \cdots & \Delta_{1,n-1}+\Pi_1 & \Pi_1 \\ \Delta_{21}+\Pi_2 & \Delta_{22}-s+\Pi_2 & \cdots & \Delta_{2,n-1}+\Pi_2 & \Pi_2 \\ \vdots & \vdots & & \vdots & \vdots \\ \Delta_{n-1,1}+\Pi_{n-1} & \Delta_{n-1,2}+\Pi_{n-1} & \cdots & \Delta_{n-1,n-1}-s+\Pi_{n-1} & \Pi_{n-1} \\ \hline 1-\sum_{i=1}^{n-1}(\Delta_{i1}+\Pi_i) & 1-\sum_{i=1}^{n-1}(\Delta_{i2}+\Pi_i) & & 1-\sum_{i=1}^{n-1}(\Delta_{i,n-1}+\Pi_i) & (1-s)-\sum_{i=1}^{n-1}\Pi_i \end{array}\right] \tag{2.194}$$

Let us notice that, denoting x_{ij} the (i,j) entry of the above determinant,

$$x_{nj} = (1 - s) - \sum_{i=1}^{n-1} x_{ij}, \quad \text{for } j=1,\ldots,n, \tag{2.195}$$

therefore,

$$\det(\underline{D}-s\,\underline{I}) = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n-1} & x_{1n} \\ \vdots & \vdots & & & \vdots \\ x_{n-1,1} & x_{n-1,2} & \cdots & x_{n-1,n-1} & x_{n-1,n} \\ (1-s) & (1-s) & \cdots & (1-s) & (1-s) \end{bmatrix},$$

or, factoring (1-s) in the bottom row,

$$\det(\underline{D}-s\underline{I}) = (1-s) \left[\begin{array}{cccc|c} (\Delta_{11}-s)+\Pi_1 & \Delta_{12}+\Pi_1 & \cdots & \Delta_{1,n-1}+\Pi_1 & \Pi_1 \\ \Delta_{21}+\Pi_1 & \Delta_{22}-s+\Pi_2 & \cdots & \Delta_{2,n-1}+\Pi_2 & \Pi_2 \\ \vdots & \vdots & & \vdots & \vdots \\ \Delta_{n-1,1}+\Pi_{n-1} & \Delta_{n-1,2}+\Pi_{n-1} & \cdots & \Delta_{n-1,n-1}-s+\Pi_{n-1} & \Pi_{n-1} \\ \hline 1 & 1 & \cdots & 1 & 1 \end{array}\right],$$

$$= (1-s)\det(\underline{\Delta} - s\,\underline{I}),$$

where the last equality is justified because the determinant does not change if the last column is subtracted from all other columns.

<u>Proof of Part 2</u>

Let $\underline{z}$ be such that

$$\underline{\Delta}\,\underline{z} = s\,\underline{z}$$

for some complex number s. Let us define $\underline{x}$ by (2.191). Then, from equation (2.165), it follows that the vector of the (n-1) first components of $\underline{D}\,\underline{x}$ is given by

$$\underline{\Delta}\,\underline{z} + \underline{\Pi}\,\underline{\nu}^T_{n-1}\,\underline{z} - \underline{\Pi}\,\underline{\nu}^T_{n-1}\underline{z} = \underline{\Delta}\,\underline{z} = s\,\underline{z}.$$

On the other hand, $\underline{\nu}^T_n\,\underline{D} = \underline{\nu}^T_n$, so that $\underline{\nu}^T_n\,\underline{D}\,\underline{x} = \underline{\nu}^T_n\,\underline{x} = 0$, and therefore, the n^{th} element of $\underline{D}\,\underline{x}$ is $(Dx)_n = -s\,\underline{\nu}^T_{n-1}\,\underline{z}$. Consequently, $\underline{D}\,\underline{x} = s\,\underline{x}$.

Proof of Part 3

Partitioning $\underline{u}$ according to (2.192) with $\underline{u}_1 \in R^{n-1}$ and $u_2 \in R$ and applying (2.165),

$$\underline{\Delta}\, \underline{u}_1 + \underline{\Pi}\, \underline{\nu}^T_{n-1}\, \underline{u}_1 + \underline{\Pi}\, u_2 = \underline{u}_1,$$

or, since

$$u_2 = F - \underline{\nu}^T_{n-1}\underline{u}_1,$$

$$\underline{\Delta}\, \underline{u}_1 + \underline{\Pi}\, \underline{\nu}^T_{n-1}\underline{u}_1 + \underline{\Pi}(1 - \underline{\nu}^T_{n-1}\underline{u}_1) = \underline{u}_1,$$

or

$$\underline{u}_1 - \underline{\Delta}\, \underline{u}_1 = F\, \underline{\Pi}.$$

Q.E.D.

Remark: Similarity of the propagation matrices corresponding to different state reductions

1) Equation (2.190) of lemma 2.12 is valid whichever labeling of the entrance-exit pairs has been adopted (section 2.3.3). Different labelings lead to different state reductions, and to different matrices $\underline{\Delta}$, but equation (2.190) shows that the eigenvalues of $\underline{\Delta}$ do not depend on the particular state reduction since they are canonically related to those of $\underline{D}$. Therefore, two matrices $\underline{\Delta}$ corresponding to two different state reductions are similar. They represent the same linear mapping in two different coordinate systems.

2) The formulas derived in lemma 2.12 also hold for $\underline{D}(k)$, $\underline{\Delta}(k)$, and $\underline{\Pi}(k)$, for any k and not only to their limiting values. This is because it uses only equation (2.129) which is valid for any k. Remark 1 concerns the various $\underline{\Delta}(k)$ matrices as well. We shall now derive a corollary relating the eigenvalues and eigenvectors of $\underline{D}$ with those of $\underline{\Delta}$.

Corollary 2.13

1) If the number 1 is not an eigenvalue of the matrix $\underline{\Delta}$, then

   a) the number 1 is a simple eigenvalue of $\underline{D}$, and

   b) the sum of the components of an eigenvector corresponding to unit eigenvalue is different from zero, so that it can be normalized to 1, and is then uniquely defined by (2.193) with F=1.

2) If in addition, the eigenvalues of $\underline{\Delta}$ all have a geometrical multiplicity equal to their algebraic multiplicity, then the eigenvectors of $\underline{D}$ with eigenvalues other than 1 are all derived from those of $\underline{\Delta}$ by (2.191).

Proof of Corollary 2.13

1a) Equation (2.190) shows that 1 is a simple eigenvalue of $\underline{D}$ if, and only if, it is not an eigenvalue of $\underline{\Delta}$.

1b) Let $\underline{D}\,\underline{u} = \underline{u}$ with $\underline{u} \neq 0$, and let us prove that $F \stackrel{\Delta}{=} \underline{\nu}_n^T\,\underline{u} \neq 0$ by contradiction. Assume $\underline{\nu}_n^T\underline{u} = 0$. Then $\underline{u}_1 \neq 0$ because if $\underline{u}_1 = 0$ and $\underline{\nu}_n^T \cdot \underline{u} = \underline{\nu}_{n-1}^T\underline{u}_1 + u_2 = 0$, then $u_2 = 0$ and $\underline{u} = 0$. However, equation (2.193), with F = 0, becomes: $(\underline{I} - \underline{\Delta})\underline{u}_1 = 0$ which, together with $\underline{u}_1 \neq 0$, contradicts the fact that 1 is not an eigenvalue of $\underline{\Delta}$.

Therefore, $\underline{\nu}_n^T\underline{u} \neq 0$ and $\underline{u}$ can be normalized by $\underline{\nu}_n^T\underline{u} = 1$. Let us call $\underline{p}$ that normalized eigenvector with eigenvalue 1.

According to equation (2.193) again, $\underline{p}$ is uniquely defined by: $\underline{p}^T = (\underline{p}_1^T, p_2)$,

$$\underline{p}_1 = (\underline{I} - \underline{\Delta})^{-1}\underline{\Pi}, \tag{2.196}$$

and

$$p_2 = 1 - \underline{\nu}_{n-1}^T\,\underline{p}_1. \tag{2.197}$$

2) Equation (2.191) shows that linearly independent eigenvectors all corresponding to the same eigenvalue of $\underline{\Delta}$ (and thus, of $\underline{D}$) give rise to as many linearly independent eigenvectors of $\underline{D}$ corresponding to the same eigenvalue. Moreover, equation (2.190) shows that an eigenvalue of $\underline{\Delta}$ has the same algebraic multiplicity as an eigenvalue of $\underline{D}$, and, therefore, from the assumption of this corollary, it has also the same geometrical multiplicity. Therefore, there may not be eigenvectors of $\underline{D}$ which are not derived from eigenvectors of $\underline{\Delta}$ by (2.191).

Corollary 2.14

If the eigenvalues of $\underline{\Delta}$ are simple and 1 is not among them, then a matrix $\underline{V}$ of linearly independent eigenvectors of $\underline{D}$ can be obtained as follows:

$$\underline{V} = \left[\begin{array}{cccc} \underline{p}_1 & \underline{z}_1 & \cdots & \underline{z}_{n-1} \\ \hline p_2 & -\underline{\nu}_{n-1}^T \underline{z}_1 & & -\underline{\nu}_{n-1}^T \underline{z}_{n-1} \end{array}\right], \quad (2.198)$$

where $\underline{z}_1, \underline{z}_2, \ldots, \underline{z}_{n-1}$ are $(n-1)$ linearly independent eigenvectors of $\underline{\Delta}$, and $\underline{p}_1$ and p_2 are given by (2.196) and (2.197), respectively.

Proof

This is an immediate consequence of corollary 2.13, and the expression of $\underline{D}$ in terms of $\underline{\Delta}$ and $\underline{\Pi}$.

Remark

The matrix $\underline{V}$ given by (2.198) can be used to compute the constant M which occurs in (2.186) to yield bounds for the perturbations. The whole problem of seeking bounds for perturbations is, therefore, completely expressed in terms of $\underline{\Delta}$ and $\underline{\Pi}$.

Preview

Having shown in theorem 2.12 that the perturbation problem in an infinitely long network reduces itself to a question about the eigenvalues of $\underline{D}$, which in turn amounts to characterizing those of $\underline{\Delta}$, we now turn our attention to $\underline{\Delta}$.

According to (2.163), the matrix $\underline{\Delta}$ depends on the data $\underline{L}$, $\underline{Y}$, $\underline{Z}$, and on $\hat{\underline{K}}$. Therefore, any further study of $\underline{\Delta}$ involves an explicit expression for $\hat{\underline{K}}$. Obtaining one is equivalent to solving analytically an algebraic Riccati equation. To do so, we follow a method of Vaughan [22], which is an adaptation of Potter's method to discrete-time systems. It uses the Hamiltonian formulation, and involves only linear operations. This method is particularly attractive because it allows us to derive immediately a simple expression for $\underline{\Delta}$ instead of substituting the value of $\hat{\underline{K}}$ in (2.163). However, Vaughan's method is directly applicable only to cost functions with no linear terms and without cross terms coupling state and control; i.e., to block-diagonal cost functions.

Let us notice though that $\underline{\Delta}$ and $\hat{\underline{K}}$ are the same if linear terms are present or not. As far as the quadratic cross term is concerned, we can define a related system with different state equations, a diagonal cost function, and the same $\underline{\Delta}$ and $\hat{\underline{K}}$ matrices. We can apply Vaughan's method to this system.

Diagonalization of cost function

We state here the result concerning the existence of an equivalent system with a diagonal cost function.

Theorem 2.15

Assume the following:

$$\text{Assumption } A_2 \quad \begin{cases} \underline{Q} \geq 0: \underline{S} \geq 0 \\ \underline{R} > 0 \\ \underline{Q} - \underline{M}^T \underline{R}^{-1} \underline{M} \geq 0 \end{cases} . \tag{2.199}$$

Let Problem 1 be: minimize

$$J_1 = \frac{1}{2} \sum_{k=1}^{N-1} [\underline{z}^T(k), \underline{v}^T(k)] \begin{bmatrix} \underline{Q} & \underline{M}^T \\ \underline{M} & \underline{R} \end{bmatrix} \begin{bmatrix} \underline{z}^T(k) \\ \underline{v}(k) \end{bmatrix} + \frac{1}{2} \underline{z}(N)^T \underline{S} \underline{z}(N) \tag{2.200}$$

over $\underline{v}(1), \underline{v}(2), \ldots, \underline{v}(N-1)$ subject to:

$$\underline{z}(i+1) = \underline{A}\, \underline{z}(i) + \underline{B}\, \underline{v}(i) \qquad (i=1,\ldots,N-1). \tag{2.201}$$

Let Problem 2 be: minimize

$$J_2 = \frac{1}{2} \sum_{k=1}^{N-1} [\underline{z}^T(k), \underline{v}^T(k)] \begin{bmatrix} (\underline{Q} - \underline{M}^T \underline{R}^{-1} \underline{M}) & 0 \\ 0 & \underline{R} \end{bmatrix} \begin{bmatrix} \underline{z}(k) \\ \underline{v}(k) \end{bmatrix} + \frac{1}{2} \underline{z}(N)^T \underline{S} \underline{z}(N) \tag{2.202}$$

over $\underline{v}(1), \underline{v}(2), \ldots, \underline{v}(N-1)$ subject to:

$$\underline{z}(i+1) = (\underline{A} - \underline{B}\, \underline{R}^{-1} \underline{M}) \underline{z}(i) + \underline{B}\, \underline{V}(i) \tag{2.203}$$

Then, Problem 1 and Problem 2 are equivalent in the sense that they have:

a) the same value function

$$V_1(\underline{z}) = \frac{1}{2} \underline{z}^T \underline{K}(i) \underline{z}, \tag{2.204}$$

with $\underline{K}(i)$ given by the recursive equations (2.133), and

$$\underline{K}(N) = \underline{S}, \tag{2.205}$$

b) the same optimal trajectory $\underline{z}^*(i)$ corresponding to the same initial condition

$$\underline{z}(1) = \underline{\xi}. \tag{2.206}$$

Proof

See appendix E. This result is often presented in the continuous-time case [26].

State-costate approach. Vaughan's method and its applications.

It is well known that, given a discrete-time optimal control problem, there exists a sequence $\underline{p}^*(i)$ (i=1,...,N) of vectors of same dimension as the state vectors, called co-states, and whose evolution is described together with that of the states by equations involving the Hamiltonian function. This is usually referred to as the discrete minimum principle [23], [24]. It is of special interest in the linear-quadratic problem as stated in the following proposition.

Proposition 2.16

In the linear-quadratic problem, with state equations

$$\underline{z}^*(i+1) = \underline{\tilde{A}}(i)\underline{z}(i) + \underline{\tilde{B}}(i)\underline{v}(i),$$

and diagonal cost function

$$J = \sum_{i=1}^{N-1} [\underline{z}^T(i) \quad \underline{v}^T(i)] \begin{bmatrix} \underline{\tilde{Q}}(i) & 0 \\ 0 & \underline{\tilde{R}}(i) \end{bmatrix} \begin{bmatrix} \underline{z}(i) \\ \underline{v}(i) \end{bmatrix} + \frac{1}{2} z^T(N) \; \underline{S} \; z(N), \tag{2.207}$$

the discrete minimum principle implies:

If $\begin{cases} \underline{S} \geq 0 \\ \underline{\tilde{Q}}(i) \geq 0; \; \underline{\tilde{R}}(i) > 0, \\ \underline{\tilde{A}}(i) \text{ invertible} \end{cases}$

then:

the optimal state $\underline{z}^*(i)$ and the optimal co-state $\underline{p}^*(i)$ at stage i are related by:

$$\underline{p}^*(i) = \underline{K}(i)\underline{z}^*(i) \qquad i=1,...,N-1, \tag{2.208}$$

where $\underline{K}(i)$ is the matrix defining the value function (2.204) at stage i.

The optimal state-costate pair at stage i is given recursively by:

$$\begin{bmatrix} \underline{z}^*(i) \\ \underline{p}^*(i) \end{bmatrix} = \underline{H}(i) \begin{bmatrix} \underline{z}^*(i+1) \\ \underline{p}^*(i+1) \end{bmatrix} \qquad (i=1,\ldots,N-1), \qquad (2.209)$$

where the Hamiltonian matrix $\underline{H}(i)$ is

$$\underline{H}(i) = \left[\begin{array}{c|c} \underline{\tilde{A}}^{-1}(i) & \underline{\tilde{A}}^{-1}(i)\,\underline{\tilde{B}}(i)\,\underline{\tilde{R}}^{-1}(i)\,\underline{\tilde{B}}^{T}(i) \\ \hline \underline{\tilde{Q}}(i)\underline{\tilde{A}}^{-1}(i) & \underline{\tilde{A}}^{T}(i) + \underline{\tilde{Q}}(i)\underline{A}^{-1}(i)\,\underline{\tilde{B}}(i)\underline{\tilde{R}}^{-1}(i)\underline{\tilde{B}}^{T}(i) \end{array}\right] \qquad (2.210)$$

Proof: see [23].

Remark

We can apply proposition 2.16 not to problem 1, but to problem 2 of theorem 2.15; that is, with

$$\begin{aligned} \underline{\tilde{A}} &= \underline{A} - \underline{B}\,\underline{R}^{-1}\underline{M}, \\ \underline{\tilde{B}} &= \underline{B}, \\ \underline{\tilde{Q}} &= \underline{Q} - \underline{M}^{T}\underline{R}^{-1}\underline{M}, \\ \underline{\tilde{R}} &= \underline{R}, \end{aligned} \qquad (2.211)$$

provided that $\underline{\tilde{A}} = \underline{A} - \underline{B}\,\underline{R}^{-1}\underline{M}$ is invertible (assumption A3).

Indeed, problem 2 has a diagonal cost function and assumptions A_2 and A_3 guarantee the assumptions of proposition 2.16.

Vaughan's method

Using the Hamiltonian approach, Vaughan [21], [22] has given an explicit expression for the sequence $\{\underline{K}(i)\}$ in the case of the discrete linear-quadratic problem with stationary coefficients. We summarize his results below.

Proposition 2.17

For the problem of proposition 2.16, if $\underline{H}$ is the Hamiltonian matrix defined by (2.210), then:

1) The inverse of any eigenvalue of $\underline{H}$ is also an eigenvalue of $\underline{H}$. Therefore, if the eigenvalues of $\underline{H}$ are simple, we have

$$\underline{W}^{-1}\underline{H}\,\underline{W} = \begin{bmatrix} \underline{\Lambda} & 0 \\ 0 & \underline{\Lambda}^{-1} \end{bmatrix}, \tag{2.212}$$

where $\underline{W}$ is a 2(n-1) x 2(n-1) matrix of linearly independent eigenvectors of $\underline{H}$ (where (n-1) is the dimension of the state $\underline{z}$); and $\underline{\Lambda}$ is an (n-1)x(n-1) diagonal matrix of eigenvalues of $\underline{H}$ of magnitude at least equal to 1.

As far as eigenvalues of magnitude 1 are concerned, we include in $\underline{\Lambda}$ those with argument between 0 and π. Automatically, their inverses are in $\underline{\Lambda}^{-1}$. That is, if $|\lambda| = 1$, then $\lambda = e^{i\omega}$ with $0 \leq \omega < 2\pi$ uniquely defined. If $0 \leq \omega < \pi$, we include λ in $\underline{\Lambda}$; if $\pi \leq \omega < 2\pi$, we include λ in $\underline{\Lambda}^{-1}$. Instead of dividing the unit circle in this fashion, we may have chosen any other partition into two half-circles. The important feature is not to include both in $\underline{\Lambda}$ or both in $\underline{\Lambda}^{-1}$ two eigenvalues which are inverses of one another.

2) If the matrix $\underline{W}$ is partitioned into four (n-1)x(n-1) submatrices:

$$\underline{W} = \begin{bmatrix} \underline{W}_{11} & \underline{W}_{12} \\ \underline{W}_{21} & \underline{W}_{22} \end{bmatrix}, \tag{2.213}$$

then the sequence $\underline{K}(j)$ is given by

$$\begin{aligned} \underline{T} &= -\,[\underline{W}_{22} - \underline{S}\,\underline{W}_{12}]^{-1}\,[\underline{W}_{21} - \underline{S}\,\underline{W}_{11}], \\ \underline{F}(j) &= \underline{\Lambda}^{-j}\underline{T}\,\underline{\Lambda}^{-1}, \\ \underline{K}(j) &= [\underline{W}_{21} + \underline{W}_{22}\underline{F}(j)][\underline{W}_{11} + \underline{W}_{12}\underline{F}(j)]^{-1}, \end{aligned} \tag{2.214}$$

and j = N-i.

<u>Proof</u> (see [21], [22]).

We shall apply Vaughan's results to problem 2 (of theorem 2.15) to derive an important expression for $\underline{\Lambda}$. To do this, we need the following corollary.

Corollary 2.18

Let the eigenvalues of the Hamiltonian matrix $\underline{H}$ be simple. If the sequence $\underline{K}(j)$ converges when $j \to +\infty$, then the limit $\hat{\underline{K}}$ is given by

$$\hat{\underline{K}} = \underline{W}_{21}\underline{W}_{11}^{-1}, \tag{2.215}$$

where $\underline{W}_{11}$ and $\underline{W}_{21}$ are the partitions of $\underline{W}$ defined by (2.213).

Proof

1) If $\underline{H}$ does not have any eigenvalue of magnitude 1, then $\lim_{j\to+\infty} \underline{\Lambda}^{-j} = 0$, and $\lim_{j\to+\infty} \underline{F}(j) = 0$, so that, according to equations (2.214),

$$\lim_{j\to+\infty} \underline{K}(j) = \underline{W}_{21}\underline{W}_{11}^{-1}.$$

2) The number 1 is not an eigenvalue of $\underline{H}$. Since the inverse of an eigenvalue is also an eigenvalue, all eigenvalues could not be simple if the number 1 were one of them.

3) If $\underline{H}$ has some eigenvalues of magnitude 1, but different from 1, then if $\lim_{j\to\infty} \underline{K}(j)$ exists, it is still given by (2.215).

Indeed, $\underline{F}(j)$ does not converge in the usual sense in that case, but $\underline{F}(j)$ does converge in the sense of Cesaro [39] to zero; that is,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} \underline{F}(j) = 0,$$

which we write $\operatorname{Clim}_{n\to\infty} \underline{F}(n) = 0$, where Clim stands for "Cesaro limit."

This is true since

$$\frac{1}{n}\sum_{j=1}^{n} \underline{F}(j) = \frac{1}{n}\left(\sum_{j=1}^{n}\underline{\Lambda}^{-j}\right)\underline{T}\,\underline{\Lambda}^{-1} = \frac{1}{n}\,\underline{\Lambda}^{-1}(\underline{\Lambda}^{-n}-\underline{I})(\underline{\Lambda}^{-1}-\underline{I})^{-1}\underline{T}\,\underline{\Lambda}^{-1},$$

and the magnitudes of the entries of $\underline{\Lambda}^{-n} - \underline{I}$ remain bounded by 2 as n grows. Therefore, it follows that $\hat{\underline{K}} = \underline{W}_{21}\,\underline{W}_{11}^{-1}$, from the following reasoning.

From (2.214), for every j,

$$\underline{K}(j)[\underline{W}_{11} + \underline{W}_{12}\underline{F}(j)] = \underline{W}_{21} + \underline{W}_{22}\underline{F}(j),$$

or

$$\underline{W}_{21} - \underline{K}(j)\,\underline{W}_{11} = \underline{K}(j)\,\underline{W}_{12}\underline{F}(j) - \underline{W}_{22}\,\underline{F}(j).$$

The left-hand side converges to $\underline{W}_{21} - \hat{\underline{K}}\,\underline{W}_{11}$. Therefore, so does the right-hand side. On the other hand, since $\lim_{j\to\infty} \underline{K}(j) = \hat{\underline{K}}$,

$$\lim_{j\to\infty} \underline{K}(j)\underline{W}_{12}\underline{F}(j) - \underline{W}_{22}\underline{F}(j) = \lim_{j\to\infty} (\hat{\underline{K}}\,\underline{W}_{12} - \underline{W}_{22})\underline{F}(j),$$

and, since the usual convergence implies the convergence in the sense of Cesaro(*) [39], the right-hand side is equal to

$$\text{Clim}\,(\hat{\underline{K}}\,\underline{W}_{12} - \underline{W}_{22})\underline{F}(j) = (\hat{\underline{K}}\,\underline{W}_{12} - \underline{W}_{22})\ \text{Clim}\,\underline{F}(j) = 0.$$

Therefore,

$$\underline{W}_{21} - \underline{K}\,\underline{W}_{11} = 0,$$

which yields (2.215). Q.E.D.

Controllability occurs here in our analysis: it enables us to show that the matrix $\underline{\Delta}$ has no eigenvalue of magnitude 1. After having established this result, we shall prove, using the Hamiltonian matrix, that those eigenvalues have, in fact, a magnitude less than 1.

Theorem 2.19

If the cost matrix $\underline{L}$ is positive-definite, the matrix _ corresponding to a controllable reduced system has no eigenvalue of magnitude 1.

Proof

Suppose λ is an eigenvalue of $\underline{\Delta}$, $|\lambda| = 1$, and that $\hat{\underline{\alpha}}$ is a corresponding eigenvector of $\underline{\Delta}^T$,

$$\underline{\Delta}^T\hat{\underline{\alpha}} = \lambda\,\hat{\underline{\alpha}}. \qquad (2.216)$$

---

*This result is usually found for scalars, but extends trivially to matrices, taking each entry separately.

Consider the steady state; i.e., when $\underline{\Delta}(i) = \underline{\Delta}$. Then, the evolution of the optimal state is described by

$$\underline{z}^*(i+1) = \underline{\Delta}\, \underline{z}^*(i) + \underline{\Pi}\, F, \tag{2.217}$$

in the linear-quadratic problem. (The rigorous formulation is: $\lim_{i\to-\infty} [\underline{z}^*(i+1) - \underline{\Delta}\, \underline{z}^*(i) - \underline{\Pi}\, F] = 0$.)

Consider now the full state

$$\underline{x}(i) = \begin{bmatrix} \underline{z}(i) \\ x_n(i) \end{bmatrix}.$$

If $F \overset{\Delta}{=} \underline{\nu}_n^T \underline{x}(i) = 0$, then equation (2.217) becomes

$$\underline{z}^*(i+1) = \underline{\Delta}\, \underline{z}^*(i), \tag{2.218}$$

which, together with (2.216), implies:

$$\hat{\underline{\alpha}}^T \underline{z}^*(i+1) = \lambda\, \hat{\underline{\alpha}}^T \underline{z}^*(i). \tag{2.219}$$

Therefore, if we define $\underline{\alpha} = \begin{bmatrix} \hat{\underline{\alpha}} \\ 0 \end{bmatrix} \varepsilon R^n$, then for any initial state $\underline{x}(1)$, such that $\underline{\nu}_n^T \underline{x}(1) = 0$, we have

$$\underline{\alpha}^T \underline{x}^*(i+1) = \lambda\, \underline{\alpha}^T \underline{x}(i), \tag{2.220}$$

thus,

$$|\underline{\alpha}^T \underline{x}^*(i+1)| = |\underline{\alpha}^T \underline{x}^*(i)|, \tag{2.221}$$

since $|\lambda| = 1$.

Therefore,

$$|\underline{\alpha}^T \underline{x}^*(i)| = |\underline{\alpha}^T\, \underline{x}(1)| \quad \text{for } i=1,2,3,..., \tag{2.222}$$

by induction.

Consider our minimization problem P_1 in its original formulation (section 2.2),

$$\min(P_1) = \min \sum_{j=1}^{N-1} \frac{1}{2} \underline{\phi}^T(j)\ \underline{L}\ \underline{\phi}(j), \tag{2.223}$$

subject to:

$$\underline{x}(j) = \underline{Y}\ \underline{\phi}(j),$$

$$\underline{x}(j+1) = \underline{Z}\ \underline{\phi}(j),$$

$$\underline{x}(1) = \underline{\xi},$$

where we choose $\underline{\xi}$ such that $\underline{\nu}_n^T\underline{\xi} = 0$.

Then, the optimal states $\underline{x}^*(j)$ satisfy constraints (2.222) if $N \to \infty$.

Now consider the auxiliary minimization problem P_2:

$$\min \frac{1}{2} \underline{\phi}^T\ \underline{L}\ \underline{\phi},$$

subject to:

$$|\underline{\phi}^T\underline{Y}^T\underline{\alpha}| = |\underline{\xi}^T\underline{\alpha}|, \tag{2.224}$$

For any j, the optimal flow perturbation vector $\underline{\phi}^*(j)$ of problem P_1 satisfies (2.224), hence is feasible for problem P_2. Therefore,

$$\frac{1}{2} \underline{\phi}^{*T}(j)\ \underline{L}\ \underline{\phi}^*(j) \geq \min\ (P_2) \text{ for any j.}$$

Consequently,

$$\min\ (P_1) = \frac{1}{2} \sum_{j=1}^{N-1} \underline{\phi}^{*T}(j)\ \underline{L}\ \underline{\phi}^*(j) \geq (N-1) \min\ (P_2),$$

or, precisely,

$$\lim_{N\to\infty} (N-1) \min\ (P_2) \leq \min\ (P_1). \tag{2.225}$$

Using the controllability of the system in $\underline{z}$, we can drive the state $\underline{z}$ to zero in some number of steps; say, k. That is, we can choose controls $\hat{\underline{v}}(1)$, $\hat{\underline{v}}(2),\ldots,\hat{\underline{v}}(k-1)$, such that $\underline{z}(k) = 0$, which implies $\underline{x}(k) = 0$, since $\underline{\nu}_n^T\underline{x}(k) = \underline{\nu}_n^T\underline{x}(1) = 0$. Also, we can choose $\underline{v}(j) = 0$ for all $j \geq k$, which implies $\underline{z}(j) = 0$ for all $j \geq k$, and $\underline{x}(j) = 0$ for all $j \geq k$. Therefore, the corresponding flow $\underline{\phi}(j) = 0$ for $j \geq k$ (see section 2.3.3).

Accordingly,

$$\min (P_1) \leq \frac{1}{2} \sum_{j=1}^{k-1} \hat{\underline{\phi}}(j)\ \underline{L}\hat{\underline{\phi}}(j), \tag{2.226}$$

when $\hat{\underline{\phi}}(1), \hat{\underline{\phi}}(2), \ldots, \hat{\underline{\phi}}(k+1)$ are the feasible flow vectors corresponding to $\hat{\underline{v}}(1), \hat{\underline{v}}(2), \ldots, \hat{\underline{v}}(k-1)$. This is because $\hat{v}(j)$ is feasible but not necessarily optimal for problem P_1. Comparing (2.225) and (2.226) yields

$$\lim_{N\to\infty} (N-1) \min (P_2) \leq \frac{1}{2} \sum_{j=1}^{k-1} \hat{\underline{\phi}}^T(j)\ \underline{L}\ \hat{\underline{\phi}}(j) < \infty, \tag{2.227}$$

which is clearly impossible unless $\min (P_2) = 0$. (It is to be emphasized that neither k nor $\min (P_2)$ depends on N.)

In fact, $\min (P_2) > 0$. Let $\tilde{\underline{\phi}}$ be the optimal flow in problem P_2, and assume

$$\min (P_2) = \frac{1}{2} \tilde{\underline{\phi}}^T \underline{L}\ \tilde{\underline{\phi}} = 0.$$

Since $\underline{L}$ is positive-definite, this implies $\tilde{\underline{\phi}} = 0$ which, in view of (2.224) is feasible, only if

$$\underline{\xi}^T \underline{\alpha} = 0.$$

However, it is possible to choose $\underline{\xi}$, such that $\underline{\xi}^T \underline{\alpha} \neq 0$, and yet, $\underline{\xi}^T \underline{\nu}_n = 0$. This is because $\underline{\alpha}$ and $\underline{\nu}_n$ are linearly independent. We have evidenced this contradiction (2.227) by assuming at the same time that the system is controllable, and that $\underline{\Delta}$ has an eigenvalue of magnitude 1. Q.E.D.

We shall now apply Vaughan's results to $\underline{\Delta}$, and combine them with those of theorem 2.19.

<u>Theorem 2.20</u>

If

1) The cost matrix $\underline{L}$ is positive-definite.

2) The matrix $\underline{Z}\ \underline{L}^{-1} \underline{Y}^T$ is invertible.

3) The Hamiltonian matrix $\underline{H}$ given by (2.210) and (2.211) has simple eigenvalues.

4) Some reduced system is controllable.

Then

a) The Hamiltonian matrix $\underline{H}$ has no eigenvalues of magnitude 1.

b) The corresponding matrix $\underline{\Delta}$ satisfies

$$\underline{\Delta} = \underline{W}_{11} \, \underline{\Lambda}^{-1} \, \underline{W}_{11}^{-1}, \tag{2.228}$$

where $\underline{\Lambda}$ and $\underline{W}_{11}$ are defined by (2.212) and (2.213).

Consequently, the eigenvalues of $\underline{\Delta}$ are those of $\underline{H}$ which lie within the open unit disk.

c) To each eigenvector of $\underline{H}$ associated to an eigenvalue of $\underline{H}$ outside of the unit disk, there corresponds an eigenvector of $\underline{\Delta}$ associated to the inverse of that eigenvalue. Its components are the (n-1) first components of that of $\underline{H}$. The columns of $\underline{W}_{11}$ are (n-1) linearly independent eigenvectors of $\underline{\Delta}$. The matrices $\underline{\Delta}(i)$ and $\underline{\Delta}$ are invertible.

d) The following equations hold, relating them to the partition $\underline{\Delta}(N-1)$ of the one-step propagation matrix $\underline{D}(N-1)$.

$$\left.\begin{aligned} \underline{\Delta}(i) &= (\underline{I} + \underline{B}\,\underline{R}^{-1}\underline{B}^T\underline{K}(i+1))^{-1}\underline{\Delta}(N-1) \\ \underline{\Delta} &= (\underline{I} + \underline{B}\,\underline{R}^{-1}\underline{B}^T\hat{\underline{K}})^{-1}\,\underline{\Delta}(N-1) \end{aligned}\right\} . \tag{2.229}$$

Before proving theorem 2.20, let us state and prove its corollary, which closes and sums up our investigation of the optimal downstream perturbations in an infinitely long stationary network.

Corollary 2.21

Under the assumptions of theorem 2.20:

1) The asymptotic propagation matrix $\underline{D}$ is invertible, and has the number 1 as a simple eigenvalue. The corresponding normalized eigenvector $\underline{p}$ is uniquely defined by

$$\left.\begin{aligned} \underline{D}\,\underline{p} &= \underline{p} \\ \underline{\nu}_n^T\underline{p} &= 1 \end{aligned}\right\} , \tag{2.230}$$

is given by:

$$\underline{p} = \begin{pmatrix} \underline{p}_1 \\ p_2 \end{pmatrix}, \tag{2.231}$$

$$\underline{p}_1 = (\underline{I} - \underline{\Delta})^{-1}\underline{\Pi}, \qquad p_2 = 1 - \underline{\nu}_{n-1}^T\underline{p}_1. \tag{2.232}$$

where the matrix $\underline{\Delta}$ is obtained from (2.163) and the vector $\underline{\Pi}$ is obtained from (2.164), using

$$\hat{\underline{K}} = \underline{W}_{21}\underline{W}_{11}^{-1}. \tag{2.233}$$

2) An initial perturbation $\underline{x}$ with total flow input $F \overset{\Delta}{=} \underline{\nu}_n^T\underline{x}$ gives rise to a sequence $\underline{x}(k)$ which converges to $F\,\underline{p}$ as k goes to infinity at an exponential rate.

3) Specifically,

$$||\underline{x}(k) - F\,\underline{p}|| \leq M\,|s|^k\,||x||, \tag{2.234}$$

where

a) $|s|$ is the largest magnitude of an eigenvalue of the Hamiltonian matrix within the open unit disk.

b) $M = (n-1)B$, (2.235)

and $B = \max_{i=1,\ldots,n} ||\underline{B}(i)||$, (2.236)

where $\underline{B}(i)$ is an nxn matrix defined by

$$B_{rs}(i) = V_{ri}(V)^{-1}_{is} \quad \begin{cases} r=1,\ldots,n, \\ s=1,\ldots,n, \end{cases} \tag{2.237}$$

$$\underline{V} = \begin{bmatrix} \underline{p}_1 & \underline{W}_{11} \\ p_2 & -\underline{\nu}_{n-1}^T\underline{W}_{11} \end{bmatrix}. \tag{2.238}$$

<u>Proof of Corollary 2.21</u>

Under the assumptions of theorem 2.20, the eigenvalues of $\underline{\Delta}$ lie within the open unit disk, so that, by lemma 2.12 (part 1), the number 1 is a simple eigenvalue of $\underline{D}$. Also, all other eigenvalues of $\underline{D}$ have magnitudes strictly less than 1. Finally, the Hamiltonian matrix $\underline{H}$ has simple eigenvalues, so that the same is true for $\underline{\Delta}$ as equation (2.228) shows. Lemma 2.12 then implies that also $\underline{D}$ has only simple eigenvalues. Therefore, all assumptions of theorem 2.12 hold. The present corollary is merely a restatement of theorem 2.12, where we adopt the approximation $\underline{x}(k) = \underline{D}^k\underline{x}$ (infinitely long network). The expression for $\underline{p}$ is derived from lemma 2.12 (part 3) taking into account the invertibility of $\underline{I} - \underline{\Delta}$. Also, the expression for $\underline{V}$ comes from corollary 2.14, and that for M from the proof of lemma 2.12.

Proof of Theorem 2.20

To apply Vaughan's results, we have first defined in theorem 2.15, a problem with diagonal cost which is equivalent to our original linear-quadratic problem. Vaughan's results are applicable to this new problem if the assumptions of proposition 2.16 hold.

1) $\underline{S} \geq 0$ since $\underline{S} = 0$,

2) $\underline{\tilde{Q}} \geq 0$,

$$\underline{\tilde{Q}} = \underline{Q} - \underline{M}^T\underline{R}^{-1}\underline{M} \text{ by (2.211).}$$

The recurrent equations for $\underline{K}(i)$ (equation (2.133) show that

$$\underline{K}(N-1) = \underline{Q} - \underline{M}^T\underline{R}^{-1}\underline{M}, \qquad (2.239)$$

since $\underline{K}(N) = 0$.

On the other hand, we have shown in lemma 2.9 that $\underline{K}(i) \geq 0$. Therefore, we have $\underline{\tilde{Q}} > 0$.

3) $\underline{\tilde{R}} = \underline{R} > 0$ was shown in lemma 2.9.

4) $\underline{\tilde{A}}$ invertible

$$\underline{\tilde{A}} = \underline{A} - \underline{B}\,\underline{R}^{-1}\underline{M},$$

by (2.211).

Equation (2.160) for $\underline{\Delta}(i)$ shows that

$$\underline{A} - \underline{B}\ \underline{R}^{-1}\underline{M} = \underline{\Delta}(N-1). \quad (2.240)$$

From lemma 2.12, with s=0, det $(\underline{\Delta}(N-1))$ = det $(\underline{D}(N-1))$.

Finally, from equations (2.60), (2.62) of section 2.2,

$$\underline{D}(N-1) = (\underline{Z}\ \underline{L}^{-1}\underline{Y}^T)(\underline{Y}\ \underline{L}^{-1}\underline{Y}^T)^{-1}. \quad (2.241)$$

Since we know that $\underline{Y}\ \underline{L}^{-1}\underline{Y}^T$ is invertible, from the positive definiteness of $\underline{L}$ (section 2.2), it follows that $\underline{D}(N-1)$ is invertible if, and only if, $\underline{Z}\ \underline{L}^{-1}\underline{Y}^T$ is. Therefore, assumption 2 of the present theorem guarantees that $\tilde{\underline{A}}$ is invertible. We can, therefore, apply Vaughan's results. Since the subsystem is controllable by assumption, we know that $\underline{K}(i)$ converges; from corollary 2.18, the limit $\hat{\underline{K}}$ satisfies

$$\hat{\underline{K}} = \underline{W}_{21}\underline{W}_{11}^{-1}. \quad (2.242)$$

Partitioning $\underline{H}$ into four (n-1)x(n-1) blocks and using (2.108),

$$z^*(i) = \underline{H}_{11}\ \underline{z}^*(i+1) + \underline{H}_{12}\underline{p}^*(i+1), \quad (2.243)$$

or, by (2.208),

$$z^*(i) = (\underline{H}_{11} + \underline{H}_{12}\underline{K}(i+1))z^*(i+1). \quad (2.244)$$

Let us now show that $\underline{H}_{11} + \underline{H}_{12}\ \underline{K}(i+1)$ is invertible. Since $\underline{K}(i+1) > 0$, $\underline{K}^{-1}(i+1) > 0$. Also $\underline{B}\ \underline{R}^{-1}\underline{B}^T \geq 0$ since $\underline{R} > 0$. Thus, $\underline{K}^{-1}(i+1) + \underline{B}\ \underline{R}^{-1}\underline{B}^T > 0$, and is, hence, invertible. The claim follows from the equality

$$\underline{H}_{11} + \underline{H}_{12}\ \underline{K}(i+1) = \tilde{\underline{A}}^{-1}(\underline{K}^{-1}(i+1) + \underline{B}\ \underline{R}^{-1}\underline{B}^T)\underline{K}(i+1), \quad (2.245)$$

which is an immediate consequence of the expression (2.210) for $\underline{H}$. Accordingly, (2.243) can be written

$$\underline{z}^*(i+1) = (\underline{H}_{11} + \underline{H}_{12}\ \underline{K}(i+1))^{-1}\underline{z}^*(i). \quad (2.246)$$

Recall that, since we are interested in $\{\underline{\Delta}(i)\}$, the problem we are now studying is the truly linear-quadratic problem, where the linear terms are absent from the cost function (in the reduced state). Therefore,

$$\underline{z}^*(i+1) = \underline{\Delta}(i)\ \underline{z}^*(i). \tag{2.247}$$

Identifying (2.246) and (2.247) yields

$$\underline{\Delta}(i) = (\underline{H}_{11} + \underline{H}_{12}\ \underline{K}(i+1))^{-1}. \tag{2.248}$$

Taking now (2.242) into account, we obtain

$$\begin{aligned}\underline{\Delta} = \lim_{i\to-\infty} \underline{\Delta}(i) &= (\underline{H}_{11} + \underline{H}_{12}\ \underline{W}_{21}\ \underline{W}_{11}^{-1})^{-1} \\ &= \underline{W}_{11}\,[\underline{H}_{11}\ \underline{W}_{11} + \underline{H}_{12}\ \underline{W}_{21}]^{-1}.\end{aligned} \tag{2.249}$$

Using (2.212),

$$\underline{H}_{11}\ \underline{W}_{11} + \underline{H}_{12}\ \underline{W}_{21} = \underline{W}_{11}\ \underline{\Lambda}, \tag{2.250}$$

so that equation (2.249) becomes

$$\underline{\Delta} = \underline{W}_{11}\ \underline{\Lambda}^{-1}\ \underline{W}_{11}^{-1}\ . \tag{2.251}$$

We have thus proved part 2 of the theorem.

On the other hand, we know from theorem 2.19 that $\underline{\Delta}$ does not have any eigenvalue of magnitude 1. Therefore, the same is true for $\underline{H}$ because equation (2.251) shows that the eigenvalues of $\underline{\Delta}$ are those of $\underline{H}$ which lie within the closed unit disk. Now, we have thus proved part 1 as well. Equation (2.251) can be rewritten

$$\underline{\Delta}\ \underline{W}_{11} = \underline{W}_{11}\ \underline{\Lambda}^{-1}, \tag{2.252}$$

which, together with (2.212) and (2.213), justifies part 3 of the theorem. To prove part 4, we just have to use (2.240), (2.245), and (2.248) together, and also to take the limit when $i \to -\infty$. Equations (2.229) is then established.

Assumption 2 implies that det $(\underline{\Delta}(N-1)) \neq 0$ and, in turn, there follows from (2.229) that det $(\underline{\Delta}(i)) \neq 0$ for all i. Q.E.D.

Remark

Under the assumptions of theorem 2.20, we have shown that $\underline{\Delta}$ is a stable matrix. Therefore, under those assumptions, $\underline{A} = \underline{\Delta}^T$ (see theorem 2.11, part 2.3) is a stable matrix and, from theorem 2.8 of section 2.3.5, the sequence $\underline{\ell}(i)$ converges, and consequently $\lim_{l\to-\infty} [\underline{C}(i) - \underline{C}(i+1)] = \alpha\nu\, \underline{\nu}_n \underline{\nu}_n^T$. Also, the sequence $\underline{\Pi}(i)$ converges and so does $\underline{D}(i)$. In the absence of the assumptions of theorem 2.20 (in particular, if the subsystem in $\underline{z}$ is not controllable), we cannot make any statement as to the convergence of $K(i)$, $\underline{\ell}(i)$, $m(i) - m(i+1)$, $\underline{C}(i) - \underline{C}(i+1)$, $\underline{\Pi}(i)$, $\underline{D}(i)$.

Discussion of Assumptions

The assumptions under which our main result is valid are those of theorem 2.20; i.e.,

a) The cost matrix $\underline{L}$ is positive-definite.

b) The one-step propagation matrix $\underline{D}(N-1)$ is invertible.

c) The Hamiltonian matrix has simple eigenvalues.

d) Some reduced system is controllable.

Assumptions (a) and (d) are fundamental. Assumptions (b) and (c) are probably of less importance (i.e., they are probably not necessary for the result to hold).* Assumption (a) is deeply rooted in our problem, and we have made it since the very beginning. Without the positive definiteness of $\underline{L}$, the minimization problem would be illposed.

The controllability assumption (d) occurs in all the proofs and is therefore fundamental. We also have examples (see section 7) which illustrate the behavior which may occur when this assumption is not met. From the geometric interpretation of controllability which is given in section 2.3.4, it is intuitive (at least in what we have called class 1) that, if the system is not controllable, not all entrances are accessible from any given entrance, and therefore some initial perturbations with zero total-flow input will not decrease toward zero.

Assumptions (b) and (c), although they are explicitly used in the proof of theorem 2.20, seem easy to elude by a density argument, which, however, is not entirely satisfactory.

Indeed, one can show [30] that, for the usual matrix norm $||\underline{M}|| = \max_{||\underline{x}||=1} ||\underline{M}\,\underline{x}||$,

* In [42], only assumptions (a) and (d) are required to prove that perturbations diminish.

the set of all matrices with simple eigenvalues is everywhere dense in the space of all matrices, and so is the set of invertible matrices.

For assumption (c), a look at equations (2.210) and (2.211) shows that the matrix $\underline{A} - \underline{B}\ \underline{R}^{-1}\underline{M}$, or its transpose or its inverse, occurs as a factor in all four partitions. That matrix is also $\underline{\Delta}(N-1)$, so that $\underline{H}$ depends continuously on $\underline{\Delta}(N-1)$ which, in turn, depends continuously on the cost matrix $\underline{L}$.

Accordingly, a slight change in $\underline{L}$ can give rise to a Hamiltonian matrix with simple eigenvalues, for which theorem 2.20 is valid. Therefore, by continuity, one shows that the original $\underline{\Delta}$ matrix has its eigenvalues within the <u>closed</u> unit disk. Applying corollary 2.19 directly to $\underline{\Delta}$, it follows that they lie within the <u>open</u> unit disk.

Likewise, the matrix $\underline{D}(N-1)$ can be approximated by one that is nonsingular, since the factor $\underline{Z}\ \underline{L}^{-1}\underline{Y}^T$ occurs in $\underline{D}(N-1)$.

The only restriction to the validity of these density arguments is the fact that $\underline{Z}$ and $\underline{Y}$ are both matrices of ones and zeroes which cannot change continuously. Therefore, it is possible a priori that, for some subnetwork, $\underline{Y}$ and $\underline{Z}$ be such that $\det(\underline{Z}\ \underline{L}^{-1}\underline{Y}^T) = 0$ for any $\underline{L}$, in which case the above arguments fail.

Finally, let us notice that, when assumption (b) is not met; i.e., $\det(\underline{D}(N-1)) = 0$, some nonzero perturbation $\underline{x}$ will be brought to zero in one step: $\underline{D}(N-1)\ \underline{x} = 0$. It is possible that, in this case, also $\det(\underline{D}) = 0$, so that there will be a nonzero $\underline{x}$ for which $\underline{D}\ \underline{x} = 0$. This is illustrated by examples in section 7.

2.3.8 <u>Illustration</u>

We illustrate below the theory of section 2.3.7 on a special class of examples.

We consider the standard two-dimensional example of Fig. 7.1 with the corresponding labeling. Let us express the Hamiltonian matrix as a function of L_1, L_2, L_3, L_4.

For $\underline{L} = \begin{pmatrix} L_1 & & & \\ & L_2 & & \\ & & L_3 & \\ & & & L_4 \end{pmatrix}$,

$\underline{A} = 1,$

$\underline{B} = (-1, 1),$

$\underline{Q} = L_1 + L_2,$

$\underline{R} = \begin{bmatrix} L_1 + L_3 & 0 \\ 0 & L_2 + L_4 \end{bmatrix}$, and $\underline{M} = \begin{pmatrix} -L_1 \\ L_2 \end{pmatrix}$,

$$\underline{B}\,\underline{R}^{-1}\underline{M} = (-1, 1) \begin{bmatrix} \frac{1}{L_1+L_3} & 0 \\ 0 & \frac{1}{L_2+L_4} \end{bmatrix} \begin{bmatrix} -L_1 \\ L_2 \end{bmatrix} = \frac{L_1}{L_1 + L_3} + \frac{L_2}{L_2 + L_4},$$

$$\tilde{\underline{A}} = \underline{A} - \underline{B}\,\underline{R}^{-1}\underline{M} = 1 - \left(\frac{L_1}{L_1+L_3} + \frac{L_2}{L_2+L_4} \right) = \frac{L_3L_4 - L_1L_2}{(L_1+L_3)(L_2+L_4)},$$

$$\tilde{\underline{Q}} = \underline{Q} - \underline{M}\,\underline{R}^{-1}\underline{M}^T = (L_1+L_2) - [-L_1, L_2] \begin{bmatrix} \frac{1}{L_1+L_3} & 0 \\ 0 & \frac{1}{L_2+L_4} \end{bmatrix} \begin{bmatrix} -L_1 \\ L_2 \end{bmatrix}$$

$$= \frac{L_1L_3}{L_1+L_3} + \frac{L_2L_4}{L_2+L_4}.$$

So we have verified that

$\tilde{\underline{Q}} \geq 0,$

$\underline{R} > 0,$

and $\underline{A}$ is invertible, provided that $L_3L_4 \neq L_1L_2$ or

$$\frac{L_1}{L_3} \neq \frac{L_4}{L_2}. \tag{2.253}$$

The Hamiltonian matrix is then given by

$$
\underline{H} = \left[\begin{array}{c|c}
\dfrac{(L_1+L_3)(L_2+L_4)}{L_3L_4-L_1L_2} & \dfrac{L_1+L_2+L_3+L_4}{L_3L_4-L_1L_2} \\
\hline
\dfrac{L_1L_3(L_2+L_4)+L_2L_4(L_1+L_3)}{L_3L_4 - L_1L_2} & \left\{\dfrac{L_3L_4-L_1L_2}{(L_1+L_3)(L_2+L_4)} + \left[\dfrac{L_1L_3}{L_1+L_3} + \dfrac{L_2L_4}{L_2+L_4}\right]\dfrac{L_1+L_2+L_3+L_4}{L_3L_4-L_1L_2}\right\}
\end{array}\right]
$$

(2.254)

In the special symmetric case when $L_1 = L_2 = 1$ and $L_3 = L_4 = a$ (see Fig. 2.3.8-1), the condition of (2.253) for $\underline{\tilde{A}}$ to be invertible becomes $a \neq 1$, and equation (2.254) reduces to

$$
\underline{H} = \begin{bmatrix} \dfrac{a+1}{a-1} & \dfrac{2}{a-1} \\ \dfrac{2a}{a-1} & \dfrac{a+1}{a-1} \end{bmatrix}, \tag{2.255}
$$

whence it follows that the eigenvalues of $\underline{H}$ are the solutions of

$$
\det(\underline{H}-s\underline{I}) = \begin{bmatrix} \dfrac{a+1}{a-1} - s & \dfrac{2}{a-1} \\ \dfrac{2a}{a-1} & \dfrac{a+1}{a-1} - s \end{bmatrix} = s^2 - \frac{2(a+1)}{a-1}s + 1 = 0,
$$

which yields

$$
\begin{cases} s_1 = \dfrac{\sqrt{a}-1}{\sqrt{a}+1}, \\ s_2 = \dfrac{\sqrt{a}+1}{\sqrt{a}-1} = \dfrac{1}{s_1}. \end{cases}
$$

The eigenvalue of magnitude less than 1 is

$$
\lambda(a) = \frac{\sqrt{a}-1}{\sqrt{a}+1}. \tag{2.256}
$$

On the other hand, the eigenvector corresponding to the eigenvalue $\frac{1}{s_1} = s_2$ is given by

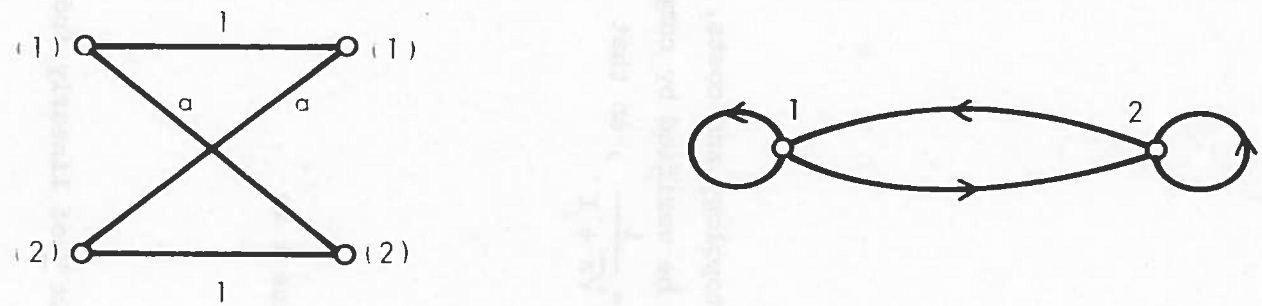

Figure 2.3.8-1 GENERAL CONTROLLABLE CASE AND CORRESPONDING ACCESSIBILITY GRAPH

$0 < a < +\infty;\ -1 < \lambda < 1$

$$\frac{-2\sqrt{a}}{a-1} W_{11} + \frac{2}{a-1} W_{21} = 0,$$

whence

$$\hat{K} = W_{21} W_{11}^{-1} = \sqrt{a}. \tag{2.257}$$

Also,

$$\Delta = \lambda_1 = \frac{\sqrt{a}-1}{\sqrt{a}+1} \; . \tag{2.258}$$

In this two-dimensional case,

$$\underline{D} = \begin{pmatrix} \Delta + \Pi & \Pi \\ 1-(\Delta+\Pi) & 1-\Pi \end{pmatrix} .$$

In addition, from the symmetry of the topology and costs, it follows that $\underline{D}$ must be symmetric: $\Delta + \Pi = 1-\Pi$. This can be verified by computing Π from (2.164), using $\hat{K} = \sqrt{a}$, therefore $\Pi = \frac{1-\Delta}{2} = \frac{1}{\sqrt{a}+1}$, so that

$$\underline{D} = \begin{pmatrix} \frac{\sqrt{a}}{\sqrt{a}+1} & \frac{1}{\sqrt{a}+1} \\ \frac{1}{\sqrt{a}+1} & \frac{\overline{a}}{\sqrt{a}+1} \end{pmatrix} , \tag{2.259}$$

and the unit-flow eigenvector with eigenvalue 1 is

$$\underline{p} = \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix} . \tag{2.260}$$

Let us compute the bounds. As a matrix $\underline{V}$ of linearly independent eigenvectors of $\underline{D}$, we may take

$$\underline{V} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}; \quad \text{then} \quad \underline{V}^{-1} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix},$$

so that $\underline{B}(0) = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix} = \underline{p}\, \underline{\nu}_2^T$, and $\quad \underline{B}(1) = \begin{pmatrix} 1/2 & -1/2 \\ -1/2 & 1/2 \end{pmatrix}$,

so that $\quad \underline{B}^T(1)\underline{B}(1) = \begin{pmatrix} 1/2 & -1/2 \\ -1/2 & 1/2 \end{pmatrix}$,

whose eigenvalues are 0 and 1, so that the maximum eigenvalue of $\underline{B}^T(1)\underline{B}(1)$ is 1, and $||\underline{B}(1)|| = 1 = B$.

Also, n=2, so that (n-1) = 1 and

$$M = (n-1)\ B = 1.$$

Therefore, $||\underline{D}^k\underline{x}|| \leq |\lambda_1|^k||\underline{x}|| = \left(\frac{\sqrt{a} - 1}{\sqrt{a} + 1}\right)^k ||\underline{x}||$,

or, in an infinitely long network,

$$||\underline{x}(k)|| \leq \left(\frac{\sqrt{a} - 1}{\sqrt{a} + 1}\right)^{k-1} ||\underline{x}(1)||. \qquad (2.261)$$

In the special case a = 2, we have

$$\lambda(a) = \frac{\sqrt{2} - 1}{\sqrt{2} + 1} = 3 - 2\sqrt{2},$$

$$K = \sqrt{2} \qquad ,$$

$$D = \begin{pmatrix} \frac{\sqrt{2}}{\sqrt{2} + 1} & \frac{1}{\sqrt{2} + 1} \\ \frac{1}{\sqrt{2} + 1} & \frac{\sqrt{2}}{\sqrt{2} + 1} \end{pmatrix} = \begin{pmatrix} 2 - \sqrt{2} & \sqrt{2} - 1 \\ \sqrt{2} - 1 & 2 - \sqrt{2} \end{pmatrix},$$

as observed numerically (section 7) and found by the Riccati equation (section 2.3.6).

Equation (2.256) gives the eigenvalues of $\underline{D}$ other than 1 as a function of the parameter a, which is plotted in Fig. 2.3.8-2. Notice that $\lambda(\frac{1}{a}) = -\lambda(a)$, so that $|\lambda(\frac{1}{a})| = |\lambda(a)|$. This is not surprising since replacing a by $\frac{1}{a}$ is equivalent, as far as the $\underline{D}$ matrix is concerned, to having the cost a along the

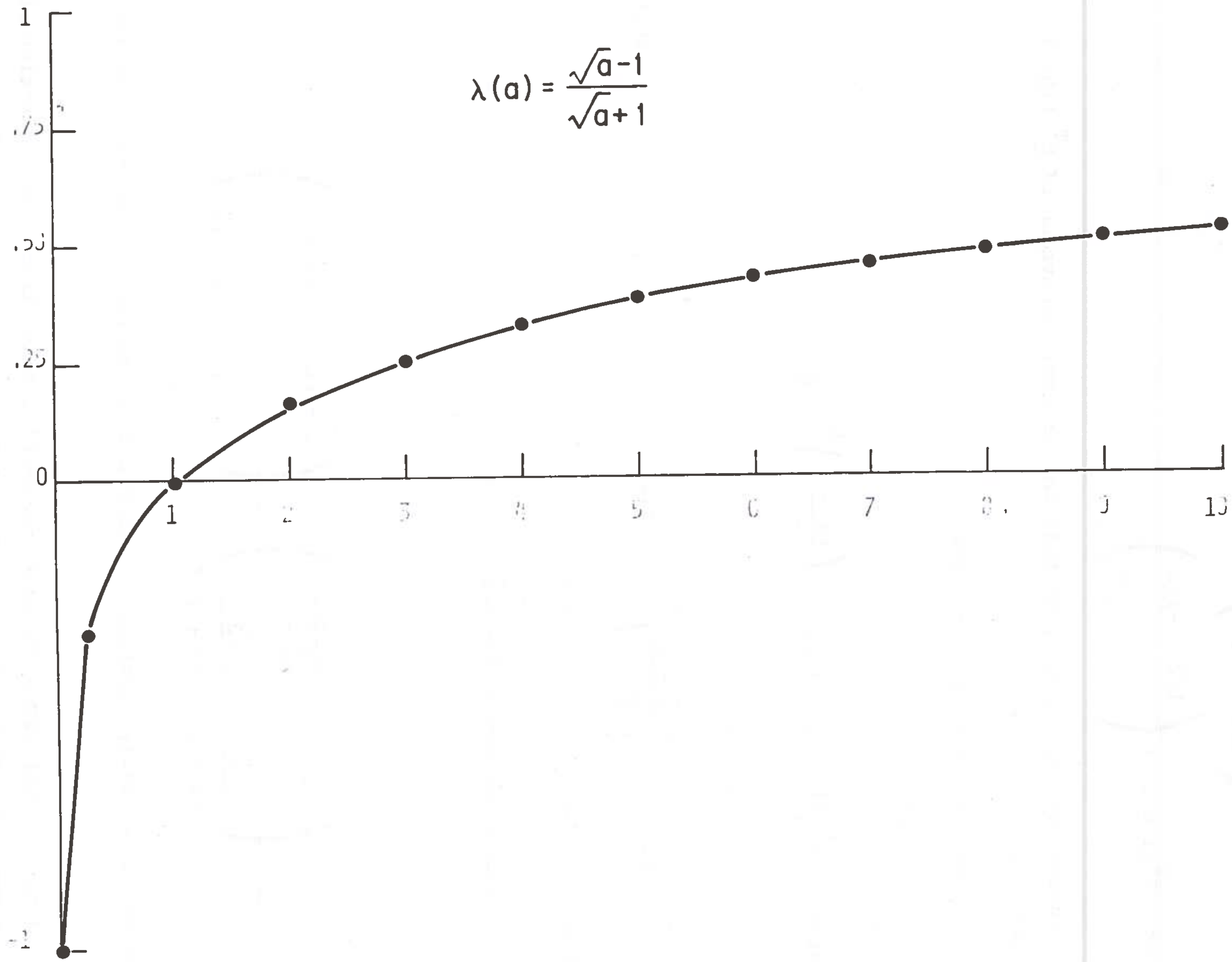

Figure 2.3.8-2 EIGENVALUE OF D AS A FUNCTION OF LINK LENGTH a

through links and the cost 1 along the cross links. Therefore, it is equivalent to interchanging costs along the cross links and the through links, which does not influence the rate at which perturbations decrease. The function $\lambda(a)$ can therefore be studied for $a > 1$ only. It is a concave, slowly increasing function with a horizontal asymptotic line at 1. For $a \to +\infty$, $\lim \lambda(a) = 1$. Letting a go to infinity is equivalent to deleting the cross links since the corresponding cost becomes so large that no traffic is sent along them. The subnetwork where the cross links are deleted is uncontrollable since it corresponds to an unconnected accessibility graph (Fig. 2.3.8-3). On the other hand, for $a \to 0$, $\lim \lambda(a) = -1$. Letting a go to zero is equivalent to deleting the through links since the $\underline{D}$ matrix corresponding to $L_1 = L_2 = 1;\ L_3 = L_4 = a$ is the same as the $\underline{D}$ matrix corresponding to $L_1 = L_2 = \frac{1}{a};\ L_3 = L_4 = 1$ ($\underline{D}$ being homogeneous of degree 0 in $\underline{L}$). The subnetwork with through links deleted is also uncontrollable (Fig. 2.3.8-4) since it is of class 3 (no choice at any node). For any finite and strictly positive value of a, the subsystems are controllable since the subnetwork is of class 1, and the accessibility graph is strongly connected and does not have any transient nodes. Let us notice that for $a = 1$, $\lambda(a) = 0$. In that case, $\underline{D} = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}$, so that a perturbation with zero total flow is driven to zero in one step: $\underline{D}\begin{pmatrix} -1 \\ +1 \end{pmatrix} = 0$.

2.3.9 Quasi-stationary Networks

The asymptotic analysis of sections 2.3.2 to 2.3.7 rests on the stationarity property. This is of course an important restriction. However, our purpose here is to show that the analysis of the previous section is extendable to a larger class of networks. Let us consider for instance, the example of Fig. 1.4.1. That network consists of two alternating types of subnetworks. The first one (type 1) is drawn in Fig. 2.3.9-1, and the other one (type 2) in Fig. 2.3.9-2. The subnetwork of Fig. 2.3.9-2 is peculiar because the flows along the various links can be expressed in terms of the incoming flows: $\underline{\phi}(k) = \underline{x}(k)$. This is because only one link originates from each entrance, so that there is no choice at any point. Let us denote by $\underline{\psi}(k)$ the vector of flows along the k^{th} subnetwork of the type 2, and by $\underline{\phi}(k)$ the vector of flows along the k^{th} subnetwork of the type 1 immediately upstream

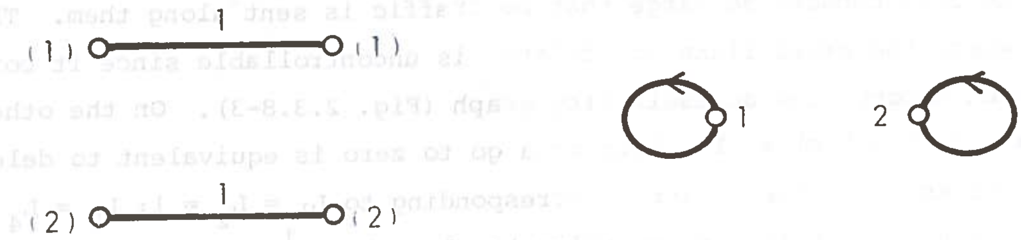

Figure 2.3.8-3 FIRST UNCONTROLLABLE CASE AND CORRESPONDING ACCESSIBILITY GRAPH:
$a = +\infty$; $\lambda = 1$

Figure 2.3.8-4 SECOND UNCONTROLLABLE CASE AND CORRESPONDING ACCESSIBILITY GRAPH:
$a = 0$; $\lambda = -1$

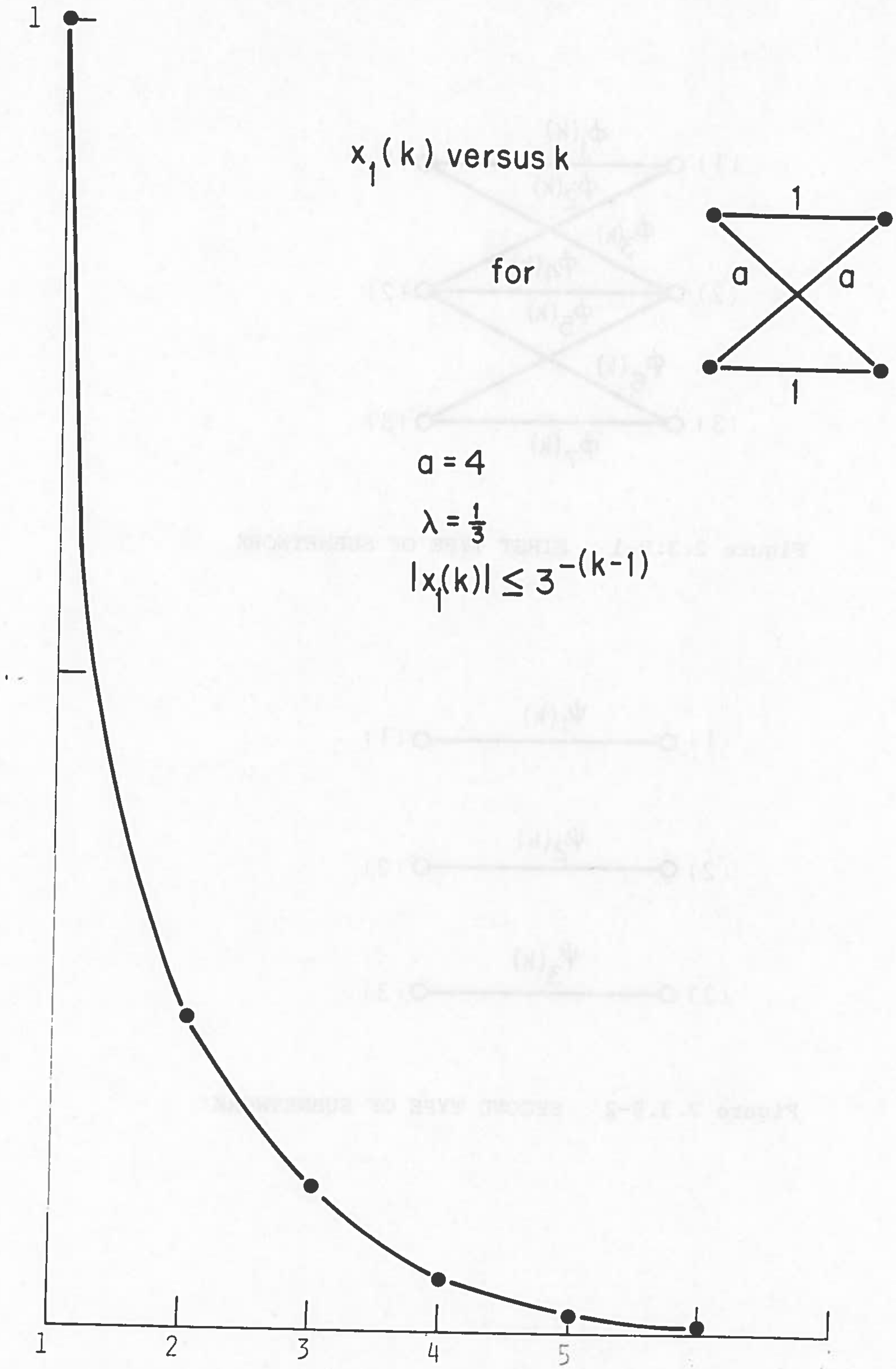

Figure 2.3.8-5 EXPONENTIAL DECAY OF DOWNSTREAM PERTURBATIONS CORRESPONDING TO AN INITIAL PERTURBATION WITH ZERO TOTAL FLOW

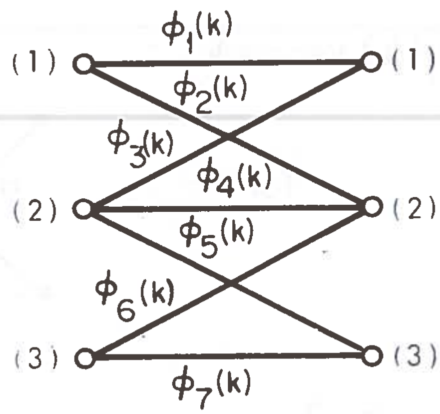

Figure 2.3.9-1 FIRST TYPE OF SUBNETWORK

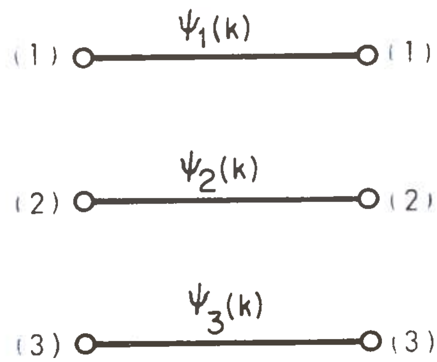

Figure 2.3.9-2 SECOND TYPE OF SUBNETWORK

as shown in Fig. 2.3.9-3.

The cost function for the subnetwork obtained when juxtaposing one of the type 1 and one of the type 2 is:

$$J_k(\underline{\phi}(k), \underline{\psi}(k)) = \frac{1}{2}\underline{\phi}^T(k)\underline{L}_1\underline{\phi}(k) + \underline{h}_1^T\underline{\phi}(k)$$

$$+ \frac{1}{2}\underline{\psi}^T(k)\underline{L}_2\underline{\psi}(k) + \underline{h}_2^T\underline{\psi}(k), \qquad (2.262)$$

where $\underline{L}_1$ and $\underline{h}_1$ come from the subnetwork of type 1, and $\underline{L}_2$ and $\underline{h}_2$ from that of type 2.

In the above equation, we may express $\underline{\psi}(k)$ in terms of $\underline{\phi}(k)$, by

$$\psi_1(k) = \phi_1(k) + \phi_3(k), \qquad (2.263)$$

$$\psi_2(k) = \phi_2(k) + \phi_4(k) + \phi_6(k), \qquad (2.264)$$

$$\psi_3(k) = \phi_5(k) + \phi_7(k), \qquad (2.265)$$

in the labeling of Figs. 2.3.9-1 and 2.3.9-2.

Thus, we can express the cost (2.262) in terms of $\underline{\phi}(k)$ only. The flow-conservation constraints are the same as before:

$$\underline{x}(k) = \underline{Y}\ \underline{\phi}(k), \qquad (2.266)$$

$$\underline{x}(k+1) = Z\ \underline{\psi}(k). \qquad (2.267)$$

In equation (2.267), we may express $\underline{\psi}(k)$ as a function of $\underline{\phi}(k)$ through (2.263), (2.264), and (2.265). Therefore, we shall have replaced the quasi-stationary network of Fig. 1.3.1 by a stationary one, where the typical sub-network is that of Fig. 2.3.9-1, and the cost function is given by (2.262) in terms of $\underline{\phi}(k)$ only. We generalize the procedure in the following theorem.

Definition 2.10

We call a network quasi-stationary if it consists of a succession of two alternating types of subnetworks, with the same number of entrance nodes, all subnetworks of the same type contributing identical terms in the cost function, and if moreover, only one link originates from each entrance node in the sub-networks of one of the two types (type 2). We denote by $\underline{Y}_1$ and $\underline{Z}_1$ the node-

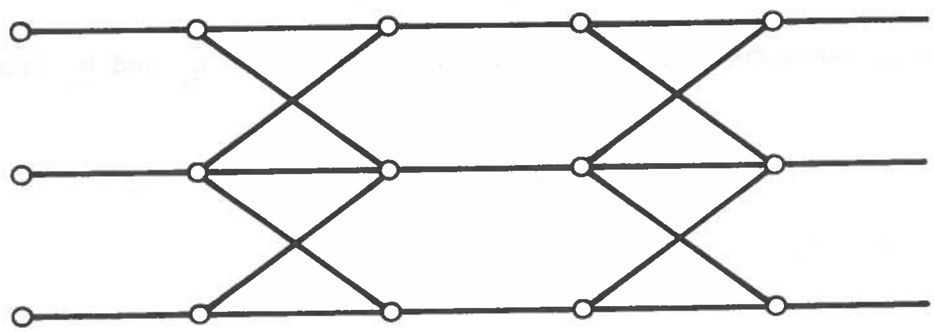

Figure 2.3.9-3 QUASI-STATIONARY NETWORK

link entrance-and-exit incidence matrices for the first type of subnetwork (that is, the type of the most upstream subnetwork), and by $\underline{Y}_2$ and $\underline{Z}_2$ the corresponding quantities for the second type. Similarly, $\underline{L}_1$, $\underline{h}_1$, $\underline{L}_2$, $\underline{h}_2$ are the cost coefficients corresponding to the two types.

Theorem 2.22

Let n be the common number of entrance or exit nodes in the two types in a quasi-stationary network. Then $\underline{Y}_2$ and $\underline{Z}_2$ are (nxn) invertible matrices, and the optimization problem is equivalent to that arising from a stationary network with incidences matrices $\underline{Y}$ and $\underline{Z}$ and cost elements $\underline{L}$ and $\underline{h}$, given by:

$$\underline{Y} = \underline{Y}_1, \tag{2.268}$$

$$\underline{Z} = \underline{Z}_2 \underline{Y}_2^{-1} \underline{Z}_1, \tag{2.269}$$

$$\underline{L} = \underline{L}_1 + \underline{Z}_1^T \underline{Y}_2^{T-1} \underline{L}_2 \ \underline{Y}_2^{-1} \underline{Z}_1, \tag{2.270}$$

$$\underline{h} = \underline{h}_1 + \underline{Z}_1^T \underline{Y}_2^{T-1} \ \underline{h}_2. \tag{2.271}$$

Remark

The matrices $\underline{Y}$ and $\underline{Z}$ (defined by (2.268) and (2.269)), are entrance-and-exit incidence matrices. Indeed, both $\underline{Z}_2$ and $\underline{Y}_2$ are permutation matrices, therefore multiplying $\underline{Z}$ by $\underline{Z}_2 \ \underline{Y}_2^{-1}$ does not change the property that each column contains exactly one 1 and zeroes. Therefore, we have demonstrated the complete equivalence between the optimization over the original quasi-stationary network and over the stationary network whose typical subnetwork is defined by $\underline{Y}$ and $\underline{Z}$. The transformation is not merely a mathematical artefact.

Proof

Let us denote by $\underline{\Psi}(k)$ the flows along the links of the type 2 subnetwork.

In the subnetworks of type 2, the number of links is equal to n since exactly one link originates from each entrance, so that $\underline{Y}_2$ and $\underline{Z}_2$ are (nxn) matrices. It is shown in section 2.2 that they have rank n, so that they are invertible.

The cost corresponding to the two adjacent subnetwork, one of type 1 and one of type 2 downstream from the previous one, is

$$J_k(\underline{\phi}(k), \underline{\psi}(k)) = (\frac{1}{2} \underline{\phi}^T(k) \underline{L}_1 \underline{\phi}(k) + \underline{h}_1^T \underline{\phi}(k))$$
$$+ (\frac{1}{2} \underline{\psi}^T(k) \underline{L}_2 \underline{\psi}(k) + \underline{h}_2^T \underline{\psi}(k)). \qquad (2.272)$$

From

$$\underline{Z}_1 \underline{\phi}(k) = \underline{Y}_2 \underline{\psi}(k), \qquad (2.273)$$

it follows that

$$\underline{\psi}(k) = \underline{Y}_2^{-1} \underline{Z}_1 \phi(k), \qquad (2.274)$$

which is the general expression corresponding to equations (2.263), (2.264), (2.265) in the case of Fig. 2.3.9-3.

If we substitute the right-hand side of (2.274) for $\underline{\psi}(k)$, we have:

$$J_k(\underline{\phi}(k)) = \frac{1}{2} \underline{\phi}^T(k)(\underline{L}_1 + \underline{Z}_1^T \underline{Y}_2^{T-1} \underline{L}_2 \underline{Y}_2^{-1} \underline{Z}_1)\underline{\phi}(k)$$
$$+ (\underline{h}_1 + \underline{Z}_1^T \underline{Y}_2^{T-1} \underline{h}_2)^T \underline{\phi}(k),$$

which is the expression that would be obtained for the cost in a stationary network with $\underline{L}$ and $\underline{h}$ given by (2.270) and (2.271).

Also,

$$\underline{x}(k) = \underline{Y}_1 \underline{\phi}(k),$$
$$\underline{x}(k+1) = \underline{Z}_2 \underline{\psi}(k) = \underline{Z}_2 \underline{Y}_2^{-1} \underline{Z}_1 \underline{\phi}(k),$$

according to (2.274), so that the flow-conservation constraints are those of a stationary network with incidence matrices $\underline{Y}$ and $\underline{Z}$ given by (2.272) and (2.273).

Example

Let us consider another quasi-stationary network (Fig. 2.3.9-4). In this example,

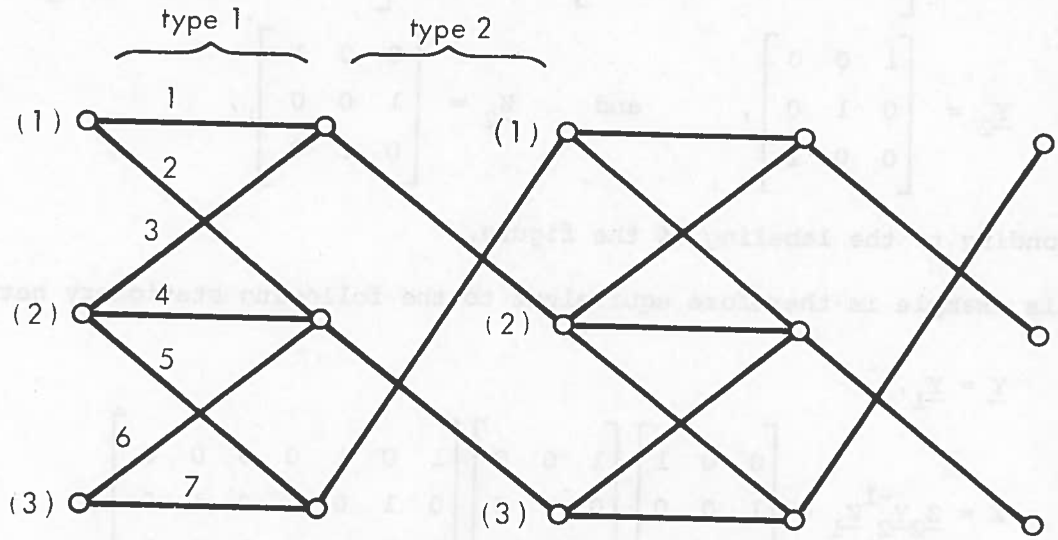

Figure 2.3.9-4 OTHER EXAMPLES OF QUASI-STATIONARY NETWORK

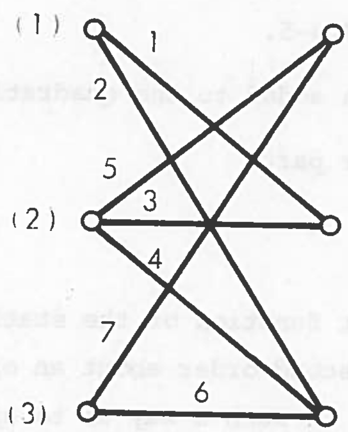

Figure 2.3.9-5 TYPICAL SUBNETWORK FOR EQUIVALENT STATIONARY NETWORK

$$\underline{Y}_1 = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} ; \quad \underline{Z}_1 = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 \end{bmatrix}$$

$$\underline{Y}_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \text{and} \quad \underline{Z}_2 = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix},$$

corresponding to the labeling of the figure.

This example is therefore equivalent to the following stationary network:

$$\underline{Y} = \underline{Y}_1,$$

$$Z = \underline{Z}_2\underline{Y}_2^{-1}\underline{Z}_1 = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 \end{bmatrix},$$

$$= \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 \end{bmatrix},$$

which is represented in Fig. 2.3.9-5.

The term $\underline{Z}_1^T \underline{L}_2 \underline{Z}_1$ has to be added to the quadratic part of the cost function, and $\underline{h}_2^T\underline{Z}_1$ to the linear part.

2.4 CONCLUDING REMARKS

We have shown that the cost function of the static optimization problem [2] can be expanded up to the second order about an optimal solution, and the network split into subnetworks, in such a way as to apply dynamic programming. Although our problem is a static one, we use dynamic programming by taking a discrete distance as stage variable. In this way, we obtain sequentially the approximate optimal downstream perturbations caused in the network by an initial perturbation. This procedure is valid for any network and easy to implement. In the case of a long stationary network, that is, one that consists of identical subnetworks contributing identical cost functions, we have completely characterized the behavior of the perturbations. The main assumption which guarantees our result is a controllability property, which can be checked either algebraically or geometrically, and is closely related to the arrangement of the links. When it holds (and two minor technical assumptions also hold), any initial perturbation with a nonzero total flow gives rise to a sequence of downstream perturbations which converge exponentially to a distribution which does not depend on the particular initial perturbation but is a characteristic of the network. We can determine that distribution, as well as bounds for the magnitudes of the deviations of the downstream perturbations from that distribution, as functions of the distance.

When the initial perturbation does not change the total arriving flow, the downstream perturbations converge exponentially to zero.

In a special network, we discuss the bounds as functions of a cost parameter, and see that the controllability assumption is met except for limiting values of that parameter.

We also show that we can apply these results to a class of nonstationary networks.

3. UPSTREAM PERTURBATIONS IN A FREEWAY CORRIDOR NETWORK

3.1 PHYSICAL DESCRIPTION OF THE PROBLEM

In the previous section, we have been interested in the downstream effects of a change in the incoming traffic pattern in the composite network of Fig. 2.2.2. We have posed the question of how the traffic assignment have to be optimally modified downstream when the incoming traffic changes. In particular, when the traffic change does not affect the total incoming flow, we have established, under certain assumptions, that the optimal downstream modifications in traffic distribution resulting from new optimal assignments decrease geometrically.

In this section, we ask a different question. Suppose an accident, or any other sudden event, occurs at subnetwork (N-k). This has the effect of shifting the traffic entering subnetwork (N-k+1) from some entrances to others. As far as subnetworks (N-k+1), (N-k+2),... are concerned, our previous analysis is applicable. However, to answer the question of how the assignments are to be changed in the "upstream" subnetworks N-k, N-k-1,...,2,1 to preserve optimality, in spite of the sudden external change in (N-k), is a different matter.

The resulting optimal variations in the traffic distribution, among the entrances of those subnetworks located upstream with respect to (N-k), are called the upstream perturbations (Fig. 3.1.1). The upstream perturbations at stage 1 are constrained to be zero since the incoming traffic is also an externally specified parameter.

We now show that, provided subnetwork (N-k) is sufficiently far from the farthest upstream network (i.e., 1), the upstream perturbations decrease geometrically, away from (N-k), in quite the same way as the downstream ones under similar assumptions. Physically, the expression "infinitely far" means "far enough for all converging quantities to have practically reached their limit." According to our numerical experience, this amounts to 5 to 10 subnetworks, even if considerable accuracy is required.

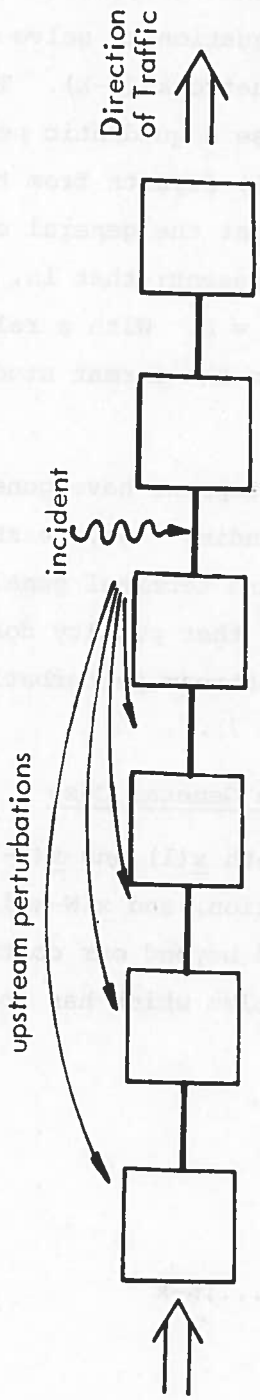

Figure 3.1.1 UPSTREAM PERTURBATIONS

3.2 MATHEMATICAL TREATMENT

For our present purposes, only the costs corresponding to the subnetworks 1,2,...(N-k) are of interest. Therefore, in section 3.2.1, we shall apply dynamic programming and set up a forward Bellman equation to solve the problem of optimal assignment with the traffic exiting subnetwork (N-k). To meet the requirement that $\underline{x}(1)$ is prescribed, we shall impose a quadratic penalty which makes the total cost exceedingly large if $\underline{x}(1)$ departs from the specified value, zero. However, for generality, we shall treat the general case when also a perturbation $\underline{\xi}$ in incoming traffic may be present; that is, $\underline{x}(1) = \underline{\xi}$. If only the incident at subnetwork (N-k) occurs, $\underline{\xi} = 0$. With a relabeling of the subnetwork, the cost function will be cast into the format studied in section 2.2.5, except for the terminal penalty.

In section 3.2.2, we will show that all the steps we have gone through in the previous section can be repeated "mutatis mutandis." So, we shall be able to apply the results of section 2.3.7 with a nonzero terminal penalty. And we shall see that, asymptotically, the presence of that penalty does not invalidate the conclusions then reached for the downstream perturbations. This is confirmed by the numerical experiments (section 7).

3.2.1 Solution by Dynamic-Programming in the General Case

Figure 3.2.1-1 illustrates the situation. Both $\underline{x}(1)$ and $\underline{x}(N-k+1)$ are now specified: $\underline{x}(1)$ is the incoming traffic perturbation, and $\underline{x}(N-k+1)$ is the traffic perturbation provoked by the accident, and beyond our control.

Under these conditions, the minimization problem which has to be solved is

Problem P_2

$$\min_{\underline{\phi}_i} \sum_{i=1}^{N-k} J_i(\underline{\phi}_i) \ , \tag{3.1}$$

subject to:

$$\left.\begin{aligned} \underline{x}(i) &= \underline{Y}(i)\ \underline{\phi}(i) \quad , \quad (3.2) \\ \underline{x}(i) &= \underline{Z}(i)\ \underline{\phi}(i) \quad , \quad (3.3) \end{aligned}\right\} \quad i = 1,\dots,N-k$$

and with boundary conditions:

$$\left.\begin{aligned} \underline{x}(i) &= \underline{\xi} \quad , \quad (3.4) \\ \underline{x}(N-k+1) &= \underline{\eta} \quad , \quad (3.5) \end{aligned}\right\}$$

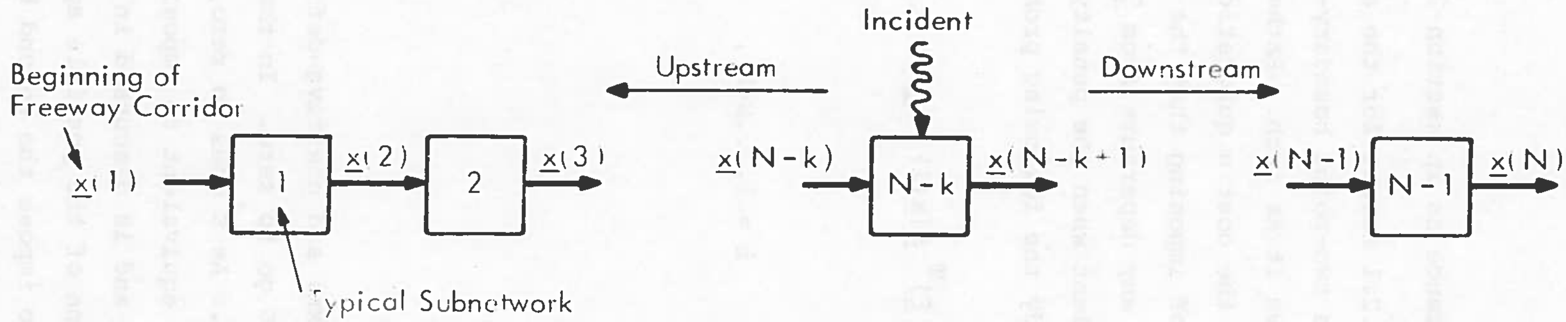

Figure 3.2.1-1 INCIDENT OCCURRING IN SUBNETWORK (N-k)

where

$$J_i(\underline{\phi}_i) = \frac{1}{2} \underline{\phi}_i^T \underline{L}_i \underline{\phi}_i + \underline{h}_i^T \underline{\phi}_i \quad , \tag{3.6}$$

and all the coefficients have the same significance as in section 2.2.

Problem P_2 is the same as problem P_1 of 2.2.1 except for the additional boundary condition (3.5). Problem P_2 is thus a two-point boundary-value optimization problem. However, we shall not solve it as such. Rather, we shall keep only 3.5 as boundary condition and add to the cost a quadratic penalty function on $(\underline{x}(1)-\underline{\xi})$. In other words, instead of imposing that the perturbation $\underline{x}(1)$ be exactly $\underline{\xi}$, we penalize strongly any departure from $\underline{\xi}$ through the cost function. Both approaches are equivalent when the penalty parameter becomes very large [26]. Specifically, we study the following problem:

<u>Problem P_3</u>

$$\min_{\underline{\phi}(i)} \sum_{i=1}^{N-k} J_i(\underline{\phi}(i)) + \frac{1}{\varepsilon} (\underline{x}(1) - \underline{\xi})^T \underline{S}(\underline{x}(1) - \underline{\xi}) . \tag{3.7}$$

Subject to: $\underline{x}(i) = \underline{Y}(i) \underline{\phi}(i)$,

$i = 1,\ldots N-k$,

$\underline{x}(i+1) = \underline{Z}(i) \underline{\phi}(i)$,

and with boundary condition:

$$\underline{x}(N-k+1) = \underline{\eta} .$$

In equation (3.7), we have added the term

$$\frac{1}{\varepsilon} (\underline{x}(1) - \underline{\xi})^T \underline{S}(\underline{x}(1) - \underline{\xi})$$

to the cost in problem P_2. The matrix $\underline{S}$ is (nxn) and positive-definite; ε is a positive scalar. We intend later on to let ε go to zero. In that manner, the cost becomes exceedingly large if $\underline{x}(1) \neq \underline{\xi}$. As ε goes to zero, the presence of the penalty function in (3.7) becomes equivalent to imposing the boundary condition (3.4). This method is classical and is discussed in [26] for continuous-time systems. It is however only one of the possible approaches to problem P_2. Another approach [26] is merely to impose the second boundary condition as such rather than replacing it by a quadratic penalty.

Solving problem (P_3) can be achieved by solving the following sequence of recursive forward Bellman equations:

$$V_{i+1}(\underline{x}(i+1)) = \min_{\underline{\phi}_i} [J_i(\underline{\phi}(i)) + V_i(\underline{x}(i))] \quad , \tag{3.8}$$

<u>subject to:</u> $\underline{Z}(i)\underline{\phi}(i) = \underline{x}(i+1)$,

and where $\underline{x}(i) = \underline{Y}(i)\underline{\phi}(i)$,

for i = N-k, N-k-1,... 2, 1 and with initial condition,

$$V_1(\underline{x}) = \frac{1}{\varepsilon} (\underline{x} - \underline{\xi})^T \underline{S}(\underline{x} - \underline{\xi}).$$

These equations (3.8) are similar to equations (2.36), (2.37) of 2.2.5 but differ from them in several respects. At stage i, $\underline{x}(i+1)$ is specified and the constraint is $Z(i)\underline{\phi}(i) = \underline{x}(i+1)$. By $\underline{x}(i) = \underline{Y}(i)\underline{\phi}(i)$, we just mean that $\underline{Y}(i)\underline{\phi}(i)$ has to be substituted for $\underline{x}(i)$ as the argument of V_i in (3.8). We obtain successively V_1, V_2,...V_{N-k+1} and the corresponding controls in closed loop. Next, we obtain successively $\underline{x}(N-k)$, $\underline{x}(N-k-1)$,...$\underline{x}(2)$, $\underline{x}(1)$ from $\underline{x}(N-k+1) = \underline{\eta}$.

<u>Notation</u>

A simple change in the labeling of the subnetworks allows us to make equations (3.8) resemble equations (2.36) of 2.2.5.

a. We replace i by N-i as the new label of the subnetwork previously labeled i. So the corresponding incidence matrices, previously called $\underline{Y}(i)$ and $\underline{Z}(i)$, become $\underline{Y}(N-i)$ and $\underline{Z}(N-i)$.

b. The vector $\underline{x}(i)$ becomes $\underline{x}(N-i+1)$.

This change of notation is illustrated in Fig. 3.2.1-2. In this new notation, problem P_3 (which is an approximation to ptoblem P_2; i.e., the closer ε is to zero, the better the approximation) can be formulated as:

<u>Problem P_3</u>

$$\min_{\underline{\phi}} \sum_{i=k}^{N-1} J_i(\underline{\phi}_i) + \frac{1}{\varepsilon} (\underline{x}(N) - \underline{\xi})^T \underline{S}(\underline{x}(N) - \underline{\xi}) \quad , \tag{3.9}$$

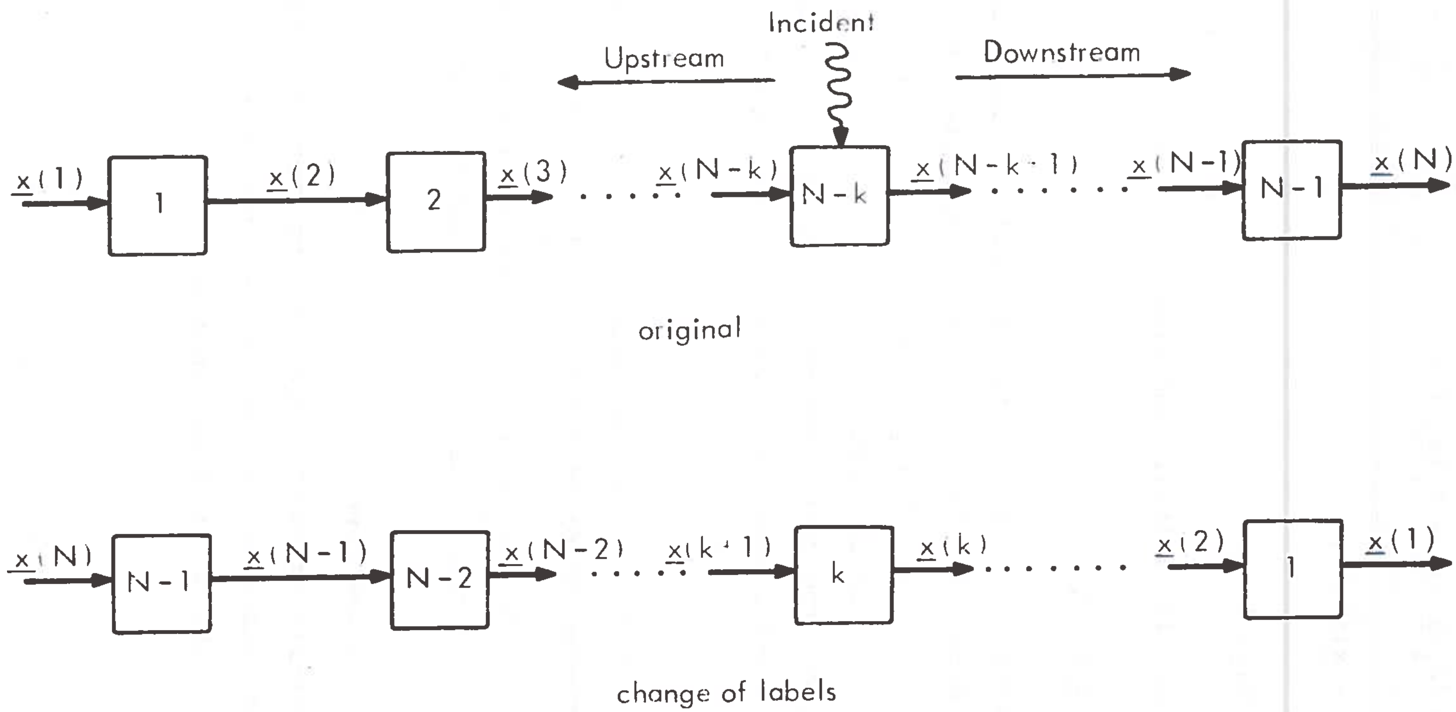

Figure 3.2.1-2 RELABELING FOR UPSTREAM PERTURBATIONS

subject to: $\underline{x}(i+1) = \underline{Y}(i)\underline{\phi}(i)$,

$i = k, k+1, \ldots N-1$,

$\underline{x}(i) = \underline{Z}(i)\underline{\phi}(i)$,

and with boundary condition: $\underline{x}(k) = \underline{\eta}$.

Comparing the expression (3.9) of problem P_3 with problem P_1 of section 2.2.5, we see that:

1) A quadratic penalty term has been added to the cost function.

2) At each stage, the $\underline{Y}(i)$ and $\underline{Z}(i)$ matrices are interchanged.

The corresponding sequence of Bellman equations is

$$V_i(\underline{x}(i)) = \min_{\underline{\phi}(i)} \Big(J_i(\underline{\phi}(i)) + V_{i+1}(\underline{x}(i+1))\Big), \tag{3.10}$$

subject to: $\underline{Z}(i)\underline{\phi}(i) = \underline{x}(i)$,

and where: $\underline{x}(i+1) = \underline{Y}(i)\underline{\phi}(i)$,

for $i = k, k+1, \ldots N-1$ and with initial condition,

$$V_N(\underline{x}) = \frac{1}{\varepsilon}(\underline{x} - \underline{\xi})^T S(\underline{x} - \underline{\xi}).$$

Again, the two differences between equations (2.36), (2.37) of section 2.2.5 and equations (3.10) are the interchange $\underline{Z}(i)$ and $\underline{Y}(i)$ and the initial condition for the value function.

<u>Solution to Problem P_3</u>

The solution to the Bellman equations is quadratic: all the calculations of section 2.2.5 can be repeated, simply interchanging $\underline{Z}(i)$ and $\underline{Y}(i)$, and will yield the corresponding results, allowing for that correction. Only the initial values of the parameters $\underline{C}(i)$, $\underline{b}(i)$, and $a(i)$ will change. We shall keep exactly the same notation as in the study of downstream perturbations because the two sets of coefficients (for the downstream and upstream perturbations, respectively) will never appear together. We formulate the corresponding results in the following theorem.

Theorem 3.1

a) The value function at stage i, given recursively by (3.10), is quadratic:

$$V_i(\underline{x}) = \frac{1}{2} \underline{x}^T \underline{C}(i) \underline{x} + \underline{b}^T (i) \underline{x} + a(i) , \tag{3.11}$$

where $\underline{C}(i) > 0$. In particular, the value function at stage N is given by

$$V_N(\underline{x}) = \frac{1}{\varepsilon} (\underline{x} - \underline{\xi})^T \underline{S}(\underline{x} - \underline{\xi}) = \frac{1}{\varepsilon} \underline{x}^T \underline{S} \underline{x} - \frac{2}{\varepsilon} \underline{\xi}^T \underline{S} \underline{x} + \frac{1}{\varepsilon} \underline{\xi}^T \underline{S} \underline{\xi}. \tag{3.12}$$

b) The coefficient $\underline{C}(i)$, $\underline{b}(i)$, and $a(i)$ are obtained recursively as follows:

$$\underline{C}(i)^{-1} = \underline{Z}(i) M(i)^{-1} \underline{Z}(i)^T , \tag{3.13}$$

$$\underline{C}(N) = \frac{1}{\varepsilon} \underline{S} , \tag{3.14}$$

$$\underline{b}(i) = \underline{C}(i) \underline{Z}(i) \underline{M}(i)^{-1} (\underline{h}(i) + \underline{Y}(i)^T \underline{b}(i+1)) , \tag{3.15}$$

$$\underline{b}(N) = - \frac{2}{\varepsilon} \underline{S} \underline{\xi} , \tag{3.16}$$

$$a(i) = a(i+1) + \frac{1}{2} (\underline{h}(i) + \underline{Y}(i)^T \underline{b}(i+1))^T (\underline{M}(i)^{-1} \underline{Z}(i)^T \cdot$$

$$\cdot \underline{C}(i) \underline{Z}(i) M(i)^{-1} - \underline{M}(i)^{-1}) (\underline{h}(i) + \underline{Y}(i)^T \underline{b}(i+1)) , \tag{3.17}$$

$$a(N) = \frac{1}{\varepsilon} \underline{\xi}^T \underline{S} \underline{\xi} , \tag{3.18}$$

where $$\underline{M}(i) = \underline{L}(i) + \underline{Y}(i)^T \underline{C}(i+1) \underline{Y}(i). \tag{3.19}$$

c) The optimal flow perturbation at stage i is given in closed loop by

$$\underline{\phi}^* (i) = \underline{M}(i)^{-1} \underline{Z}^T(i) \underline{C}(i) \underline{x}^* (i) + (\underline{M}(i)^{-1} \underline{Z}(i)^T \underline{C}(i) \underline{M}(i)^{-1}$$

$$- \underline{M}(i)^{-1}) (\underline{h}(i) + \underline{Y} (i)^T \underline{b}(i+1)). \tag{3.20}$$

Proof

Taking into account (3.12) and the previous remarks, we just have to interchange $\underline{Y}(i)$ and $\underline{Z}(i)$ in equations (2.32) of section 2.2.5 to obtain the corresponding recursive equations for the coefficients. The proof that $\underline{C}(i) > 0$ is carried out exactly in the same way as in section 2.2.5, using now the fact that $\underline{Z}\ \underline{L}^{-1}\ \underline{Z}^T$ is invertible (instead of $\underline{Y}\ \underline{L}^{-1}\ \underline{Y}^T$). We do not then use explicitly the fact that $\underline{C}(N)$ is zero, but just the property $\underline{C}(N) \geq 0$, which still holds under (3.14). Q.E.D.

We can now state the main results about upstream perturbations in the following corollary.

Corollary 3.2 The sequence of optimal upstream perturbations, $\underline{x}(k+1)$, $\underline{x}(k+2), \ldots$, is given recursively by:

$$x^*(i+1) = \underline{D}(i)\ \underline{x}^*(i) \quad , \tag{3.21}$$

$$\underline{x}(k) = \underline{\eta} \quad , \tag{3.22}$$

where $\underline{\eta}$ is the initial perturbation, and

$$\underline{D}(i) = \underline{Y}(i)\ \underline{M}(i)^{-1}\ \underline{Z}(i)^T\ \underline{C}(i). \tag{3.23}$$

and the sequence $\underline{C}(i)$ is recursively obtained from equations (3.13) and (3.14).

Proof

By the same reasoning as done in section 2.2.5 (i.e., because we start with an optimal solution), the constant term in (3.20) is zero, so that $\underline{x}^*(i+1) = \underline{Y}(i)\ \phi^*(i) = \underline{Y}(i)\ \underline{M}(i)^{-1}\ \underline{Z}(i)^T\ \underline{C}(i)$.

Remark. In the limit, when $\varepsilon \to 0$, the sequence $\underline{D}(i)$ becomes such that $\underline{x}^*(N) = \underline{\xi}$, or in other words, $\underline{D}(i) = \underline{D}(i, \varepsilon)$ and $\underline{x}^*(i) = x^*(i, \varepsilon)$. The true upstream perturbations are in fact $\lim_{\varepsilon \to 0} \underline{x}^*(i, \varepsilon)$. And $\lim_{\varepsilon \to 0} x^*(N, \varepsilon) = \underline{\xi}$. For a given $\varepsilon > 0$, the perturbations $\underline{x}^*(i, \varepsilon)$ correspond to an incoming flow perturbation of $\underline{x}^*(N, \varepsilon)$. Since the specified incoming flow perturbation is in fact $\underline{\xi}$ (in particular, $\underline{\xi} = 0$ if there is no such perturbation), it is only for $\varepsilon \to 0$ that the sequence $\underline{x}^*(i, \varepsilon)$ corresponds to the real traffic situation.

We have thus solved completely the problem of upstream perturbations in the general, nonstationary case. (Numerically, taking ε very small is a good approximation). We shall now investigate asymptotic properties, in the special stationary case.

3.2.2 Asymptotic Sensitivity Analysis in the Stationary Case

We can now again perform step by step the analysis of the previous section, interchanging everywhere the matrices $\underline{Y}$ and $\underline{Z}$.

We reduce the states and controls to (n-1)-dimensional states and (m-n) dimensional controls, as indicated in section 2.3.3.

Controllability of a reduced system is reduced to questions about the accessibility graph G corresponding to a traffic flow from the right to the left hand in Fig. 3.2.1-1 if we take the definition of G, in 2.3.4, literally (since $\underline{Y}$ and $\underline{Z}$ have been interchanged). Alternatively, we may now define G as follows: arc (i, j) exists if, and only if, there is a path from exit i to entrance j in the subnetwork (without taking the arrows into account). Let us denote this new graph by G(u) (for upstream) and the previous one by G(d) (for downstream). Figure 3.2.2 illustrates the definition of G(u) in an example.

In the reduced state/reduced control formulation, equation (2.150) is still valid:

$$\underline{K}(i) = \underline{C}_{11}(i) - \underline{C}_{12}(i)\,\underline{\nu}_{n-1}^{T} - \underline{\nu}_{n-1}C_{21}(i) + C_{22}(i)\,\underline{\nu}_{n-1}\underline{\nu}_{n-1}^{T}. \qquad (3.24)$$

Therefore, $\underline{K}(i) > 0$ since $\underline{C}(i) > 0$. (The proof is the same as in lemma 2.9.) It also follows from the terminal condition (3.14) that

$$\underline{K}(N) = \frac{1}{\varepsilon}(\underline{S}_{11} - \underline{S}_{12}\,\underline{\nu}_{n-1}^{T} - \underline{\nu}_{n-1}S_{21} + \underline{S}_{22}\,\underline{\nu}_{n-1}\underline{\nu}_{n-1}^{T}) \triangleq \frac{1}{\varepsilon}\underline{\Sigma}\,, \qquad (3.25)$$

where $\underline{S}$ has been partitioned like $\underline{C}(i)$.

The sequence $\underline{K}(i)$ satisfies equations (2.133) and converges to a positive semi-definite solution $\hat{\underline{K}}$ of (2.136) if the reduced system is controllable. Indeed, the proof given in [21] does not use the fact that $\underline{K}(N) = 0$ but only the property $\underline{K}(i) \geq 0$.

Equations (3.13) and (3.23) show that $\underline{D}(i)$ and $\underline{C}(i)$ do not depend on the sequence $\{\underline{h}(i)\}$ in the cost function. Therefore, we may choose any sequence $\{\underline{h}(i)\}$ that we please so as to make the equations for a(i) and $\underline{b}(i)$ the simplest. We need not choose all the $\underline{h}(i)$ the same.

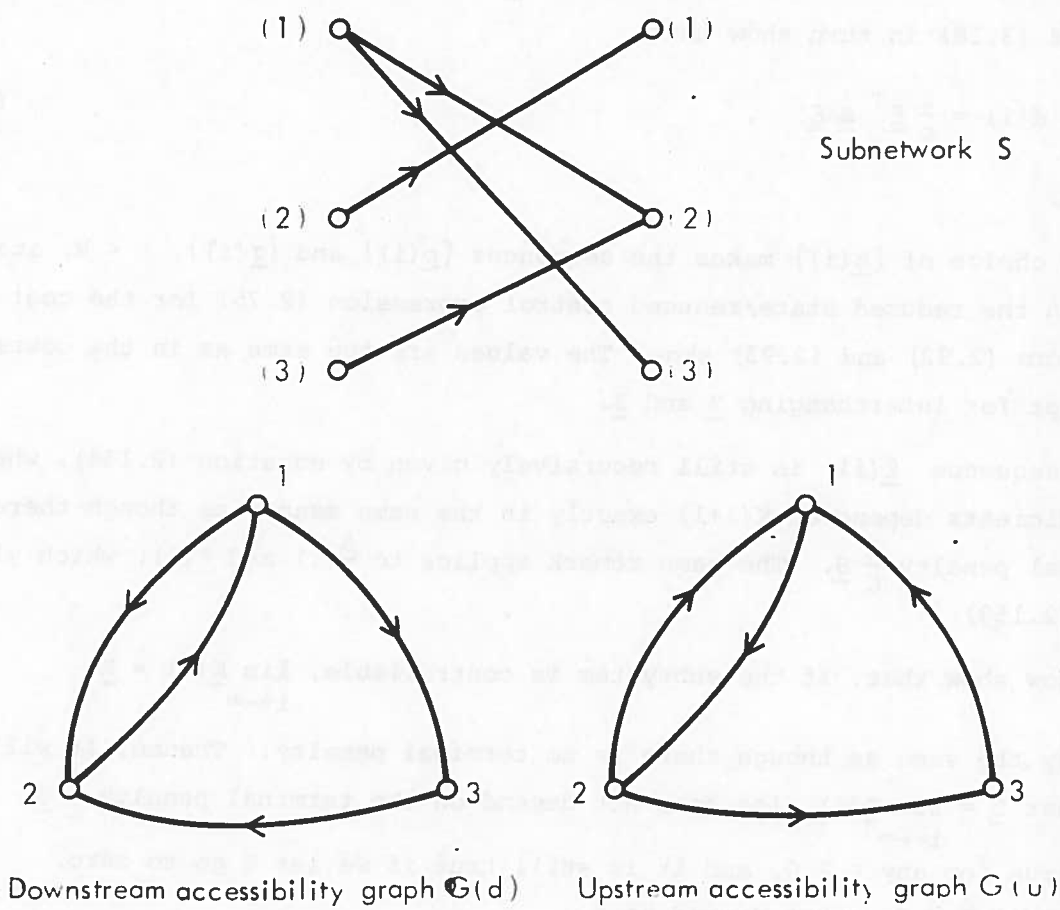

Figure 3.2.2 DOWNSTREAM AND UPSTREAM ACCESSIBILITY GRAPHS FOR ONE SAME SUBNETWORK

If we choose

$$\underline{h}(N-1) = \frac{2}{\varepsilon} \underline{Y}^T_{(N-1)} \underline{S}\, \underline{\xi} \quad , \tag{3.26}$$

and $\quad \underline{h}(i) = 0 \quad \text{for } i=1,\ldots N-2 \quad ,$

then equations (3.15) and (3.16) show that $\underline{b}(i) = 0$ for all $i < N$, and equations (3.17) and (3.18) in turn show that

$$a(i) = \frac{1}{\varepsilon} \underline{\xi}^T \underline{S}\, \underline{\xi} \quad , \tag{3.27}$$

for all i.

This choice of $\{\underline{h}(i)\}$ makes the sequences $\{\underline{p}(i)\}$ and $\{\underline{q}(i)\}$, $i < N$, stationary in the reduced state/reduced control expression (2.75) for the cost as equations (2.92) and (2.93) show. The values are the same as in the downstream case except for interchanging $\underline{Y}$ and $\underline{Z}$.

The sequence $\underline{\ell}(i)$ is still recursively given by equation (2.134), where the coefficients depend on $\underline{K}(i+1)$ exactly in the same manner as though there were no terminal penalty $\frac{1}{\varepsilon}\underline{S}$. The same remark applies to $\underline{\Delta}(i)$ and $\underline{\pi}(i)$, which yield $\underline{D}(i)$ by (2.159).

We now show that, if the subsystem is controllable, $\lim_{i\to-\infty} \hat{\underline{K}}(i) = \hat{\underline{K}}$ is exactly the same as though there is no terminal penalty. Thence, it will follow that $\underline{D} = \lim_{i=-\infty} \underline{D}(i)$ also does not depend on the terminal penalty $\frac{1}{\varepsilon}\underline{S}$. This is true for any $\varepsilon > 0$, and it is still true if we let ε go to zero.

$\lim_{i\to-\infty} \underline{K}(i) \stackrel{\Delta}{=} \hat{\underline{K}}$ does not depend on $\underline{S}$.

In section 2.3.7, we display an expression for $\underline{K}(i)$ (equation (2.214)) in the general linear-quadratic optimal control problem with quadratic terminal penalty. This is a result of Vaughan, which we can use provided that the problem is equivalent to one with a diagonal cost function. The diagonalization is possible if the assumptions A_2 (equation (2.199)) of theorem 2.15 are met. Those assumptions still hold for the upstream problem:

1) $\underline{S} \geq 0$ by assumption.

2) $\underline{R} > 0$ comes directly from the expression 2.91 of $\underline{R}$, where $\underline{Y}$ has been replaced by $\underline{Z}$.

3) $\underline{Q} - \underline{M}\,\underline{R}^{-1}\,\underline{M}^T \geq 0$ because, if $\underline{S} = 0$, then $\underline{Q} - \underline{M}\,\underline{R}^{-1}\,\underline{M}^T = \underline{K}(N-1) > \underline{0}$. It does not matter if $\underline{S}$ is zero or not in our particular problem because we know that if $\underline{S}$ is equal to zero, the corresponding $\underline{C}(N-1)$ and $\underline{K}(N-1)$ matrices are positive-definite as well, and in that case, one has: $\underline{Q} - \underline{M}\,\underline{R}^{-1}\,\underline{M}^T = \underline{K}(N-1)$.

4) $Q = (Q - M R^{-1} M^T) + M R^{-1} M^T \geq 0$.

5) Since $\underline{\Delta}(N-1) = \underline{A} - \underline{B}\,\underline{R}^{-1}\,\underline{M}$ and $\det(\underline{D}(N-1) = \det(\underline{\Delta}(N-1))$, it follows that $(\underline{A} - \underline{B}\,\underline{R}^{-1}\,\underline{M}$ is invertible if, and only if, $\underline{D}(N-1)$ is invertible. Now,

$$\underline{D}(N-1) = (\underline{Y}\,L^{-1}\,\underline{Z}^T)\,(\underline{Z}\,L^{-1}\,\underline{Z}^T)^{-1}, \tag{3.28}$$

for the upstream case, so that the invertibility of $\underline{A} - \underline{B}\,\underline{R}^{-1}\,\underline{M}$ is equivalent to the invertibility of $(\underline{Y}\,\underline{L}^{-1}\,\underline{Z}^T)$, which is the same condition as for the downstream case.

We may therefore use Vaughan's method under the assumptions of theorem 2.20.

From equation (2.214), we notice that the penalty matrix $\underline{S}$ occurs only in $\underline{T}$. However, $\lim_{j\to+\infty} \underline{F}(j) = 0$ (at least in the sense of Cesaro) whatever $\underline{T}$. Consequently the result of corollary 2.18 is still valid for any $\underline{S}$; i.e., if $\lim_{j\to\infty} \underline{K}(j)$ exists, the limit $\hat{\underline{K}}$ is given by equation (2.215). Therefore, theorem 2.20 and collary 2.21 are valid "mutatis mutandis:" $\underline{Y}$ and $\underline{Z}$ are to be interchanged throughout.

Remarks

Remark 1. We do not claim that $\underline{K}(i)$ and $\underline{D}(i)$ are independent of $\underline{S}$ for every i, but only that their limiting values are. Letting i go to $-\infty$, or N go to $+\infty$, is equivalent to requiring that the incident take place far from the entrance of the network at least far enough, so that $\underline{K}(i) \simeq \hat{\underline{K}}$ and $\underline{D}(i) \simeq \underline{D}$. Numerically, we have observed the independence of $\underline{D}$ and $\hat{\underline{K}}$ on $\underline{S}$, trying various penalty matrices $\underline{S}$. However, when ε becomes very small, numerical inaccuracy arises. It seems plausible that the value of the penalty $\frac{\underline{S}}{\varepsilon}$ affects the rate of convergence of K(i) and D(i).

Remark 2. One may wonder if the original system and the one derived from it by interchanging the matrices Y and Z both behave in the same way with respect to controllability. One of those two systems may contain a controllable (n-1)-dimensional sybsystem, while the other one does not. According to section 2.3.4, this question is reduced to a comparison between the graphs G(d) and G(u). The graphs G(d) and G(u) can be derived from one another by reversing the arrows*. Indeed, the existence of an arc (i, j) in G(d) means that of a path from entrance i to exit j in the subnetwork, which is equivalent to that of arc (j,i) in G(u). The properties of uniqueness of a final class as well as its aperiodicity are not in general preserved by reversing the arrows. However, our graphs G are not arbitrary graphs: they are built in a specific manner. If G(d) is strongly connected, so is G(u). Indeed, given nodes i and j, there will be a path from i to j in G(u) since we know there is one from j to i in G(d).

Moreover, if G(d) and G(u) are strongly connected, they both have the same period. Indeed, the period only depends on the lengths of all the cycles through a given node, and there is a one-to-one correspondence between those cycles in one graph and in the other, which preserves the length of the cycles. Each cycle in G(d) is indeed transformed into a cycle of same length in G(u) by reversing the arrows. Therefore, in view of theorem 2.5 on controllability, we have:

1) For a subnetwork of class 1

If G is strongly connected**, both the upstream and downstream dynamical systems give rise to controllable subsystems.

2) For a subnetwork of class 2

If G is strongly connected** and aperiodic, both the upstream and downstream dynamical systems give rise to controllable subsystems.

* In graph terminology, G(u) is called the dual graph of G(d)

** The strong connectedness implies that no node is transient. (Appendix A)

3.3 CONCLUDING REMARKS

We have displayed a method to calculate optimal <u>upstream</u> perturbations in a freeway network split into subnetworks. This method is dynamic programming over space. It is quite general and can be numerically implemented fairly easily just as in the case of downstream perturbations. In the special stationary case, we have shown that an analysis entirely comparable to that of section 2.3.7 for the downstream perturbations can be carried out. The main result is that the asymptotic propagation matrix for the upstream perturbations is the same as the asymptotic propagation matrix for the downstream perturbations in the same network, where the direction of traffic has been reversed (i.e., where the Y and Z matrices are interchanged). Consequently, the upstream perturbations provoked by an initial perturbation with zero total flow decrease exponentially with the distance from the incident site, when this takes place far enough from the entrances, and do so when the same connectivity properties hold as for the downstream perturbations.

4. CHARACTERIZATION OF NETWORK STEADY-STATE CONSTANTS BY A VARIATIONAL PROPERTY

4.1 INTRODUCTION

In a stationary network, we have proved, both for the downstream perturbations (section 2) and for the upstream perturbations (section 3), the following behavior. We assume the imposed perturbation occurs sufficiently far from the exits and the entrances of the network. Then, both the sequences of downstream and upstream perturbations converge (exponentially with the distance) to a stationary distribution, repeated from one subnetwork to the next. The upstream and downstream limiting distributions do not depend on the imposed perturbation, but are characteristics of the network (i.e., are determined by the structure and the costs). Also once the perturbation has redistributed itself among the entrance nodes according to the stationary distribution, the cost of the perturbation grows linearly with the distance. The rate of growth (the asymptotic cost increase rate per unit of flow) is also a characteristic of the network. In sections 2 and 3, we have expressed these steady-state characteristics through the solution of a Riccati equation, involving the matrix partitionings of section 2.3.3. Here we wish to relate those two steady-state characteristics by showing that the asymptotic cost increase rate satisfies a minimality property. At the same time, this approach allows the direct derivation of these characteristic quantities from the data of the network (structure and cost).

4.2 NOTATION AND PRELIMINARIES

We need to recall some notation from the previous sections, for we shall use it continuously in the sequel of this section. We shall concentrate on the downstream case in the proofs: as far as the upstream case is concerned, we know from section 3 that all the asymptotic results are obtained by interchanging the incidence matrices $\underline{Y}$ and $\underline{Z}$. We here make the assumptions of theorem 2.20: some subsystem of the dynamical system in $\underline{x}$ is controllable, the Hamiltonian matrix $\underline{H}$ of (2.210) has simple eigenvalues, and the matrix $\underline{Y}\,\underline{L}^{-1}\underline{Z}^T$ is invertible. We then know that the eigenvalue 1 of the limiting propagation matrix $\underline{D}$ is simple, because the matrix $\underline{\Delta}$ has no eigenvalue of magnitude 1 (2.190). Therefore, the stationary distribution $\underline{p}$ is uniquely defined by

$$\underline{D}\,\underline{p} = \underline{p}, \tag{4.1}$$

$$\underline{\nu}_n^T\,\underline{p} = 1. \tag{4.2}$$

As observed several times, the matrix $\underline{D}$ does not change if the linear terms $\underline{h}^T\,\underline{\Phi}_k$ are omitted in the cost function. When these linear terms are not present, the value functions are purely quadratic. The Bellman equation (2.36) of section 2 at stage i is therefore:

$$\frac{1}{2}\,\underline{x}^T\;\underline{C}(i)\,\underline{x} = \min_{\underline{\phi}}\,[\frac{1}{2}\,\underline{\phi}^T\,\underline{L}\,\underline{\phi} + \frac{1}{2}\,\underline{\phi}^T\,\underline{Z}^T\,\underline{C}(i+1)\;\underline{Z}\,\underline{\phi}], \tag{4.3}$$

$$\text{s.t.}\;\; \underline{Y}\,\underline{\phi} = \underline{x}, \tag{4.4}$$

where $\underline{C}(i+1)$ is given, and $\underline{C}(i)$ is derived from it by equation (4.3). The perturbation imposed on subnetwork i is $\underline{x}$. In the following, i will vary, but $\underline{x}$ is considered given and constant.

We denote the unique minimizing vector by $\underline{\Phi}^*(i)$. It is a linear function of $\underline{x}$ given by equation (2.43). By definition of the propagation matrix $\underline{D}(i)$, we have

$$\underline{Z}\,\underline{\Phi}^*(i) = \underline{D}\,(i)\;\underline{x}, \tag{4.5}$$

because, from (2.32), $\underline{Z\Phi}$ is the resulting perturbation after stage i. Therefore, we know that, in the minimization problem (4.3), the minimizing vector $\underline{\Phi}^*(i)$ satisfies the constraint (4.5). Accordingly, if that constraint is added to (4.4), the solution of the new minimization problem is still $\underline{\Phi}^*(i)$, and the minimal value is the same. Thus,

$$\frac{1}{2}\,\underline{x}^T\;C(i)\;\underline{x} = \min_{\underline{\Phi}}\,[\frac{1}{2}\,\underline{\Phi}^T\;L\;\underline{\Phi} + \frac{1}{2}\,\underline{\Phi}^T\underline{Z}^T\;\underline{C}(i+1)\;\underline{Z}\,\underline{\Phi}], \tag{4.6}$$

$$\text{s.t.}\;\; \underline{Y}\,\underline{\Phi} = \underline{x}, \tag{4.7}$$

$$\underline{Z}\,\underline{\Phi} = \underline{D}(i)\underline{x}, \tag{4.8}$$

and the solution to (4.6), (4.7), (4.8) is $\Phi^*(i)$. However, for all Φ's satisfying the constraint (4.8), the second term on the right-hand side of (4.6) is constant. Therefore, we can rewrite:

$$\frac{1}{2}\,\underline{x}^T\,\underline{C}(i)\,\underline{x} - \frac{1}{2}\,\underline{x}^T\,\underline{D}^T(i)\,\underline{C}(i+1)\,\underline{D}(i)\,\underline{x} = \min_{\underline{\Phi}}\ \frac{1}{2}\,\underline{\Phi}^T\,\underline{L}\,\underline{\Phi}] \tag{4.9}$$

$$\text{s.t. } \underline{Y}\,\underline{\Phi} = \underline{x},$$

$$\underline{Z}\,\underline{\Phi} = \underline{D}(i)\,\underline{x},$$

and the unique minimizing vector for this problem is $\underline{\Phi}^*(i)$.

Remark. Consider the sequence of problems (4.3) obtained by letting i vary and keeping the entering traffic flow distribution at stage i always equal to $\underline{x}$. For each value of i, the corresponding minimizing vector $\underline{\Phi}^*(i)$ satisfies constraints (4.7) and (4.8). On the other hand, according to the reduced state, reduced control formulation of section 2, $\underline{\Phi}^{*T}(i) = (\underline{u}^*(i), \underline{v}^*(i))$ and from equation (2.144), we know that $v^*(i)$ converges when $i \to -\infty$ because $\underline{K}(i)$ and $\underline{\ell}(i)$ converge under the hypotheses of theorem 2.20. Accordingly, $\underline{u}^*(i) = \underline{x} - \underline{\tilde{Y}}\,\underline{v}^*(i)$ (equation (2.99)) converges and so does $\underline{\Phi}^*(i)$. It is essential in this argument that $\underline{x}$ be kept fixed.

Let

$$\underline{\Phi}^* = \lim_{i \to -\infty} \underline{\Phi}^*(i). \tag{4.10}$$

Since $\lim_{i \to -\infty} \underline{D}(i) = \underline{D}$, we see that $\underline{\Phi}^*$ satisfies:

$$\underline{Y}\,\underline{\Phi}^* = \underline{x}, \tag{4.11}$$

$$\underline{Z}\,\underline{\Phi}^* = \underline{D}\,\underline{x}. \tag{4.12}$$

In particular, if we choose $\underline{x} = \underline{p}$ (defined by (4.1) and (4.2)) to obtain the sequence of optimization problems, the corresponding Φ^* will satisfy:

$$\underline{Y}\,\underline{\Phi}^* = \underline{p}, \tag{4.13}$$

$$\underline{Z}\,\underline{\Phi}^* = \underline{p}, \tag{4.14}$$

since $\underline{D}\,\underline{p} = \underline{p}$.

We shall prove (in lemma 4.2) that the fact that $\underline{\Phi}^*(i)$ is the solution to (4.6), (4.7) and (4.8) for all i implies that $\underline{\Phi}^*$ is the solution to the same problem with the limiting constraints (4.11) and (4.12). To do so, we need an approximation lemma.

4.3 APPROXIMATION LEMMA

Lemma 4.1. Let $\underline{x}$ be a vector in R^n. Given any positive scalar ε and any m-dimensional vector $\underline{\Phi}$ satisfying

$$\underline{Y}\,\underline{\Phi} = \underline{x},$$

$$\underline{Z}\,\underline{\Phi} = \underline{D}\,\underline{x},$$

there exists an integer i_o and a sequence of m-dimensional vectors $\hat{\underline{\Phi}}(i)$, such that:

i) $||\hat{\underline{\Phi}}(i) - \underline{\Phi}||$, for $i \le i_o$,

ii) $\underline{Y}\,\hat{\underline{\Phi}} = \underline{x}$,

iii) $\underline{Z}\,\hat{\underline{\Phi}} = \underline{D}(i)\,\underline{x}$.

Proof

Consider the 2n x m matrix $\left(\frac{\underline{Y}}{\underline{Z}}\right)$. It is the matrix of a linear mapping from R^m to R^{2n}. Let r be its rank. Let X_2 be its kernel(*) and X_1, the orthogonal complement(**) of X_2. Let W be the image space of the mapping. Therefore, X_1 and W are both of dimension r and the restriction of $\left(\frac{\underline{Y}}{\underline{Z}}\right)$ to X_1 is a one-to-one mapping from X_1 to W [40]. Let us denote by Ω that restriction. The mapping Ω can be represented by an rxr invertible matrix after vector bases in X_1 and W have been chosen. Every vector $\underline{f}$ in R^m is uniquely decomposed into $\underline{f} = \underline{f}_1 + \underline{f}_2$, with $\underline{f}_1 \in X_1$ and $\underline{f}_2 \in X_2$.

For every i, let $\underline{f}(i)$ be some vector that satisfies:

$$\underline{Y}\,\underline{f}(i) = \underline{x},$$

---

* By definition, the kernel of a linear mapping from R^s to R^t is the set of those vectors in R^s which are mapped to zero; i.e., if $\underline{f}: R^s \rightarrow R^t$ is a linear mapping, the kernel of f is

$$\text{Ker}\ \underline{f} = \{\underline{x} \in R^s : \underline{f}(\underline{x}) = 0\}.$$

** The orthogonal complement V_2 of a vector subspace V_1 in a vector space V with scalar product $<\underline{x},\underline{y}>$ is the set of those vectors which are orthogonal to every vector in V_1; i.e., $V_2 = \{\underline{x} \in V : <\underline{x},\underline{y}> = 0\ \forall\ \underline{y} \in V_1\}$; here, the euclidean scalar product is used: $<\underline{x},\underline{y}> = \Sigma_{i=1}^{n} x_i y_i$.

$$\underline{Z}\ \underline{f}(i) = \underline{D}(i)\ \underline{x}.$$

Such a vector $\underline{f}(i)$ exists since, in particular, $\underline{\Phi}^*(i)$ satisfies those constraints. Define

$$\hat{\underline{\Phi}}(i) = \underline{f}_1(i) + \underline{\Phi}_2.$$

Equivalently, $\hat{\underline{\Phi}}_1(i) = \underline{f}_1(i)$ and $\hat{\underline{\Phi}}_2(i) = \underline{\Phi}_2$, where the subscripts 1 and 2 refer to the decomposition according to the subspaces X_1 and X_2, explained in part A. Then

$$\underline{Y}\ \hat{\underline{\Phi}}_1(i) = \underline{Y}\ \underline{f}_1(i) = \underline{x},$$

$$\underline{Z}\ \hat{\underline{\Phi}}_1(i) = \underline{Z}\ \underline{f}_1(i) = \underline{D}(i)\ \underline{x}.$$

Indeed, from the definition of X_2, $\underline{Y}\ \underline{f}_2 = \underline{Y}\ \underline{\Phi}_2 = 0$, and $\underline{Z}\ \underline{f}_2 = \underline{Z}\ \underline{\Phi}_2 = 0$. Also,

$$\underline{Y}(\hat{\underline{\Phi}}_1(i) - \underline{\Phi}_1) = 0.$$

Since

$$\underline{Y}\ \underline{\Phi}_1 = \underline{Y}\ \underline{\Phi} = \underline{x}.$$

However

$$Z(\underline{\Phi}_1(i) - \underline{\Phi}_1) = \underline{D}(i)\ \underline{x} - \underline{D}\ \underline{x}.$$

Since $\underline{\Omega}^{-1}$ exists, there is some linear mapping $\underline{M}$ from W to X_1 such that, for $\underline{a} \in X_1$ and $\underline{b} \in W$,

$$\begin{pmatrix}\underline{Y}\\ \underline{Z}\end{pmatrix}\underline{a} = \underline{b} \text{ if and only if } \underline{M}\ \underline{b} = \underline{a}.$$

Then, since Ω^{-1} is uniformly continuous, for all $\underline{f}$, $\underline{g}$ in W, there is some positive scalar η such that

$$||\underline{f} - \underline{g}|| < \eta \text{ implies } ||\underline{M}\ \underline{f} - \underline{M}\ \underline{g}|| < \varepsilon.$$

Also, there is an integer i_o such that

$$i \leq i_o,\ ||\underline{D}(i)\ \underline{x} - D\ \underline{x}|| < \eta.$$

Therefore,

$$||M\begin{pmatrix}Y\\ Z\end{pmatrix}(\hat{\underline{\Phi}}_1(i) - \underline{\Phi}_1)|| = ||\hat{\underline{\Phi}}_1(i) - \underline{\Phi}_1|| = ||\hat{\underline{\Phi}}(i) - \underline{\Phi}|| < \varepsilon$$

$$\text{for } i \leq i_o.$$

The sequence $\hat{\underline{\Phi}}(i)$ thus has all the specified properties. Q.E.D.

4.4 MAIN RESULT

We now use the above lemmas to establish that $\underline{\Phi}^*$ is the minimizing vector for the limiting constraints. Our central result will follow.

Lemma 4.2

For any integer i, let $\underline{\Phi}^*(i)$ be defined as in section 4.2, and let

$$\underline{\Phi}^* = \lim_{i \to -\infty} \underline{\Phi}^*(i).$$

Then, $\underline{\Phi}^*$ is the minimizing vector in the following optimization problem:

$$\min_{\underline{\Phi}} \frac{1}{2} \underline{\Phi}^T \underline{L}\, \underline{\Phi},$$

s.t.

$$\underline{Y}\, \underline{\Phi} = \underline{x},$$

$$\underline{Z}\, \underline{\Phi} = \underline{D}\, \underline{x}.$$

Proof. We shall establish the proof by contradiction. Suppose that there exists some vector $\hat{\underline{\Phi}}$ which satisfies

$$\frac{1}{2} \hat{\underline{\Phi}}^T \underline{L}\, \hat{\underline{\Phi}} < \frac{1}{2} \underline{\Phi}^{T*} \underline{L}\, \underline{\Phi}^*,$$

and

$$\underline{Y}\, \hat{\underline{\Phi}} = \underline{x}$$

$$\underline{Z}\, \hat{\underline{\Phi}} = \underline{D}\, \underline{x}$$

Let $\delta = \frac{1}{2} (\underline{\Phi}^{T*} \underline{L}\, \underline{\Phi}^* - \hat{\underline{\Phi}}^T \underline{L}\, \hat{\underline{\Phi}}) > 0.$

Since the function $\underline{\Phi} \to \frac{1}{2} \underline{\Phi}^T \underline{L}\, \underline{\Phi}$ is continuous, there exists an $\eta > 0$, such that, for any $\underline{\Phi}$ in R^m,

$$||\underline{\Phi} - \hat{\underline{\Phi}}|| < \eta \text{ implies } \frac{1}{2} |\underline{\Phi}^T \underline{L}\, \underline{\Phi} - \hat{\underline{\Phi}}^T \underline{L}\, \hat{\underline{\Phi}}| < \frac{\delta}{2}.$$

According to lemma 4.1, there exists a sequence $\hat{\underline{\Phi}}(i)$ such that:

$$||\hat{\underline{\Phi}}(i) - \hat{\underline{\Phi}}|| < \eta \quad \text{for } i \le i_o,$$

and

$$\underline{Y}\ \hat{\underline{\Phi}}(i) = \underline{x},$$

$$\underline{Z}\ \hat{\underline{\Phi}}(i) = \underline{D}(i)\ \underline{x}.$$

Therefore,

$$\frac{1}{2}\left|\hat{\underline{\Phi}}^T \underline{L}\ \hat{\underline{\Phi}} - \hat{\underline{\Phi}}^T(i)\ \underline{L}\ \hat{\underline{\Phi}}(i)\right| < \frac{\delta}{2} \quad \text{for } i \leq i\ 0,$$

whence, for $i \leq i_o$,

$$\frac{1}{2}\ \hat{\underline{\Phi}}(i)^T \underline{L}\ \hat{\underline{\Phi}}(i) < \frac{\delta}{2} + \frac{1}{2}\ \hat{\underline{\Phi}}^T\ \underline{L}\ \hat{\underline{\Phi}} = \frac{1}{2}\ \underline{\Phi}^{T*}\ \underline{L}\ \underline{\Phi}^* - \frac{\delta}{2}\ , \tag{4.15}$$

where the equality is just the definition of δ. Since $\hat{\underline{\Phi}}(i)$ satisfies the constraints of the problem for which $\underline{\Phi}^*(i)$ is optimal (section 4.2),

$$\frac{1}{2}\ \Phi^{T*}(i)\ \underline{L}\ \underline{\Phi}^*(i) \leq \frac{1}{2}\ \hat{\Phi}^T(i)\ \underline{L}\ \hat{\Phi}\ (i).$$

Therefore,

$$\frac{1}{2}\ \underline{\Phi}^{T*}(i)\ \underline{L}\ \underline{\Phi}^*(i) < \frac{1}{2}\ \underline{\Phi}^{T*} \underline{L}\ \underline{\Phi}^* - \frac{\delta}{2}\ , \quad \text{for all } i \leq i_o$$

from (4.15).

Taking limit for $i \to -\infty$ gives:

$$\frac{1}{2}\ \underline{\Phi}^{T*} \underline{L}\ \underline{\Phi}^* \leq \frac{1}{2}\ \underline{\Phi}^{T*} \underline{L}\ \underline{\Phi}^* - \frac{\delta}{2}\ ,$$

which is a contradiction since $\delta > 0$. Q.E.D.

We can now establish the main theorem of this section.

Theorem 4.3

If the assumptions of theorem 2.20 hold, i.e., if

a. The matrix $\underline{L}$ is positive-definite.

b. Some reduced system of the dynamical system in $\underline{x}$ is controllable.

c. The Hamiltonian matrix $\underline{H}$ of (2.210) has only simple eigenvalues.

d. The matrix $\underline{Y}\underline{L}^{-1}\underline{Z}^T$ is invertible.

Then, the asymptotic cost increase rate per unit of flow, α, defined by

$$\lim_{i \to -\infty} \underline{C}(i) - \underline{C}(i+1) = \alpha \, \underline{\nu}_n \, \underline{\nu}_n^T, \tag{4.16}$$

and given by equation (2.157) and the stationary distribution $\underline{p}$ uniquely defined by (4.1) and (4.2), are related as follows:

1) $$\alpha = \min_{\underline{r}} \ J(\underline{r}), \tag{4.17}$$

$$\text{s.t. } \underline{\nu}_n^T \, \underline{r} = 1 \tag{4.18}$$

where, for any $\underline{r}$ in R^n,

$$J(\underline{r}) = \min_{\underline{\Phi}} (\underline{\Phi}^T \, \underline{L} \, \underline{\Phi}), \tag{4.19}$$

$$\text{s.t. } \underline{Y} \, \underline{\Phi} = \underline{r}, \tag{4.20}$$

$$\underline{Z} \, \underline{\Phi} = \underline{r}. \tag{4.21}$$

2) The minimum of J is reached at $\underline{r} = \underline{p}$; i.e.,

$$J(\underline{p}) = \alpha.$$

Note. If for some $\underline{r}$, the constraints (4.20) and (4.21) define an empty set, $J(\underline{r})$ is defined as $+\infty$.

Proof of theorem (4.3)

The assumptions of the theorem guarantee the convergence in (4.16) and the uniqueness of $\underline{p}$ (see section 2).

a) We first prove that $J(\underline{p}) = \alpha$. First of all, the constraints (4.20) and (4.21) define a nonempty set of $\underline{\Phi}$'s when $\underline{r} = \underline{p}$. Indeed, the remark of section 4.2 shows that the $\underline{\Phi}^*$ corresponding to an incoming flow distribution $\underline{x} = \underline{p}$ satisfies those two constraints, since it satisfies (4.11) and (4.12) with $\underline{x} = \underline{p}$, and $\underline{D} \, \underline{p} = \underline{p}$.

Therefore, $J(\underline{p}) < \infty$.

Equation (4.9), with $\underline{x} = \underline{p}$, yields:

$$\frac{1}{2} \, \underline{p}^T (\underline{C}(i) - \underline{D}^T(i) \, C(i+1) \, \underline{D}(i)) \underline{p} = \frac{1}{2} \, \underline{\Phi}^{T*}(i) \, \underline{L} \, \underline{\Phi}^*(i),$$

whence,

$$\lim_{i \to -\infty} \underline{p}^T(\underline{C}(i) - \underline{D}^T(i)\underline{C}(i+1)\ \underline{D}(i))\ \underline{p} = \underline{\Phi}^{T*}\ L\ \Phi^* = J(\underline{p})$$

where the second equality is lemma (4.1) applied to $\underline{x} = \underline{p}$.

On the other hand, the limit of the left-hand side is also

$$\lim_{i \to -\infty} \underline{p}^T\ (\underline{C}(i) - \underline{C}(i+1))\ \underline{p}.$$

Since $\underline{D}\ \underline{p} = \underline{p}$.

According to (4.16), this limit is equal to

$$\alpha(\underline{p}^T\ \underline{\nu}_n)(\underline{\nu}_n^T\ p) = \alpha,$$

from (4.2). Therefore, $J(\underline{p}) = \alpha$.

b) We now prove that α is less than or equal to the minimum of $J(r)$ over all r such that $\underline{\nu}_n^T \cdot \underline{r} = 1$. Let $\underline{q}$ be such that the minimum is reached at $\underline{r} = \underline{q}$, and let $\check{\underline{\Phi}}$ such that $J(\underline{q}) = \check{\underline{\Phi}}^T\ L\ \check{\underline{\Phi}}$.

Thus

$$\underline{Y}\ \check{\underline{\Phi}} = \underline{Z}\ \check{\underline{\Phi}} = \underline{q}.$$

Let us now consider the minimization problem (4.3), (4.4) with $\underline{x} = \underline{q}$, and let us call $\tilde{\underline{\Phi}}(i)$ the corresponding minimizing vector. Therefore,

$$\frac{1}{2}\underline{q}^T\ \underline{C}(i)\ \underline{q} = \frac{1}{2}\tilde{\underline{\Phi}}^T(i)\ \underline{L}\ \tilde{\underline{\Phi}}(i) + \frac{1}{2}\tilde{\underline{\Phi}}^T(i)\ \underline{Z}^T\ \underline{C}(i+1)\ \underline{Z}\ \tilde{\underline{\Phi}}(i),$$

and

$$\underline{Y}\ \tilde{\underline{\Phi}}(i) = \underline{q}.$$

Hence,

$$\underline{q}^T\ \underline{C}(i)\ \underline{q} \leq \check{\underline{\Phi}}^T\ \underline{L}\ \check{\underline{\Phi}} + \check{\underline{\Phi}}^T\ \underline{Z}^T\ \underline{C}(i+1)\ \underline{Z}\ \check{\underline{\Phi}} = \check{\underline{\Phi}}^T\ \underline{L}\ \check{\Phi} + \underline{q}^T\ \underline{C}(i+1)\ \underline{q}, \tag{4.22}$$

where the inequality follows from the fact that $\underline{\Phi}$ is feasible, but not necessarily optimal, for (4.3), (4.4). From (4.22),

$$\underline{q}^T(\underline{C}(i) - \underline{C}(i+1))\underline{q} \leq \check{\underline{\Phi}}^T\ \underline{L}\ \check{\Phi}. \tag{4.23}$$

Equation (4.23) holds for every i, and $\check{\underline{\Phi}}$ does not depend on i. Letting i go to $-\infty$ and using (4.16), one obtains:

$$\alpha(\underline{\nu}_n^T \cdot \underline{q})^2 \leq \check{\underline{\Phi}}^T \underline{L} \check{\underline{\Phi}},$$

or, since $\underline{q}$ satisfies the constraint (4.18),

$$\alpha \leq \check{\underline{\Phi}}^T \underline{L} \check{\underline{\Phi}} = J(\underline{q}) = \min_{\underline{r}} J(\underline{r}),$$

$$\text{s.t. } \underline{\nu}_n^T \cdot r = 1.$$

c) Combining the results of parts A and B, we have:

$$\alpha = J(\underline{p}) \geq J(\underline{q}) \geq \alpha,$$

whence,

$$\alpha = J(\underline{p}) = J(\underline{q}) = \min_{\underline{r}} J(\underline{r}),$$

$$\text{s.t. } \underline{\nu}_n^T \cdot \underline{r} = 1,$$

which proves parts 1 and 2 of the theorem. Q.E.D.

Remarks

1 The interpretation of theorem 4.3 is as follows. At points sufficiently far from the site of an imposed perturbation, the resulting downstream perturbation redistributes itself according to the stationary distribution $\underline{p}$ of the asymptotic propagation matrix $\underline{D}$. The cost per unit of flow and per subnetwork is then constant and equal to α. What theorem 4.3 establishes is that the asymptotic propagation matrix $\underline{D}$ is such that its stationary distribution $\underline{p}$ leads to the least possible cost per subnetwork, α.

However, theorem 4.3 does not fully characterize the matrix $\underline{D}$. This matrix indeed contains additional information: the eigenvalues other than 1, that is, the eigenvalues of $\underline{\Delta}$, teach us how fast the convergence of the sequence of downstream perturbations to $\underline{p}$ occurs. No information on these eigenvalues is contained in theorem 4.3.

$\underline{2}$ The upstream analog of this theorem is obtained by interchanging the $\underline{Y}$ and $\underline{Z}$ matrices (see section 3). Note that interchanging $\underline{Y}$ and $\underline{Z}$ leaves the minimization problem (4.17) unchanged since it leaves $J(\underline{r})$ unchanged for all $\underline{r}$. Therefore, the constant α is the same for the upstream as for the downstream perturbations. Moreover, if $\underline{p}$ is the unique point at which the minimum of J is reached (under the constraint (4.18)), then the stationary distribution for the upstream perturbations is the same as for the downstream ones. Thus, although the asymptotic propagation matrices for the upstream and downstream perturbations are generally different, they have the same eigenvector corresponding to the eigenvalue 1.

4.5 EXPLICIT EXPRESSION FOR α AND p IN A SPECIAL CASE

We here minimize $J(\underline{r})$ under the constraint (4.18) in a special case: when $\underline{L}$ is diagonal and some symmetry is present in the subnetwork. When $\underline{L}$ is diagonal, each diagonal entry can be thought of as a cost associated to one specific link. Therefore, it is possible to store all information about the subnetwork; i.e., $\underline{L}$, $\underline{Y}$ and $\underline{Z}$, in a single (nxn) matrix, that we denote $\underline{\mathscr{L}} = [\mathscr{L}_{ij}]$.

$$\mathscr{L}_{ij} = \begin{cases} \text{cost corresponding to link (i,j) if this link exists,} \\ +\infty \text{ otherwise.} \end{cases}$$

For the network of Fig. 7.2 corresponding to

$$\underline{L} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{bmatrix},$$

we have

$$\underline{\mathscr{L}} = \begin{bmatrix} 1 & +\infty \\ 2 & 4 \end{bmatrix},$$

since there is no link from entrance-node 1 to exit-node 2. Replacing also the (mxl) vector $\underline{\phi}$ by an (nxn) matrix $\underline{\Phi}$, where Φ_{ij} is the flow along link (i,j), the flow-conservation constraints become:

$$\sum_{j=1}^{n} \Phi_{ij}(k) = x_i(k), \qquad i = 1,\ldots,n, \tag{4.24}$$

$$\sum_{i=1}^{n} \phi_{ij}(k) = x_j(k+1), \qquad j = 1,\ldots,n, \tag{4.25}$$

and the cost function is

$$\frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \mathscr{L}_{ij} \, \Phi_{ij}^2, \tag{4.26}$$

where the sum is over those subscripts for which $\mathscr{L}_{ij} < \infty$.

In this notation, the expression for $J(\underline{r})$ becomes:

$$J(\underline{r}) = \min_{\Phi} \left(\sum_{i=1}^{n} \sum_{j=1}^{n} \mathscr{L}_{ij} \, \Phi_{ij}^2 \right),$$

$$\text{s.t.} \quad \sum_{j=1}^{n} \Phi_{ij} = r_i, \qquad i = 1,\ldots,n, \tag{4.27}$$

$$\sum_{i=1}^{n} \Phi_{ij} = r_j, \qquad j = 1,\ldots n. \tag{4.28}$$

However, let us take only the constraints (4.27) into account. Then the Lagrangean function is:

$$F(\underline{\Phi}, \underline{\lambda}, \underline{r}) = \sum_{i=1}^{n} \sum_{j=1}^{n} \mathscr{L}_{ij} \, \Phi_{ij}^2 + \sum_{i=1}^{n} \lambda_i \left(p_i - \sum_{j=1}^{n} \Phi_{ij} \right).$$

Equating to zero the derivative of F with respect to $\underline{r}$ yields:

$$2 \mathscr{L}_{ij} \, \Phi_{ij} - \lambda_i = 0.$$

The constraint (4.27) implies:

$$\frac{1}{2}\lambda_i \sum_{j=1}^{n} \frac{1}{\mathscr{L}_{ij}} = r_i,$$

therefore,

$$\Phi_{ij} = \frac{r}{\mathscr{L}_{ij} \sum_{k=1}^{n} \frac{1}{\mathscr{L}_{ik}}} \tag{4.29}$$

is optimal. Substituting the right-hand side of (4.29) for $\underline{\Phi}$ in (4.26) yields

$$J(\underline{r}) = \sum_i \sum_j \mathscr{L}_{ij} \left(\frac{r_i^2}{\left(\sum_i \frac{1}{\mathscr{L}_{ik}}\right)^2 \mathscr{L}_{ij}^2} \right) = \sum_{i=1}^{n} \left[\frac{r_i^2}{\sum_i \frac{1}{\mathscr{L}_{ij}}} \right] .$$

We now minimize $J(\underline{r})$ under constraint (4.18). The new Lagrangean is

$$\Psi(\underline{r}, \mu) = \sum_{i=1}^{n} \left[\frac{r_i^2}{\sum_k \frac{1}{\mathscr{L}_{ik}}} \right] + \mu \left[1 - \sum_{i=1}^{n} p_i \right] .$$

Taking the derivative of Ψ with respect to $\underline{r}$ and setting it to zero yields:

$$\frac{2\, r_i}{\sum_k \frac{1}{\mathscr{L}_{ik}}} - \mu = 0,$$

which, taking constraint (4.18) into account, yields uniquely the optimal $\underline{r}$,

$$r_i = \frac{\sum_k \frac{1}{\mathscr{L}_{ik}}}{\sum_r \sum_s \frac{1}{\mathscr{L}_{rs}}} . \tag{4.30}$$

The corresponding optimal $\underline{\Phi}$; i.e, $\underline{\check{\Phi}}$, is given by

$$\underline{\check{\Phi}}_{ij} = \frac{1}{\mathscr{L}_{ij} \sum_r \sum_s \frac{1}{\mathscr{L}_{rs}}} . \tag{4.31}$$

Now from (4.31),

$$\sum_{i=1}^{n} \check{\phi}_{ij} = \frac{\sum_{i=1}^{n} \frac{1}{\mathscr{L}_{ij}}}{\sum_r \sum_s \frac{1}{\mathscr{L}_{rs}}} .$$

Therefore, if $\mathscr{L}$ is symmetric; i.e., if $\mathscr{L}_{ij} = \mathscr{L}_{ji}$ for all i,j, then

$$\sum_{i=1}^{n} \check{\phi}_{ij} = r_j ,$$

and constraint (4.28) is also satisfied. Accordingly, in that symmetric case, the solutions are $\check{\underline{\Phi}}$, given by (4.31), and $\underline{p}$, given by (4.30); i.e.,

$$p_i = \frac{\sum_k \frac{1}{\mathscr{L}_{ik}}}{\sum_r \sum_s \frac{1}{\mathscr{L}_{rs}}} . \tag{4.32}$$

The corresponding minimum of $\underline{J}(\underline{r})$ is

$$\alpha = \sum_i \sum_j \mathscr{L}_{ij} \check{\phi}_{ij}^2 = \sum_i \sum_j \mathscr{L}_{ij} \left[\frac{\left(\sum_k \frac{1}{\mathscr{L}_{ik}} \right)^2}{\left(\sum_r \sum_s \frac{1}{\mathscr{L}_{rs}} \right)^2} \right] = \frac{1}{\sum_r \sum_s \frac{1}{\mathscr{L}_{rs}}} . \tag{4.33}$$

<u>Example</u>. For the standard two-dimensional example, with cost matrix

$$\underline{L} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix} ,$$

or

$$\underline{\mathscr{L}} = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} ,$$

since $\underline{L}$ is diagonal and $\underline{\mathscr{L}}$ symmetric, equations (4.32) and (4.33) are applicable and yield:

$$\alpha = (1 + \frac{1}{2} + \frac{1}{2} + 1)^{-1} = \frac{1}{3},$$

$$p_1 = p_2 = \frac{1}{2},$$

which are the values of α and $\underline{p}$ found both numerically (see section 7) and by the method of section 2 (i.e., by solving a Riccati equation). When $\underline{L}$ is diagonal, but $\underline{\mathscr{L}}$ not symmetric, problem (4.17) can still be solved analytically, but does not lead to compact expressions. In the minimization problem which defines $J(\underline{r})$, the optimal Lagrange multipliers are not unique; they depend on a scalar parameter, and we re-optimize the cost function with respect to it. For instance, in the standard two-dimensional example, with

$$\underline{L} = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix}$$

(example 7.1.4 of section 7), one finds:

$$\underline{p}^T = \frac{11}{29} \quad \frac{18}{29} \quad \text{and } \alpha = \frac{14}{29} \simeq 0.4827586,$$

which are exactly values observed by implementing the recursive equations of section 2.2 on a digital computer. We have not explicitly solved the minimization problem for nondiagonal matrices $\underline{L}$.

5. EXAMPLE OF NONSTATIONARY NETWORK

5.1 INTRODUCTION

In section 2.2, we have mentioned that a general freeway corridor network can be split into subnetworks. Correspondingly, the average travel time cost function used in [2] can be broken up into a sum of as many terms as there are subnetworks, where each term depends only on the flows of a single subnetwork. This allows one to expand the cost function to the second order about an optimal solution, and to apply dynamic programming to determine the optimal perturbations (section 2.2).

Here, we actually show how to split a network into subnetworks, and how to obtain the corresponding expression for the cost function. We also demonstrate the way to handle positivity and capacity constraints. We then illustrate the general technique on an example of freeway corridor network which is presented in [2]. We compare the values obtained for downstream perturbations by applying the method of section 2.2 with those found by solving globally the flow-assignment problem over the whole network, and by applying the accelerated gradient projection algorithm to slightly different incoming traffic distributions, as has been done in [2]. The agreement is satisfactory.

5.2 SPLITTING OF NETWORK AND COST FUNCTION

5.2.1 Notation

We shall here adopt again the notation of section 2.2; i.e., we denote traffic flows by $\underline{\phi}(k)$ and flow perturbations by $\delta\underline{\phi}(k)$. Let us briefly recall the expression for the average travel time cost function, already given in 2.2.

$$J(\underline{\phi}) = \sum_i \ell_i \rho_i(\phi_i) + \sum_{i \notin A} \frac{\phi_i^2}{E_i(E_i - \phi_i)} , \tag{5.1}$$

where the first sum is over all links, and the second sum only over links of class B (entrance ramps) and C (signalized arterials) but not A (see 2.2). Also,

$$E_i = \begin{cases} g_i \phi_i \max, & \text{if } i \in C, & (5.2) \\ \phi_i \max\,(1 - \phi_j/\phi_j \max), & \text{if } i \in B. & (5.3) \end{cases}$$

In equation (5.3), link j is that portion of the freeway the ramp i impinges upon (Fig. 5.2.1).

The function $\rho_i(\cdot)$ is the inverse of the fundamental diagram [15]. In fact, to approximate the inverse of the fundamental diagram (corresponding to $\phi_i \leq E_i$), a seventh-degree polynomial is chosen; in the usual fundamental diagram, ϕ attains a maximum, ϕ_{max}, for a certain value of ρ, and $\frac{d\phi}{d\rho}$ is zero at that value. To invert the fundamental diagram perfectly, one requires that $\frac{d\rho}{d\phi}$ go to infinity at $\phi = \phi_{max}$. Instead, one has adopted here a large finite slope. Thus, we take

$$\rho_i(\phi) = \frac{\phi}{b_i} + k_i\,\phi^7, \tag{5.4}$$

where

$$k_i = \frac{.3\rho_{imax} - \phi_{imax}/b_i}{(\phi_{imax})^7}\ , \tag{5.5}$$

which implies that

$$\lim_{\phi \to \phi_{imax}} \rho_i(\phi) = 0.3\rho_{imax}. \tag{5.6}$$

Also

$$\lim_{\phi \to 0} \rho_i(\phi) = 0,$$

and

$$\lim_{\phi \to 0} \frac{\phi}{\rho_i(\phi)} = b_i.$$

Indeed, $v_i(\phi) = \dfrac{\phi}{\rho_i(\phi)}$ is the speed along link i corresponding to a flow ϕ, and $\lim_{\phi \to 0} v_i(\phi)$ is the free speed, in the absence of flow, denoted by b_i.

In the examples, the free speed b_i is taken equal to 55 mph. The maximum number of vehicles per mile, ρ_{imax}, is taken equal to 225 times the number of lanes of link i.

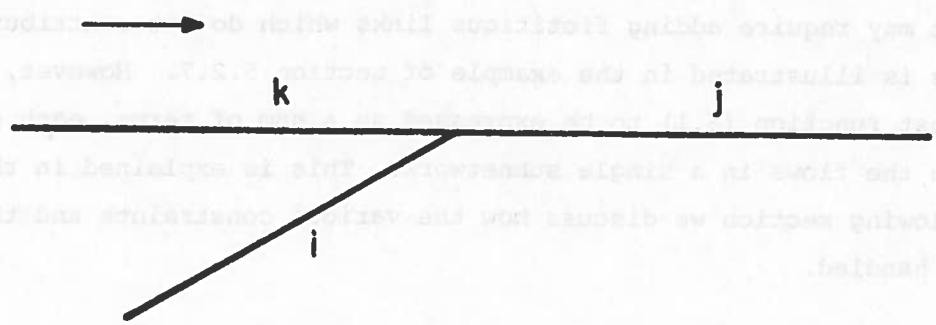

Figure 5.2.1 FREEWAY-ENTRANCE RAMP

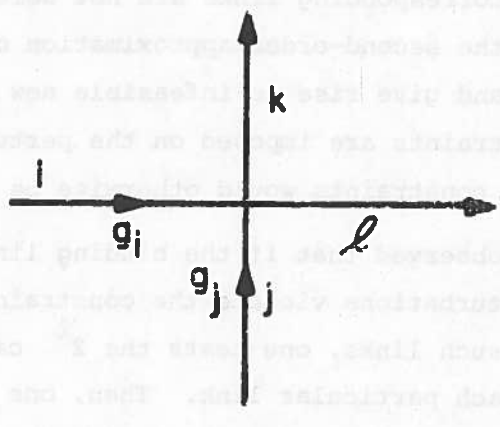

Figure 5.2.2 SIGNALIZED INTERSECTION

5.2.2 Decomposition of the Network

There are many ways to decompose a freeway-corridor network into subnetworks. The only restriction is that there be no internal nodes, so that the flow-conservation constraints may be expressed by equations (2.9,2.10) via the incidence matrices $\underline{Y}(k)$ and $\underline{Z}(k)$. To fulfill that condition is always possible, although it may require adding fictitious links which do not contribute to the cost. This is illustrated in the example of section 5.2.7. However, we also need the cost function (5.1) to be expressed as a sum of terms, each one depending only on the flows in a single subnetwork. This is explained in the sequel. In the following section we discuss how the various constraints and the green splits are handled.

5.2.3. Constraints. Green-Split Variables

a. As already mentioned in section 2.2., we delete, for the sensitivity analysis, those links on which the positivity or capacity constraints (2.3) are binding in the nominal optimal solution found by the global nonlinear optimization technique. The reason for this procedure is that it is extremely likely that, for a small change in the initial conditions (i.e., the incoming traffic vector), the constraints will still be binding in the new optimal solution. Therefore, if the corresponding links are not deleted, the optimal flow perturbations obtained by the second-order approximation of section 2 will violate those constraints and give rise to infeasible new flow assignments. In effect, inequality constraints are imposed on the perturbations by deleting the links on which the constraints would otherwise be violated.

Numerically, we have observed that if the binding links are not deleted, the corresponding flow perturbations violate the constraints. To be entirely rigorous, if there are ℓ such links, one tests the 2^{ℓ} cases corresponding to deleting or not deleting each particular link. Then, one makes the choice which leads to feasible perturbations with the lowest cost. Each time a link is deleted in subnetwork k, the number m_k of components of $\underline{\phi}(k)$ is reduced by 1. Also, the corresponding row and column of $\underline{L}(k)$ are deleted.

b. In section 2.2, the green splits are not included in our analysis. We show here how to include them. A traffic light is always located at an intersection between two different directions of traffic (Fig. 5.2.3). The green splits along the two different directions satisfy the relation

$$g_i + g_j = \alpha , \quad g_i \geq 0 , \quad g_j \geq 0 , \tag{5.7}$$

with $0 < \alpha \leq 1$. The number α is part of the data. Therefore, g_i and g_j cannot be chosen independently. If we decide to include link i and link j of Fig. 5.2.3 in the same subnetwork, we can define only one green-split variable per signalized arterial, instead of two. In other words, we replace g_j by $(\alpha - g_i)$.

Now, in the vector $\underline{\phi}(k)$ corresponding to subnetwork k, we shall include not only the flows along the various links of subnetwork k, but also one green split variable per signalized intersection in the subnetwork.

The flow-conservation constraints apply to flows, and obviously not to green splits. However, equations (2.9,2.10) can be easily modified, taking into account the fact that some components of $\underline{\phi}(k)$ are green-split variables, not flow variables. To each green-split variable, there will correspond a column of zeroes in the incidence matrices $\underline{Y}(k)$ and $\underline{Z}(k)$. This does not alter the ranks of $\underline{Y}(k)$ and $\underline{Z}(k)$, which are still n_k and n_{k+1} respectively, since $\underline{Y}(k)$ and $\underline{Z}(k)$ have more columns than rows.

5.2.4 Decomposition of the Cost Function

Taking into account equations (5.2) and (5.3), we can rewrite the cost function (5.1) in the following form:

$$J(\underline{\phi}) = \sum_i J_i(\phi_i) + \sum_{i \varepsilon B} J_i'(\phi_i, \phi_{j_i}) + \sum_{i \varepsilon C} J_i''(\phi_i, g_i) , \tag{5.8}$$

where

$$J_i(\phi_i) = \ell_i \rho_i(\phi_i) , \tag{5.9}$$

$$J_i'(\phi_i, \phi_{j_i}) = \frac{\phi_i^2}{\phi_{imax}(1-\phi_j/\phi_{jmax})[\phi_{imax}(1-\phi_j/\phi_{jmax})-\phi_i]} \tag{5.10}$$

$$J_i''(\phi_i, g_i) = \frac{\phi_i^2}{g_i\phi_{imax}(g_i\phi_{imax}-\phi_i)} . \tag{5.11}$$

In equation (5.8), the second-term $\sum_{i \varepsilon B} J_i'(\phi_i, \phi_{j_i})$ is to be understood as follows: for any link i of class B, there is one link j_i uniquely specified, whose relationship with i is explained in Fig. 5.2.1. Link i is a ramp which impinges upon the freeway-link j_i. We make explicit the dependence of J_i' on ϕ_{j_i} as well as on ϕ_i because both variables have to be considered when differentiating J. Likewise, the notation $J_i''(\phi_i, g_i)$ emphasizes the fact that J_i'' depends on the green split g_i as well as on the flow ϕ_i. The green split is also a variable with respect to which J shall be differentiated when expanding J to the second order.

Below, we now expand the cost function J, given by (5.8), (5.9), (5.10), and (5.11), about a nominal optimal solution $\underline{\phi}^*$, and show how the network has to be split into subnetworks so as to decompose J properly.

5.2.5 Expansion to the Second Order

We examine separately the three sums occurring in (5.8). Every link i contributes a cost $J_i(\phi_i)$ which depends on the variable ϕ_i only. Therefore, it contributes in the quadratic expansion of J about the nominal solution $\underline{\phi}^*$ the first derivative $\partial J_i/\partial\phi_i$ (ϕ_i) and the second derivative $\partial^2 J_i/\partial\phi_i^2$ (ϕ_i^*).

To each link i in class B, there correspond two terms in (5.8) which are: $J_i(\phi_i)$ and $J_i'(\phi_i, \phi_{j_i})$. Each time we encounter such a link, we shall put links i and k of Fig. 5.2.1 into the same subnetwork, and link j in the subnetwork immediately downstream. In order to distinguish between flows along links of different subnetworks, in the expansion of J, we replace ϕ_{j_i} by $\phi_i + \phi_k$ in (5.10). Therefore, we replace $J_i'(\phi_i, \phi_j)$ by $J_i'(\phi_i, \phi_i + \phi_k)$ and differentiate it which gives rise to five terms:

$$\frac{\partial J_i'}{\partial\phi_i}, \quad \frac{\partial J_i'}{\partial\phi_k}, \quad \frac{\partial^2 J_i'}{\partial\phi_i^2}, \quad \frac{\partial^2 J_i'}{\partial\phi_i\partial\phi_k}, \quad \frac{\partial^2 J_i'}{\partial\phi_k^2} \quad .$$

To each link i in class C, there also correspond two terms in (5.8): they are $J_i(\phi_i)$ and $J_i''(\phi_i, g_i)$. For the second term, the derivative gives rise to five terms, i.e.:

$$\frac{\partial J_i''}{\partial\phi_i}, \quad \frac{\partial J_i''}{\partial g_i}, \quad \frac{\partial^2 J_i''}{\partial\phi_i^2}, \quad \frac{\partial^2 J_i''}{\partial\phi_i\partial g_i}, \quad \frac{\partial^2 J_i''}{\partial g_i^2} \quad .$$

Corresponding to link i we also have link j, in the same subnetwork, leading to the same traffic light. Link j contributes the term $J''_j(\phi_j, \alpha - g_i)$ in (5.8), in view of (5.7). This term in turn contributes five terms in the quadratic expansion of J.

By these operations, we obtain a quadratic expansion of the cost-function J about the optimal solution $\underline{\phi}^*$, of the type:

$$J(\underline{\phi}^* + \delta\underline{\phi}) - J(\underline{\phi}^*) = \sum_{k=1}^{N-1} \left[\frac{1}{2}\,\delta\underline{\phi}^T(k)\underline{L}(k)\delta\underline{\phi}(k) + \underline{h}^T(k)\delta\underline{\phi}(k)\right], \quad (5.12)$$

where $\delta\underline{\phi}(k)$ is the vector of flow perturbations and perturbations in green-split variables for subnetwork k. We have been careful when defining the sub-network to express J as a sum of terms, each of which depends only on the vector $\underline{\phi}(k)$ corresponding to one subnetwork k.

5.2.6 Analytical Expressions for the Derivatives

To implement the decomposition method described in sections 5.2.4 and 5.2.5 we need analytical expressions for the first and second derivatives of the terms which appear in (5.9), (5.10), and (5.11).

a. Costs common to all links

$$\frac{\partial J_i}{\partial \phi_i} = \frac{1}{b_i} + 7k_i\phi_i^{6}, \quad (5.13)$$

$$\frac{\partial^2 J_i}{\partial \phi_i} = 42k_i\phi_i^{5} \quad . \quad (5.14)$$

b. Cost special to links of class B

$$\frac{\partial J'_i}{\partial \phi_i}(\phi_i, \phi_i + \phi_k) = \frac{2\phi_i x - \phi_i^2}{x(x-\phi_i)^2} + \frac{\phi_{i\,max}}{\phi_{j\,max}}\,\frac{(2x - \phi_i)\phi_i^2}{x^2(x-\phi_i)^2}, \quad (5.15)$$

$$\frac{\partial J_{i'}}{\partial \phi_k}(\phi_i, \phi_i + \phi_k) = \frac{\phi_{i\,max}}{\phi_{j\,max}}\,\frac{(2x - \phi_i)\phi_i^2}{x^2(x-\phi_i)^2}, \quad (5.16)$$

$$\frac{\partial^2 J_{i'}}{\partial \phi_i^2}(\phi_i, \phi_i + \phi_k) = U + 2V + W \quad , \tag{5.17}$$

$$\frac{\partial^2 J_{i'}}{\partial \phi_i \partial \phi_k}(\phi_i, \phi_i + \phi_k) = V + W \quad , \tag{5.18}$$

$$\frac{\partial^2 J_{i'}}{\partial \phi_k^{\ 2}}(\phi_i, \phi_i + \phi_k) = W \quad , \tag{5.19}$$

where

$$U \triangleq \frac{2x}{(x - \phi_i)^3} \tag{5.20}$$

$$V \triangleq \frac{\phi_{i\ max}}{\phi_{j\ max}} \phi_i \frac{(4x^2 - 3x\phi_i) + (4x^2 - 3x\phi_i + \phi_i^{\ 2})}{x^2 \ (x - \phi_i)^3} \quad , \tag{5.21}$$

$$W \triangleq \frac{2\phi_{i\ max}}{\phi_{j\ max}} \phi_i^{\ 2} \frac{(3x\phi_i - 3x^2 - \phi_i^{\ 2})}{x^3 (x - \phi_i)^3} \quad , \tag{5.22}$$

$$x \triangleq \phi_{i\ max} \left(1 - \frac{\phi_i + \phi_k}{\phi_{j\ max}} \right) \quad . \tag{5.23}$$

c. Cost special to links of class C

$$\frac{\partial J_{i''}}{\partial \phi_i} = \frac{2\phi_{i\ max} g_i \phi_i - \phi_i^{\ 2}}{\phi_{i\ max} g_i \ (\phi_{i\ max} g_i - \phi_i)^2} \quad , \tag{5.24}$$

$$\frac{\partial J_{i''}}{\partial g_i} = \phi_{i\ max} \frac{\phi_i^{\ 2}(\phi_i - 2\phi_{i\ max} g_i)}{(\phi_{i\ max} g_i)^2 (\phi_{i\ max} g_i - \phi_i)^2} \quad , \tag{5.25}$$

$$\frac{\partial^2 J_{i''}}{\partial \phi_i^{\ 2}} = \frac{2\phi_{i\ max} g_i}{(\phi_{i\ max} g_i - \phi_i)^3} \quad , \tag{5.26}$$

$$\frac{\partial^2 J_{i''}}{\partial \phi_i \partial g_i} = \frac{\phi_{i\ max} \ \phi_i \ (\phi_i^{\ 2} - 3\phi_i \ \phi_{i\ max} g_i + 4\phi_i^{\ 2})}{(\phi_{i\ max} g_i)^2 \ (\phi_i - \phi_{i\ max} g_i)^3} \quad , \tag{5.27}$$

$$\frac{\partial^2 J_{i''}}{\partial g_i^{\,2}} = \frac{2(\phi_{i\,max}g_i)^2\,\phi_i^{\,2}[(\phi_{i\,max}g_i)^2 + (\phi_{i\,max}g_i - \phi_i)(2\phi_{imax}g_i - \phi_i)]}{(\phi_{i\,max}\,g_i)^3\,(\phi_{i\,max}g_i - \phi_i)^3} . \tag{5.28}$$

<u>Remark</u>

The <u>D</u>(k) matrices of section 2.2. depend only on the <u>L</u>(k) matrices of equation (5.12), therefore only on the second derivatives of J.

5.2.7 <u>Example</u>

This example, whose network appears in Fig. 5.2.7-1, is solved in [2] by the accelerated projected gradient algorithm, for various incoming traffic distributions. The squares represent signalized intersections, and the nodes represent entrances.

We denote by g_1 the green split for traffic on link 7, and by g_2 for traffic on 15.

$$g_1 + g_2 = 1 \quad .$$

Likewise, g_3 is the green split for traffic coming along 8 and g_4 for traffic coming along 9.

$$g_3 + g_4 = 1.$$

Links 1,2,3,4,5,6,8,10,16 belong to class A.

Links 11,12,13,14, are of class B.

Links 7,15, 8, 9 are of class C.

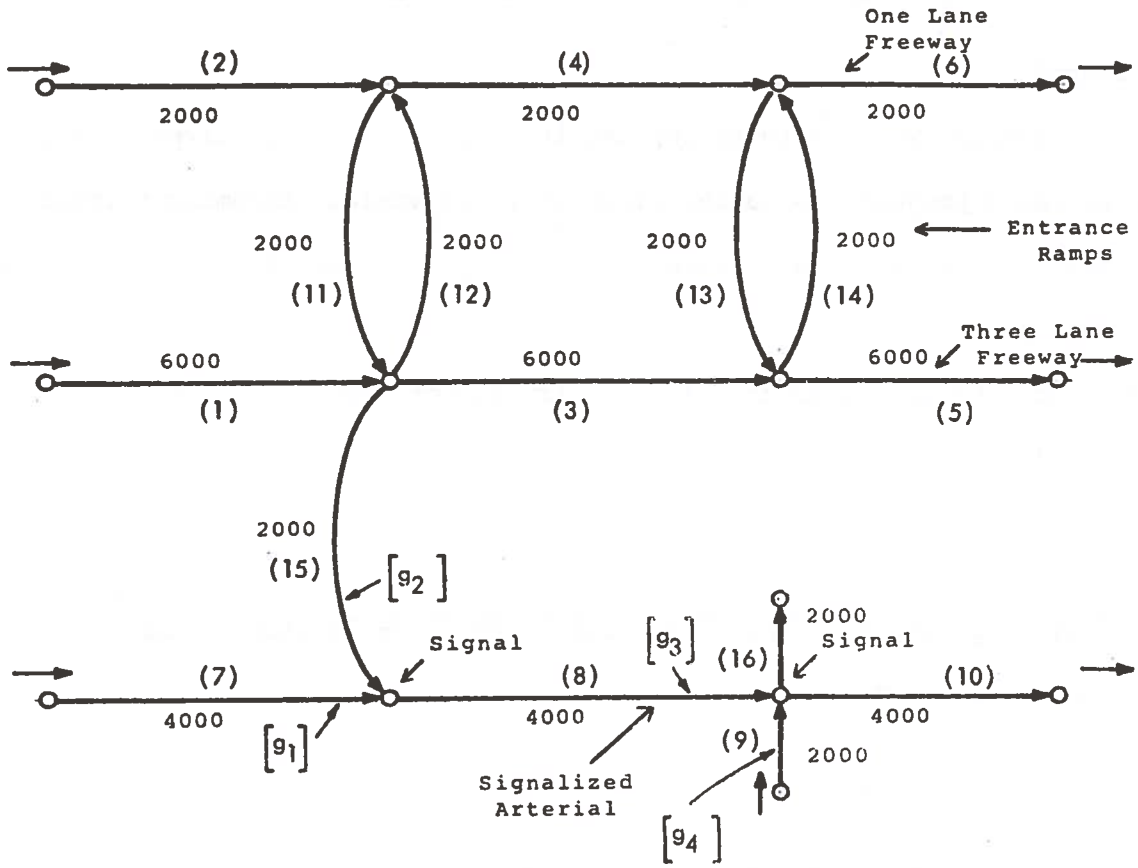

Figure 5.2.7-1 EXAMPLE OF NONSTATIONARY NETWORK

The optimal flows corresponding to the initial condition $\underline{x}_i^T = (1000, 5000, 1000)$ are listed in Table 5.1. The various parameters, i.e., maximum flows $\phi_{i\ max}$, lengths ℓ_i, as well as numbers of lanes and maximum capacities $\rho_{i\ max}$ are also given in Table 5.1.

To apply the method of sections 3 and 4, we have to split the network into subnetworks. We shall define four subnetworks by introducing fictitions links with zero length as shown in Fig. 5.2.7-2. (They are represented by dotted lines.)

This network is not stationary, but the dimension of the state is constant: $n_k = 3$ for each k, since each subnetwork consists of three entrances and three exits. However, the dimension m_k of $\phi(k)$ is not consistent with k.

The positivity constraint is binding on link (1,2) of subnetwork 2 and on link (1,2) of subnetwork 4. Therefore, according to section 5.2.3, we delete those links in our analysis and replace the network of Fig. 5.2.7-2 by that of Fig. 5.2.7-3.

For the new network, the incidence matrices are as follows:

a. $\underline{Y}(1) = \underline{Z}(1) = \underline{Y}(3) = \underline{Z}(3) = \underline{I}$ (i..e, the 3 by 3 identity).

b. $\underline{Y}(2) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$, and $\underline{Z}(2) = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \end{bmatrix}$,

corresponding to the labeling of Fig. 5.2.7-4.

The control $\phi(2)$ is 6-dimensional. Its first five components are the flows along the five links of subnetwork 2, and its sixth component is the green-split g_i (Fig. 5.2.7-1). To that sixth component corresponds a sixth column of zeroes in $\underline{Y}(2)$ and $\underline{Z}(2)$.

c. $\underline{Y}(4) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$, and $\underline{Z}(4) = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$,

correspond to the labeling of Fig. 5.2.7-5.

The control (4) is 5-dimensional. Its first four components are the flows along the four links of subnetwork 4 and its fifth component is the green-split g_3 (Fig. 5.2.7-1). To that fifth component corresponds a fifth column of zeroes in $\underline{Y}(4)$ and $\underline{Z}(4)$.

TABLE 5.1 DATA OF THE NETWORK

i	ρ_i	$\rho_{i\,max}$	$\phi_{i\,max}$	number of lanes	nominal optimal flow ϕ_i^*
1	0.5	675	6000	3	5000
2	0.5	225	2000	1	1000
3	0.5	675	6000	1	1000
4	0.5	225	2000	1	1256
5	0.5	675	6000	3	4100
6	0.5	225	2000	1	1295
7	0.5	450	4000	2	1000
8	0.5	450	4000	2	1605
9	0.05	225	2000	1	500
10	0.5	450	4000	2	1605
11	0.1	225	2000	1	0
12	0.1	225	2000	1	256
13	0.1	225	2000	1	0
14	0.1	225	2000	1	39
15	0.15	225	2000	1	605
16	0.05	225	2000	1	500

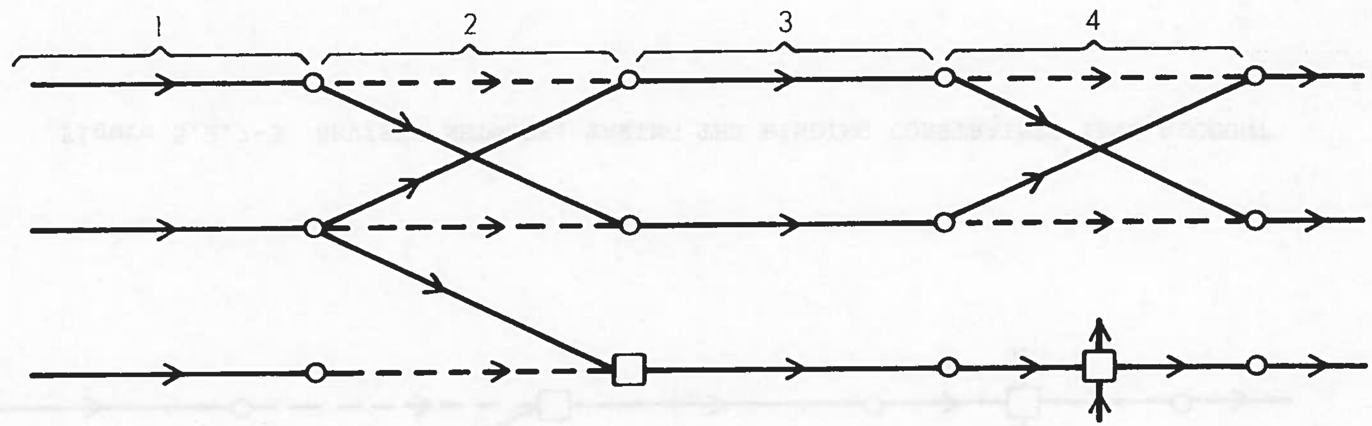

Figure 5.2.7-2 DECOMPOSITION INTO SUBNETWORKS

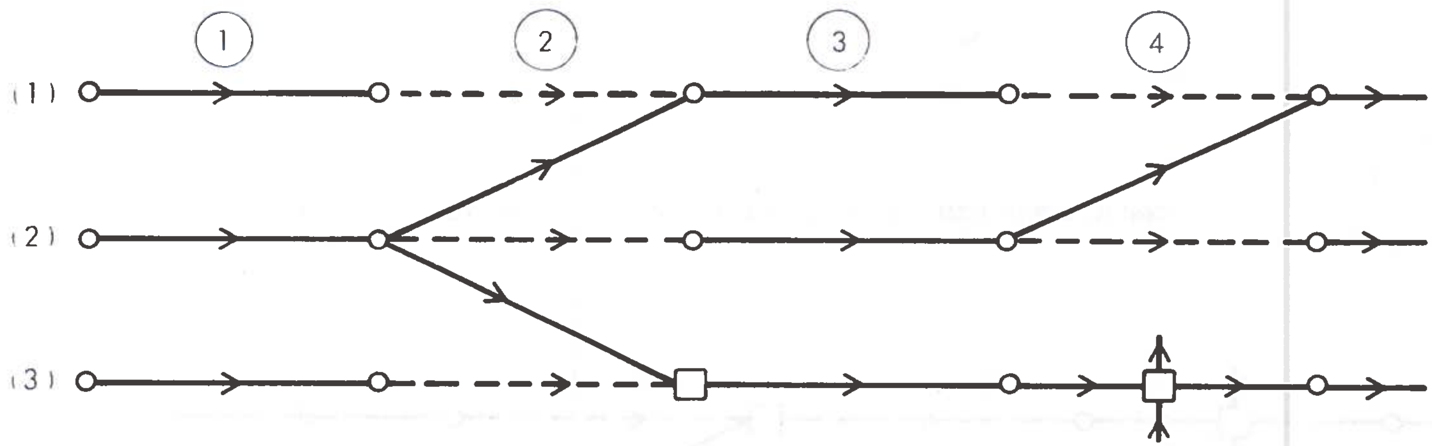

Figure 5.2.7-3 REVISED NETWORK, TAKING THE BINDING CONSTRAINTS INTO ACCOUNT

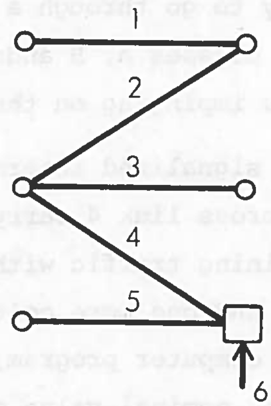

Figure 5.2.7-4 LABELING OF LINKS AND SIGNALIZED INTERSECTIONS IN SUBNETWORK 2

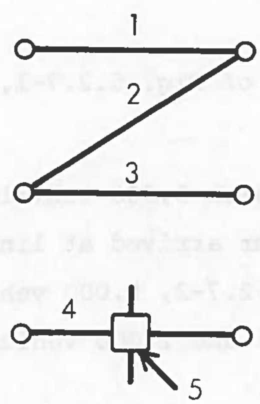

Figure 5.2.7-5 LABELING OF LINKS AND SIGNALIZED INTERSECTIONS IN SUBNETWORK 3

The cost matrices $\underline{L}(k)$ are calculated as explained in section 2.2.4, using the equations of 5.2.6. Given the changes in the traffic entering subnetwork 1, the computer program uses the equations (2.59) of section 2.2.5 with initial condition (2.58) to obtain the perturbations. To make the program operational, it is necessary to go through a finer classification of the links than the distinction among classes A, B and C, particularly, to include in the same subnetwork both links impinging on the same traffic light.

In the present example, the signalized intersection of subnetwork 4 has a peculiarity. The links which cross link 4 carry a flow of traffic which does not interfere with the remaining traffic within subnetwork 4. We might have added one component to $\underline{\phi}(4)$ and one more column of zeroes to $\underline{Y}(4)$ and $\underline{Z}(4)$. Instead, we have devoted, in the computer program, a memory for the length, maximum capacity, number of lanes, nominal value of the optimal flow and the green split, and have stored into it the corresponding quantities for that link (see Appendix F).

Numerical Results

We consider, for the network of Fig. 5.2.7-1, three different cases of incoming traffic distribution.

The nominal case is one in which 5,000 vehicles per hour entered on link 1 and 1,000 other vehicles per hour arrived at link 2 and at link 7. Alternatively, in the notation of Fig. 5.2.7-2, 5,000 vehicles per hour entered through entrance 2 of subnetwork 1 and 1,000 vehicles per hour arrived at entrances 1 and 3 of subnetwork 1.

Case 1 is the same as the nominal case except that 1010 veh/hr. enter through entrance 1 to subnetowrk 1 instead of 1,000. In case 2, 5010 veh/hr enter through entrance 2 to subnetwork 1 instead of 5,000. In case 3, 1010 veh/hr enter through entrance 3 instead of 1,000.

Cases 1, 2, and 3 will give rise to downstream optimal distributions of traffic which are different from the nominal case. For instance, in case 1 the optimal traffic flows through exits 1, 2, and 3 of subnetwork 4 are 1274.4, 4047.3 and 1638.3, respectively (as calculated by the global optimization technique of [2]) instead of 1295, 4100, and 1605, respectively, in the nominal case.

Table 5.2 lists the optimal traffic-flow distributions among the exits of subnetwork 4 corresponding to all four cases (nominal and three perturbed cases) as obtained in [2]. It is to be noticed that if $\underline{x}_I(\alpha)$ and $\underline{X}_I(\beta)$ represent the three-dimensional vectors of input flows (i.e., at the entrance of subnetwork 1) in cases α and β, and $\underline{x}_o(\alpha)$, $\underline{x}_o(\beta)$ represent the output flow vectors at the exits of subnetwork 4, then:

$$||\underline{x}_o(\alpha) - \underline{x}_o(\beta)|| < || \underline{x}_I(\alpha) - \underline{x}_I(\beta)|| \quad ,$$

where $||\underline{Y}|| \triangleq \sum_{i=1}^{3} Y_i^{2})^{1/2}$ for any three-dimensional vector $\underline{Y}$. This damping of perturbations is analogous to what we have demonstrated rigorously in the stationary case. Instead of applying the accelerated gradient-projection method to all four cases to obtain $\underline{x}_o(1)$, $\underline{x}_o(2)$, $\underline{x}_o(3)$ from $\underline{x}_I(1)$, $\underline{x}_I(2)$ and $\underline{x}_I(3)$ respectively, which has been done in [2] and yields the figures of Table 5.2, we can apply the quadratic sensitivity analysis as presented in section 2.2.

That is, a numerical optimization technique is used only to compute traffic assignments in the nominal case. They are given in Table 5.1. For the other three cases, we focus on the differences in incoming traffic flows between each particular perturbed case and the nominal case. We have thus the incoming traffic-flow perturbations $\delta\underline{x}^T(1) = [\delta x_1(1), \delta x_2(1), \delta x_3(1)]$. In case 1, $\delta\underline{x}(1)^T = (10, 0, 0)$.

TABLE 5.2 EFFECT OF VARIATIONS IN INPUT FLOWS

Case	$x_1(4)$	$x_2(4)$	$x_3(4)$
nominal	1271.9	4092.8	1635.3
1	1273.7	4097.1	1639.3
2	1274.7	4097.3	1638.3
3	1273.2	4096.0	1640.8

TABLE 5.3 TRAFFIC-FLOW PERTURBATIONS, BY THE TWO METHODS

Case	$\delta x_1(4)$		$\delta x_2(4)$		$\delta x_3(4)$	
	G	Q	G	Q	G	Q
1	2.58	2.60	4.48	4.46	2.94	2.93
2	1.80	1.64	4.27	4.20	3.94	4.15
3	1.33	1.19	3.16	3.04	5.52	5.77

Table 5.4a

Cost Matrices

L(1)

$$\left[\begin{array}{c:c:c} .18605 \times 10^{-4} & 0 & 0 \\ 0 & .81184 \times 10^{-5} & 0 \\ 0 & 0 & .85010 \times 10^{-6} \end{array}\right]$$

L(2)

$$\left[\begin{array}{c:c:c:c:c} -.41440 \times 10^{-7} & .46514 \times 10^{-6} & 0 & 0 & 0 \\ .46514 \times 10^{-6} & .57448 \times 10^{-5} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & .99490 \times 10^{-5} & -.32570 \times 10^{-1} \\ 0 & 0 & 0 & -.32570 \times 10^{-1} & .23518 \times 10^{3} \end{array}\right]$$

L(3)

$$\left[\begin{array}{c:c:c} .15967 \times 10^{-4} & 0 & 0 \\ 0 & .85119 \times 10^{-5} & 0 \\ 0 & 0 & .85010 \times 10^{-6} \end{array}\right]$$

L(4)

$$\left[\begin{array}{c:c:c:c:c:c} -.31624 \times 10^{-5} & .36358 \times 10^{-5} & 0 & 0 & 0 & 0 \\ .36358 \times 10^{-5} & .23239 \times 10^{-4} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & .17895 \times 10^{-4} & 0 & 0 \\ 0 & 0 & 0 & 0 & .72143 \times 10^{-5} & -.21235 \times 10^{-1} \\ 0 & 0 & 0 & 0 & -.21235 \times 10^{-1} & .94209 \times 10^{2} \end{array}\right]$$

Table 5.4b

Propagation Matrices

D(2)

$$\begin{bmatrix} .51422 & .11137 & .80553 \times 10^{-1} \\ .19238 & .47348 & .34247 \\ .29340 & .41516 & .57698 \end{bmatrix}$$

D(3)

$$\begin{bmatrix} 1.0 & -.27756 \times 10^{-16} & 0 \\ .55511 \times 10^{-16} & 1.0 & 0 \\ 0 & 0 & 1.0 \end{bmatrix}$$

D(4)

$$\begin{bmatrix} .41266 & .25004 & 0 \\ .58735 & .74996 & 0 \\ 0 & 0 & 1.0 \end{bmatrix}$$

D(5)

$$\begin{bmatrix} 1.0 & 0 & 0 \\ 0 & 1.0 & 0 \\ 0 & 0 & 1.0 \end{bmatrix}$$

6. CONJECTURES

6.1 INTRODUCTION

In the stationary case, the approach of sections 2 and 3 is somewhat artificial in that we have to reduce the dimensionality of the state and the control to derive the results, and to reformulate those results in the initial parameterization $\underline{x}$, $\underline{\phi}$. In section 4, we have characterized one asymptotic property of the network directly in terms of the initial formulation, and have interpreted the steady-state constants in that manner. We try here to follow a similar approach for the eigenvalues of the asymptotic propagation matrix which measures the speed at which a perturbation settles down. We formulate a conjecture, explain how it arises, and illustrate it by examples in which it is observed to be true. Likewise, we conjecture that the propagation matrices at all stages are transposes of stochastic matrices, and present evidence for that property in view of what has been proved.

6.2 IS $\underline{D}^T(k)$ A STOCHASTIC MATRIX?

In section 2.3.2, we have observed that $\underline{D}^T(k)$ is a stochastic matrix in all numerical experiments. In section 2.3.6, we have proved that the sum of the entries of a column of $\underline{D}(k)$ is equal to 1. We have not proved that $\underline{D}(k)$ has only non-negative entries. This property is never needed in our argument, but we have shown that the limit $\underline{D}^T$ has spectral properties possessed by stochastic matrices. That is what enabled us to use, in section 2.3.7, a proof found in textbooks [32] on stochastic matrices to establish the convergence of $\underline{D}^k\underline{x}(1)$. Those spectral properties are:

a. The eigenvalues of $\underline{D}^T$ have magnitudes not greater than 1.

b. The number 1 is an eigenvalue of $\underline{D}^T$, and $\underline{v}_n^T = (1,1,\ldots,1)$ is the transpose of the corresponding eigenvector.

Moreover, when some reduced system is controllable, we have shown (in 2.3.7) that: <u>the number 1 is a simple eigenvalue, and is the only eigenvalue of $\underline{D}^T$ of magnitude 1</u>. This property implies that, if $\underline{D}^T$ is a stochastic matrix, then it corresponds to a Markov chain with one single final class, and that final class is aperiodic [32]. This property is to be compared with the conditions which we saw in 2.3.4 ensure controllability of a reduced system. They are: the accessibility graph G has only one final class, and that final class is aperiodic.

Now, let us relate the $\underline{D}^T(k)$ matrices (for any k) to the accessibility graph G. We have

$$\underline{x}^*(k+1) = \underline{D}(k)\underline{x}^*(k), \tag{6.1}$$

or

$$x_j^*(k+1) = \sum_{\ell=1}^{n} D_{j\ell}(k)\, x_\ell^*(k). \tag{6.2}$$

In particular, if

$$x_\ell(k) = \begin{cases} 0 & \text{for } \ell \neq i, \\ 1 & \text{for } \ell = i, \end{cases} \tag{6.3}$$

then

$$x_j^*(k+1) = D_{ji}(k) = D_{ij}^T(k), \tag{6.4}$$

so that the (i,j) entry of $\underline{D}^T(k)$ is the amount of flow perturbation which comes out of subnetwork k through exit j, given that one unit of flow perturbation enters through entrance i and nothing else enters through any other entrance. Therefore, if arc (i,j) does not exist in the graph G, there is no path leading from entrance i to exit j within the typical subnetwork and, by the above interpretation of (6.1),

$$D_{ij}^T(k) = 0, \text{ for all } k.$$

This does not imply that, if arc (i,j) exists in G, then $D_{ij}^T(k) \neq 0$. However, it seems reasonable that it will not be optimal to have $D_{ij}^T(k) = 0$ in spite of the existence of arc (i,j) in G; i.e., not to make use of the paths from entrance i to entrance j. We thus have

Conjecture 6.1: $D_{ij}(k) \geq 0$, for all (i,j),

and

Conjecture 6.2: $D^T_{ij}(k) = 0$, if, and only if, (i,j) is not an arc in G.

Combining these two is equivalent to stating

Conjecture 6.3: $\underline{D}^T(k)$ is a stochastic matrix adapted to the graph G (Appendix A).

This conjecture implies that the conditions on the final classes of G are equivalent to: Condition C. "the number 1 is a simple eigenvalue of $\underline{D}^T(k)$, and $\underline{D}^T(k)$ has no other eigenvalue of magnitude 1." Therefore, we conjecture that condition C holds, for $\underline{D}^T(k)$, for each k, in a stationary network whose accessibility graph is strongly connected and aperiodic (Conjecture 5.4). In the absence of the stationarity assumption, to each subnetwork k would correspond an accessibility graph G_k. By the same reasoning, Conjecture 5.4 applies to each $\underline{D}(k)$ corresponding to an accessibility graph G_k that is strongly connected and aperiodic.

If we could prove that $\underline{D}$ is a non-expanding mapping for the norm

$$||\underline{x}||_1 \triangleq \sum_{i=1}^{n} |x_i| \; ; \tag{6.5}$$

i.e., that

$$||\underline{D}\,\underline{x}||_1 \leq ||x||_1, \tag{6.6}$$

for any $\underline{x} \in R^n$, it follows, that $D_{ij} \geq 0 \;\forall\; (i,j)$ and, therefore, $\underline{D}^T$ is a stochastic matrix. Indeed, let us take $\underline{x}$ as in (6.3). Then,

$$\underline{\nu}_n^T \underline{D}\,\underline{x} = \underline{\nu}_n^T \underline{x}$$

becomes

$$\sum_{j=1}^{n} D^T_{ij} = 1.$$

Here, $||x||_1 = 1$, and $||\underline{D}\,\underline{x}||_1 = \sum_{j=1}^{n} |D^T_{ij}|$.

Let $J_1 = \{j: D^T_{ij} \geq 0\}$, and $J_2 = \{j: D^T_{ij} < 0\}$. Then,

$$||\underline{D}\ \underline{x}||_1 = \sum_{j\varepsilon J_1} D^T_{ij} - \sum_{j\varepsilon J_2} D^T_{ij} = \left(\sum_{j=1}^{n} D^T_{ij} - \sum_{j\varepsilon J_2} D^T_{ij}\right) - \sum_{j\varepsilon J_2} D^T_{ij} \geq \sum_{j=1}^{n} D^T_{ij} = 1,$$

since

$$- D^T_{ij} > 0 \text{ for } j \ \varepsilon\ J_2.$$

Thus, unless J_2 is empty, we have

$$||\underline{D}\ \underline{x}||_1 > ||\underline{x}||_1 \text{ for that } \underline{x}, \text{ contradicting (6.6).}$$

It should be noticed that condition (6.6) is also <u>necessary</u> for $\underline{D}^T$ to be a stochastic matrix [30]. Therefore, further investigation may follow these lines; i.e., try to prove (6.6).

6.3 RANDOM WALK-LIKE EXPRESSION FOR $\underline{D}^T$

In section 4, we have given a way to determine directly both the unit eigenvector $\underline{p}$ of $\underline{D}$ corresponding to eigenvalue 1 and the asymptotic cost increase rate per unit of flow, α. It will be useful also to have a compact expression for $\underline{D}$. To this end, we formulate here another conjecture which gives the matrix $\underline{D}$ as an analytical function of the one-step propagation matrix $\underline{D}(N-1)$. In spite of the rather unorthodox reasoning we are about to present, we obtain an equation which holds in all the numerical examples studied. To this end, we will concentrate on the case of a diagonal $\underline{L}$ matrix, and we will make use of the $\mathscr{L}$ formulation of (section 4.5). The minimization of

$$\sum_{i=1}^{n} \sum_{j=1}^{n} \mathscr{L}_{ij} \quad \phi^2_{ij}$$

under constrain (4.27) but without constraint (4.28) is in fact the general optimization problem of section 2.2 for one step (i.e., one single subnetwork), corresponding to a general incoming traffic-flow distribution $\underline{r}$. The propagation matrix for that problem is $\underline{D}(N-1)$.

Therefore, if $r_k = \begin{cases} 0 & \text{for } k \neq i \ , \\ 1 & \text{for } k = i \ , \end{cases}$

then, the outgoing traffic at exit j is $D^T_{ij}(N-1)$. It is also $\sum_{k=1}^{n} \phi_{kj}$. Accordingly, from (4.29),

$$D^T_{ij}(N-1) = \frac{1}{\mathscr{L}_{ij} \sum_{k=1}^{n} \frac{1}{\mathscr{L}_{ik}}} \ . \qquad (6.7)$$

This is the special case of diagonal $\underline{L}$ with the $\mathscr{L}$ - formulation. The general expression for $\underline{D}(N-1)$ is

$$\underline{D}(N-1) = (\underline{Z}\ \underline{L}^{-1}\ \underline{Y}^T)(\underline{Y}\ \underline{L}^{-1}\ \underline{Y}^T)^{-1} \qquad (6.8)$$

(see sections 2.2 and 2.37). For simplicity, let us denote $\underline{D}(N-1)$ by $\underline{T}$. To find a compact experssion giving $\underline{D}$ as a function of $\underline{T}$, consider the following probabilistic reasoning. If $\underline{D}^T$ is really a stochastic matrix, then, D^T_{ij} represents the proportion of the flow perturbation entering a subnetwork through entrance i which will leave it through exit j. This does not mean the proportion of that flow which will follow link (i,j) however, because there may be many other paths within the subnetwork which lead from entrance i to exit j, and we allow negative flow perturbations (2.3.4). Also, we may say that D^T_{ij} is the probability that a unit of flow perturbation entering through entrance i will leave the subnetwork through exit j. In this statement, we are thinking of the downstream perturbations, but the reasoning applies as well to upstream perturbations (section 3) by interchanging $\underline{Y}$ and $\underline{Z}$.

We assume that a "particle of flow" goes back and forth between entrances and exists until the time it leaves the subnetwork; i.e., goes from an exit, which is also an entrance to the immediately downstream subnetwork, forward to an exit of that downstream subnetwork.

We now make the following assumption: at each step forward within a subnetwork, the probability that the particle of flow at entrance i follows link (i,j) to exit j is given by T_{ji}. In other words, we assume that the percentage of the flow perturbation entering through node i which will follow link (i,j) is given by the j, i entry of the one-step propagation matrix $\underline{T}$.

When the flow particle comes backward from an exit to an entrance, we use the one-step propagation matrix equivalent to (6.7) but for the upstream direction of traffic; i.e.,

$$Q_{ij} = \frac{1}{\mathscr{L}_{ji} \sum_k \frac{1}{\mathscr{L}_{ki}}} , \tag{6.9}$$

and the corresponding optimal entering-flow distribution $\underline{q}$:

$$q_i = \frac{\sum_k \frac{1}{\mathscr{L}_{ki}}}{\sum_r \sum_s \frac{1}{\mathscr{L}_{rs}}} , \tag{6.10}$$

where we have replaced $\underline{\mathscr{L}}$ by $\underline{\mathscr{L}}^T$. This is the particular expression for a diagonal $\underline{L}$ matrix of

$$\underline{Q} = (\underline{Y}\ \underline{L}^{-1}\ \underline{Z}^T)(\underline{Z}\ \underline{L}^{-1}\ \underline{Z}^T)^{-1} , \tag{6.11}$$

where $\underline{Y}$ and $\underline{Z}$ have been interchanged in (6.8). By reasoning analogously, we assume that, at each step backward, the probability that the particle of flow being at exit j, next follows link (i,j) backward to entrance i, is Q_{ij}. In the symmetric case; i.e., when $\underline{\mathscr{L}} = \underline{\mathscr{L}}^T$, then $\underline{Q} = \underline{T}$. We assume in that case that, each time the particle of flow has reached an exit, either it goes forward to an exit of the next subnetwork or it comes back to an entrance of the same one both events being equally likely. With those assumptions, the probability that an entering particle of flow will leave the subnetwork after exactly n steps (i.e., after having followed n links, some forward and other backward) is the same as the probability of a gambler to be ruined in exactly n trials of a fair game(*) [33]. (In the so called "gambler's ruin" problem, it is assumed that the gambler must bet $1 each game, has an infinitely rich opponent,

---

* That is, at each trial, he has a probability $\frac{1}{2}$ to win 1 and a probability $\frac{1}{2}$ to lose 1.

and starts with a given capital which is here \$1.) This probability is known to be given by

$$a_n = \begin{cases} 0, \text{ if } n \text{ is even,} \\ \frac{1}{n} \begin{pmatrix} n \\ \frac{n+1}{2} \end{pmatrix} 2^{-n}, \quad \text{if } n \text{ is odd,} \end{cases} \tag{6.12}$$

with $a_n \geq 0$ and $\sum_{n=0}^{\infty} a_n = 1$.

According to our heuristic approach, we should therefore have:

$$D_{ji} = \sum_{n=1}^{\infty} a_n T_{ji}^{(n)}. \tag{6.13}$$

To obtain equation (6.13), we argue as follows. The probability that the particle, entering at i, leaves the subnetwork through j, is

$$D_{ji} = \sum_{n=1}^{\infty} P[j, n|i],$$

where $P[j,n|i]$ is the probability that the particle leaves the subnetwork after exactly n steps and through exit j, given that it enters through i. Also,

$$P[j, n|i] = P[n|i] \quad P[j|n, i]$$

Let $P(n|i)$ denote the probability that exactly n steps are required to leave the subnetwork, and $P(j|n, i)$ is the probability that the subnetwork is left through exit j, given that n steps are required and the particle entered through i. Then if

$$P[n|i] = a_n \qquad \text{for all } i,$$

$$P[j|n, i] = T_{ji}^{(n)},$$

equation (6.13) follows.

Since $a_n = 0$ for n even, equation (6.13) is equivalent to:

$$D_{ji} = \sum_{n=0}^{\infty} a_{2n+1} T_{ji}^{(2n+1)} . \tag{6.14}$$

From (6.7), it is clear that $T_{ji} \geq 0$; therefore, $\underline{T}^T$ is a stochastic matrix, and so is $(\underline{T}^T)^n$ for every positive integer n. Therefore, if (6.14) is established rigorously, it proves that $\underline{D}^T$ is a stochastic matrix (in the case of diagonal $\underline{L}$ and symmetric $\underline{\mathscr{L}}$). In matrix form, equation (6.14) reads:

$$\underline{D} = \sum_{n=0}^{\infty} a_{2n+1} \underline{T}^{2n+1} . \tag{6.15}$$

This will be transformed into the corresponding relations for the eigenvalues λ_T and λ_D of $\underline{T}$ and $\underline{D}$:

$$\lambda_D = \sum_{n=0}^{\infty} a_{2n+1} (\lambda_T)^{2n+1} . \tag{6.16}$$

And, given the particular sequence a_n (equation 6.12), this series is equialent to:

$$\lambda_D = \frac{1 - \sqrt{1-\lambda_T^2}}{\lambda_T} . \tag{6.17}$$

We know that $|\lambda_T| \leq 1$ since λ_T is the eigenvalue of a stochastic matrix; therefore, the series in (6.16) converges. However, λ_T may be complex since $\underline{T}$ is not necessarily symmetric although $\mathscr{L}$ is. In the examples, we have always found λ_T real, so that $1-\lambda_T^2 \geq 0$. It is clear that the eigenvalue 1 of $\underline{T}$ is transformed into 1 by (6.17). Moreover, equation (6.17) has been checked for the other eigenvalues in all our symmetric examples (i.e., $\underline{L}$ diagonal and $\underline{\mathscr{L}} = \underline{\mathscr{L}}^T$) with considerable accuracy.

Consider the standard two-dimensional example with $\underline{L}$ diagonal and

$$\underline{\mathscr{L}} = \begin{bmatrix} 1 & a \\ a & 1 \end{bmatrix} .$$

Equation (6.9) gives

$$\underline{T} = \begin{bmatrix} \frac{a}{a+1} & \frac{1}{a+1} \\ \frac{1}{a+1} & \frac{a}{a+1} \end{bmatrix},$$

so that the eigenvalues of $\underline{T}$ are 1 and $\frac{a-1}{a+1}$. Applying (6.17) with $\lambda_T = \frac{a-1}{a+1}$ gives

$$\lambda_D = \frac{1 - \sqrt{1 - \left(\frac{a-1}{a+1}\right)^2}}{\frac{a-1}{a+1}} = \frac{a+1 - \sqrt{(a+1)^2-(a-1)^2}}{a-1} = \frac{a+1 -2\sqrt{a}}{a-1} = \frac{\sqrt{a} - 1}{\sqrt{a} + 1},$$

which is exactly the expression found in section 2.3.8 when applying the hamiltonian matrix technique to that class of examples. It is much easier to obtain the results in the present manner because it avoids the laborious and artificial partitioning presented in section 2. However, the present method is not at all rigorously established.

Consider now a symmetric three-dimensional example (example 7.3 of section 7).

$$\underline{\mathscr{L}} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 2 \\ 3 & 2 & 1 \end{bmatrix}.$$

Equation (6.7) yields

$$\underline{T} = \begin{bmatrix} \frac{6}{11} & \frac{3}{11} & \frac{2}{11} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{2}{11} & \frac{3}{11} & \frac{6}{11} \end{bmatrix}.$$

So that the eigenvalues of $\underline{T}$ are 1, $\frac{4}{11}$ and $\frac{5}{22}$. Applying (6.17) to $\lambda_T = \frac{4}{11}$ yields

$$\lambda_1(D) = \frac{1 - \sqrt{1 - \left(\frac{4}{11}\right)^2}}{\frac{4}{11}} \simeq 0.18826231,$$

and, with $\lambda_T = \frac{5}{22}$, one obtains

$$\lambda_2(D) = \frac{1 - \sqrt{1 - \left(\frac{5}{22}\right)^2}}{\frac{5}{22}} \simeq 0.11514294.$$

On the other hand, the computer program shows that

$$\underline{D} = \begin{bmatrix} (D_{11} - D_{13}) & (D_{12} - D_{13}) \\ (D_{21} - D_{23}) & (D_{22} - D_{23}) \end{bmatrix} = \begin{bmatrix} 0.18826231 & 0.036559683 \\ 0 & 0.11514299 \end{bmatrix},$$

so that the two eigenvalues of $\underline{\Delta}$, which are the eigenvalues of $\underline{D}$ other than 1, are indeed those found from equation (6.17).

In the nonsymmetric case (when $\underline{L}$ diagonal but $\mathscr{L}$ not symmetric), equation (6.17) does not hold, but, following the same heuristics, we now argue that $\underline{Q}$, given by (6.9) must be used instead of $\underline{T}$ for backward transitions. Also, the gambler's ruin problem may be such that the probability of a forward and a backward transition are not the same. If α is the probability of a forward, and $\beta = 1 - \alpha$ that of a backward transition, a_n of equation (6.13) becomes

$$a_n = \begin{cases} 0, & \text{for } n \text{ even}, \\ \frac{1}{n}\binom{n}{\frac{n+1}{2}} \alpha^{\frac{n-1}{2}} \beta^{\frac{n+1}{2}}, & \text{for } n \text{ odd}, \end{cases} \qquad (6.18)$$

we replace $(\underline{T}^T)^{2n+1}$ in (6.15) by $\underline{T}^T(\underline{Q}^T\underline{T}^T)^n$.

Therefore,

$$D^T = \sum_{n=0} an\, T^T(Q^T\, T^T)^n, \qquad (6.19)$$

which, given equation (6.18) for a_n, is formally equivalent to

$$\underline{D}^T = \underline{T}^T \frac{(I - \sqrt{I - 4\,\alpha\beta\, \underline{Q}^T\, \underline{T}^T})(\underline{Q}^T\, \underline{T}^T)^{-1}}{2\,\alpha} . \qquad (6.20)$$

If true, equation (6.19) again implies that $\underline{D}^T$ is stochastic. That $\underline{D}$ will now be given as a function of two matrices does not make as simple the translation into eigenvalue relations. However, the equation:

$$\lambda_D = \frac{1 - \sqrt{1 - 4\,\alpha\beta\, \lambda_Q\, \lambda_T}}{2\alpha\, \lambda_Q} \qquad (6.21)$$

is found to hold in all two-dimensional examples, with α and β chosen as follows:

$$\alpha + \beta = 1, \qquad \left\{ \frac{\alpha}{\beta} = \frac{p_1 p_2}{q_1 q_2} . \right. \qquad (6.22)$$

where p_1 and p_2 are the two components of $\underline{p}$ given by (4.32), and q_1 and q_2 those of $\underline{q}$ given by (6.10).

<u>Example</u>. For the standard two-dimensional example,

$$\text{with } \underline{L} = \begin{bmatrix} 2 & 1 \\ 3 & 4 \end{bmatrix}, \text{ or } \underline{\mathscr{L}} = \begin{bmatrix} 2 & 3 \\ 4 & 1 \end{bmatrix}, \quad \underline{\mathscr{L}} \neq \underline{\mathscr{L}}^T.$$

$$\text{Equation (4.32) gives } \underline{p} = \begin{bmatrix} 2/5 \\ 3/5 \end{bmatrix}, \text{ and (6.10) gives } \underline{q} = \begin{bmatrix} 9/25 \\ 16/25 \end{bmatrix} .$$

So, according to (6.22), we have $\frac{\alpha}{\beta} = \frac{2/5 \quad 3/5}{9/25 \quad 16/25}$ and $\alpha + \beta = 1$,

whence $\alpha = \frac{144}{294}$, $\beta = \frac{150}{294}$.

From (6.7), $\underline{T} = \begin{bmatrix} 3/5 & 2/5 \\ \frac{1}{5} & \frac{4}{5} \end{bmatrix}$. $\underline{T}$ has eigenvalues 1 and 2/5.

From (6.9), $\underline{Q} = \begin{bmatrix} 2/3 & 1/3 \\ \frac{1}{3} & 3/4 \end{bmatrix}$. $\underline{Q}$ has eigenvalues 1 and 5/12.

If we now apply (6.21) with the values found for α and β, first of all we see that $\alpha < \beta$, so that $\alpha < 1/2$; therefore, for $\lambda Q = 1$ and $\lambda_T = 1$,

$$\sqrt{1 - 4\alpha\beta} = \sqrt{1 - 4\alpha(-\alpha)} = \sqrt{(1-2\alpha)^2} = 1 - 2\alpha,$$

and $\lambda_D = \frac{1 - (1-2\alpha)}{2\alpha} = 1$ is found.

Corresponding to $\lambda_T = 2/5$ and $\lambda Q = 5/12$, we find

$$\lambda_D = \frac{1 - \sqrt{1 - 4 \cdot \frac{144}{294} \cdot \frac{150}{294} \cdot \frac{2}{5} \cdot \frac{5}{12}}}{2 \cdot \frac{144}{294} \cdot \frac{5}{12}} \simeq 0.21337308.$$

On the other hand, either from the implementation of the equations of section 2 or from the computer results (section 7), we see that $\Delta \equiv D_{11} - D_{12} = 0.21337308$ is the eigenvalue of $\underline{D}$ other than 1. In three-dimensional nonsymmetric examples, we have tried $\frac{\alpha}{\beta} = \frac{p_1 p_2 p_3}{q_1 q_2 q_3}$, but the results obtained, although close to the true values, are not as satisfactory. Those equations are, of course, only conjectural, but the coincidenc is striking. Perhaps, the equations derived in 2.3.7 for $\underline{\Delta}$ and $\underline{K}$ can be used together with (6.8) and (6.11) to derive them at least for the range of cases to which they apply.

7. NUMERICAL EXPERIMENTATION

7.1 INTRODUCTION

In this section, we survey the examples of stationary network patterns to which we have applied the equations of section 2.2.

For each stationary example, we sketch the typical subnetwork and its associated accessibility graph G. We also give the incidence matrices $\underline{Y}$ and $\underline{Z}$ and the cost matrix $\underline{L}$.

For some examples, we refer to the corresponding table which lists the computed values for

$$\underline{DC}(k) = \underline{C}(k) - \underline{C}(k+1),$$

$$\underline{D}(k) = \frac{\partial x^*(k+1)}{\partial x^*(k)},$$

$\underline{\delta x}(k)$ (denoted DELTA(k)), the perturbation at stage k.

For others, we just indicate the limiting values, and ocasionally show a graph.

7.2 VARIOUS STATIONARY NETWORKS

Example 7.1 Standard two-dimensional example.

The network and its accessibility graph G are given in Fig. 7.1. The numbers are link and node labels. Since the graph G is strongly connected and aperiodic, any subsystem is controllable. The same is still true when the arrows are reversed in G; i.e., for the upstream case. The incidence matrices are:

$$\underline{Y} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} \text{ and } \underline{Z} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix},$$

corresponding to the labeling of the links indicated on Fig. 7.1. For the upstream perturbations, $\underline{Y}$ and $\underline{Z}$ are interchanged. We have studied this example for various cost matrices. For some cases, we present the upstream perturbations as well as the downstream ones. The penalty function used is as indicated.

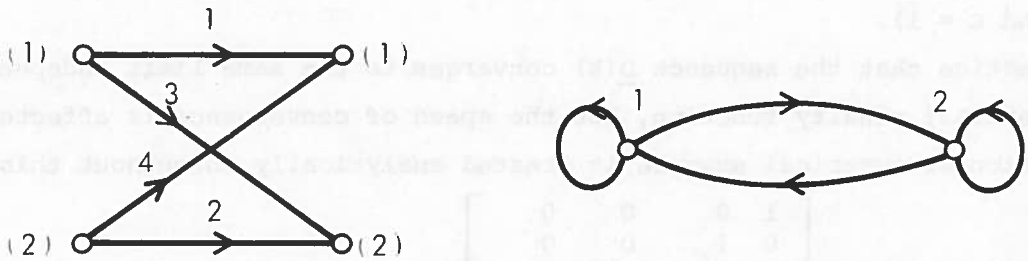

Figure 7.1 STANDARD TWO-DIMENSIONAL EXAMPLE AND CORRESPONDING ACCESSIBILITY GRAPH

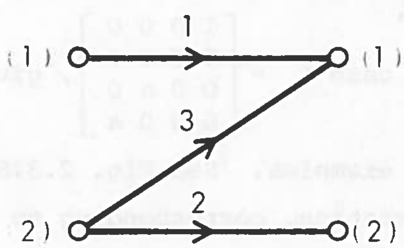

Figure 7.2 LIMITING CASE OF STANDARD TWO-DIMENSIONAL EXAMPLE (SAME ACCESSIBILITY GRAPH)

$$\underline{7.1.1} \quad \underline{L} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}.$$

See table 7.1.1-a (downstream); table 7.1.1-b (upstream with penalty $\frac{1}{\varepsilon}\underline{S}$, where $\underline{S} = \underline{I}$, and $\varepsilon = 10^{-6}$); table 7.1.1-c (upstream, with penalty $\frac{1}{\varepsilon}\underline{S}$, where $\underline{S} = \underline{I}$, and $\varepsilon = 1$).

We notice that the sequence $\underline{D}(k)$ converges to the same limit independently of the terminal penalty function, but the speed of convergence is affected. This particular numerical example is treated analytically throughout this work.

$$\underline{7.1.2} \quad \underline{L} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 10^{-10} & 0 \\ 0 & 0 & 0 & 10^{-10} \end{bmatrix}.$$

See table 7.1.2 (downstream perturbations).

$$\underline{7.1.3} \quad \underline{L} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

See table 7.1.3 (downstream perturbations).

The expressions for $\underline{\Delta}$ and $\hat{\underline{K}}$ in the special case $L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & a & 0 \\ 0 & 0 & 0 & a \end{bmatrix}$, given in section 2.3.7, have been checked in a series of examples. See Fig. 2.3.8-5 for the decay of one component of a zero-flow perturbation, corresponding to $a = 4$. Also, Fig. 2.3.8-1 shows that the eigenvalue $\lambda(a)$ other than 1 increases slowly with a.

Example 7.1.2 is the uncontrollable case corresponding to $a \to 0$. We have approximated $a = 0$ by $a = 10^{-10}$ because $\underline{L}$ has to be inverted. One can observe that the perturbations do not decrease in that case. Rather, they alternate in sign, which is to be expected from the structure of the subnetworks. In example 7.1.3, $a = 1$, which makes det $(\underline{D}(N-1))$ and det $(\underline{D})$ equal to zero; one can observe that a zero total-flow perturbation is driven to zero in one step (apart from numerical errors) as predicted in section 2.3.7. After stage 3, the numbers are just rounding errors.

$\underline{7.1.4}$ We consider the same structure, but $\underline{L} = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix}$.

Table 7.1.1a RESULTS OF EXAMPLE 7.1, DOWNSTREAM

DC(K)=C(K-1)-C(K)

K	DC11	DC12	DC13
19	0.66666666666666670D 00	0.00000000000000000D 00	0.66666666666666670D 00
18	0.35294117647058820D 00	0.31372549019607850D 00	0.35294117647058820D 00
17	0.33392751039809830D 00	0.33273915626856770D 00	0.33392751039809830D 00
16	0.33335083941669690D 00	0.33331582724306890D 00	0.33335083941669690D 00
15	0.33333384867744620D 00	0.33333241798921950D 00	0.33333384867744640D 00
14	0.33333334850365830D 00	0.33333331816300690D 00	0.33333334850365830D 00
13	0.33333333377990490D 00	0.33333333288675960D 00	0.33333333377990520D 00
12	0.33333333334647900D 00	0.33333333332018710D 00	0.33333333334647870D 00
11	0.33333333333371980D 00	0.33333333333274580D 00	0.33333333333371960D 00
10	0.33333333333334330D 00	0.33333333333332060D 00	0.33333333333334370D 00
9	0.33333333333333170D 00	0.33333333333333100D 00	0.33333333333333170D 00
8	0.33333333333332990D 00	0.33333333333333020D 00	0.33333333333333040D 00
7	0.33333333333333150D 00	0.33333333333333130D 00	0.33333333333333080D 00
6	0.33333333333333020D 00	0.33333333333333040D 00	0.33333333333333060D 00
5	0.33333333333333060D 00	0.33333333333333060D 00	0.33333333333333060D 00
4	0.33333333333332900D 00	0.33333333333332900D 00	0.33333333333332900D 00
3	0.33333333333333060D 00	0.33333333333333060D 00	0.33333333333333060D 00
2	0.33333333333332700D 00	0.33333333333332730D 00	0.33333333333332750D 00
1	0.33333333333332820D 00	0.33333333333332790D 00	0.33333333333332770D 00

D(K)=MATRIX DERIVATIVE OF X(K+1) W. RESPECT TO X(K)

K	D11	D12	D13
20	0.66666666666666667D 00	0.33333333333333333D 00	
19	0.58823529411764710D 00	0.41176470588235300D 00	
18	0.58585858585858570D 00	0.41414141414141400D 00	
17	0.58578856152512980D 00	0.41421143847486980D 00	
16	0.58578650014867650D 00	0.41421349985132300D 00	
15	0.58578643946737390D 00	0.41421356053262580D 00	
14	0.58578643768108270D 00	0.41421356231891630D 00	
13	0.58578643762849970D 00	0.41421356237150020D 00	
12	0.58578643762695150D 00	0.41421356237304790D 00	
11	0.58578643762690570D 00	0.41421356237309350D 00	
10	0.58578643762690420D 00	0.41421356237309480D 00	
9	0.58578643762690420D 00	0.41421356237309470D 00	
8	0.58573643762690450D 00	0.41421356237309480D 00	
7	0.58578643762690430D 00	0.41421356237309480D 00	
6	0.58578643762690450D 00	0.41421356237309490D 00	
5	0.58578643762690450D 00	0.41421356237309490D 00	
4	0.58578643762690440D 00	0.41421356237309480D 00	
3	0.58578643762690450D 00	0.41421356237309470D 00	
2	0.58578643762670460D 00	0.41421356237309470D 00	

K	D21	D22	D23
20	0.33333333333333333D 00	0.66666666666666667D 00	
19	0.41176470588235290D 00	0.58823529411764710D 00	
18	0.41414141414141400D 00	0.58585858585858580D 00	
17	0.41421143847487000D 00	0.58578856152513000D 00	
16	0.41421349985132320D 00	0.58578650014867660D 00	
15	0.41421356053262580D 00	0.58578643946737390D 00	
14	0.41421356231891660D 00	0.58578643768108300D 00	
13	0.41421356237147990D 00	0.58578643762849960D 00	
12	0.41421356237304770D 00	0.58578643762695170D 00	
11	0.41421356237309330D 00	0.58578643762690620D 00	
10	0.41421356237309470D 00	0.58578643762690480D 00	
9	0.41421356237309460D 00	0.58578643762690480D 00	
8	0.41421356237309460D 00	0.58578643762690490D 00	
7	0.41421356237309450D 00	0.58578643762690460D 00	
6	0.41421356237309510D 00	0.58578643762690460D 00	
5	0.41421356237309490D 00	0.58578643762690490D 00	
4	0.41421356237309470D 00	0.58578643762690460D 00	
3	0.41421356237309450D 00	0.58578643762690460D 00	
2	0.41421356237309450D 00	0.58578643762690440D 00	

DELTA(K)=PERTURBATION AT STAGE K

K	DELTA1	DELTA2	DELTA3
2	0.17157287525380990D 00	-0.17157287525380990D 00	
3	0.29437251522859370D-01	-0.29437251522859500D-01	
4	0.50506338833465090D-02	-0.50506338833466570D-02	
5	0.86655177722001340D-03	-0.86655177722016260D-03	
6	0.14867677907387460D-03	-0.14867677997402360D-03	
7	0.25508902623533980D-04	-0.25504902623683040D-04	
8	0.43766357676274340D-05	-0.43766357677765050D-05	
9	0.75091198252875470D-06	-0.75091198267762960D-06	
10	0.12883612784324800D-06	-0.12883612799232330D-06	
11	0.22104784828884770D-07	-0.22104784977960170D-07	
12	0.37925814282115760D-08	-0.37925815772869830D-08	
13	0.65070403853550220D-09	-0.65070418761091010D-09	
14	0.11164310115225420D-09	-0.11164325022766220D-09	
15	0.19154966528840000D-10	-0.19155015604248020D-10	
16	0.32863961716017350D-11	-0.32865452470097510D-11	
17	0.56380865158765330D-12	-0.56395772692566900D-12	
18	0.96753888740014740D-13	-0.96902964148030530D-13	
19	0.17012831668466680D-13	-0.17161907076482480D-13	
20	0.56212520868169620D-14	-0.57703274948327570D-14	

Table 7.1.1b RESULTS OF EXAMPLE 7.1, UPSTREAM, LARGE PENALTY

DC(K)=C(K-1)-C(K)

K	DC11	DC12	DC13
19	-0.4999992916737375D 06	0.4999999583263250D 06	-0.4999992916737375D 06
18	0.3124698813626310D 00	0.3541364873090060D 00	0.3124698814353906D 00
17	0.3327076647983631D 00	0.3338981393171707D 00	0.3327076647983631D 00
16	0.3332875091000460D 00	0.3333225230599055D 00	0.3332875091145979D 00
15	0.3333122397889383D 00	0.3333132704574382D 00	0.3333122397598345D 00
14	0.3333126841753256D 00	0.3333127145160688D 00	0.3333126841607736D 00
13	0.3333235369791510D 00	0.3333235378813697D 00	0.3333235369937029D 00
12	0.3333135729626520D 00	0.3333135729772039D 00	0.3333135729481000D 00
11	0.3333001393184531D 00	0.3333001393475570D 00	0.3333001393621089D 00
10	0.3332971142226597D 00	0.3332971142081078D 00	0.3332971142081078D 00
9	0.3333183756039944D 00	0.3333183755894424D 00	0.3333183755748905D 00
8	0.3333084114710800D 00	0.3333084114710800D 00	0.3333084114710800D 00
7	0.3333019167330349D 00	0.3333019167766906D 00	0.3333019168057945D 00
6	0.3333127697551390D 00	0.3333127697405871D 00	0.3333127697114833D 00
5	0.3333236227190355D 00	0.3333236226899317D 00	0.3333236226608278D 00
4	0.3333067193889292D 00	0.3333067194034811D 00	0.3333067194471369D 00
3	0.3333141028706450D 00	0.3333141028851969D 00	0.3333141028851969D 00
2	0.3333076082053594D 00	0.3333076082053594D 00	0.3333076081908075D 00
1	0.3333045830368064D 00	0.3333045830368064D 00	0.3333045830368064D 00

D(K)=MATRIX DERIVATIVE OF X(K+1) W. RESPECT TO X(K)

K	D11	D12	D13
20	0.5000001249863999D 00	0.4999998749844963D 00	
19	0.5833333367536397D 00	0.4166666631517728D 00	
18	0.5857142857730651D 00	0.4142857141441709D 00	
17	0.5857843137073360D 00	0.4142156862653792D 00	
16	0.5857863750816250D 00	0.4142136248692623D 00	
15	0.5857864357640210D 00	0.4142135642041467D 00	
14	0.5857864375529971D 00	0.4142135623842478D 00	
13	0.5857864376066573D 00	0.4142135623778813D 00	
12	0.5857864375911959D 00	0.4142135623596914D 00	
11	0.5857864376039288D 00	0.4142135623578724D 00	
10	0.5857864376312136D 00	0.4142135623587819D 00	
9	0.5857864375957433D 00	0.4142135623660579D 00	
8	0.5857864375811914D 00	0.4142135623314971D 00	
7	0.5857864376039288D 00	0.4142135623733338D 00	
6	0.5857864376412181D 00	0.4142135623833383D 00	
5	0.5857864375911959D 00	0.4142135623305876D 00	
4	0.5857864375930149D 00	0.4142135623460490D 00	
3	0.5857864375893769D 00	0.4142135623424110D 00	
2	0.5857864375993813D 00	0.4142135623496870D 00	
K	D21	D22	D23
20	0.4999998749990482D 00	0.5000001249863999D 00	
19	0.4166666631790577D 00	0.5833333368063904D 00	
18	0.4142857141559944D 00	0.5857142857712461D 00	
17	0.4142156862653792D 00	0.5857843137200685D 00	
16	0.4142136248847237D 00	0.5857863750861725D 00	
15	0.4142135642177891D 00	0.5857864357676590D 00	
14	0.4142135624133516D 00	0.5857864375720965D 00	
13	0.4142135623696959D 00	0.5857864376012003D 00	
12	0.4142135623551440D 00	0.5857864376157522D 00	
11	0.4142135623405920D 00	0.5857864375866484D 00	
10	0.4142135623696959D 00	0.5857864376012003D 00	
9	0.4142135623696959D 00	0.5857864376012003D 00	
8	0.4142135623696959D 00	0.5857864376012003D 00	
7	0.4142135623551440D 00	0.5857864376012003D 00	
6	0.4142135623696959D 00	0.5857864376012003D 00	
5	0.4142135623696959D 00	0.5857864376012003D 00	
4	0.4142135623842478D 00	0.5857864376157522D 00	
3	0.4142135623405920D 00	0.5857864376012003D 00	
2	0.4142135623551440D 00	0.5857864376157522D 00	

DELTA(K)=PERTURBATION AT STAGE K

K	DELTA1	DELTA2	DELTA3
2	0.1715728752496943D 00	-0.1715728752606083D 00	
3	0.2943725151645833D-01	-0.2943725152971295D-01	
4	0.5050633876556619D-02	-0.5050633889356096D-02	
5	0.8665517707877339D-03	-0.8665517834402167D-03	
6	0.1486767736330278D-03	-0.1486767862626571D-03	
7	0.2550889630086316D-04	-0.2550890893279119D-04	
8	0.4376629450469479D-05	-0.4376642081933499D-05	
9	0.7509056665082570D-06	-0.7509182979802378D-06	
10	0.1288298121341763D-06	-0.1288424435754242D-06	
11	0.2209846917099289D-07	-0.2211110061224009D-07	
12	0.3786265779605271D-08	-0.3798897221495311D-08	

Table 7.1.1c RESULTS OF EXAMPLE 7.1, UPSTREAM, SMALL PENALTY

DC(K)=C(K-1)-C(K)

K	DC11	DC12	DC13
19	0.1904761904761905D 00	0.4761904761904762D 00	0.1904761904761905D 00
18	0.3298490127758416D 00	0.3368176518908242D 00	0.3298490127758418D 00
17	0.3332312821036154D 00	0.3334353845630499D 00	0.3332312821036154D 00
16	0.3333303296718693D 00	0.3333363369947959D 00	0.3333303296718691D 00
15	0.3333332449141817D 00	0.3333334217524842D 00	0.3333332449141815D 00
14	0.3333333307305157D 00	0.3333333359361490D 00	0.3333333307305162D 00
13	0.3333333332567119D 00	0.3333333334099515D 00	0.3333333332567121D 00
12	0.3333333333107710D 00	0.3333333333355877D 00	0.3333333333107766D 00
11	0.3333333333332658D 00	0.3333333333333995D 00	0.3333333333332658D 00
10	0.3333333333333284D 00	0.3333333333333324D 00	0.3333333333333286D 00
9	0.3333333333333306D 00	0.3333333333333308D 00	0.3333333333333306D 00
8	0.3333333333333315D 00	0.3333333333333313D 00	0.3333333333333313D 00
7	0.3333333333333330D 00	0.3333333333333328D 00	0.3333333333333326D 00
6	0.3333333333333306D 00	0.3333333333333306D 00	0.3333333333333306D 00
5	0.3333333333333290D 00	0.3333333333333293D 00	0.3333333333333293D 00
4	0.3333333333333306D 00	0.3333333333333304D 00	0.3333333333333304D 00
3	0.3333333333333282D 00	0.3333333333333294D 00	0.3333333333333286D 00
2	0.3333333333333277D 00	0.3333333333333282D 00	0.3333333333333282D 00
1	0.3333333333333266D 00	0.3333333333333264D 00	0.3333333333333264D 00

D(K)=MATRIX DERIVATIVE OF X(K+1) W. RESPECT TO X(K)

K	D11	D12	D13
20	0.5714285714285714D 00	0.4285714285714286D 00	
19	0.5853658536585364D 00	0.4146341463414633D 00	
18	0.5857740585774054D 00	0.4142259414225940D 00	
17	0.5857860732232588D 00	0.4142139267767407D 00	
16	0.5857864268998644D 00	0.4142135731001354D 00	
15	0.5857864373111301D 00	0.4142135626888692D 00	
14	0.5857864376176090D 00	0.4142135623823903D 00	
13	0.5857864376266310D 00	0.4142135623733687D 00	
12	0.5857864376268964D 00	0.4142135623731028D 00	
11	0.5857864376269040D 00	0.4142135623730947D 00	
10	0.5857864376269044D 00	0.4142135623730949D 00	
9	0.5857864376269043D 00	0.4142135623730947D 00	
8	0.5857864376269049D 00	0.4142135623730952D 00	
7	0.5857864376269047D 00	0.4142135623730949D 00	
6	0.5857864376269045D 00	0.4142135623730949D 00	
5	0.5857864376269044D 00	0.4142135623730948D 00	
4	0.5857864376269043D 00	0.4142135623730945D 00	
3	0.5857864376269043D 00	0.4142135623730947D 00	
2	0.5857864376269041D 00	0.4142135623730944D 00	

K	D21	D22	D23
20	0.4285714285714286D 00	0.5714285714285715D 00	
19	0.4146341463414634D 00	0.5853658536585366D 00	
18	0.4142259414225940D 00	0.5857740585774058D 00	
17	0.4142139267767407D 00	0.5857860732232589D 00	
16	0.4142135731001354D 00	0.5857864268998643D 00	
15	0.4142135626888696D 00	0.5857864373111302D 00	
14	0.4142135623823902D 00	0.5857864376176091D 00	
13	0.4142135623733683D 00	0.5857864376266311D 00	
12	0.4142135623731028D 00	0.5857864376268967D 00	
11	0.4142135623730948D 00	0.5857864376269043D 00	
10	0.4142135623730947D 00	0.5857864376269046D 00	
9	0.4142135623730947D 00	0.5857864376269046D 00	
8	0.4142135623730947D 00	0.5857864376269049D 00	
7	0.4142135623730949D 00	0.5857864376269049D 00	
6	0.4142135623730947D 00	0.5857864376269049D 00	
5	0.4142135623730947D 00	0.5857864376269049D 00	
4	0.4142135623730949D 00	0.5857864376269049D 00	
3	0.4142135623730947D 00	0.5857864376269046D 00	
2	0.4142135623730947D 00	0.5857864376269044D 00	

DELTA(K)=PERTURBATION AT STAGE K

K	DELTA1	DELTA2	DELTA3
2	0.1715728752538097D 00	-0.1715728752538097D 00	
3	0.2943725152285932D-01	-0.2943725152285937D-01	
4	0.5050633883346539D-02	-0.5050633983346610D-02	
5	0.8665517772200498D-03	-0.8665517772201245D-03	
6	0.1486767739739117D-03	-0.1486767749973786D-03	
7	0.2550890262357795D-04	-0.2550890262364611D-04	
8	0.4376635767664389D-05	-0.4376635767739555D-05	
9	0.7509119825657078D-06	-0.7509119826408765D-06	
10	0.1288361278802919D-06	-0.1288361279553700D-06	
11	0.2210478486583752D-07	-0.2210478494100556D-07	
12	0.3792581465162249D-08	-0.3792581540331298D-08	
13	0.6507040754740780D-09	-0.6507041506431286D-09	
14	0.1116431380203720D-09	-0.1116432131893729D-09	
15	0.1915490298595614D-10	-0.1915497815500693D-10	
16	0.3286430233516332D-11	-0.3285505402567129D-11	
17	0.5638287532434296D-12	-0.5639039722942182D-12	
18	0.9669262404582642D-13	-0.9676779309466151 7D-13	
19	0.1647732913310673D-13	-0.1655249018389540D-13	
20	0.2321688854391492D-14	-0.2396857905180255D-14	

Table 7.1.2 RESULTS OF EXAMPLE 7.1.2

DELTA(K)=PERTURBATION AT STAGE K

K	DELTA1	DELTA2	DELTA3
2	-0.99999999260000005D 00	0.99999999260000003D 00	
3	0.99999998560000008D 00	-0.99999998560000008D 00	
4	-0.99999997900000013D 00	0.99999997900000011D 00	
5	0.99999997280000016D 00	-0.99999997280000016D 00	
6	-0.99999996700000019D 00	0.99999996700000019D 00	
7	0.99999996160000022D 00	-0.99999996160000022D 00	
8	-0.99999995660000026D 00	0.99999995660000026D 00	
9	0.99999995200000029D 00	-0.99999995200000029D 00	
10	-0.99999994780000032D 00	0.99999994780000032D 00	
11	0.99999994400000034D 00	-0.99999994400000035D 00	
12	-0.99999994060000037D 00	0.99999994060000037D 00	
13	0.99999993760000038D 00	-0.99999993760000038D 00	
14	-0.99999993500000042D 00	0.99999993500000040D 00	
15	0.99999993280000042D 00	-0.99999993280000042D 00	
16	-0.99999993100000043D 00	0.99999993100000042D 00	
17	0.99999992960000043D 00	-0.99999992960000043D 00	
18	-0.99999992860000043D 00	0.99999992860000043D 00	
19	0.99999992800000045D 00	-0.99999992800000045D 00	
20	-0.99999992780000045D 00	0.99999992780000045D 00	

Table 7.1.3 RESULTS OF EXAMPLE 7.1.3

DELTA(K)=PERTURBATION AT STAGE K

K	DELTA1	DELTA2	DELTA3
2	0.83266726846886740D-16	-0.44408920985006270D-15	
3	-0.18041124150158881D-15	-0.18041124150158881D-15	
4	-0.18041124150158883D-15	-0.18041124150158883D-15	
5	-0.18041124150158883D-15	-0.18041124150158883D-15	
6	-0.18041124150158883D-15	-0.18041124150158884D-15	
7	-0.18041124150158884D-15	-0.18041124150158884D-15	
8	-0.18041124150158886D-15	-0.18041124150158886D-15	
9	-0.18041124150158888D-15	-0.18041124150158888D-15	
10	-0.18041124150158889D-15	-0.18041124150158889D-15	
11	-0.18041124150158887D-15	-0.18041124150158888D-15	
12	-0.18041124150158887D-15	-0.18041124150158888D-15	
13	-0.18041124150158887D-15	-0.18041124150158887D-15	
14	-0.18041124150158889D-15	-0.18041124150158888D-15	
15	-0.18041124150158888D-15	-0.18041124150158888D-15	
16	-0.18041124150158888D-15	-0.18041124150158888D-15	
17	-0.18041124150158888D-15	-0.18041124150158888D-15	
18	-0.18041124150158888D-15	-0.18041124150158889D-15	
19	-0.18041124150158889D-15	-0.18041124150158889D-15	
20	-0.18041124150158889D-15	-0.18041124150158889D-15	

See table 7.1.4-a for the downstream perturbations and table 7.1.4-b for the upstream perturbations. The entering flow perturbation contributes a total flow of 1. One can observe that it is quickly redistributed according to $\underline{p}^T = (\frac{11}{29}, \frac{18}{29}) = (0.37931, 0.62069)$ obtained in section 4.3, and also that $\alpha = \frac{14}{29}$.

In table 7.1.4-b, we used, as a terminal penalty, $\frac{1}{\varepsilon}\underline{S}$, with $\varepsilon = 1$ and $\underline{S} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \geq 0$. The same cost-increase rate and distribution $\underline{p}$ are observed, and also the same propagation matrix $\underline{D}$.

7.2 We have also considered the subnetwork of Fig. 7.2. The accessibility graph G is still the same as in example 7.1 because the arrows in the network are not taken into account to construct G. Accordingly, this example brings nothing new which is not contained already in example 7.1. It is the limiting case of ex. 7.1 when the cost corresponding to link (1,2) goes to infinity. The incidence matrices are

$$\underline{Y} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}, \text{ and } \underline{Z} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

We have noticed almost no difference with respect to the case where the structure of 7.1 is used with a very large cost (10^{10}) along link (1,2).

Example 7.3 See Fig. 7.3 and table 7.3.

In this example, every entrance is connected to every exit by a link. The accessibility graph G is complete. The incidence matrices are:

$$\underline{Y} = \begin{bmatrix} 1&1&1&0&0&0&0&0&0 \\ 0&0&0&1&1&1&0&0&0 \\ 0&0&0&0&0&0&1&1&1 \end{bmatrix}, \text{ and } \underline{Z} = \begin{bmatrix} 1&0&0&1&0&0&1&0&0 \\ 0&1&0&0&1&0&0&1&0 \\ 0&0&1&0&0&1&0&0&1 \end{bmatrix}.$$

We consider

$$\underline{L} \doteq \begin{bmatrix} 1&0&0&0&0&0&0&0&0 \\ 0&2&0&0&0&0&0&0&0 \\ 0&0&3&0&0&0&0&0&0 \\ 0&0&0&2&0&0&0&0&0 \\ 0&0&0&0&1&0&0&0&0 \\ 0&0&0&0&0&2&0&0&0 \\ 0&0&0&0&0&0&3&0&0 \\ 0&0&0&0&0&0&0&2&0 \\ 0&0&0&0&0&0&0&0&1 \end{bmatrix}, \text{ or } \mathscr{L} = \begin{bmatrix} 1&2&3 \\ 2&1&2 \\ 3&2&1 \end{bmatrix},$$

and input the perturbation $\underline{\delta x}^T = (-1, 2, 1)$.

Table 7.1.4a RESULTS OF EXAMPLE 7.1.4, DOWNSTREAM

DC(K)=C(K-1)-C(K)

K	DC11	DC12	DC13
19	0.1200000000000000D 01	0.0000000000000000D 00	0.8000000000000000D 00
18	0.5244444444444443D 00	0.4355555555555554D 00	0.5244444444444442D 00
17	0.4800211137503249D 00	0.4795783187437316D 00	0.4376220638692045D 00
16	0.4831626995913391D 00	0.4824187214186908D 00	0.4835957148097816D 00
15	0.4824922822734432D 00	0.4827011751274591D 00	0.4829230910264114D 00
14	0.4827006594292664D 00	0.4827471019724456D 00	0.4827943431766535D 00
13	0.4827462017980286D 00	0.4827561945927126D 00	0.4827662237488708D 00
12	0.4827559684807083D 00	0.4827581044685643D 00	0.4827602421055113D 00
11	0.4827580546742505D 00	0.4827585106076544D 00	0.4827589666164285D 00
10	0.4827584999123447D 00	0.4827585972041093D 00	0.4827586944993052D 00
9	0.4827585949198085D 00	0.4827586156786086D 00	0.4827586364385652D 00
8	0.4827586151908387D 00	0.4827586196204374D 00	0.4827586240509434D 00
7	0.4827586195163565D 00	0.4827586204615142D 00	0.4827586214066721D 00
6	0.4827586204393040D 00	0.4827586206409753D 00	0.4827586208426466D 00
5	0.4827586206362335D 00	0.4827586206792651D 00	0.4827586207222967D 00
4	0.4827586206782537D 00	0.4827586206874357D 00	0.4827586206966175D 00
3	0.4827586206872192D 00	0.4827586206891781D 00	0.4827586206911372D 00
2	0.4827586206891366D 00	0.4827586206895542D 00	0.4827586206899719D 00
1	0.4827586206895433D 00	0.4827586206896322D 00	0.4827586206897212D 00

D(K)=MATRIX DERIVATIVE OF X(K+1) W. RESPECT TO X(K)

K	D11	D12	D13
20	0.6000000000000000D 00	0.2000000000000000D 00	
19	0.5111111111111110D 00	0.2888888888888888D 00	
18	0.5106888361045130D 00	0.2969121140142517D 00	
17	0.5114805616098129D 00	0.2980891074053423D 00	
16	0.5116896544638288D 00	0.2983157423419973D 00	
15	0.5117360970070073D 00	0.2983629835456049D 00	
14	0.5117460898016916D 00	0.2983730127017629D 00	
13	0.5117482257831728D 00	0.2983751503387118D 00	
12	0.5117486817165765D 00	0.2983756063474857D 00	
11	0.5117487790083412D 00	0.2983757036426816D 00	
10	0.5117487997681416D 00	0.2983757244026384D 00	
9	0.5117488041977399D 00	0.2983757288322439D 00	
8	0.5117488051428979D 00	0.2983757297774023D 00	
7	0.5117488053445696D 00	0.2983757299790736D 00	
6	0.5117488053876005D 00	0.2983757300221049D 00	
5	0.5117488053967825D 00	0.2983757300312867D 00	
4	0.5117488053987415D 00	0.2983757300332454D 00	
3	0.5117488053991595D 00	0.2983757300336641D 00	
2	0.5117488053992487D 00	0.2983757300337531D 00	

K	D21	D22	D23
20	0.3999999999999999D 00	0.8000000000000000D 00	
19	0.4888888888888887D 00	0.7111111111111110D 00	
18	0.4893111163895486 9D 00	0.7030378859857482D 00	
17	0.4885194383901869D 00	0.7019108925946575D 00	
16	0.4883103455361709D 00	0.7016842576580023D 00	
15	0.4882639029929910D 00	0.7016370164543948D 00	
14	0.4882539101983077D 00	0.7016269872982368D 00	
13	0.4882517742168269D 00	0.7016248496612880D 00	
12	0.4882513182834228D 00	0.7016243936525139D 00	
11	0.4882512209916582D 00	0.7016242963573180D 00	
10	0.4882512002318578D 00	0.7016242755973612D 00	
9	0.4882511958022591D 00	0.7016242711677554D 00	
8	0.4882511948571016D 00	0.7016242702225974D 00	
7	0.4882511946554304D 00	0.7016242700209259D 00	
6	0.4882511946123991D 00	0.7016242699778950D 00	
5	0.4882511946032171D 00	0.7016242699687130D 00	
4	0.4882511946012580D 00	0.7016242699667536D 00	
3	0.4882511946008401D 00	0.7016242699663355D 00	
2	0.4882511946007513D 00-	0.7016242699662465D 00	

DELTA(K)=PERTURBATION AT STAGE K

K	DELTA1	DELTA2	DELTA3
2	0.7251218807647443D 00	0.2748781192352561D 00	
3	0.4530972157472498D 00	0.5469027842527502D 00	
4	0.3950544763967798D 00	0.6049455236032195D 00	
5	0.3826697185969730D 00	0.6173302814030258D 00	
6	0.3800271447283895D 00	0.6199728552716091D 00	
7	0.3794632905721361D 00	0.6205367094278602D 00	
8	0.3793429790750895D 00	0.6206570209249083D 00	
9	0.3793173068958041D 00	0.6206826931041929D 00	
10	0.3793118247143556D 00	0.6206881752856409D 00	
11	0.3793106342045426D 00	0.6206893657954534D 00	
12	0.3793102828879079D 00	0.6206897171120874D 00	
13	0.3793097519462118D 00	0.6206902480537833D 00	
14	0.3793075016485351D 00	0.6206924983514593D 00	
15	0.3792970061316175D 00	0.6207029938683768D 00	
16	0.3792478283963971D 00	0.6207521716035968D 00	
17	0.3790174530107352D 00	0.6209825469892585D 00	
18	0.3779372013562122D 00	0.6220627986437811D 00	
19	0.3728749336347113D 00	0.6271250663652814D 00	
20	0.3491499734538834D 00	0.6508500265461100D 00	

Table 7.1.4b RESULTS OF EXAMPLE 7.1.4, UPSTREAM

DC(K)=C(K-1)-C(K)

K	DC11	DC12	DC13
19	0.3749999999999998D 00	0.9374999999999998D 00	-0.1562500000000000D 00
18	0.5446310935441372D 00	0.5104578342621869D 00	0.4217308959156782D 00
17	0.4994917184027030D 00	0.4867166570992522D 00	0.4714891245951245D 00
16	0.4864968334014352D 00	0.4835175460063656D 00	0.4804266662970369D 00
15	0.4835639831231000D 00	0.4829166781055176D 00	0.4822642926378239D 00
14	0.4829308161087256D 00	0.4827921695651343D 00	0.4826532917187334D 00
13	0.4827953786250596D 00	0.4827657710905442D 00	0.4827361530252512D 00
12	0.4827664645750647D 00	0.4827601450272981D 00	0.4827538270000855D 00
11	0.4827602943969209D 00	0.4827589461390036D 00	0.4827575978592584D 00
10	0.4827589778152377D 00	0.4827586901310255D 00	0.4827584024458196D 00
9	0.4827586968907083D 00	0.4827586355065396D 00	0.4827585741223255D 00
8	0.4827586369489101D 00	0.4827586238511767D 00	0.4827586107534412D 00
7	0.4827586241589403D 00	0.4827586213642363D 00	0.4827586185695325D 00
6	0.4827586214299049D 00	0.4827586208335901D 00	0.4827586202372753D 00
5	0.4827586208476009D 00	0.4827586207203636D 00	0.4827586205931260D 00
4	0.4827586207233547D 00	0.4827586206962056D 00	0.4827586206690564D 00
3	0.4827586206968435D 00	0.4827586206910504D 00	0.4827586206852574D 00
2	0.4827586206911845D 00	0.4827586206899486D 00	0.4827586206887129D 00
1	0.4827586206899783D 00	0.4827586206897148D 00	0.4827586206894510D 00

D(K)=MATRIX DERIVATIVE OF X(K+1) W. RESPECT TO X(K)

K	D11	D12	D13
20	0.5624999999999999D 00	0.4062500000000000D 00	
19	0.5283267457180501D 00	0.3175230566534913D 00	
18	0.5155516844145990D 00	0.3022955241493634D 00	
17	0.5125723970195298D 00	0.2992046444400347D 00	
16	0.5119250920019471D 00	0.2985522589723460D 00	
15	0.5117864454583558D 00	0.2984133811259450D 00	
14	0.5117568379238408D 00	0.2983837630506600D 00	
13	0.5117505193760738D 00	0.2983774440334473D 00	
12	0.5117491711181568D 00	0.2983760957537019D 00	
11	0.5117488834339441D 00	0.2983758080684959D 00	
10	0.5117488220497757D 00	0.2983757466842822D 00	
9	0.5117488089520420D 00	0.2983757335865464D 00	
8	0.5117488061573382D 00	0.2983757307918424D 00	
7	0.5117488055610236D 00	0.2983757301955281D 00	
6	0.5117488054337861D 00	0.2983757300682905D 00	
5	0.5117488054066373D 00	0.2983757300411415D 00	
4	0.5117488054008442D 00	0.2983757300353486D 00	
3	0.5117488053996081D 00	0.2983757300341127D 00	
2	0.5117488053993447D 00	0.2983757300338487D 00	
K	**D21**	**D22**	**D23**
20	0.4374999999999999D 00	0.5937500000000002D 00	
19	0.4716732542819498D 00	0.6824769433465085D 00	
18	0.4844483155854007D 00	0.6977044758506364D 00	
17	0.4874276029804700D 00	0.7007953555599651D 00	
16	0.4880749079980528D 00	0.7014477410276539D 00	
15	0.4882135545416438D 00	0.7015866188740549D 00	
14	0.4882431620761590D 00	0.7016162369393399D 00	
13	0.4882494806239257D 00	0.7016225559665523D 00	
12	0.4882508288818432D 00	0.7016239042462977D 00	
11	0.4882511165660551D 00	0.7016241919315039D 00	
10	0.4882511779502236D 00	0.7016242533157174D 00	
9	0.4882511910479573D 00	0.7016242664134531D 00	
8	0.4882511938426613D 00	0.7016242692081571D 00	
7	0.4882511944389758D 00	0.7016242698044717D 00	
6	0.4882511945662134D 00	0.7016242699317090D 00	
5	0.4882511945933625D 00	0.7016242699588582D 00	
4	0.4882511945991554D 00	0.7016242699646513D 00	
3	0.4882511946003916D 00	0.7016242699658872D 00	
2	0.4882511946006554D 00	0.7016242699661512D 00	

DELTA(K)=PERTURBATION AT STAGE K

K	DELTA1	DELTA2	DELTA3
2	0.7251218807648405D 00	0.2748781192351595D 00	
3	0.4530972157477188D 00	0.5469027842522810D 00	
4	0.3950544763989827D 00	0.6049455236010168D 00	
5	0.3826697186072980D 00	0.6173302813927013D 00	
6	0.3800271447767782D 00	0.6199728552232205D 00	
7	0.3794632907989172D 00	0.6205367092010810D 00	
8	0.3793429801379183D 00	0.6206570198620794D 00	
9	0.3793173118768854D 00	0.6206826881231117D 00	
10	0.3793118480588244D 00	0.6206881519411772D 00	
11	0.3793107436113578D 00	0.6206892563886383D 00	
12	0.3793107956368344D 00	0.6206892043631614D 00	
13	0.3793121550091359D 00	0.6206878449908594D 00	
14	0.3793187639079374D 00	0.6206812360920578D 00	
15	0.3793497881397624D 00	0.6206502118602325D 00	
16	0.3794951979769045D 00	0.6205048020230797D 00	
17	0.3801766819470758D 00	0.6198233180529187D 00	
18	0.3833705435637729D 00	0.6166294564362215D 00	
19	0.3983389815154180D 00	0.6016610184845762D 00	
20	0.4684904658617316D 00	0.5315095341382125D 00	

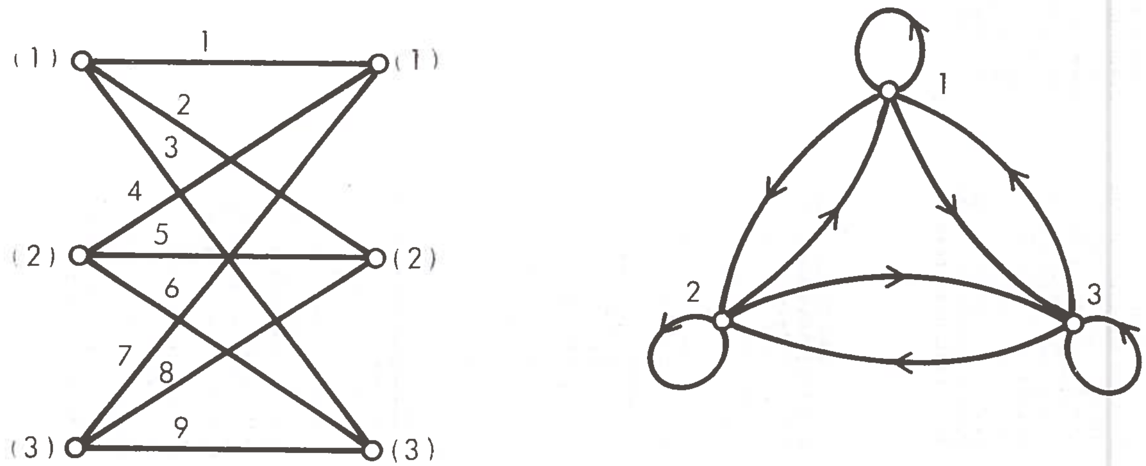

Figure 7.3 EXAMPLE OF SUBNETWORK WITH ALL LINKS PRESENT; CORRESPONDING ACCESSIBILITY GRAPH

Table 7.3 RESULTS OF EXAMPLE 7.3

DC(K)=C(K-1)-C(K)

K	DC11	DC12	DC13
19	0.5454545454545455D 00	0.0000000000000000D 00	0.0000000000000000D 00
18	0.1983305876641964D 00	0.1717921527041358D 00	0.1597143366585648D 00
17	0.1772134242497264D 00	0.1764077398267617D 00	0.1757963141210793D 00
16	0.1764961874178161D 00	0.1764697548511150D 00	0.1764458981991501D 00
15	0.1764714854943863D 00	0.1764705771863469D 00	0.1764697030295998D 00
14	0.1764706199029247D 00	0.1764705880888085D 00	0.1764705567274665D 00
13	0.1764705893559078D 00	0.1764705882333517D 00	0.1764705871167988D 00
12	0.1764705882749884D 00	0.1764705882352684D 00	0.1764705881956283D 00
11	0.1764705882367001D 00	0.1764705882352935D 00	0.1764705882338873D 00

K	DC22	DC23	DC33
19	0.5000000000000000D 00	0.0000000000000000D 00	0.5454545454545455D 00
18	0.1850477200424180D 00	0.1717921527041358D 00	0.1983305876641964D 00
17	0.1765858103176036D 00	0.1764077398267617D 00	0.1772134242497261D 00
16	0.1764721161062998D 00	0.1764697548511149D 00	0.1764961874178159D 00
15	0.1764706084916285D 00	0.1764705771863468D 00	0.1764714854943863D 00
14	0.1764705885038518D 00	0.1764705880888088D 00	0.1764706199029253D 00
13	0.1764705882388544D 00	0.1764705882333518D 00	0.1764705893559075D 00
12	0.1764705882353408D 00	0.1764705882352682D 00	0.1764705882749884D 00
11	0.1764705882352948D 00	0.1764705882352935D 00	0.1764705882366999D 00

D(K)=MATRIX DERIVATIVE OF X(K+1) W. RESPECT TO X(K)

K	D11	D12	D13
9	0.4379799088799248D 00	0.2862772831877294D 00	0.2497176003698245D 00
8	0.4379799088799251D 00	0.2862772831877297D 00	0.2497176003698248D 00
7	0.4379799088799247D 00	0.2862772831877297D 00	0.2497176003698244D 00
6	0.4379799088799245D 00	0.2862772831877297D 00	0.2497176003698247D 00
5	0.4379799088799247D 00	0.2862772831877297D 00	0.2497176003698242D 00
4	0.4379799088799242D 00	0.2862772831877296D 00	0.2497176003698245D 00
3	0.4379799088799246D 00	0.2862772831877296D 00	0.2497176003698240D 00
2	0.4379799088799247D 00	0.2862772831877295D 00	0.2497176003698242D 00

K	D21	D22	D23
9	0.3123024907502506D 00	0.4274454336245403D 00	0.3123024907502505D 00
8	0.3123024907502509D 00	0.4274454336245410D 00	0.3123024907502511D 00
7	0.3123024907502508D 00	0.4274454336245411D 00	0.3123024907502509D 00
6	0.3123024907502502D 00	0.4274454336245400D 00	0.3123024907502502D 00
5	0.3123024907502498D 00	0.4274454336245402D 00	0.3123024907502502D 00
4	0.3123024907502501D 00	0.4274454336245401D 00	0.3123024907502500D 00
3	0.3123024907502505D 00	0.4274454336245398D 00	0.3123024907502503D 00
2	0.3123024907502505D 00	0.4274454336245399D 00	0.3123024907502501D 00

K	D31	D32	D33
9	0.2497176003698249D 00	0.2862772831877303D 00	0.4379799088799255D 00
8	0.2497176003698249D 00	0.2862772831877300D 00	0.4379799088799254D 00
7	0.2497176003698244D 00	0.2862772831877300D 00	0.4379799088799255D 00
6	0.2497176003698250D 00	0.2862772831877303D 00	0.4379799088799255D 00
5	0.2497176003698249D 00	0.2862772831877298D 00	0.4379799088799252D 00
4	0.2497176003698250D 00	0.2862772831877304D 00	0.4379799088799254D 00
3	0.2497176003698249D 00	0.2862772831877305D 00	0.4379799088799254D 00
2	0.2497176003698250D 00	0.2862772831877304D 00	0.4379799088799253D 00

DELTA(K)=PERTURBATION AT STAGE K

K	DELTA1	DELTA2	DELTA3
2	0.3842922578653587D 00	0.8548908672490794D 00	0.7608168748855611D 00
3	0.6030374872984897D 00	0.7230396317919130D 00	0.6739228809095963D 00
4	0.6393985298196795D 00	0.7078578925197637D 00	0.6527435776605553D 00
5	0.6456889040521388D 00	0.7061098223820181D 00	0.6482012735658407D 00
6	0.6468092355367546D 00	0.7059085444420072D 00	0.6472822200212358D 00
7	0.6470127930706998D 00	0.7058853687076603D 00	0.6471018302216384D 00
8	0.6470502679844590D 00	0.7058827001854048D 00	0.6470670318301367D 00
9	0.6470572255379071D 00	0.7058823929238987D 00	0.6470603915381951D 00
10	0.6470585241495976D 00	0.7058823575449049D 00	0.6470591183054982D 00
11	0.6470587673357876D 00	0.7058823534712638D 00	0.6470588791929501D 00
12	0.6470588129696503D 00	0.7058823530022134D 00	0.6470588340281380D 00
13	0.6470588215436187D 00	0.7058823529482050D 00	0.6470588255081595D 00
14	0.6470588231558231D 00	0.7058823529419864D 00	0.6470588239021926D 00
15	0.6470588234591094D 00	0.7058823529412707D 00	0.6470588235796231D 00
16	0.6470588235161808D 00	0.7058823529411888D 00	0.6470588235426347D 00
17	0.6470588235267224D 00	0.7058823529411796D 00	0.6470588235319031D 00
18	0.6470588235287440D 00	0.7058823529411786D 00	0.6470588235298829D 00
19	0.6470588235293222D 00	0.7058823529411785D 00	0.6470588235295050D 00
20	0.6470588235293804D 00	0.7058823529411785D 00	0.6470588235294469D 00

The sequence of downstream perturbations converges to $\underline{p}$, the fixed point of $\underline{D}$. One verifies that $\underline{p}$ is given by equation (4.32); i.e.,

$$\underline{p}^T = (\frac{11}{34}, \frac{12}{34}, \frac{11}{34}). \quad \text{Also, } \alpha = \left(\sum_i \sum_j \frac{1}{\mathscr{L}_{ij}}\right)^{-1} = \frac{3}{17} \cong 0.17647059.$$

7.4 See Fig. 2.3.4-8 for this example. The accessibility graph is again complete.

The incidence matrices are

$$\underline{Y} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}, \text{ and } \underline{Z} = \begin{bmatrix} 0 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 0 \end{bmatrix}.$$

We consider:

$$\underline{L} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 7 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 5 \end{bmatrix},$$

or

$$\underline{\mathscr{L}} = \begin{bmatrix} +\infty & 1 & 3 \\ 7 & +\infty & 2 \\ 3 & 5 & +\infty \end{bmatrix}.$$

Table 7.4.1 shows the downstream perturbations and table 7.4.2 the upstream perturbations. We have chosen $\underline{S} = \underline{I}$ as a penalty. The initial perturbation is $\underline{\delta x}^T = (-1, 1, 1)$. The same stationary distribution is converged to by the downstream and upstream perburtabions although the limiting propagation matrices differ. In fact, we observe:

$$\underline{D}(d) = \begin{bmatrix} 0.14656750 & 0.29141406 & 0.56558282 \\ 0.53676130 & 0.18965133 & 0.30141036 \\ 0.31667120 & 0.51893461 & 0.13300682 \end{bmatrix},$$

and

$$\underline{D}(u) = \begin{bmatrix} 0.14364343 & 0.51897671 & 0.33009378 \\ 0.29248921 & 0.19781874 & 0.54214274 \\ 0.56386736 & 0.28320455 & 0.12776348 \end{bmatrix}.$$

In this example, the graph G is complete although not every entrance is connected with every exit. In particular, G is strongly connected and aperiodic, so that

Table 7.4.1 RESULTS OF EXAMPLE 7.4, DOWNSTREAM

DELTA(K)=PERTURBATION AT STAGE K

K	DELTA1	DELTA2	DELTA3
2	0.7104293879217859D 00	-0.4556960334472256D-01	0.3352702154229361D 00
3	0.2804314238386069D 00	0.4737179249017126D 00	0.2458506512596787D 00
4	0.3181991009095605D 00	0.3144679056174490D 00	0.3673329934729881D 00
5	0.3460352415798521D 00	0.3411541904542912D 00	0.3128105679658532D 00
6	0.3270549414530845D 00	0.3447230080668679D 00	0.3282220504800427D 00
7	0.3340294901501544D 00	0.3398571838152441D 00	0.3261133760345956D 00
8	0.3324411087145208D 00	0.3420422882856918D 00	0.3255166027997800D 00
9	0.3325078233547761D 00	0.3414245358792564D 00	0.3260676407659596D 00
10	0.3326480262709051D 00	0.3415093967176719D 00	0.3258425770114128D 00
11	0.3325695774255298D 00	0.3415298245971230D 00	0.3259005979773357D 00
12	0.3325907331927208D 00	0.3415227051745090D 00	0.3258865516327578D 00
13	0.3325809564330137D 00	0.3414912331036085D 00	0.3259278104633647D 00
14	0.3326622464548221D 00	0.3415476158664865D 00	0.3257901376786776D 00
15	0.3323116313010061D 00	0.3416319819274472D 00	0.3260563867715324D 00
16	0.3332548567682132D 00	0.3407036470233169D 00	0.3260414962084552D 00
17	0.3319121870493715D 00	0.3443736383772510D 00	0.3237141745733623D 00
18	0.3298389209232848D 00	0.3352180652213070D 00	0.3349430138553927D 00
19	0.3517796968404032D 00	0.3458999072031239D 00	0.3023203959564572D 00
20	0.2658168935179243D 00	0.3772049211139738D 00	0.3569781853680860D 00

Table 7.4.2 RESULTS OF EXAMPLE 7.4, UPSTREAM

K	DELTA1	DELTA2	DELTA3
2	0.7054270580719675D 00	0.4474722716426983D 00	-0.1528973297146672D 00
3	0.2830865323117460D 00	0.2119540417406477D 00	0.5049585259476043D 00
4	0.3173468681682628D 00	0.3984380174237583D 00	0.2341651144084460D 00
5	0.3461919341314446D 00	0.3257069836375000D 00	0.3281010822310518D 00
6	0.3270666517869722D 00	0.3435659752186064D 00	0.3293673729944166D 00
7	0.3340058572035689D 00	0.3421913955242869D 00	0.3238027472871389D 00
8	0.3324523920738449D 00	0.3409322080958364D 00	0.3266153998303127D 00
9	0.3325039794791941D 00	0.3417539834307213D 00	0.3257420311900774D 00
10	0.3326507378736575D 00	0.3414571348226595D 00	0.3258921273036749D 00
11	0.3325637469599148D 00	0.3415240999757805D 00	0.3259121530642959D 00
12	0.3325991844737062D 00	0.3415252649468579D 00	0.3258755505794266D 00
13	0.3325934266600118D 00	0.3414915959553972D 00	0.3259149774345812D 00
14	0.3325261444722295D 00	0.3415985324729484D 00	0.3258733230543118D 00
15	0.3328487412235594D 00	0.3413502105078465D 00	0.3258010482685834D 00
16	0.3319168417900292D 00	0.3416000590918803D 00	0.3264330991180794D 00
17	0.3334664643254876D 00	0.3426296346180175D 00	0.3239039308564836D 00
18	0.3344462269513075D 00	0.3354805190572891D 00	0.3300732539913918D 00
19	0.3158609578773939D 00	0.3603109708708518D 00	0.3238230712517524D 00
20	0.3956493111605510D 00	0.3061571865930860D 00	0.2981935022463508D 00

any subsystem is controllable. Note that the subnetwork is of class 2 (section 2.3.4).

We present a graph for the sequence of perturbations in entrance node 1 (Fig. 7.4.1). The behavior demonstrated by that graph is very clear and in accordance with the theoretical predictions. The initial perturbation settles down very quickly to the stationary distribution, and stays at that constant level up to the last stages when an end effect takes place at the exit. The end effect is simply the result of the important difference between the one-step propagation matrix and the other propagation matrices.

Example 7.5 See Fig. 7.5.1 and tables 7.5.1 (downstream), 7.5.2 (upstream with = 1 and $\underline{S} = \underline{I}$). The initial perturbation is $\underline{\delta x}^T = (-1, 2, 1)$. Although several links are missing in the network, the accessibility graph G is still complete.

$$\underline{Y} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}, \text{ and } \underline{Z} = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \end{bmatrix},$$

We used the cost matrix,

$$\underline{L} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 6 & 0 & 0 \\ 0 & 0 & 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 10 \end{bmatrix}.$$

Again, the downstream and upstream stationary distributions are the same:

$$\underline{p}^T = (0.1313, 0.4425, 0.2441),$$

as apparent from tables 7.5.1 and 7.5.2.

The asymptotic cost-increase rate is

= 0.43540328, and the propagation matrices are

$$\underline{D}(d) = \begin{bmatrix} 0.65909001 & 0.65205236 & 0.65205236 \\ 0.29043477 & 0.08898132 & 0.08898132 \\ 0.05047522 & 0.25896632 & 0.25896632 \end{bmatrix}.$$

$$= \begin{bmatrix} 0.65824275 & 0.86512167 & 0.27034476 \\ 0.22026586 & 0.08779220 & 0.46865304 \\ 0.12149131 & 0.04708563 & 0.26100220 \end{bmatrix}.$$

We show a graph for the convergence of one component of the downstream perturbations (Fig. 7.5.2). The same behaviors are observed as in the graph of Fig. 7.4.2.

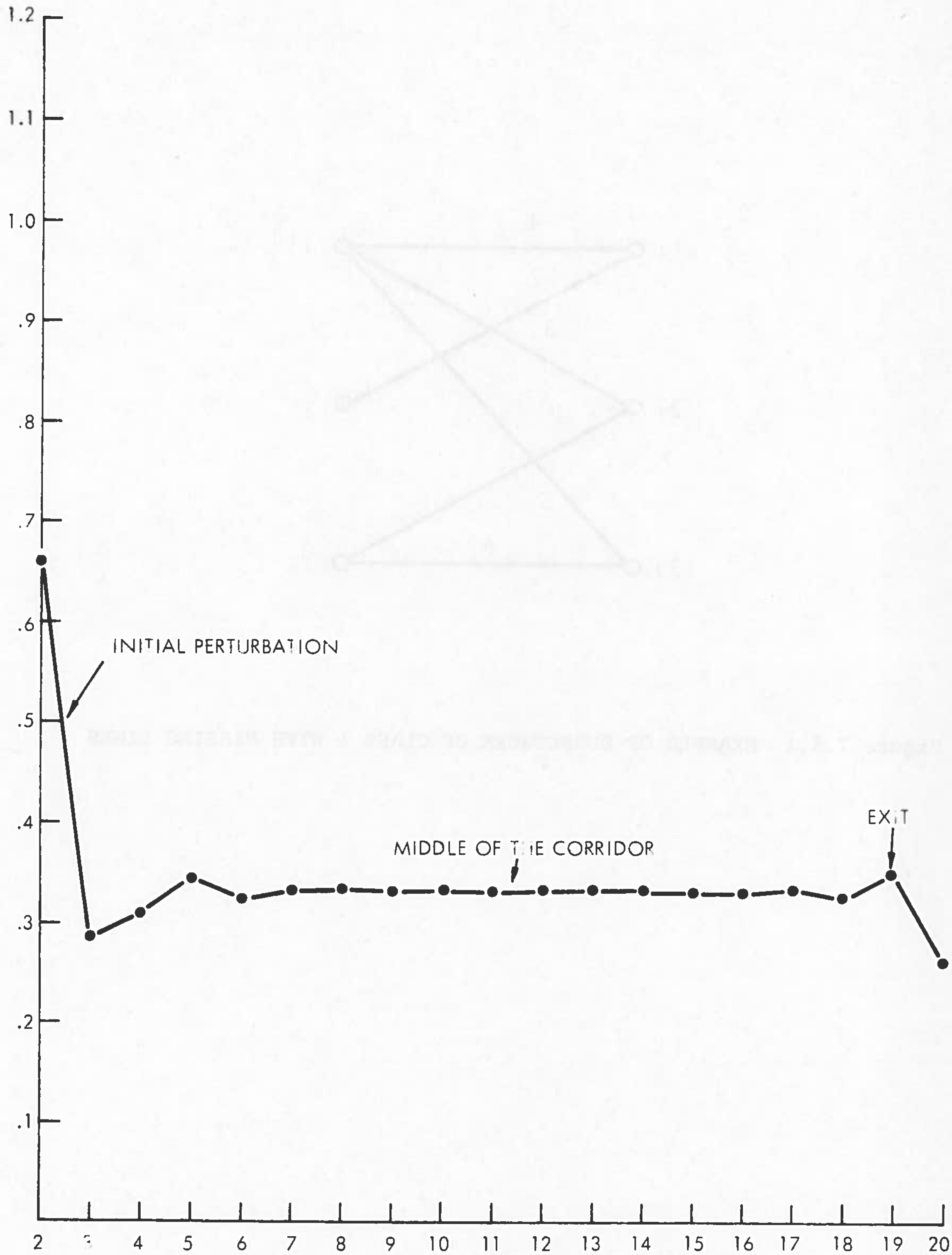

Figure 7.4.1 $x_1(k)$ VERSUS k. INITIAL PERTURBATION $\underline{x}(1) = (-1,1,1)$

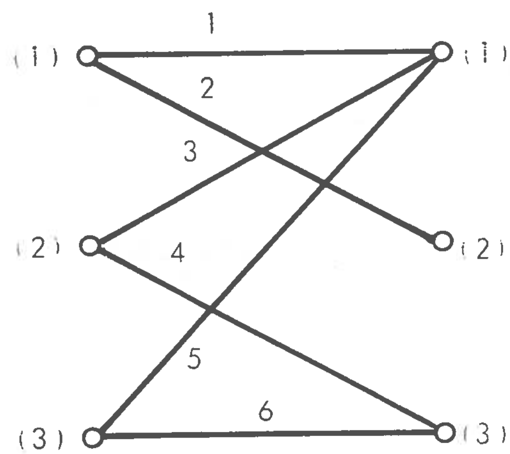

Figure 7.5.1 EXAMPLE OF SUBNETWORK OF CLASS 1 WITH MISSING LINKS

DELTA(K)=PERTURBATION AT STAGE K

K	DELTA1	DELTA2	DELTA3
2	0.1297067086786037D 01	-0.2349081689253385D-01	0.7264237301064928D 00
3	0.1313233032296180D 01	0.4392612747598714D 00	0.2475056929439435D 00
4	0.1313346802501559D 01	0.4425179602704473D 00	0.2441352372279862D 00
5	0.1313347603176014D 01	0.4425408796709340D 00	0.2441115171530427D 00
6	0.1313347608810876D 01	0.4425410409695658D 00	0.2441113502195466D 00
7	0.1313347608850530D 01	0.4425410421047277D 00	0.2441113490447275D 00
8	0.1313347608850808D 01	0.4425410421127164D 00	0.2441113490364592D 00
9	0.1313347608850808D 01	0.4425410421127720D 00	0.2441113490364009D 00
10	0.1313347608850807D 01	0.4425410421127720D 00	0.2441113490364002D 00
11	0.1313347608850806D 01	0.4425410421127716D 00	0.2441113490363998D 00
12	0.1313347608850804D 01	0.4425410421127710D 00	0.2441113490363996D 00
13	0.1313347608850803D 01	0.4425410421127701D 00	0.2441113490363990D 00
14	0.1313347608850907D 01	0.4425410421127021D 00	0.2441113490363609D 00
15	0.1313347608865861D 01	0.4425410421031260D 00	0.2441113490309823D 00
16	0.1313347610990757D 01	0.4425410407424663D 00	0.2441113482667449D 00
17	0.1313347912923628D 01	0.4425408474023179D 00	0.2441112396740210D 00
18	0.1313390815459812D 01	0.4425133751279548D 00	0.2440958094121989D 00
19	0.1319486963918795D 01	0.4386097582811349D 00	0.2419032778000355D 00
20	0.1333333333333310D 01	0.4398289879729316D 00	0.2268376786937234D 00

TABLE 7.5.2 RESULTS OF EXAMPLE 7.5, UPSTREAM

DELTA(K)=PERTURBATICN AT STAGF K

K	DELTA1	DELTA2	DELTA3
2	0.1342345359689770D 01	0.4239725738516808D 00	0.2336820614585465D 00
3	0.1313551684762520D 01	0.4424103638374166D 00	0.2440379514000578D 00
4	0.1313349045064886D 01	0.4425401224453037D 00	0.2441108324898038D 00
5	0.1313347618958381D 01	0.4425410356404807D 00	0.2441113454011292D 00
6	0.1313347608921946D 01	0.4425410420672246D 00	0.2441113490108175D 00
7	0.1313347608851312D 01	0.4425410421124528D 00	0.2441113490362207D 00
8	0.1313347608850813D 01	0.4425410421127710D 00	0.2441113490363994D 00
9	0.1313347608850809D 01	0.4425410421127726D 00	0.2441113490364007D 00
10	0.1313347608850808D 01	0.4425410421127725D 00	0.2441113490364002D 00
11	0.1313347608850807D 01	0.4425410421127719D 00	0.2441113490364001D 00
12	0.1313347608850805D 01	0.4425410421127718D 00	0.2441113490363999D 00
13	0.1313347608850804D 01	0.4425410421127652D 00	0.2441113490364064D 00
14	0.1313347608850772D 01	0.4425410421118615D 00	0.2441113490373411D 00
15	0.1313347608846286D 01	0.4425410419834963D 00	0.2441113491701899D 00
16	0.1313347608209093D 01	0.4425410237437962D 00	0.2441113680470816D 00
17	0.1313347517668575D 01	0.4425384320108135D 00	0.2441140503205819D 00
18	0.1313334652498131D 01	0.4421701649937172D 00	0.2444951825081200D 00
19	0.1311506602318823D 01	0.3898420135973533D 00	0.2986513840837909D 00
20	0.1051753537009650D 01	0.5603274554033690D 00	0.3879100075869478D 00

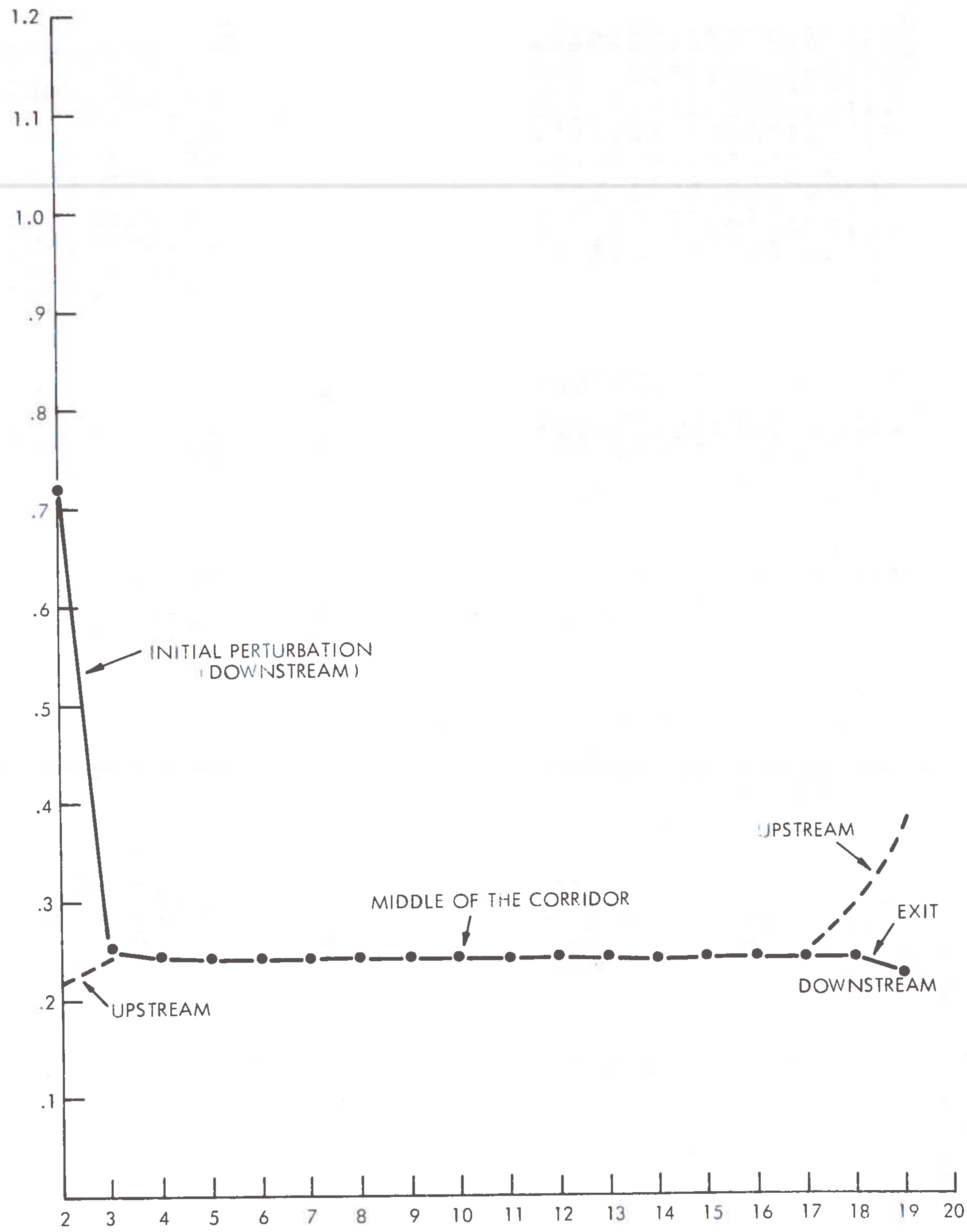

Figure 7.5.2 $x_3(k)$ VERSUS k (DOWNSTREAM AND UPSTREAM PERTURBATIONS). INITIAL PE
TURBATION $\underline{x}(1)$ = (-1,2,1)

We have included both a downstream and an upstream perturbation which contribute the same total flow. The transient evolution (stage 1) and the end effect (exit) are different, but both lead to the same stationary distribution.

The component of the upstream perturbation is plotted with a dotted line and the downstream perturbation with a full line.

Example 7.6 See Figs. 2.3.4-10 and 2.3.4-11 and Table 7.6 (downstream perturbations).

The graph has only one final class, but it is periodic with period 2. This is the uncertain case (section 2.3.4). In this specific example, the subsystems described in 2.3.4 are not controllable. Indeed, to find a vector $\underline{d}$ as explained in appendix B, we can choose either $\underline{\zeta}^T = (1,1,0)$ or $\underline{\zeta}^T = (0,1,1)$ but not $\underline{\zeta}^T = (1,0,1)$ because there is no cycle direct from nodes 1 to 3. If we take $\underline{\zeta}^T = (1,1,0)$ for instance, we are able to spread all the flow evenly between 1 and 2 for controllability to hold. However, if all the flow perturbation enters through entrance 3, then after one step, it is all concentrated in 2; after two steps, some part of it is in 1 and some in 3, and after three steps, everything is back in 2. In general, after an odd number of steps, all the flow will be concentrated in 2. This is because every path of odd length with origin 3 in G has 2 as its extremity. The periodicity subclasses (appendix A) of G are:

$$G = \{1,3\} \text{ and } G_1 = \{2\}.$$

For the same reason, choosing $\underline{\zeta}^T = (0,1,1)$ does not permit a state reduction leading to a controllable system. One can, for instance, verify that an initial perturbation $\delta\underline{x}^T(1) = (\alpha,0,\alpha)$ will not converge to a constant, as confirmed by a numerical experiment (7.6). The incidence matrices are:

$$\underline{Y} = \begin{bmatrix} 1&0&0&0 \\ 0&1&1&0 \\ 0&0&0&1 \end{bmatrix}, \text{ and } \underline{Z} = \begin{bmatrix} 0&1&0&0 \\ 1&0&0&1 \\ 0&0&1&0 \end{bmatrix},$$

we have chosen

$$\underline{L} = \begin{bmatrix} 2&0&0&0 \\ 0&3&0&0 \\ 0&0&4&0 \\ 0&0&0&3 \end{bmatrix}.$$

Example 7.7. See Figs. 2.3.4-6 and 2.3.4-7 and Table 7.7. This example has been discussed in sections 2.3.2 and 2.3.4.

In this case, G has two final classes: {1,2} and {3}. Therefore, no subsystem is controllable and that component of the perturbation which enters

Table 7.6 RESULTS OF EXAMPLE 7.6

DC(K)=C(K-1)-C(K)

K	DC11	DC12	DC13
19	0.2000000000000000D 01	0.0000000000000000D 00	0.0000000000000000D 00
18	0.1714285714285714D 01	0.0000000000000000D 00	0.1714285714285714D 01
17	0.1202380952380952D 01	0.0000000000000000D 00	0.1202380952380952D 01
16	0.1714285714285714D 01	0.0000000000000000D 00	0.1714285714285714D 01
15	0.1202380952380954D 01	0.0000000000000000D 00	0.1202380952380954D 01
14	0.1714285714285715D 01	0.0000000000000000D 00	0.1714285714285714D 01
13	0.1202380952380953D 01	0.0000000000000000D 00	0.1202380952380952D 01
12	0.1714285714285714D 01	0.0000000000000000D 00	0.1714285714285714D 01
11	0.1202380952380954D 01	0.0000000000000000D 00	0.1202380952380953D 01
10	0.1714285714285711D 01	0.0000000000000000D 00	0.1714285714285710D 01
9	0.1202380952380943D 01	0.0000000000000000D 00	0.1202380952380951D 01
8	0.1714285714285709D 01	0.0000000000000000D 00	0.1714285714285704D 01
7	0.1202380952380953D 01	0.0000000000000000D 00	0.1202380952380953D 01
6	0.1714285714285701D 01	0.0000000000000000D 00	0.1714285714285705D 01
5	0.1202380952380953D 01	0.0000000000000000D 00	0.1202380952380953D 01
4	0.1714285714285694D 01	0.0000000000000000D 00	0.1714285714285694D 01
3	0.1202380952380956D 01	0.0000000000000000D 00	0.1202380952380956D 01
2	0.1714285714285712D 01	0.0000000000000000D 00	0.1714285714285708D 01
1	0.1202380952380942D 01	0.0000000000000000D 00	0.1202380952380942D 01

K	DC22	DC23	DC33
19	0.1714285714285714D 01	0.0000000000000000D 00	0.3000000000000000D 01
18	0.1202380952380953D 01	0.0000000000000000D 00	0.1714285714285714D 01
17	0.1714285714285715D 01	0.0000000000000000D 00	0.1202380952380953D 01
16	0.1202380952380953D 01	0.0000000000000000D 00	0.1714285714285713D 01
15	0.1714285714285715D 01	0.0000000000000000D 00	0.1202380952380954D 01
14	0.1202380952380953D 01	0.0000000000000000D 00	0.1714285714285714D 01
13	0.1714285714285714D 01	0.0000000000000000D 00	0.1202380952380952D 01
12	0.1202380952380954D 01	0.0000000000000000D 00	0.1714285714285714D 01
11	0.1714285714285714D 01	0.0000000000000000D 00	0.1202380952380953D 01
10	0.1202380952380951D 01	0.0000000000000000D 00	0.1714285714285708D 01
9	0.1714285714285705D 01	0.0000000000000000D 00	0.1202380952380953D 01
8	0.1202380952380953D 01	0.0000000000000000D 00	0.1714285714285705D 01
7	0.1714285714285705D 01	0.0000000000000000D 00	0.1202380952380953D 01
6	0.1202380952380949D 01	0.0000000000000000D 00	0.1714285714285708D 01
5	0.1714285714285705D 01	0.0000000000000000D 00	0.1202380952380949D 01
4	0.1202380952380949D 01	0.0000000000000000D 00	0.1714285714285698D 01
3	0.1714285714285705D 01	0.0000000000000000D 00	0.1202380952380956D 01
2	0.1202380952380953D 01	0.0000000000000000D 00	0.1714285714285705D 01
1	0.1714285714285708D 01	0.0000000000000000D 00	0.1202380952380945D 01

D(K)=MATRIX DERIVATIVE OF X(K+1) W. RESPECT TO X(K)

K	D11	D12	D13
9	0.0000000000000000D 00	0.5833333333333335D 00	0.0000000000000000D 00
8	0.0000000000000000D 00	0.5833333333333330D 00	0.0000000000000000D 00
7	0.0000000000000000D 00	0.5833333333333333D 00	0.0000000000000000D 00
6	0.0000000000000000D 00	0.5833333333333333D 00	0.0000000000000000D 00
5	0.0000000000000000D 00	0.5833333333333333D 00	0.0000000000000000D 00
4	0.0000000000000000D 00	0.5833333333333335D 00	0.0000000000000000D 00
3	0.0000000000000000D 00	0.5833333333333335D 00	0.0000000000000000D 00
2	0.0000000000000000D 00	0.5833333333333333D 00	0.0000000000000000D 00

K	D21	D22	D23
9	0.9999999999999995D 00	0.0000000000000000D 00	0.9999999999999997D 00
8	0.9999999999999995D 00	0.0000000000000000D 00	0.9999999999999995D 00
7	0.9999999999999990D 00	0.0000000000000000D 00	0.9999999999999995D 00
6	0.9999999999999995D 00	0.0000000000000000D 00	0.9999999999999997D 00
5	0.9999999999999992D 00	0.0000000000000000D 00	0.9999999999999992D 00
4	0.9999999999999992D 00	0.0000000000000000D 00	0.9999999999999995D 00
3	0.9999999999999997D 00	0.0000000000000000D 00	0.9999999999999995D 00
2	0.9999999999999992D 00	0.0000000000000000D 00	0.9999999999999995D 00

K	D31	D32	D33
9	0.0000000000000000D 00	0.4166666666666665D 00	0.0000000000000000D 00
8	0.0000000000000000D 00	0.4166666666666667D 00	0.0000000000000000D 00
7	0.0000000000000000D 00	0.4166666666666667D 00	0.0000000000000000D 00
6	0.0000000000000000D 00	0.4166666666666665D 00	0.0000000000000000D 00
5	0.0000000000000000D 00	0.4166666666666665D 00	0.0000000000000000D 00
4	0.0000000000000000D 00	0.4166666666666663D 00	0.0000000000000000D 00
3	0.0000000000000000D 00	0.4166666666666663D 00	0.0000000000000000D 00
2	0.0000000000000000D 00	0.4166666666666667D 00	0.0000000000000000D 00

DELTA(K)=PERTURBATION AT STAGE K

K	DELTA1	DELTA2	DELTA3
2	0.0000000000000000D 00	0.1999999999999999D 01	0.0000000000000000D 00
3	0.1166666666666666D 01	0.0000000000000000D 00	0.8333333333333320D 00
4	0.0000000000000000D 00	0.1999999999999997D 01	0.0000000000000000D 00
5	0.1166666666666665D 01	0.0000000000000000D 00	0.8333333333333317D 00
6	0.0000000000000000D 00	0.1999999999999996D 01	0.0000000000000000D 00
7	0.1166666666666664D 01	0.0000000000000000D 00	0.8333333333333315D 00
8	0.0000000000000000D 00	0.1999999999999994D 01	0.0000000000000000D 00
9	0.1166666666666663D 01	0.0000000000000000D 00	0.83333333333331[illegible]6D 00
10	0.0000000000000000D 00	0.1999999999999993D 01	0.0000000000000000D 00
11	0.1166666666666662D 01	0.0000000000000000D 00	0.8333333333333295D 00
12	0.0000000000000000D 00	0.1999999999999991D 01	0.0000000000000000D 00

Table 7.7 RESULTS OF EXAMPLE 7.7

DC(K)=C(K-1)-C(K)

K	DC11	DC12	DC13
9	0.3333333333333326D 00	0.3333333333333315D 00	0.0000000000000000D 00
8	0.3333333333333310D 00	0.3333333333333313D 00	0.0000000000000000D 00
7	0.3333333333333302D 00	0.3333333333333304D 00	0.0000000000000000D 00
6	0.3333333333333324D 00	0.3333333333333319D 00	0.0000000000000000D 00
5	0.3333333333333324D 00	0.3333333333333322D 00	0.0000000000000000D 00
4	0.3333333333333322D 00	0.3333333333333322D 00	0.0000000000000000D 00
3	0.3333333333333315D 00	0.3333333333333317D 00	0.0000000000000000D 00
2	0.3333333333333266D 00	0.3333333333333279D 00	0.0000000000000000D 00
1	0.3333333333333297D 00	0.3333333333333297D 00	0.0000000000000000D 00

K	DC22	DC23	DC33
9	0.3333333333333319D 00	0.0000000000000000D 00	0.3000000000000000D 01
8	0.3333333333333315D 00	0.0000000000000000D 00	0.3000000000000000D 01
7	0.3333333333333304D 00	0.0000000000000000D 00	0.3000000000000000D 01
6	0.3333333333333317D 00	0.0000000000000000D 00	0.3000000000000000D 01
5	0.3333333333333322D 00	0.0000000000000000D 00	0.3000000000000000D 01
4	0.3333333333333322D 00	0.0000000000000000D 00	0.3000000000000000D 01
3	0.3333333333333317D 00	0.0000000000000000D 00	0.3000000000000000D 01
2	0.3333333333333277D 00	0.0000000000000000D 00	0.3000000000000000D 01
1	0.3333333333333295D 00	0.0000000000000000D 00	0.3000000000000000D 01

D(K)=MATRIX DERIVATIVE OF X(K+1) W. RESPECT TO X(K)

K	D11	D12	D13
9	0.5857864376269045D 00	0.4142135623730949D 00	0.0000000000000000D 00
8	0.5857864376269044D 00	0.4142135623730947D 00	0.0000000000000000D 00
7	0.5857864376269049D 00	0.4142135623730949D 00	0.0000000000000000D 00
6	0.5857864376269046D 00	0.4142135623730951D 00	0.0000000000000000D 00
5	0.5857864376269047D 00	0.4142135623730949D 00	0.0000000000000000D 00
4	0.5857864376269047D 00	0.4142135623730947D 00	0.0000000000000000D 00
3	0.5857864376269041D 00	0.4142135623730947D 00	0.0000000000000000D 00
2	0.5857864376269050D 00	0.4142135623730949D 00	0.0000000000000000D 00

K	D21	D22	D23
9	0.4142135623730951D 00	0.5857864376269047D 00	0.0000000000000000D 00
8	0.4142135623730944D 00	0.5857864376269044D 00	0.0000000000000000D 00
7	0.4142135623730949D 00	0.5857864376269046D 00	0.0000000000000000D 00
6	0.4142135623730951D 00	0.5857864376269051D 00	0.0000000000000000D 00
5	0.4142135623730947D 00	0.5857864376269046D 00	0.0000000000000000D 00
4	0.4142135623730954D 00	0.5857864376269049D 00	0.0000000000000000D 00
3	0.4142135623730945D 00	0.5857864376269049D 00	0.0000000000000000D 00
2	0.4142135623730947D 00	0.5857864376269044D 00	0.0000000000000000D 00

K	D31	D32	D33
9	0.0000000000000000D 00	0.0000000000000000D 00	0.9999999999999998D 00
8	0.0000000000000000D 00	0.0000000000000000D 00	0.1000000000000000D 01
7	0.0000000000000000D 00	0.0000000000000000D 00	0.1000000000000000D 01
6	0.0000000000000000D 00	0.0000000000000000D 00	0.9999999999999998D 00
5	0.0000000000000000D 00	0.0000000000000000D 00	0.9999999999999998D 00
4	0.0000000000000000D 00	0.0000000000000000D 00	0.9999999999999998D 00
3	0.0000000000000000D 00	0.0000000000000000D 00	0.9999999999999998D 00
2	0.0000000000000000D 00	0.0000000000000000D 00	0.9999999999999998D 00

DELTA(K)=PERTURBATION AT STAGE K

K	DELTA1	DELTA2	DELTA3
2	0.5857864376269050D 00	0.4142135623730947D 00	-0.9999999999999998D 00
3	0.5147186257614289D 00	0.4852813742385697D 00	-0.9999999999999998D 00
4	0.5025253169416723D 00	0.4974746830583262D 00	-0.9999999999999998D 00
5	0.5004332758886090D 00	0.4995667241113889D 00	-0.9999999999999997D 00
6	0.5000743383899857D 00	0.4999256616100121D 00	-0.9999999999999997D 00
7	0.5000127544513106D 00	0.4999872455486869D 00	-0.9999999999999997D 00
8	0.5000021883178821D 00	0.4999978116821143D 00	-0.9999999999999997D 00
9	0.5000003754559891D 00	0.4999996245440068D 00	-0.9999999999999997D 00
10	0.5000000644180616D 00	0.4999999355819337D 00	-0.9999999999999997D 00
11	0.5000000110523895D 00	0.4999999889476049D 00	-0.9999999999999997D 00
12	0.5000000018962875D 00	0.4999999981037062D 00	-0.9999999999999997D 00
13	0.5000000003253486D 00	0.4999999996746445D 00	-0.9999999999999995D 00
14	0.5000000000558177D 00	0.4999999999441747D 00	-0.9999999999999995D 00
15	0.5000000000095733D 00	0.4999999999904147D 00	-0.9999999999999995D 00
16	0.5000000000016390D 00	0.4999999999983527D 00	-0.9999999999999995D 00
17	0.5000000000002776D 00	0.4999999999997139D 00	-0.9999999999999995D 00
18	0.5000000000000439D 00	0.4999999999999472D 00	-0.9999999999999995D 00
19	0.5000000000000041D 00	0.4999999999999870D 00	-0.9999999999999995D 00
20	0.4999999999999984D 00	0.4999999999999927D 00	-0.9999999999999994D 00

through entrance 3 is propagated all the way through.

The part of the perturbation which enters through nodes 1 and 2 constantly remains within those nodes and is quickly redistributed according to the vector $\underline{p}^T = (\frac{1}{2}, \frac{1}{2})$ since the cost matrix $\underline{\mathscr{L}}_{12}$ corresponding to the partial subnetwork with nodes 1 and 2 is symmetric: $\underline{\mathscr{L}}_{12} = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$.

8. CONCLUSIONS AND SUGGESTIONS FOR FUTURE RESEARCH

8.1 SUMMARY OF THIS WORK

In the introduction, we raise the question of how a traffic perturbation occurring at some point of a freeway-corridor network influences the optimal assignments away from that point. We are particularly, but not exclusively, concerned with perturbations which leave the total flow unchanged. Not only the sensitivity of the network to an incident, but also the meaningfulness of a static optimization in a long corridor, depend on the answer to that question.

For networks consisting of identical subnetworks (stationary networks), we can prove, under a major condition and two technical ones, that zero-flow perturbations decay exponentially, and calculate the bounds analytically. The main question of the introduction is, therefore, answered, and a static optimization is meaningful even for very long freeway corridors. Also, if a static assignment is used online for traffic control, a reoptimization has to be carried out only within some neighborhood of an accident. This neighborhood can be determined in advance, and it depends on the level of accuracy required as well as the structure of the network.

In addition, we have also established what changes are caused in a very long corridor by perturbations which contribute a non-zero total change in flow. They give rise to a sequence of downstream and upstream perturbations which converge exponentially to a distribution which is a characteristic of the network, and which we can determine analytically. The major condition which guarantees the results is that there exist enough cross links between main and side roads so as to allow traffic perturbations to be spread out all over the network. This condition can be expressed mathematically either as a controllability property of a certain system or in the language of graph and Markov chain theory. Consequently for a class of linear discrete dynamical systems, we have characterized controllability by means of graph concepts.

Our general method of analysis is twofold:

a. Quadratic expansion of the cost function about a nominal optimal solution.

b. Splitting of the network into subnetworks and application of discrete dynamic programming over distance.

This method is applicable to any freeway-corridor network and any cost function locally convex about the nominal solution.

From this, we establish the above results in the stationary case. For non-stationary networks, the only available results are obtained from the numerical implementation of this method. In all cases checked, the magnitudes of the perturbations have been found to decrease, but we have not formulated conditions which ensure that behavior in the non-stationary case.

We present below some possible extensions of this work both to traffic-engineering and optimal control theory.

8.2 FURTHER RESEARCH IN TRAFFIC-ENGINEERING

a. Numerical implementation of the general method to real systems: One may also try to extend to more general types of networks the approach which we have followed for freeway corridors.

b. In the case of a stationary network, we have proved the qualitative property that, if some links are removed, the new network may no longer be described by controllable systems, and therefore, perturbations in it may no longer decrease. It seems reasonable that a related quantitative property will hold. If some links are removed without invalidating controllability, the perturbations will still decrease in the new network, but not as fast as in the previous one. Also, it may be possible to identify those links whose presence is essential in keeping the speed of decay high or, alternatively, the eigenvalues of the propagation matrix small. Such a knowledge will prove useful in the design of new road links.

c. An objection to the usefulness of our results is that stationary networks do not exist in real freeway-corridor systems. However, studying such systems provides insight into the more general situation. For example, it is qualitatively reasonable that the perturbations die away in a non-stationay network, where enough cross links exist, because they do so in a stationary one.

However, it may be extremely useful to derive quantitative properties of a given non-stationary network from those of its "pure components;" that is, the stationary networks corresponding to its various subnetworks. A timid attempt in that direction is shown in section 2.3.9, where we have studied quasi-stationary networks. Perhaps, a more general approach uses a relationship between the one-step and the limiting propagation matrices, such as conjectured in section 6.

d. One may consider including origin-destination constraints; e.g., the possibility for cars to arrive or depart at intermediate points in the network. This may further cause perturbations to decay.

e. Also, the same problem can be studied from a "user optimization" [6] rather than system-optimization viewpoint.

8.3 FURTHER RESEARCH IN OPTIMAL CONTROL THEORY

a. It must be emphasized that many problems other than traffic regulation can be described as optimal assignment of flows within a network.

The network may convey a flow of information, of money, of products, etc. The nodes may represent warehouses, power stations, locations, and service points. The links may stand for communication channels or transmission lines: they may also represent various relationships like demand and supply. The literature is abundant on those matters, but in the majority of cases, linear cost functions have been considered. The general results of linear-programming have then been successfully particularized [41] to the case of networks. This leads to important simplifications. However, linear cost functions are not necessarily satisfactory for they always lead to optimal solutions with extreme components. In our traffic context for instance, this means that some roadways will be either saturated or unused.

It seems, on the other hand, that a quadratic expansion of the cost function about an optimal solution is a natural way to study perturbations.

Analogously to what has been done already for more than a decade in the case of linear cost functions, the need arises for a <u>particularization of many results of optimal control theory to the case of flows in networks.</u>

For instance, in our case, the most natural way to define a state vector is by describing the evolution of the flow distribution among the various nodes. However, we have seen that this formulation does not lead far in the analytical explanation of the observed behavior because the existing body of optimal control theory requires a certain standard form. Thus, we have to artificially define other state variables. In section 4, we have tried a more direct way which deals with the natural formulation. We indicate below possible applications of optimal control theory to networks with illustrations in the context of traffic flow.

b. For a class of linear discrete dynamical systems which describe flow conservation in networks, we have constructed a controllability condition which uses graph concepts instead of the standard rank test. This may be extended by associating networks with some class of linear systems much in the same way as one can describe certain matrices by graphs.

c. Discrete duality theory and the minimum principle will probably be helpful as specialized to the network context. For instance in our case, the downstream-and-upstream perturbation problems are related to one another by a "time-reversal" (where the "time" parameter is the label of the subnetwork).

This property is translated into a correspondence between the respective assessibility graphs which are dual of one another (section 3).

Perhaps in the case of network systems, duality can be described in a fruitful way by means of graphs.

d. We have stressed several times that Markov chain concepts emerge quite naturally in the traffic-network context in spite of the deterministic character of the problem. This may seem paradoxical. However, it becomes intuitive if one considers the microscopic level. In our traffic example, if we observe a particular car chosen at random, instead of dealing with the aggregate traffic flow, then the route followed by the car in the network can be viewed as a realization of a discrete stochastic process: at each node, the car has the choice between several links, and the successive choices can be viewed as a random process. The question is if there is a rigorous way to define the microscopic level, and to establish correspondences between the probability distributions of the stochastic process alluded to and the optimal control parameters.

We have made an attempt in that direction in section 6, and although we do not pretend at rigor, we have derived a result confirmed by numerical experiments. In any case, this curious type of complementarity between a micro and a macro level is worth investigating.

REFERENCES

[1] Paul K. Houpt and Michael Athans, "Dynamic Stochastic Control of Freeway Corridor Systems; Volume I: Summary," M.I.T. Electronic Systems Laboratory, Report ESL-R-608, 1975.

[2] Stanley B. Gershwin, "Dynamic Stochastic Control of Freeway Corridor Systems; Volume II: Steady State Optimal Traffic Assignment Using the Accelerated Gradient Projection Method," M.I.T. Electronic Systems Laboratory, Report ESL-R-609, 1975.

[3] Douglas Looze, Paul K. Houpt, and Michael Athans, "Dynamic Stochastic Control of Freeway Corridor Systems; Volume III: Dynamic Centralized and Decentralized Control Strategies," M.I.T. Electronic Systems Laboratory, Report ESL-R-610, 1975.

[4] Daniel Orlhac, Michael Athans, Jason L. Speyer and Paul K. Houpt, "Dynamic Stochastic Control of Freeway Corridor Systems; Volume IV: Estimation of Traffic Variables Via Extended Kalman Filter Methods," M.I.T. Electronic Systems Laboratory, Report ESL-R-611, 1975.

[5] Marcel Singleton and John E. Ward, "Dynamic Stochastic Control of Freeway Corridor Systems, Volume V: A Comparative Study of Various Types of Vehicle Detectors," M.I.T. Electronic Systems Laboratory, Report ESL-R-612, 1975.

[6] John G. Wardrop, "Some Theoretical Aspects of Road Traffic Research," Proc. Inst. Civil Engineers, 1952, Part 2 , p. 325-378.

[7] Peter A. Steenbrink, <u>Optimization of Transport Network</u>, London:Wiley, 1974.

[8] Harold J. Payne and Warren A. Thompson, "Traffic Assignment on Transportation Network with Capacity Constraints and Queuing," OSRA/TIMS Meeting, 1975.

[9] S. Dafermos and F.T. Sparrow, "The Traffic Assignment Problem for a General Network," Journal of Research of the U.S. National Bureau of Standards, Vol. 73B, No. 2, p. 91-118, April-June 1969.

[10] Sam Yagar, "Dynamic Traffic Assignment by Individual Path Minimization and Queuing," Transportation Research, Vol. 5, p. 179-196, 1971.

[11] Sam Yagar, "A Study of the Travel Patterns in a Corridor with Respect to the Assignment Principles of Wardrop," in <u>Traffic Flow and Transportation</u>, p. 165-181, G.F. Newell, Ed., New York: American Elsevier, 1972.

[12] Paul K. Houpt, Decentralized Stochastic Control of Finite-State Systems with Application to Vehicular Traffic Control, Ph.D. Thesis, M.I.T. Electronic Systems Laboratory, Report ESL-R-579.

[13] Harold J. Payne and Warren A. Thompson, "Allocation of Freeway Ramp Metering Volumes to Optimize Corridor Performance," IEEE Trans. Auto. Control, Vol. AC-19, p. 177-185, June 1974.

[14] Nathan Gartner, John D.C. Little and Henry Gabbay, "Optimization of Traffic Signal Settings in Networks by Mixed-Integer Linear Programming", Technical Report #91, Operations Research Center, M.I.T., March 1974.

[15] M.J. Lighthill and G.B. Whitham, "On Kinematic Waves II, A Theory of Traffic Flow on Long Crowded Roads," Proc. Roy. Soc. London, Series A, Vol. 229, p. 317-345, 1955.

[16] John G. Wardrop, "Experimental Speed/Flow Relations in a Single Lane," in Proceedings of the Second International Symposium on the Theory of Traffic Flow and Transportation, p. 104-119, 1963.

[17] Harold Greenberg, "An Analysis of Traffic Flow," Operations Research Vol. 7, p. 79-85, 1959.

[18] F.V. Webster, "Traffic Signal Settings," Road Research Laboratory Tech. Report No. 39, London, H.M. Stationary Office, 1958.

[19] Leonard Shaw, "On Optimal Ramp Control of Traffic Jam Queues," IEEE Trans. Auto. Control, Vol. AC-17, No. 5, p. 630-637, October 1972.

[20] Leonard Shaw and William R. McShane, "Optimal Ramp Control for Incident Response," Transpn. Res. Vol. 7, p. 393-411, 1973.

[21] P. Dorato and A.H. Levis, "Optimal Linear Regulators" The Discrete-Time Case," IEEE Trans. Auto. Control, Vol. AC-17, p. 613-240, December 1971.

[22] D.R. Vaughan, "A Nonrecursive Algebraic Solution of the Discrete Riccati Equation," IEEE Trans. Auto. Control, Vol. AC-15, p. 597-599, October 1970.

[23] D.L. Kleinman and M. Athans, "The Discrete Minimum Principle with Application to the Linear Regulator Problem," Electronic Systems Laboratory, M.I.T., Report ESL-R-260, 1966.

[24] J.M. Holtzman and H. Halkin, "Directional Convexity and the Maximum Principle for Discrete Systems," p. 263-275, SIAM Journal on Control, Vol. 4, 1966.

[25] R.W. Brockett, Finite-dimensional Linear Systems, New York: Wiley, 1970.

[26] Arthur E. Bryson, Jr., and Yu-Chi Ho, Applied Optimal Control, Waltham, Mass.: Ginn and Company, 1969.

[27] M. Athans and P.L. Falb, Optimal Control, New York: McGraw-Hill, 1966.

[28] "Special Issue on Linear Quadratic-Gaussian Problem," IEEE Trans. Auto. Control, Vol. AC-16, December 1971.

[29] R. Bellman, "Notes on the Theory of Dynamic Programming-Transportation Models," Man. Sci., Jan. 1958, Vol. 4, No. 2, p. 191-195.

[30] R. Bellman, Introduction to Matrix Analysis, New York: McGraw-Hill, 1960.

[31] R. Fortet, "Algèbre des Tenseurs et des Matrices, Problèmes de Valeurs Propres" (C.D.U., Paris, 1961).

[32] P. Gordon, "Théorie des Chaînes de Markov Finiès et ses Applications," Dunod, Paris, 1965.

[33] W. Feller, An Introduction to Probability Theory and Its Applications, New York: Wiley, 1950.

[34] K.L. Chung, Markov Chains with Stationary Transition Probabilities, Berlin, Springer Verlag, 1961.

[35] R.B. Busacker and T.L. Saaty, Finite Graphs and Networks: An Introduction with Applications, New York: McGraw-Hill, 1965.

[36] F. Harary, Graph Theory, Reading, Mass.: Addison-Wesley, 1971.

[37] C. Berge, Graphs and Hypergraphs; translated by Edward Minieka, New York, American Elseview, 1973.

[38] J.N. Franklin, Matrix Theory, Englewood Cliffs, N.J.:Prentice-Hall, 1968.

[39] R.E. Powell and S.M. Shah, Summability Theory and Applications, London: Van Nostrand, 1972.

[40] R.H. Brinkmann and E.A. Klotz, Linear Algebra and Analytic Geometry, Addison Wesley, Reading, Mass., 1971.

[41] L.R. Ford and D.R. Fulkerson, Flows in Networks, Princeton, N.J., Princeton University Press, 1962.

[42] S.B. Gershwin, Pierre Dersin, and Michael Athans, "Sensitivity Analysis of Optimal Static Traffic Assignment in a Freeway Corridor System," Proceedings of the 1977 Joint Automatic Control Conference, June 22-24, 1977, page 457-465; also to appear in IEEE Transactions on Automatic Control.

APPENDIX A

GRAPH-THEORETICAL CONCEPTS AND BASIC RESULTS

A.1 INTRODUCTION

We collect here the main definitions about graphs, as well as the basic results which we need in the main course of this work at least as far as the terminology is not standard.

The graphs which we consider are those which are usually associated with Markov chains. We make use of many of the special properties of these graphs. We wish to consider those properties from the viewpoint of graph theory without embodying them too early into the algebraic apparatus of Markov chains. Graph theorists who study this class of graphs [35] and Markov chains specialists [32], [33] generally are primarily concerned with algebraic properties of Markov chains, and consider graphs to be useful tools for this. Here we must follow the reverse order: the Markov chain will be a useful tool to study the graph. The only approach which we have found which does this is by Gordon [32] in French. We translate his graph terminology here, and we use Feller [33] for the standard Markov chain notions, and Busacker and Saaty [35] for standard graph concepts.

A.2 TWO SUCCESSIVE PARTITIONS OF A DIRECTED GRAPH

Definition A.1

A directed graph is a pair $G = (\Gamma, S)$, where S is a set, and Γ a multi-valued mapping from this set into itself (i.e., for $i \in S$, $\Gamma(i)$ is a subset of S). The elements of S are called nodes (or vertices).

If $j \in \Gamma(i)$ for i in S, an arrow is drawn from i to j in the pictorial representation of G. The nodes j such that $j \in \Gamma(i)$ are called the followers of j. We consider finite graphs; i.e., graphs with a finite number of nodes. The oriented pairs (i,j) with $j \in \Gamma(i)$ are called arcs. In particular, it is possible that $s \in \Gamma(s)$. Arc (s,s) is then called a loop.

Note: A freeway network can be described by a succession of bipartite graphs; i.e., graphs where the set of nodes is partitioned into 2 disjoint sets: $S = S_1 \cup S_2$, and only arcs (i,j) with i in S_1 and j in S_2 exist. In this context, S_1 will be the set of entrances and S_2 the set of exits. We call the arcs of a freeway network "links" and reserve the word "arc" for other types of graphs.

Definition A2:

A path is an ordered set of arcs where every two consecutive arcs have a common node: (s_1, s_2), (s_2, s_3), ..., (s_{n-1}, s_n), denoted $(s_1, s_2, \ldots, s_n)$.

A path can go several times through the same node; for example, (1,2,3,2,3,2,3,2,1,2). Node s is called the origin of the path and s_n its extremity. If a path exists whose origin is i and whose extremity is j,node j is said to be accessible from node i.

Definition A3:

A cycle is a path whose origin and extremity are the same.

A.2.1 First Partition

The notion of cycle enables one to define an equivalence relation on a directed graph: two nodes i and j are said to be $\mathcal{R}_1$-equivalent, or $i\mathcal{R}_1 j$, if, and only if, there is a cycle containing both nodes, or alternatively, if there is a path from node i to node j and a path from node j to node i. One can verify that this relation has all the properties of an equivalence relation: it is reflexive, transitive, and symmetric.

Therefore, any directed graph can be partitioned into subgraphs defined by the $\mathcal{R}_1$-equivalence classes of nodes. Each $\mathcal{R}_1$-equivalence class $\mathcal{N}$ defines a subgraph H of G. The nodes of H are the elements of $\mathcal{N}$, and the arcs of H are those arcs of G which connect two elements of $\mathcal{N}$.

Each of these subgraphs is called a strongly connected component (abbreviated s.c.c.).

Definition A4:

The graph G is said to be strongly connected if it contains only one strongly connected component. This amounts to the property that each node is accessible from any other one; i.e., given any two nodes i and j, there is a path with origin i and extremity j.

From these definitions follows:

Proposition A.1: The strongly connected components of a graph are strongly connected graphs.

Final Classes: We derive a new graph g from G. Each strongly connected component of G is represented by a node in g. Arc (i,j) exists in g if, and only if, there is an arc in G from some node of the i^{th} s.c.c. to some node of the j^{th} s.c.c. We call g the reduced graph of G.

Definition A5:

A strongly connected component of G is called a final class if it corresponds to a node without follower in g. An example of graph G with its reduced graph g is shown in Fig. A.1. The s.c.c.'s are indicated by Roman numerals and identified with dotted lines. The final classes are III and IV in this example.

Definition A6:

The nodes in a final class are called final nodes. All other nodes are called transient.

Proposition A2:

a. Two nodes in the same final class are accessible from each other.

b. If two nodes are in two different final classes, neither is accessible from the other.

c. No transient node is accessible from a final node.

d. Any transient node is the origin of a path whose extremity is in some final class.

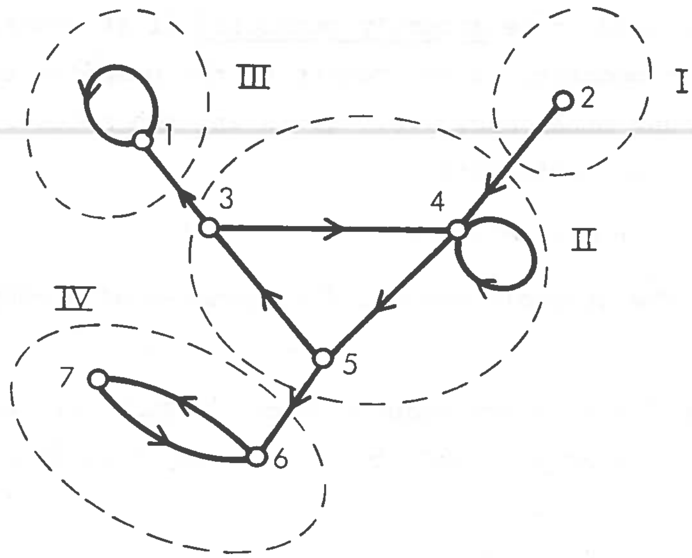

Figure A.1-a EXAMPLE OF DIRECTED GRAPH G

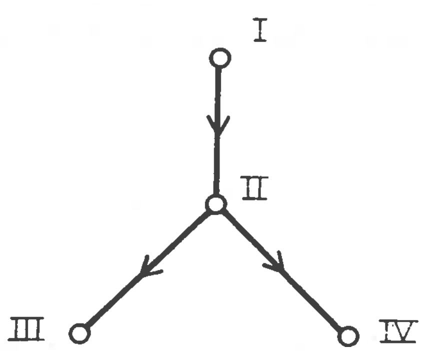

Figure A.1-b REDUCED GRAPH g CORRESPONDING TO GRAPH G

Proof

Parts a and b are true because final classes are s.c.c.'s. Part c comes from the definition of the reduced graph. If a transient node were accessible from a node in a final class, this class would correspond to a node in g with followers, contradicting definition A5.

Part d follows from the fact that, by definition (A5) of a final class, every node i is the origin of a path with extremity in some final class.

A.2.2 SECOND PARTITION

We now shall partition each s.c.c. by means of a new equivalence relation R_2.

Definition A7:

The length of a path $(s_0, s_1, \ldots, s_n)$ is n, the number of arcs it consists of. If the path contains cycles, one has to specify how many times it goes around those cycles to define its length.

For instance in Fig. A2, the path (1,2,3,2,4) has length 4 because it consists of the 4 arcs (1,2), (2,3), (3,2), (2,4). However, the path (1,2,3,2,3,2,4) has length 6 since it goes through arcs (1,2), (2,3), (3,2), (2,3), (3,2), (2,4) successively.

Proposition A3: The greatest common divisor of the lengths of all cycles through a given node, in a strongly connected graph, does not depend on the particular node considered.

Proof

See [32], [35].

This number is, therefore, a characteristic of the graph.

Definition A8:

The period of a strongly connected graph is the greatest common divisor of its cycles.

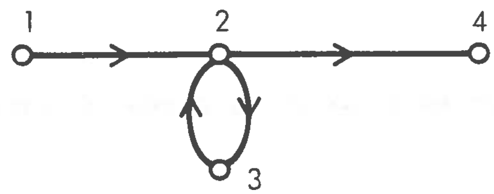

Figure A.2 EXAMPLE OF PATH CONTAINING A CYCLE

Definition A9

A strongly connected graph is called aperiodic if its period is equal to 1. In particular, proposition A3 as well as defintions A8 and A9 applies to the s.c.c.'s of an arbitrary directed graph since they are strongly connected graphs.

In the same spirit as proposition A3, it is possible to prove:

Proposition A4: Given any two nodes i and j in a strongly connected graph G, the lengths of all paths from i to j are congruent modulo the period of G [32], [35].

Thus if L is the length of a path from i to j,

$$L = qt + r,$$

where t is the period of G, q and r are integers, $r < t$, and r is the same for all paths from i to j.

Definition A10:

This remainder r is called the distance from i to j and denoted d(i,j). Proposition A4 makes it possible to define a new equivalence relation $\mathcal{R}_2$ on each s.c.c.

Definition A11:

Two nodes i and j are said to be $\mathcal{R}_2$-equivalent if $d(i,j) = 0$. (This amounts to the same thing as $d(j,i) = 0$ although $d(i,j) \neq d(j,i)$, in general, because the length of a cycle through i is a multiple of the period.)

The corresponding $\mathcal{R}_2$-equivalcne classes are easily obtained in the following manner. Let us choose an arbitrary node i_0, and define

$$G_k = \{j \in G : d(i_0,j) = k\}, \text{ for } k = 0, 1, ..., t-1,$$

where t is the period of G.

Indeed, if $j \in G_k$ and $i \in G_k$, then $d(i_0, i) = k$. Thus, the length of a path from i_0 to i is congruent to k modulo the period. Adding a path from i to j to the previous one yields a path from i_0 to j, also of length k modulo the period since $d(i_0, j) = k$. Therefore, the length of a path from i to j, is a multiple of the period; i.e., $d(i,j) = 0$.

Definition A12:

The $\mathcal{R}_2$-equivalence classes $\mathcal{G}_k$ are called periodicity subclasses. These classes are independent of the choice of i_0. If we have chosen a different node to begin with, we will have obtained the same classes but with different labels [33]. Figure A3 shows an example of such partitioning for a graph of period 5.

Proposition A5: The followers of a node in $\mathcal{G}_k$ are in $\mathcal{G}_{k+1}$ (for $k+1 < t$). The followers of a node in $\mathcal{G}_t$ are in $\mathcal{G}_0$.

This is straightforward from the definition of the $\mathcal{G}_\ell$'s. Thus, there is a cyclical movement through these classes. This can be verified in the example of Fig. A3.

To sum up, we have first partitioned the nodes of a directed graph into strongly connected components (A.2.1), among which we distinguish between final nodes and transient nodes. Subsequently, we have partitioned each s.c.c. into periodic subclasses (A.2.2).

A.3 CORRESPONDENCE WITH MARKOV CHAINS

Proposition A6: A directed graph is the graph of possible transitions associated with some Markov chain if, and only if, each node has at least one follower [32].

Proof

Starting from the Markov chain, one represents each state by a node and each possible transition by a directed arc. Accordingly, each node in G has at least one follower (which may be itself). Conversely to any directed graph which has that property, a Markov chain can be associated by defining states corresponding to the nodes and transition probabilities p_{ij}, such that $p_{ij} > 0$ if, and only if, arc (i,j) exists in the graph. The graph is then the graph of possible transitions of such a Markov chain.

Q.E.D.

The two partitions defined above can equally well be applied to the classification of states of the associated Markov chain. The irreducible subchains

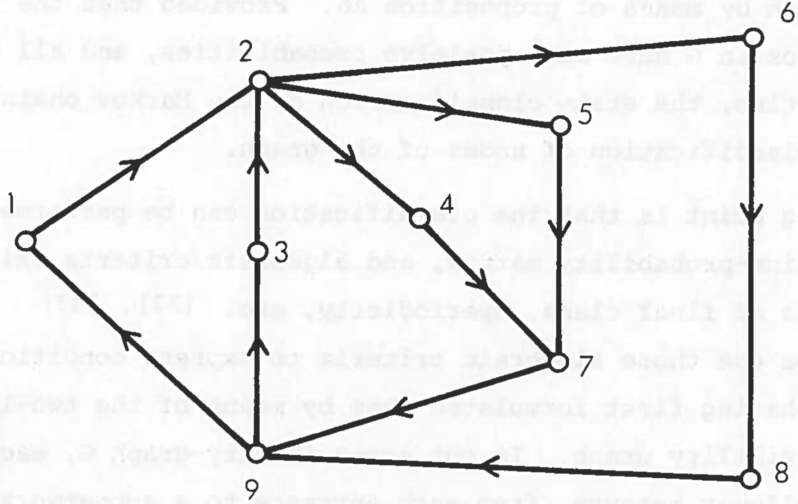

Figure A.3 PERIODICITY SUBCLASSES

If $i_0 = 1$, $G_0 = \{1,3\}$, $G_1 = \{2\}$,

$G_2 = \{4,5,6\}$, $G_3 = \{7,8\}$, $G_4 = \{9\}$

of persistent states [33] are precisely the final classes of the graph. The transient states are those represented by transient nodes in G. Likewise, the period concept is the same as in Markov chains. These facts show that the traditional classification of Markov chains [33] depends only on which probabilities are positive and which are zero. One can associate a Markov chain with a directed graph by means of proposition A6. Provided that the transitions corresponding to arcs in G have some positive probabilities, and all others have zero probabilities, the state classification of the Markov chain will be equivalent to the classification of nodes of the graph.

The interesting point is that the classification can be performed by looking only at the transition-probability matrix, and algebraic criteria exist to guarantee uniqueness of final class, aperiodicity, etc. [32], [33]. Thus in section 2.3.4, we use those algebraic criteria to express conditions for controllability after having first formulated them by means of the two-level partition of the accessibility graph. In our accessibility-graph G, each node has at least one follower because, from each entrance to a subnetwork, it is possible to reach at least one exit.

A.4 LENGTHS OF PATHS

We establish here a property used in the proof of theorem 2.5 (section 2.3.4).

Proposition A7: Given r nodes 1, 2, ...,r in a strongly connected aperiodic directed graph G, there exist r paths of the same length from nodes 1, 2, ...,r, respectively, to node r.

Proof

a. For any positive integers m and $k < m$ and any node i in G, there is a path from node i to node r of length congruent to k modulo m. This statement is a consequence of the property that there are paths from i to r with relatively prime lengths (aperiodicity).

b. Let us then choose a particular cycle through r, with length L, say, and a particular path from 1 to r with length denoted by $\ell(1)$.

There exists a path from node 2 to node r with length $\ell(2)$ congruent to

$\ell(1)$ modulo L.

$$\ell(2) = \ell(1) + k_2 L,$$

for some positive integer k_2. Likewise from any node i (i = 2,3, ...,r), there is a path from i to r with length $\ell(i)$, and

$$\ell(i) = \ell(1) + k_i L \quad , \tag{A.4.1}$$

for some positive integer k_i.

Let m be the least common multiple of k_2, ..., k_r.

$$m = m_i k_i \quad , \qquad i = 2, \ldots, r. \tag{A.4.2}$$

For each i = 2, ..., r, we extend the path from i to r of length $\ell(i)$ into a new path of length

$$\ell'(i) = \ell(i) + (m_i - 1) k_i L \quad , \tag{A.4.3}$$

by adding $(m_i-1)k_i$ times the cycles through r (Fig. A.4). For i=1, we add m times that cycle

From (A4.1) and (A4.3),

$$\ell'(i) = \ell(1) + k_i L + (m_i - 1) k_i L = \ell(1) + m_i k_i L \quad ,$$

or

$$\ell'(i) = \ell(1) + mL \; , \qquad (i = 1, \ldots r) \quad , \tag{A.4.4}$$

using (A4.2).

Equation (A.4.4) shows that all new paths from 1, 2, ...,r, respectively, to r have the same length.

Q.E.D.

A.5 ADJACENCY MATRIX AND NUMBER OF PATHS

We prove here a property used in section 2.3.4 when algebraic tests for controllability are given.

<u>Definition A.13</u>:

An <u>undirected bipartite graph</u> with origin node set S_1 and extremity node set S_2 is a graph whose nodes are those of S_1 and those of S_2, and where every arc connects a node in S_1 with a node in S_2.

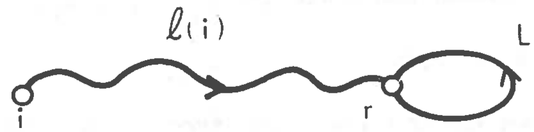

Figure A.4 CONSTRUCTION OF A NEW PATH

In particular, we have described subnetworks by undirected bipartite graphs having as their two node sets the set of entrances and the set of exits.

Definition A.14:

The adjacency matrix of a graph with m nodes is a mxm matrix $\underline{X}$, where

$$x_{ij} = \begin{cases} 1 \text{ if } (i,j) \text{ is an arc,} \\ 0 \text{ otherwise.} \end{cases}$$

The adjacency matrix of an undirected graph is symmetric.

Proposition A.8: In an undirected bipartite graph with n origins and n extremities, one can label the origins 1,2, ...,n and also the extremities 1,2,...,n.

Then, the adjacency matrix is a 2n by 2n matrix of the type:

$$\underline{X} = \begin{pmatrix} \underline{0} & \underline{U} \\ \underline{U}^T & \underline{0} \end{pmatrix} \qquad (A.5.1)$$

where $\underline{0}$ is the nxn zero matrix, and $\underline{U}$ is the nxn matrix defined by:

$$U_{ij} = \begin{cases} 1 \text{ if there is an arc from origin } i \text{ to extremity } j, \\ 0 \text{ otherwise.} \end{cases} \qquad (A.5.2)$$

Proof

In a bipartite graph, no arc connects the nodes belonging both to the set of origins or to the set of extremities. Therefore, the upper left-hand and lower right-hand nxn partitions of $\underline{X}$ consist of zeroes. On the other hand, the upper right-hand nxn partition corresponds to the definition given for $\underline{U}$. The lower left=hand partition is $\underline{U}^T$ because $\underline{X}$ is symmetric (the graph being undirected).

Proposition A.9: In an undirected bipartite graph if $\underline{U}$ is defined by (A.5.2), the (i,j) entry of $(\underline{U}\,\underline{U}^T)^k$ is the number of paths of length 2k connecting origin i and origin j.

Proof

It is known [35] that the (i,j) entry of $\underline{X}^{2k}$ is the number of paths of length 2k connecting nodes i and j. On the other hand from (A.5.1),

$$\underline{X}^2 = \begin{pmatrix} \underline{0} & \underline{U} \\ \underline{U}^T & \underline{0} \end{pmatrix} \begin{pmatrix} \underline{0} & \underline{U} \\ \underline{U}^T & \underline{0} \end{pmatrix} = \begin{pmatrix} \underline{U}\ \underline{U}^T & \underline{0} \\ \underline{0} & \underline{U}^T\ \underline{U} \end{pmatrix},$$

therefore by induction,

$$\underline{X}^{2k} = \begin{pmatrix} (\underline{U}\ \underline{U}^T)^k & \underline{0} \\ \underline{0} & (\underline{U}^T\ \underline{U})^k \end{pmatrix} .$$

The claim follows since the upper left-hand partition of $\underline{X}^{2k}$ corresponds to paths of length 2k connecting origins with origins.

Q.E.D.

APPENDIX B

PECULIAR CLASSES OF SUBNETWORKS

B.1 INTRODUCTION

This appendix is devoted to extending properly to subnetworks of class 2 the results which have been presented for subnetworks of class 1. Most often, the results for class 2 have been stated, and we prove them here. Also in the section on controllability (section 2.3.4), we have omitted the case when the accessibility graph contains transient nodes. We treat here that situation which may occur for subnetworks of any class.

B.2 SYSTEM TRANSFORMATION

We prove here part 2 of theorem 2.3. To this end, we need the following definition and lemma.

<u>Definition B.2.1:</u>

To a subnetwork (represented by a <u>directed</u> bipartite graph with 2n nodes), we associate a directed graph $\vec{G}$ defined as follows:

The nodes of $\vec{G}$ correspond to the entrance-exit pairs of S. Arc (i,j) exists in $\vec{G}$ if, and only if, there is a link from entrance i to exit j in S.

<u>Note:</u>

This graph $\vec{G}$ is different from the accessibility-graph G which we define in the same way but after having removed the arrows from the subnetwork S.

<u>Lemma B.2.2:</u>

Unless the subnetwork S is of class 3, there exists a cycle passing through less than n nodes in $\vec{G}$.

<u>Proof:</u>

We shall prove that, if the subnetwork is not of class 3, there is at least one cycle passing through less than n nodes in $\vec{G}$. Recall that the subnetwork is of class 3 if, and only if, from each entrance, exactly one link originates, and it does not lead to the corresponding exit. This amounts to the property that, from each node in G, exactly one arc originates, and this arc is not a loop.

Let us notice that $\vec{G}$ is such that at least one arc originates from each node, and at least one arc leads to each node. Indeed, at least one link originates from each entrance, and at least one link leads to each exit in the subnetwork.

If S is not of class 3, either some entrance is connected by a link with the corresponding exit, or there is some entrance from which at least two links originate. In the former case, $\vec{G}$ contains a loop; that is, a cycle with only one node, and the lemma is proved since $n > 1$. In the latter case, there is at least one node in $\vec{G}$ from which at least two arcs originate. Let us call such a node i, and consider arcs (i,j) and (i,k) (Fig. B.2.1). Now, consider the graph $\vec{G}_k$ obtained by removing from G node k and all arcs adjacent to it. The graph $\vec{G}_k$ contains n-1 nodes. If $\vec{G}_k$ contains a cycle, this is a cycle of less than n nodes in $\vec{G}$, and the lemma is proved. If $\vec{G}_k$ contains no cycle, then $\vec{G}_k$ can be ordered [32]; i.e., it is possible to define the nodes of generation 0 (without followers), the nodes of generation 1 (those that would be of generation 0 if the nodes of generation 0 were removed together with the adjacent arcs) and so on up to the last generation of nodes in $\vec{G}_k$.

A node of generation g has followers only in generations g-1, g-2, ..., 0. Therefore, no arc leads to a node of the last generation. On the other hand, a node in generation g is the extremity of some arc with origin in generation g+1. Therefore recurrently, every node in a generation other than the last one is accessible from some node in the last generation. Let ℓ be a node of the last generation from which i is accessible, if i is not in the last generation, and let ℓ=i if it is.

No arc originating from a node in $\vec{G}_k$ has ℓ as its extremity. However, we know that there is at least one arc in $\vec{G}$ which has ℓ as its extremity. Therefore, that arc can only have node k as its origin; i.e., (k,ℓ) is an arc in $\vec{G}$. Accordingly, if $(\ell, \ell_1, \ell_2, \ldots, i)$ is a path from ℓ to i, then $(\ell, \ell_1, \ell_2, \ldots i,k,\ell)$ is a cycle in $\vec{G}$ which has less than n nodes since it does not go through j.

Q.E.D.

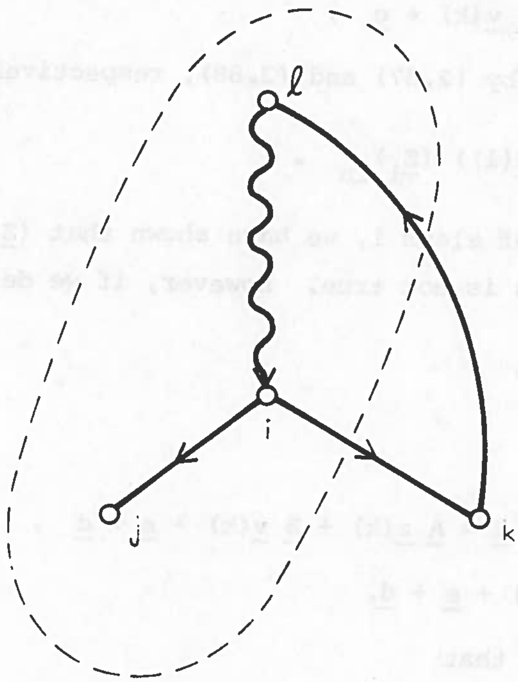

Figure B.2.1 THE GRAPH $\vec{G}_k$ IS LOCALIZED BY A DOTTED LINE. ARCS ARE REPRESENTED BY STRAIGHT LINES AND PATHS, BY TWISTED CURVES

Proof of Theorem 2.3 (part 2)

We have shown in the proof of part 1 that the equations satisfied by $\underline{z}(k)$ (defined by (2.71)) are

$$\underline{z}(k+1) = \underline{A}\ \underline{z}(k) + \underline{B}\ \underline{v}(k) + \underline{e} \quad , \tag{B.1}$$

where $\underline{A}$ and $\underline{B}$ are given by (2.87) and (2.88), respectively, and

$$\underline{e} = F(\underline{Z}_1)_{1n} = (\underline{\nu}_n^T\ \underline{x}(1))\ (\underline{Z}_1)_{1n} \quad . \tag{B.2}$$

When the subnetwork is of class 1, we have shown that $(\underline{Z}_1)_{1n} = 0$. For a subnetwork of class 2, this is not true. However, if we define

$$\underline{z}'(k) = \underline{z}(k) + \underline{d} \quad , \tag{B.3}$$

then,

$$\underline{z}'(k+1) = \underline{z}(k+1) + \underline{d} = \underline{A}\ \underline{z}(k) + \underline{B}\ \underline{v}(k) + \underline{e} + \underline{d} \quad ,$$

$$= \underline{A}\ (\underline{z}'(k) - \underline{d}) + \underline{B}\ \underline{v}(k) + \underline{e} + \underline{d}.$$

Therefore, if $\underline{d}$ is such that

$$\underline{A}\ \underline{d} - \underline{d} = \underline{e} \ , \tag{B.4}$$

then, the sequence $\underline{z}'(k)$ satisfies

$$\underline{z}'(k+1) = \underline{A}\ \underline{z}'(k) + \underline{B}\ \underline{v}(k) \quad . \tag{B.5}$$

We now show that, for a subnetwork of class 2, it is possible to find such a vector $\underline{d}$, and derive the corresponding expression for the cost function.

Let $\underline{\delta}_\ell$ be the (n-1)-dimensional vector whose ℓ^{th} component is 1, and whose all other components are zero. Also recall the partition $\underline{\phi}^T = (\underline{u}^T, \underline{v}^T)$ of section 2.3.3, where $\underline{u} \in R^n$ and $\underline{v} \in R^{m-n}$.

Equation (2.79) together with the flow-conservation constraints (2.10) shows that u_j, for $j<n$, is the flow along link (j,k) if, and only if,

$$\underline{\tilde{Z}}_1 \underline{\delta}_j = \underline{\delta}_k \quad , \tag{B.6}$$

and u_n is the flow along link (n,k) if, and only if,

$$(\underline{Z}_1)_{1n} = \underline{\delta}_k \quad . \tag{B.7}$$

Using lemma B.2.2, we know that G contains a cycle which goes through fewer than n nodes. Consider such a cycle $(i_1, i_2, \ldots, i_r)$ with $r < n$. This cycle corresponds, in the succession of subnetworks identical with S which constitute the stationary network, to a path which visits successively nodes $i_1, i_2, \ldots, i_r$ (Fig. B.2.2 and B.2.3).

Let u_n be the flow along link (n,i). Then,

$$(\underline{Z}_1)_{1n} = \underline{\delta}_i \quad . \tag{B.8}$$

Node i is different from n. Otherwise, the subnetwork will be of class 1. From (B.6), it follows that, if we define u_{i_s} to be the flow along link (i_s, i_{s+1}), for s = 1, 2, ..., r-1 (and u_r to be the flow along link (i_r, i_1)), we have:

$$\tilde{\underline{Z}}_1 \underline{\delta}_{i_1} = \underline{\delta}_{i_2} ; \quad \tilde{\underline{Z}}_1 \underline{\delta}_{i_2} = \underline{\delta}_{i_3} ; \quad \ldots ; \quad \tilde{\underline{Z}}_1 \underline{\delta}_{i_r} = \underline{\delta}_{i_1} \quad . \tag{B.9}$$

Now, let

$$\underline{\zeta} = \underline{\delta}_{i_1} + \underline{\delta}_{i_2} + \ldots + \underline{\delta}_{i_r} \quad . \tag{B.10}$$

From (B.9),

$$\tilde{\underline{Z}}_1 \underline{\zeta} = \underline{\zeta} \quad . \tag{B.11}$$

Choosing now

$$\underline{d} = \frac{-F\underline{\zeta}}{\underline{\nu}_{n-1}^T \underline{\zeta}} \tag{B.12}$$

it follows that (B.4) is satisfied (taking B.8 into account).

Indeed from (2.87), $\underline{A} = \tilde{\underline{Z}}_1 - (\underline{Z}_1)_{1n} \underline{\nu}_{n-1}^T$, so that

$$\underline{A}\,\underline{d} - \underline{d} = (\tilde{\underline{Z}}_1 - \underline{\delta}_i \underline{\nu}_{n-1}^T - \underline{I})\underline{d} = \frac{F}{\underline{\nu}_{n-1}^T \underline{\zeta}} (\underline{I} + \underline{\delta}_i \underline{\nu}_{n-1}^T - \tilde{\underline{Z}}_1)\, \underline{\zeta}$$

$$= F\underline{\delta}_i = \underline{e}.$$

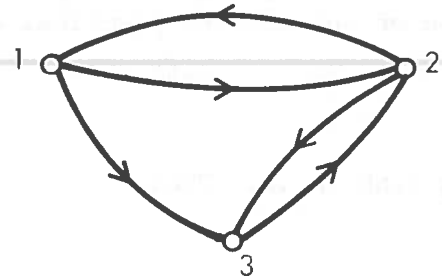

Figure B.2.2 GRAPH G CONTAINING A CYCLE WITH LESS THAN n=3 NODES

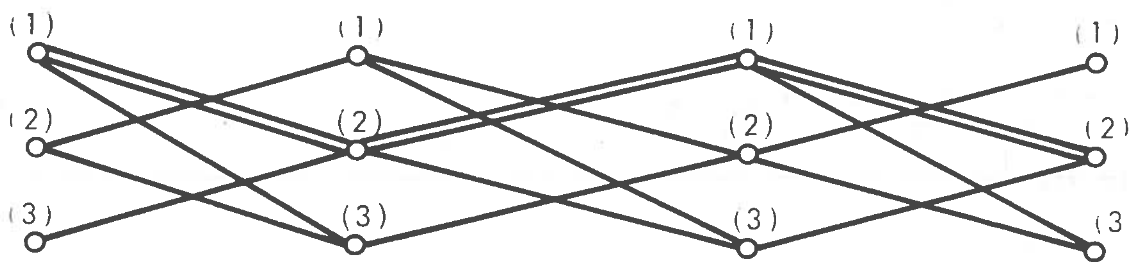

Figure B.2.3 CORRESPONDING PATH IN THE NETWORK

To express the cost function in terms of $\underline{z}'(k)$ and $\underline{v}(k)$, we just have to replace $\underline{z}(k)$ by $\underline{z}'(k) - \underline{d}$, where $\underline{d}$ is given by B.12, and we see that only the coefficients of the linear terms, $\underline{p}$, $\underline{q}$, $\underline{r}$, are changed into $\underline{p}'$, $\underline{q}'$, $\underline{r}'$:

$$\underline{p}' = \underline{p} - 2\,\underline{Q}\,\underline{d} \quad , \tag{B.13}$$

$$\underline{q}' = \underline{q} - 2\,\underline{M}\,\underline{d} \quad , \tag{B.14}$$

$$r' = r + \underline{d}^T\,\underline{Q}\,\underline{d} - \underline{p}^T\underline{d} \quad . \tag{B.15}$$

Q.E.D.

<u>Remark</u>

In subnetworks of class 1, one has the choice between n entrance-exit pairs as to which shall be labeled n. Similarly in subnetworks of class 2, one can choose among all cycles of less than n nodes to define $\underline{\zeta}$ and $\underline{d}$. Once such a cycle has been chosen, the entrance-exit pair to be labeled n must be chosen, such that $n \notin \{i_1, i_2, \ldots, i_r\}$. This also gives several possibilities if $r < n-1$.

B.3 CONTROLLABILITY OF REDUCED SYSTEMS

We first examine the controllability of the reduced systems describing flow conservation in a stationary network consisting of subnetworks of class 2 (without transient nodes). The results have been stated, but not proved, in section 2.3.4. Next, we turn to the case when transient nodes are present (for subnetworks of class 1 or 2). For that category, the results have not been studied in the main work.

B.3.1 <u>Subnetworks of Class 2 without Transient Nodes</u>

<u>Proof of Theorem 2.4 (part 2)</u>

The nodes $i_1, i_2, \ldots, i_r$ mentioned in the theorem are those of section B.2 (i.e., $(i_1, i_2, \ldots, i_{r-1}, i_r, i_1)$ is the cycle with $r < n$ nodes in $\vec{G}$ which has been chosen to define the reduced system in $\underline{z}'$). Controllability of the reduced system in $\underline{z}'$ means, by definition, that it is possible to drive the reduced-state $\underline{z}'$ in some number of steps, say k, to $\underline{z}'(k) = 0$. From B.3, $\underline{z}'(k) = 0$ is equivalent to $\underline{z}(k) = -\underline{d}$, or, according to B.12, to

$$\underline{z}(k) = \frac{F\underline{\zeta}}{\nu^T_{n-1}\underline{\zeta}} \quad . \tag{B.16}$$

From the way $\underline{\zeta}$ is defined (B.10), equation (B.16) amounts to:

$$z_j(k) = \begin{cases} \frac{F}{r} & , \text{ if } j \in \{i_1, i_2, \ldots, i_r\} \ , \\ 0 & \text{otherwise} \ . \end{cases} \tag{B.17}$$

On the other hand, $n \notin \{i_1, \ldots, i_r\}$, so that (B.7) is equivalent to:

$$x_j(k) = \begin{cases} \frac{F}{r} & , \text{ if } j \in \{i_1, \ldots, i_r\} \ , \\ 0 & \text{otherwise.} \end{cases} \tag{B.18}$$

Q.E.D.

Proof of Theorem 2.5 (part 2)

Proof of Part 2.2

The nodes of $i_1, i_2, \ldots, i_r$ are all in the same strongly connected component of G. Indeed, they are successively visited by a cycle in $\vec{G}$, which is also a cycle in G. Let $C(i_1, \ldots, i_r)$ be the strongly connected component to which nodes $i_1, \ldots, i_r$ belong. If G is not strongly connected, then $C(i_1, \ldots, i_r) \neq G$. If i is a node in G and $i \notin C(i_1, \ldots, i_r)$, then there is no path in G from node i to any of the nodes $i_1, \ldots, i_r$. Therefore, it is impossible for a perturbation input in entrance i to be driven to the set of entrances $i_1, \ldots, i_r$, and it is impossible that condition B.18 be fulfilled for any k.

Proof of Part 2.1

Since the graph G is strongly connected and aperiodic by assumption, we can find paths of equal length with origin 1, 2, ..., n, respectively, and extremity j_0, where j_0 is some specific node. (Proposition A.7 of appendix A, where we have replaced n by j_0.) Therefore, we can drive all the components of the flow perturbation to entrance j_0 in a number of steps (i.e., subnetworks) equal to the common length of the n paths. It is possible to find paths of equal length with origin j_0 and with extremities $i_1, i_2, \ldots, i_r$, respectively. To see that it is true, it suffices to apply proposition A_1 to $i_1, \ldots, i_r$ and j_0 in the graph derived from G by reversing all the arrows. In that graph, there are paths of equal length from $i_1, \ldots, i_r$ (and j_0), respectively, to j_0 since that graph is strongly connected and aperiodic as well as G.

Using those r paths in G from j_0 to $i_1, \ldots, i_r$, respectively, one can carry a flow perturbation $\frac{F}{r}$ along each of these paths up to the corresponding node i_j, so that ultimately the total incoming-flow perturbation is spread out evenly among $i_1, \ldots, i_r$; i.e., (B.18) holds.

Q.E.D.

This operation is illustrated in Fig. B.3.1 for the network of Fig. B.2.3. (It is easily verified that the accessibility-graph G is complete, and thus, strongly connected and aperiodic in that example.)

Remark:

As noticed in theorem 2.4, part 2 is only a sufficient condition; i.e., at least we have not proved the necessity. This is because we have not proved that it is necessary first to drive the total flow perturbation to a specific entrance node before spreading it evenly among the nodes $i_1, i_2, \ldots, i_r$.

B.3.2 Subnetworks with Transient Nodes

We extend the results of section 2.3.4 to the case where the accessibility graph associated with the typical subnetwork may contain transient nodes (Fig. 2.3.4-4). In fact, the results of section 2.3.4 are generalized in a straightforward manner with the only restriction that n (for a subnetwork of class 1) or $i_1, i_2, \ldots, i_r$ (for class 2) must not be transient nodes.

Proposition B.3.1:

a. For a subnetwork of class 1

1) If the entrance-exit pair labeled n corresponds to a transient node

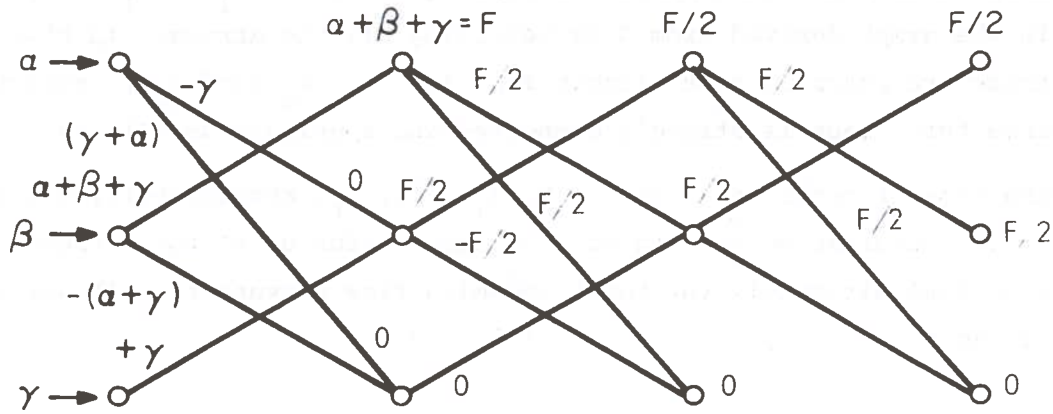

Figure B.3.1 CONTROLLABILITY IN A NETWORK OF CLASS 2

Here, r = 2; i_s = 1, i_2 = 2; n = 3.
The total flow perturbation is sent to 1 in the first step, then divided into two equal parts sent to 1 and 2 respectively, in the second step. The distribution then obtained can be repeated at will in the subsequent sub-networks.

in the accessibility-graph G, the corresponding subsystem in $\underline{z}$ is not controllable.

2) If the entrance-exit pair labeled n corresponds to a final node in G, a necessary and sufficient condition for controllability of the subsystem in $\underline{z}$ is that the graph G have one single final class.

b. For a subnetwork of class 2

1) If the entrance-exit pairs labeled $i_1, i_2, \ldots, i_r$ (defined in section B.2) correspond to transient nodes in the accessibility-graph G, the corresponding subsystem in $\underline{z}'$ is not controllable.

2) If the entrance-exit pairs labeled $i_1, \ldots, i_r$ correspond to final nodes in G, a sufficient condition for controllability of the reduced system in $\underline{z}'$ is that the graph G have one single final class, and that this final class be aperiodic.

Proof:

Part (1) Graph G contains at least one final node. If n is a transient node, it is not accessible from final nodes. Therefore, a perturbation input through an entrance i corresponding to a final node in G cannot be driven to n, and the subsystem is not controllable (theorem 2.4).

Part (2) From proposition A_2(d) (appendix A), we know that from each entrance corresponding to a transient node, a path originates which eventually leads to a node in a later subnetwork which corresponds to a final node in G. Since transient nodes are not accessible from final ones, it is possible to drive all the incoming-flow perturbation to the entrances corresponding to final nodes and to have it stay within them. Therefore, the problem of controllability is reduced to the same problem over the network obtained by deleting the entrances corresponding to transient nodes and the links adjacent to them. The subsystems describing flow conservation in that network are controllable if, and only if, its accessibility graph is strongly connected (theorem 2.4), which amounts to the property that the accessibility graph of the original subnetwork has one single final class.

Part 2 We have already noticed, in B.3.1, that $i_1, i_2, \ldots, i_r$ are in the same strongly connected component of G. Therefore, either they are all transient, or they are all final.

Part (1) If $i_1, \ldots, i_r$ are all transient, they are not accessible from final nodes. Therefore, a perturbation input through an entrance corresponding to a final node in G cannot be spread out among $i_1, \ldots, i_r$, and the subsystem is not controllable (theorem 2.4 and B.3.1).

Part (2) By the same reasoning as in part (2), we conclude that the controllability of subsystems in $\underline{z}'$ is equivalent to the controllability of the subsystems describing flow conservation in the network obtained by deleting the transient nodes and the links adjacent to them. Those subsystems are controllable if the accessibility graph of that subnetwork is strongly connected and aperiodic (theorem 2.5.2). This is equivalent to requiring that the accessibility graph of the original subnetwork have only one final class, and that this final class be aperiodic.

Q.E.D.

We now give an algebraic test to determine the transient nodes.

Having defined the entrance-exit adjacency matrix $\underline{U}$ of the subnetwork (section 2.3.4), the path matrix $\underline{V}$ and the stochastic matrix $\underline{P}$ adapted to the graph G, we know that G has only one final class if, and only if, the number 1 is a simple eigenvalue of $\underline{P}$. If $\underline{t}$ is the corresponding stationary distribution; i.e.,

$$\underline{t}^T \underline{P} = t^T \quad , \text{ and } \quad \underline{t}^T \underline{\nu}_n = 1 \quad ,$$

then, the following property holds.

Proposition B.3.2:

The node i is transient if, and only if, $t_i = 0$.

Proof: in [33]

Taking into account this property, as well as proposition B.3 and the other algebraic tests given in section 2.3.4, we give in Fig. B.3.2 a flowchart which sums up the algebraic tests for controllability in the most general case.

We now go through sections 2.3.5 and 2.3.6, pointing out how little the results there depend on the class to which the subnetwork belongs.

B.4 EXPRESSION FOR THE MATRIX DERIVATIVE

In section 2.3.5, we show that, if the system ($\underline{A}$, $\underline{B}$) is controllable, then the sequence $\underline{K}(i)$ converges. $\underline{K}(i)$ depends only on the quadratic coefficients of the cost functions, which are the same if that function is expressed in terms of $\underline{z}$ or $\underline{z}'$.

This property is also true for $\underline{\Lambda}(i)$, which determines the convergence of the sequence $\underline{\ell}(i)$. Therefore, in spite of the fact that $\underline{z}(k)$ satisfies

$$z(k+1) = \underline{A}\ \underline{z}(k) + \underline{B}\ \underline{v}(k) + \underline{e} \quad , \qquad \text{(B.19)}$$

and that it is the sequence $\underline{z}'(k)$ which satisfies the same system without the constant term $\underline{e}$, we express the cost function in terms of $\underline{z}$ and not $\underline{z}'$. This yields the identification of section 2.3.5, which is valid if the subsystem is of class 1 or 2.

In section 2.3.6, we see from equation (2.160) that $\underline{\Delta}(k)$ is the same for a system in $\underline{z}'$ as for a system in $\underline{z}$. However, $\underline{\Pi}(k)$ will be different. Indeed, from equation (2.168),

$$\underline{z}^{*\prime}(k+1) = \underline{\Delta}(k)\ \underline{z}'^{*}(k) + \underline{\Pi}(k)F \quad , \qquad \text{(B.20)}$$

where

$$\underline{\Pi}'(k) = -\ \underline{B}[\underline{R} + \underline{B}^T\underline{K}(k+1)\underline{B}]^{-1}\ [\underline{\Lambda}_n^T + \frac{2M\underline{\zeta}}{\underline{\nu}_{n-1}^T\underline{\zeta}} + \underline{B}^T\underline{\ell}'(k+1)]F \quad , \qquad \text{(B.21)}$$

where we have used equation (B.14) for $\underline{q}'$.

Equations (2.169) and (2.170) have to be modified. Indeed, (B.20) can be written

$$x_i^*(k+1) + d_i = \sum_{j=1}^{n-1} \Delta_{ij}(k)\ \ [x_j^*(k) + d_j] + \Pi_i'(k)\ \ (\sum_{j=1}^{n} x_j^*(k)) \quad ,$$

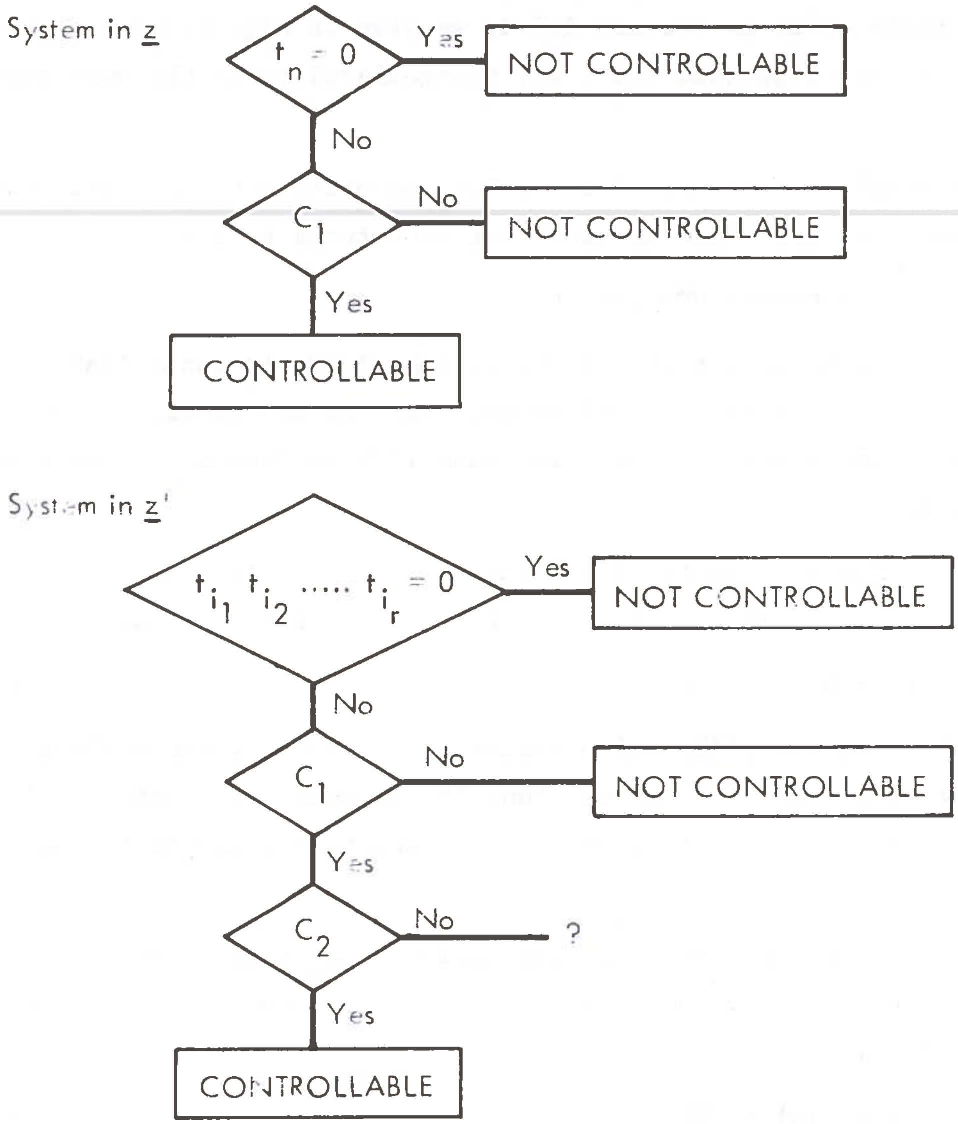

Figure B.3.2 CONTROLLABILITY FLOWCHART IN THE GENERAL CASE

Condition C_1 is: "the number 1 is a simple eigenvalue of the Stochastic Matrix $\underline{P}$ (of section 2.3.4):"

Condition C_2 is: "the Matrix $\underline{P}$ has no other eigenvalue of magnitude 1 than 1."

for i=1,2,...,n-1, from which one obtains the equivalent of equation (2.159) for $\underline{D}(k)$, with $\underline{\Pi}''(k)$ instead of $\underline{\Pi}(k)$, where

$$\underline{\Pi}''(k) = \underline{\Pi}'(k) + (\underline{I} - \underline{\Delta}(k)) \frac{\underline{\zeta}}{\underline{\nu}_{n-1}^T \underline{\zeta}}, \tag{B.22}$$

and $\Pi'(k)$ is given by (B.21).

We can relate $\underline{\Pi}'' \stackrel{\Delta}{=} \lim_{k \to -\infty} \underline{\Pi}''(k)$ to the corresponding limit $\underline{\Pi}$.

Indeed, comparing equation 2.134 for $\underline{\ell}(i)$ with the corresponding equation for $\underline{\ell}'(i)$, and taking into account equations (B.13), (B.14) for $\underline{p}'$ and $\underline{q}'$, respectively, one obtains:

$$\hat{\underline{\ell}}' = \hat{\underline{\ell}} + 2(\underline{I}-\underline{\Delta}^T)^{-1} [\underline{Q} - (\underline{M}^T + \underline{A}^T\hat{\underline{K}}\,\underline{B})\,(\underline{R} + \underline{B}^T\,\hat{\underline{K}}\,\underline{B})^{-1}\underline{M}]\, \frac{\underline{\zeta}}{\underline{\nu}_{n-1}^T \underline{\zeta}}, \tag{B.23}$$

whence (B.24)

$$\underline{\Pi}' = \underline{\Pi} - \underline{B}(\underline{R} + \underline{B}^T\hat{\underline{K}}\,\underline{B})^{-1} \left\{ \frac{2\underline{M}\underline{\zeta}}{\underline{\nu}_{n-1}^T\underline{\zeta}} + 2\,\underline{B}^T(\underline{I}-\underline{\Delta}^T)^{-1}\,[\underline{Q}-(\underline{M}^T+\underline{A}^T\hat{\underline{K}}\,\underline{B})\,(\underline{R}+\underline{B}^T\hat{\underline{K}}\,\underline{B})^{-1}\underline{M}]\,\frac{\underline{\zeta}}{\underline{\nu}_{n-1}^T\underline{\zeta}} \right\},$$

and

$$\underline{\Pi}'' = \underline{\Pi}' + (I - \underline{\Delta}) \frac{\underline{\zeta}}{\underline{\nu}_{n-1}^T \underline{\zeta}}.$$

In spite of this change in $\underline{\Pi}(k)$ and $\underline{\Pi}$, lemma 2.12, relating the eigenvalues of $\underline{D}$ and those of $\underline{\Delta}$, is still valid because it does not depend on what $\underline{\Pi}$ is but only on the structure (2.165) of the matrix $\underline{D}$.

Vaughan's results can then be applied to the linear system in $\underline{z}'(i)$, and lead to the same relation for $\underline{\Delta}$ as before. Also, the controllability of the system in $\underline{z}'$ implies that $\underline{\Delta}$ has no eigenvalue on the unit circle.

The proof of theorem 2.19 is still valid since if F vanishes, $\underline{d}$ also vanishes (equation (B.12).

The only change caused by the change in $\underline{\Pi}$ concerns $\underline{p}$, the eigenvector of $\underline{D}$ with eigenvalue 1. Also, in the bound

$$||\underline{D}^K \underline{x}|| \leq M|s|^K ,$$

valid when $\nu_n^T \underline{x} = 0$, the ratio $|s|$ is still the same since it is the maximum magnitude of an eigenvalue of $\underline{\Delta}$, but M differs from the case of a reduction in $\underline{z}$ since it involves $\underline{p}$ and thus $\underline{\Pi}$.

APPENDIX C

CLASS OF DISCRETE LINEAR RECURSIVE EQUATIONS

We present here a result on convergence of sequences defined by recursive linear relations. This result is needed in section 2.3 to prove that the sequence $\underline{\ell}(i)$ converges under certain conditions.

Definition C.1

The spectral radius [25] of a square matrix $\underline{A}$, denoted $\rho(\underline{A})$, is the largest magnitude of an eigenvalue of $\underline{A}$.

Definition C.2

A regular vector norm [38] on R^n is a continuous mapping $\underline{x} \rightarrow ||\underline{x}||$ from R^n to R that has the following properties:

a. $||\underline{x}|| \geq 0$ and $||x|| = 0$ if, and only if, $\underline{x} = 0$.

b. $||a\,\underline{x}|| = |a| \cdot ||\underline{x}||$ for any scalar a.

c. $||\underline{x} + \underline{y}|| \leq ||\underline{x}|| + ||\underline{y}||$ (triangle inequality).

d. There are positive constants α and β, such that

$$\alpha \max_{k=1,\ldots,n} |x_k| \leq ||\underline{x}|| \leq \beta \max_{k=1,\ldots,n} |x_k|.$$

Remark

From part (d) of definition C.2, it follows that

$$\lim_{m\to\infty} ||\underline{x}(m) - \underline{x}|| = 0$$

is equivalent to: $\lim_{m\to\infty} |x_k(m) - x_k| = 0$ for $k = 1,\ldots,n$.

Definition C.3

The matrix norm induced by a regular vector norm is the mapping from the space R^{nxn} of (n x n) matrices to R defined by:

$$||\underline{A}|| = \max_{||\underline{x}||=1} ||\underline{A}\,\underline{x}|| = \max_{\underline{x}\neq 0} \frac{||\underline{A}\,\underline{x}||}{||\underline{x}||}.$$

Lemma C.4

For every matrix norm induced by a regular vector norm:

a. $\rho(\underline{A}) \leq ||\underline{A}||$.

b. Given any positive number ε, a regular vector norm can be defined on R^n for which the induced matrix norm satisfies the inequality

$$||\underline{A}|| < \rho(\underline{A}) + \varepsilon.$$

Proof: See [38].

Lemma C.5

Every matrix norm induced by a regular vector norm has the following properties:

a. $||A|| \geq 0$.

b. $||c\,\underline{A}|| = |c| \cdot ||\underline{A}||$ for any scalar c.

c. $||\underline{A} + \underline{B}|| \leq ||\underline{A}|| + ||\underline{B}||$.

d. $||\underline{A} \cdot \underline{B}|| \leq ||A||\ ||\underline{B}||$,

and thus,

$$||\underline{A}^i|| \leq ||\underline{A}||^i \quad \text{for all positive integers } i.$$

e. $||\underline{A}\,\underline{x}|| \leq ||\underline{A}|| \cdot ||\underline{x}||$ for any matrix $\underline{A}$ and vector $\underline{x}$.

f. There exist positive constants, a and b, such that:

$$a \max_{i,j} |A_{ij}| \leq ||\underline{A}|| \leq b \max_{i,j} |A_{ij}| ,$$

for any matrix $\underline{A}$.

Proof: See [38].

Remark

From part (f) of lemma C.5, it follows that

$$\lim_{m\to\infty} ||\underline{A}(m) - \underline{A}|| = 0,$$

is equivalent to

$$\lim_{m\to\infty} |A_{ij}(m) - A_{ij}| = 0, \quad \text{for } i = 1,\ldots,n, \; j = 1,\ldots,n,$$

we can now state the theorem that we need.

Theorem C.6

a. Let $\underline{A}(i)$ be a converging sequence of matrices and $\underline{b}(i)$ a converging sequence of vectors.

$$\lim_{i\to\infty} \underline{A}(i) = \underline{A} , \tag{C.1}$$

$$\lim_{i\to\infty} \underline{b}(i) = \underline{b} , \tag{C.2}$$

or alternatively,

$$\lim_{i\to\infty} A_{jk}(i) = A_{jk}, \quad \text{for } j = 1,\ldots,n; \;\; k = 1,\ldots,n,$$

$$\lim_{i\to\infty} b_j(i) = b_j, \quad \text{for } j = 1,\ldots,n.$$

b. Let $\underline{A}$ be a stable matrix; i.e., a matrix with spectral radius less than 1.

$$\rho(\underline{A}) < 1.$$

c. Consider the following recursive sequence $\underline{x}(i)$ defined from $\underline{x}(1)$ by:

$$\underline{x}(i+1) = \underline{A}(i)\, \underline{x}(i) + \underline{b}(i). \tag{C.3}$$

Then, $(\underline{I}-\underline{A})$ is invertible, the sequence $\underline{x}(i)$ converges, and the limit $\hat{\underline{x}}$ is given by

$$\hat{\underline{x}} = (\underline{I}-\underline{A})^{-1}\, b \;. \tag{C.4}$$

Proof:

Since $\rho(\underline{A})$, the maximum magnitude of an eigenvalue of $\underline{A}$, is less than 1, the number 1 is not an eigenvalue, so that $(\underline{I}-\underline{A})$ is invertible. The component-wise convergence is equivalent to the convergence in any regular norm (or in the induced matrix norm).

By lemma C.4, we choose a regular vector norm, such that, for the induced matrix norm,

$$||\underline{A}|| < \rho(\underline{A}) + \varepsilon < 1 \;. \tag{C.5}$$

a. Let us first study the sequence $\underline{y}(i)$ defined recursively from $\underline{y}(1)$ by

$$\underline{y}(i+1) = \underline{A}\,\underline{y}(i) + \underline{b}. \tag{C.6}$$

Equation (C.6) implies

$$\underline{y}(i+1) = \underline{A}^i\,\underline{y}(1) + (\underline{I}+\underline{A} + \ldots + \underline{A}^{i-1})\underline{b} \tag{C.7}$$

Since $||A^i|| \leq ||A||^i$, it follows, from (C.5), that

$$\lim_{i\to\infty} \underline{A}^i = 0, \text{ and}$$

$$\lim_{i\to\infty} \underline{y}(i) = \lim_{i\to\infty} (\underline{I}+\underline{A} + \ldots + \underline{A}^{i-1})\underline{b} = (\underline{I}-\underline{A})^{-1}\,\underline{b} = \hat{\underline{x}}.$$

b. Consider now $\underline{x}(i)$ defined by (C.3).

$$\underline{x}(i+1) - \hat{\underline{x}} = \underline{A}(i)\,\underline{x}(i) - \hat{\underline{x}} + \underline{b}(i),$$

$$= \underline{A}(i)\,\underline{x}(i) - \underline{A}\,\hat{\underline{x}} + \underline{b}(i) - \underline{b}, \tag{C.8}$$

since $\hat{\underline{x}} = \underline{A}\,\hat{\underline{x}} + \underline{b}$ from (C.4).

Equation (C.8) can be written

$$\underline{x}(i+1) - \hat{\underline{x}} = \underline{A}(i)\,(\underline{x}(i)-\hat{\underline{x}}) + (\underline{A}(i)-\underline{A})\,\hat{\underline{x}} + (\underline{b}(i)-\underline{b}), \tag{C.9}$$

whence there follows, letting

$$\underline{u}(i) \triangleq \underline{x}(i) - \hat{\underline{x}},$$

$$||\underline{u}(i+1)|| \leq ||\underline{A}(i)||\cdot||u(i)|| + ||\underline{A}(i)-\underline{A}||\cdot||\underline{x}|| + ||\underline{b}(i)-\underline{b}||, \tag{C.10}$$

where we have used the properties of regular norms (def. C.2 and lemma C.5). Because of (C.1) and (C.2), for any positive number η, there exists an integer N, such that, for $i \geq N$,

$$||\underline{A}(i)-\underline{A}||\cdot||\hat{\underline{x}}|| + ||\underline{b}(i)-\underline{b}|| < \eta$$

$$||\underline{A}(i) - \underline{A}|| < \eta.$$

We choose η, such that $||\underline{A}|| + \eta < 1$.

From (C.10),

$$||\underline{u}(i+1)|| \leq ||\underline{A}(i)||\cdot||u(i)|| + \eta \quad \text{for all } i \geq N \;, \qquad \text{(C.11)}$$

also,

$$||\underline{A}(i)|| \leq ||\underline{A}|| + ||\underline{A}(i) - \underline{A}|| < ||\underline{A}|| + \eta \quad \text{for all } i \geq N \;,$$

so that (C.11) becomes:

$$||\underline{u}(N+k+1)|| \leq (||\underline{A}||+\eta)\; ||\underline{u}(N+k|| \;+\; \eta, \qquad \text{(C.12)}$$

for all positive integers k.

And, by induction

$$||\underline{u}(N+k)|| \leq (||\underline{A}||+\eta)^k \; ||\underline{u}(N)||$$

$$+ \; \eta[1+(||\underline{A}||+\eta) + (||\underline{A}||+\eta)^2 + \ldots + (||\underline{A}||+\eta)^{k-1}]$$

$$\leq (||\underline{A}||+\eta)^k ||\underline{u}(N)|| + \frac{\eta}{1-(||\underline{A}||+\eta)} \;.$$

Since $||\underline{A}|| + \eta < 1$, it follows that

$$\lim_{k\to\infty} ||u(N+k)|| \leq \frac{\eta}{1-(||\underline{A}||+\eta)} \;.$$

This is true for any η satisfying $0 < \eta < 1 - ||\underline{A}||$. Therefore, it is true for $\eta \to 0$, and

$$\lim_{k\to\infty} ||\underline{u}(N+k)|| = \lim_{i\to\infty} ||u(i)|| = 0 \;,$$

or

$$\lim_{i\to\infty} \underline{x}(i) = \hat{\underline{x}} \;.$$

Q.E.D.

APPENDIX D

POWERS OF A CERTAIN CLASS OF MATRICES

We prove here the first part of theorem 2.12. To do so, let us use the following expansion, valid for any (nxn) matrix and every positive integer k [31].

Theorem D.1

Let $\underline{D}$ be an (nxn) matrix. Then, for any positive integer k,

$$\underline{D}^k = \underline{E}_k + \sum_{i=0}^{r} s_i^k \underline{B}_i(k) \quad \text{(D.1)}$$

where

a. $\underline{E}_k$, $\underline{B}_0$, $\underline{B}_1,\ldots,\underline{B}_r$ are (nxn) matrices.

b. The matrix $\underline{E}_k$ is different from zero if and only if 0 is an eigenvalue of $\underline{D}$ and $k < n$.

c. $\underline{D}$ has $r + 1 \leq n$ distinct eigenvalues: $s_0, s_1,\ldots,s_r$ other than 0.

d. $\underline{B}_i(k)$ is a polynomial in k, with degree $\delta_i = \alpha_i - \beta_i$, where α_i is the geometrical multiplicity. In particular, if those two multiplicities are equal, $\underline{B}_i(k)$ is independent of k.

See [31] for the proof of this general theorem. We can apply this theorem to our case, taking into account the assumptions of theorem 2.12. Because of assumption 1, $\underline{E}_k=0$ for all k. We can take $s_0=1$ since $\underline{D}$ has 1 as an eigenvalue in our case. Because of assumption (d) of theorem 2.12, the $\underline{B}_i$ matrices are independent of k.

Therefore, (D.1) can be written

$$\underline{D}^k = \underline{B}_0 + \sum_{i=1}^{r} s_i^k \underline{B}_i$$

where $|s_i| < 1$ for i=1,2,...,r (by assumption c). It remains to show that

$$\underline{B}_0 = \underline{p}\, \underline{\nu}_n^T \quad .$$

We shall use assumption (b) for that purpose.

Since $|s_i| < 1$ for i = 1,2,...,r,

$$\lim_{k\to\infty} s_i^k = 0$$

and

$$\lim_{k\to\infty} \underline{D}^k = \underline{B}_0 \tag{D.2}$$

From $\underline{D}^{k+1} = \underline{D}\,\underline{D}^k$ and (D.2), it follows that

$$\underline{D} \cdot \underline{B}_0 = \underline{B}_0 \tag{D.3}$$

or

$$\underline{D}\,\underline{b}(j) = \underline{b}(j) \qquad j = 1,\ldots,n \tag{D.4}$$

where $\underline{b}(j)$ is the jth column of $\underline{B}_0$. On the other hand, we know (2.177) that

$$\underline{\nu}_n^T \cdot \underline{D} = \underline{\nu}_n^T$$

therefore,

$$\underline{\nu}_n^T \cdot \underline{B}_0 = \underline{\nu}_n^T$$

from (D.2). Or

$$\underline{\nu}_n^T \cdot \underline{b}(j) = 1 \qquad j = 1,\ldots,n\ . \tag{D.5}$$

Finally, (D.4) and (D.5) together imply that

$$\underline{b}(j) = p \qquad j = 1,\ldots,n \tag{D.6}$$

because $\underline{p}$ is the only vector solution of the system (D.4), (D.5) by assumption (b).

Equation (D.6) is equivalent to:

$$\underline{B}_0 = \underline{p}\,\underline{\nu}_n^T\ . \tag{D.7}$$

Q.E.D.

<u>Remark</u>

Theorem 2.12 is usually found in Markov chain textbooks. However, no positivity assumption on the entries of $\underline{D}$ is needed as we have seen. Only the assumptions of theorem 2.12 have been used. Those assumptions are usually satisfied by stochastic matrices.

APPENDIX E

DIAGONALIZATION THEOREM FOR THE DISCRETE-TIME LINEAR-QUADRATIC PROBLEM

We prove here Theorem 2.15. What is there called Problem 1 is a discrete-time optimal control problem.

To any sequence $\underline{v}(i)$, i=1,2,...,N-1, we can associate the sequence $\tilde{\underline{v}}(i)$, i = 1,2,...,N-1, defined by

$$\tilde{\underline{v}}(i) = \underline{v}(i) + \underline{R}^{-1} \underline{M}\, \underline{z}(i) \quad . \tag{E.1}$$

The sequences of states $\underline{z}(i)$ and of controls $\underline{v}(i)$ satisfy (2.201) if, and only if, the sequences of states $\underline{z}(i)$ and of new controls $\tilde{v}(i)$ satisfy:

$$\begin{aligned} \underline{z}(i+1) &= \underline{A}\, \underline{z}(i) + \underline{B}[\underline{v}(i) - \underline{R}^{-1}\underline{M}\, \underline{z}(i)], \\ &= (\underline{A} - \underline{B}\, \underline{R}^{-1}\underline{M})\underline{z}(i) + \underline{B}\, \tilde{\underline{v}}(i), \end{aligned} \tag{E.2}$$

which is equation (2.203).

A sequence $\underline{v}(i)$ and a sequence $\tilde{\underline{v}}(i)$ related by (E.1) give rise to the same sequence $\underline{z}(i)$ of states, given the same initial state $\underline{z}(1)$.

Now let us show that choosing $\underline{v}(i)$ to minimize the cost function (2.200) while satisfying (2.201) is equivalent to choosing $\tilde{\underline{v}}(i)$ to minimize the cost function (2.202) while satisfying (2.203).

To this end, it suffices to express the cost (2.200) in terms of $\underline{z}(i)$ and $\underline{v}(i)$, using (E.1). One obtains:

$$\begin{aligned} [\underline{z}^T(k)\ \underline{v}^T(k)] \begin{bmatrix} \underline{Q} & \underline{M}^T \\ \underline{M} & \underline{R} \end{bmatrix} \begin{bmatrix} \underline{z}(k) \\ \underline{v}(k) \end{bmatrix} &= \underline{z}^T(k)\ \underline{Q}\ \underline{z}(k) + 2\underline{z}^T(k)\underline{M}^T\ \underline{v}(k) + \underline{v}^T(k)\underline{R}\ \underline{v}(k), \\ &= \underline{z}^T(k)\underline{Q}\ \underline{z}(k) + 2\underline{z}^T(k)\underline{M}^T(\tilde{\underline{v}}(k) - R^{-1}\underline{M}\ \underline{z}(k)), \\ &\quad + (\tilde{\underline{v}}(k) - \underline{R}^{-1}\underline{M}\ \underline{z}(k))^T\underline{R}(\tilde{\underline{v}}(k) - \underline{R}^{-1}\underline{M}\ \underline{z}(k)), \\ &= \underline{z}^T(k)(\underline{Q} - \underline{M}^T\underline{R}^{-1}\underline{M})\underline{z}(k) + \tilde{\underline{v}}^T(k)\underline{R}\ \tilde{\underline{v}}(k) \quad . \end{aligned} \tag{E.3}$$

Therefore, to the optimal sequence $\underline{v}(i)$ for the cost function (2.200) of Problem 1, there corresponds the optimal sequence $\underline{\tilde{v}}(i)$ for the cost function (2.202) of Problem 2.

The value of the cost function (2.200) corresponding to a given choice of $\underline{v}(i)$ is equal to the value of the cost function (2.202) for the corresponding choice of $\underline{\tilde{v}}(i)$, so that the minimal value is the same for both. The two corresponding choices lead to the same sequence of optimal states $\underline{z}^*(i)$.

APPENDIX F

COMPUTER PROGRAM

The program is written in WATFIV (*) using double precision. See table F.1.

To begin with, the number of subnetworks, N, is read. Next, a parameter NSTAT is read. NSTAT is equal to zero if the network is nonstationary, and different from zero if it is stationary.

F.1 STATIONARY NETWORK

The data are: dimension NN of the state, dimension MM of the control, initial perturbation vector PERT, cost matrix $\underline{L}$, and incidence matrices $\underline{y}$ and $\underline{Z}$. After echoing the data, the program jumps to line 59 where it reads the terminal penalty $\underline{C}$(N). Then, the dynamic-programming equations of section 2.2 are implemented in two steps. In a first step (DO loop over k from line 66 to line 103), the sequence $\underline{C}$(i) is computed backward rom $\underline{C}$(N), and a sequence of matrices $\underline{A}$(k) is kept in memory.

In a second step (DO loop over M from line 127 to line 153), the sequence $\underline{D}$(i) is computed by means of the $\underline{A}$(k). ($\underline{D}$ is denoted DER for derivative.)

At the same time, the sequence of perturbations x_i(k) (denoted DD(I,K)) is computed.

The printing of the results is done by the subroutine IMP, which selects different formats according to the dimension. TRANS is a subroutine which transposes matrices. Built-in subroutines for matrix addition (DMAGG), multiplication (DMMGG), and inversion (DMIG) are used.

F.2 NON-STATIONARY NETWORK

In this case (i.e., NSTAT=0), the same steps are followed as in the stationary case, but the incidence matrices $\underline{Y}$(k) and $\underline{Z}$(k) have to be read at each stage k. Also, the cost matrices $\underline{L}$(k) are not given as such in the data, but have to be computed, at each stage, by the subroutine COST (called at line

* This is the University of Waterloo (Waterloo, Ontario, Canada), version of IBM 370 Fortran.

71). The data of subroutine COST, at stage k, are: NSTOP, NCLASS, NRAMP, LIGHT, G, PHI, XLONG, XLANE.

a. NSTOP is the number of traffic lights within subnetwork k.

b. NCLASS is an MM-dimensional integer vector (where MM, read at stage k, is the sum of the number of links and of the number of traffic lights is subnetwork k.)

The ith component of NCLASS is the class number of the link or traffic light to which the ith component of $\underline{\phi}(k)$ corresponds. We distinguish between five classes.

Class 1 is that of freeway links (called class A in Section 5). Class 2 is that of entrance ramps (called class B in Section 5). Among the links leading to traffic lights (called class C in Section 5), we distinguish between class 3 and class 4. This has no physical meaning but comes from the method adopted in Section 5: we include only one green-split variable per light and replace the other one in terms of the first one ($g_j=\alpha-g_i$, see Section 5).

Thus, we call links of class 3 those which lead to a traffic light and whose corresponding green split has been chosen as a component of $\underline{\phi}(k)$; we call links of class 4 the other links which lead to traffic lights. For instance in a corridor, we may adopt, for class 3, the links oriented in the overall direction of traffic, and, for class 4, the side roads which lead to traffic lights. Finally, class 5 consists of the traffic lights.

c. NRAPM is an MM-dimensional integer-vector. The ith component of NRAMP is the label of the freeway link which link i impinges upon if i corresponds to an entrance ramp, and zero otherwise. More precisely, if link i is an entrance ramp, NRAMP(I) is the corresponding link k is Fig. 5.2.1.

d. LIGHT is an MM-dimensional integer vector. If the ith component of $\underline{\phi}(k)$ corresponds to a link leading to a signal, LIGHT(I) is the label of that signal; likewise, if the ith component of $\underline{\phi}(k)$ corresponds to a green split, LIGHT(I) = I. Otherwise, LIGHT(I) = 0.

e. G is a vector with NSTOP components; i.e., as many as there are traffic lights. Its ith component is the otpimal value of that green split variable selected for light i (as a component of $\underline{\phi}(k)$), found by the global nonlinear otpimization technique.

f. PHI is an MM-dimensional vector. PHI(I) = $\phi_i^*(k)$ is the optimal value of the flow along link i found by the global optimization technique. PHI has also somponents corresponding to green-split variables. A link may cross a freeway at a green traffic light without interfering anywhere else with the remainder of the network (for instance, links 9 and 16 in Fig. 5.2.7-1). If j is the label of the corresponding traffic light, PHI(J) will be the (nominal) optimal flow along the corresponding crossing link. If there is no crossing link at some traffic light, the corresponding component of PHI is of no use and we may set it equal to an arbitrary value.

g. XLONG is an MM-dimensional vector. If the ith component of $\underline{\phi}(k)$ is the flow along some link, then XLONG(I) is the length of that link. If the ith component of $\underline{\phi}(k)$ is a green-split variable, then XLONG(I) is the length of the corresponding crossing link if such a link exists, and zero otherwise.

h. XLANE is an MM-dimensional vector. If the ith component of $\underline{\phi}(k)$ is the flow along some link, XLANE(I) is the number of lanes that link consists of; if the ith component of $\underline{\phi}(k)$ is a green-split variable, XLANE(I) is the number of lanes of the corresponding crossing link if such a link exists, and zero otherwise. From XLANE(I), PHIMAX(I) and RHOMAX(I) (i.e., $\phi_{i\ max}$ and $\rho_{i\ max}$ of Secion 5) are computed.

After having read the data (of subnetwork k), the subroutine COST reviews all the links and lights in a DO loop (lines 292 to 339), and computes the corresponding terms in the cost function from the equations of Section 5. In doing so, it uses three external function; i.e., FUNCT 1, FUNCT2, and FUNCT 3.

```
$JOB           DERSIN,NOSUBCHK
          IMPLICIT  REAL*8(A-H,O-Z,$)
         DOUBLE PRECISION L,LI
          INTEGER M,N
         DIMENSION X(10,10),AUX(10)
         DIMENSION DD(10,30)
         DIMENSION RAT(30)
         DIMENSION Y(10,10),Z(10,10),Q(10,10),L(10,10),F(10,10),H(10,
         DIMENSION U(10,10),V(10,10),E(10,10),P(10,10),D1(30),C(10,10
         DIMENSION R(10,10),D(10),T(10),W(10,10)
         DIMENSION PERT(10),DELTA(30),LI(10),A(10,10,30),DC(10,10,30)
         DIMENSION DER(10,10,30)
         NCOUNT=0
         PRINT 1001
          PRINT 900
          PRINT 900
    100  READ,N
         IF(N.EQ.0) GO TO 1111
         NCOUNT=NCOUNT+1
         PRINT 1101,NCOUNT
         READ,NSTAT
         IF(NSTAT.EQ.0) GO TO 81
         PRINT 999
         READ,NN,MM
         READ,(PERT(I),I=1,NN)
         DO 1000 I=1,MM
         READ,(L(I,J),J=1,MM)
   1000  CONTINUE
         READ,(LI(J),J=1,MM)
         READ,B,((Y(I,J),J=1,MM),I=1,NN)
         READ,((Z(I,J),J=1,MM),I=1,NN)
         PRINT 900
         PRINT 1102,NN
         PRINT 1103,MM
         PRINT 1002,N
         PRINT 1003,B
         PRINT 1004
         PRINT,(LI(J),J=1,MM)
         PRINT 1005
         PRINT 1006
         DO 106 I=1,NN
         PRINT,(Y(I,J),J=1,MM)
    106  CONTINUE
         PRINT 1007
         DO 107 I=1,NN
         PRINT,(Z(I,J),J=1,MM)
    107  CONTINUE
         PRINT 1008
         PRINT,(PERT(I),I=1,NN)
         PRINT 900
  81     IF(NSTAT.NE.0) GO TO 88
         PRINT 9999
         READ,NN
         READ,(PERT(I),I=1,NN)
         PRINT 1002,N
```

```
      PRINT 1102,NN
      PRINT 1008
      PRINT,(PERT(I),I=1,NN)
   88 CONTINUE
      DO 71 I=1,NN
      READ,(R(I,J),J=1,NN)
   71 CONTINUE
      DO 17 I=1,NN
      DO 34 J=1,NN
   34 C(I,J,1)=R(I,J)
   17 CONTINUE
      IER=0
      NMI=N-1
      DO 1 K=1,NMI
      IF(NSTAT.NE.0) GO TO 82
       READ,MM,((Y(I,J),J=1,MM),I=1,NN)
       READ,((Z(I,J),J=1,MM),I=1,NN)
      PRINT 1019,K
      CALL COST(L,MM)
      PRINT 1006
      DO 836 I=1,NN
      PRINT,(Y(I,J),J=1,MM)
  836 CONTINUE
      PRINT 1007
      DO 807 I=1,NN
      PRINT,(Z(I,J),J=1,MM)
  807 CONTINUE
      PRINT 900
   82 CONTINUE
      CALL DMMGG(R,10,Z,10,NN,NN,MM,E,10,IER)
      CALL TRANS(Z,X,MM,NN)
      CALL DMMGG(X,10,E,10,MM,NN,MM,Q,10,IER)
      CALL DMAGG(L,10,Q,10,MM,MM,3,F,10,IER)
      CALL DMIG(F,10,MM,AUX,IER)
      DO 33I=1,MM
      DO 666 M=1,MM
      A(I,M,K)=F(I,M)
  666 CONTINUE
   33 CONTINUE
      CALL TRANS(Y,X,MM,NN)
      CALLDMMGG(F,10,X,10,MM,MM,NN,P,10,IER)
      CALL DMMGG(Y,10,P,10,NN,MM,NN,H,10,IER)
      CALL DMIG(H,10,NN,AUX,IER)
      DO 22 I=1,NN
      DO 44 J=1,NN
      C(I,J,K+1)=H(I,J)
      DC(I,J,K+1)=C(I,J,K+1)-C(I,J,K)
      R(I,J)=H(I,J)
   44 CONTINUE
   22 CONTINUE
    1 CONTINUE
       NMI=N-1
      PRINT 499
      CALL IMP(C,NMI,1,NN)
      IF(NN.LT.3) GO TO 16
```

```
      PRINT 498
      CALL IMP(C,NMI,2,NN)
      IF(NN.LT.4) GO TO 16
      PRINT 497
      CALL IMP(C,NMI,3,NN)
   16 PRINT 900
      PRINT 1009
      PRINT 599
      CALL IMP(DC,NMI,1,NN)
      IF(NN.LT.3) GO TO 20
      PRINT 598
      CALL IMP(DC,NMI,2,NN)
      IF(NN.LT.4) GO TO 20
      PRINT 597
       CALL IMP(DC,NMI,3,NN)
   20 PRINT 900
      DO 30 I=1,NN
   30 T(I)=PERT(I)
      NMK=N-1
      DO2M=1,NMK
      K=N-M
      DO 13 I=1,NN
      DO 26 J=1,NN
   26 H(I,J)=C(I,J,K+1)
   13 CONTINUE
      IF(NSTAT.NE.0) GO TO 83
       READ,MM,((Y(I,J),J=1,MM),I=1,NN)
       READ,((Z(I,J),J=1,MM),I=1,NN)
  83  CONTINUE
      CALL TRANS(Y,X,MM,NN)
      CALL DMMGG(X,10,H,10,MM,NN,NN,U,10,IER)
      DO 31 I=1,MM
      DO 62 J=1,MM
   62 F(I,J)=A(I,J,K)
   31 CONTINUE
      CALL DMMGG(F,10,U,10,MM,MM,NN,V,10,IER)
      CALL DMMGG(Z,10,V,10,NN,MM,NN,W,10,IER)
      DO 19 I=1,NN
      DO 38 J=1,NN
   38 DER(I,J,K)=W(I,J)
   19 CONTINUE
      CALL DMMGG(W,10,T,10,NN,NN,1,D,10,IER)
      DO 14 I=1,NN
      DD(I,K)=D(I)
   14 T(I)=D(I)
    2 CONTINUE
      PRINT 1010
      PRINT 1011
      CALL IMP(DER,N,4,NN)
      PRINT 1016
      CALL IMP(DER,N,5,NN)
      IF(NN.LT.3) GO TO 15
      PRINT 1017
      CALL IMP(DER,N,6,NN)
      IF(NN.LT.4) GO TO 15
```

```
      PRINT 1018
      CALL IMP(DER,N,7,NN)
   15 PRINT 900
      PRINT 1012
      PRINT 1013
      NMI=N-1
      DO 222 M=1,NMI
      K=N-M
      NK=M+1
      PRINT,NK,(DD(I,K),I=1,NN)
  222 CONTINUE
      PRINT 900
      GO TO 100
  499 FORMAT(11X,'K',17X,'C11',27X,'C12',27X,'C13')
  498 FORMAT(29X,'C22',27X,'C23',27X,'C33')
  497 FORMAT(29X,'C33',27X,'C34',27X,'C44')
  599 FORMAT(11X,'K',16X,'DC11',26X,'DC12',26X,'DC13')
  598 FORMAT(28X,'DC22',26X,'DC23',26X,'DC33')
  597 FORMAT(28X,'DC33',26X,'DC34',26X,'DC44')
  900 FORMAT('---------------------------------------------------------')
  999 FORMAT(' STATIONARY CASE')
 9999 FORMAT(' TIME-VARYING CASE')
 1001 FORMAT(' DOWNSTREAM PROPAGATION')
 1101 FORMAT(' EXAMPLE NO.',I2)
 1102 FORMAT(' DIMENSION OF THE STATE:',I2)
 1103 FORMAT(' DIMENSION OF THE CONTROL:',I2)
 1002 FORMAT(' NUMBER OF STAGES=',I2)
 1003 FORMAT(' B=',E10.4)
 1004 FORMAT(' DIAGONAL ELEMENTS OF COST MATRIX L:')
 1005 FORMAT(' OTHER ELEMENTS EQUAL ZERO')
 1006 FORMAT(' MATRIX Y:')
 1007 FORMAT(' MATRIX Z:')
 1008 FORMAT(' INITIAL PERTURBATION:')
 1009 FORMAT(' DC(K)=C(K-1)-C(K)')
 1010 FORMAT(' D(K)=MATRIX DERIVATIVE OF X(K+1) W. RESPECT TO X(K)')
 1011 FORMAT(11X,'K',26X,'D11',26X,'D12',26X,'D13',26X,'D14')
 1012 FORMAT(' DELTA(K)=PERTURBATION AT STAGE K')
 1013 FORMAT(11X,'K',24X,'DELTA1',24X,'DELTA2',24X,'DELTA3')
 1016 FORMAT(11X,'K',26X,'D21',26X,'D22',26X,'D23',26X,'D24')
 1017 FORMAT(11X,'K',26X,'D31',26X,'D32',26X,'D33',26X,'D34')
 1018 FORMAT(11X,'K',26X,'D41',26X,'D42',26X,'D43',26X,'D44')
 1019 FORMAT(' K=',I3)
 1020 FORMAT(' H(K)=MATRIX DERIVATIVE OF PHI(K) W.RESP. TO X(K)')
 1021 FORMAT(24X,'COLUMN 1',22X,'COLUMN 2',22X,'COLUMN 3')
 1111 CONTINUE
      STOP
      END
```

```
      SUBROUTINE TRANS(A,B,II,JJ)
      DOUBLE PRECISION A(10,10),B(10,10)
      DO 4 I=1,II
      DO 5 J=1,JJ
      B(I,J)=A(J,I)
    5 CONTINUE
    4 CONTINUE
      RETURN
      END

      SUBROUTINE IMP(A,NUMBER,L,NDIM)
      DOUBLE PRECISION A(10,10,30)
      IF(L.NE.1) GO TO 1
      IF(NDIM.NE.2) GO TO 22
      DO 10 K=1,NUMBER
      M=K+1
      NK=NUMBER-K+1
      PRINT,NK,A(1,1,M),A(1,2,M),A(2,2,M)
   10 CONTINUE
      RETURN
   22 DO 9 K=1,NUMBER
      M=K+1
      NK=NUMBER-K+1
      PRINT,NK,(A(1,J,M),J=1,NDIM)
    9 CONTINUE
      RETURN
  1      IF(L.NE.2) GO TO 2
      IF(NDIM.NE.3) GO TO 33
      DO 8 K=1,NUMBER
      M=K+1
      NK=NUMBER-K+1
      PRINT,NK,A(2,2,M),A(2,3,M),A(3,3,M)
    8 CONTINUE
      RETURN
   33 DO 7 K=1,NUMBER
      M=K+1
      NK=NUMBER-K+1
      PRINT,NK,(A(2,J,M),J=1,NDIM)
    7 CONTINUE
      RETURN
    2 IF(L.NE.3) GO TO 3
      DO 6 K=1,NUMBER
      M=K+1
      NK=NUMBER-K+1
      PRINT,NK,A(3,3,M),A(3,4,M),A(4,4,M)
    6 CONTINUE
      RETURN
    3  I=L-3
      NUM=NUMBER-1
      DO 101 K=1,NUM
      M=K+1
      NK=NUMBER-K+1
      PRINT,NK,(A(I,J,K),J=1,NDIM)
  101 CONTINUE
      RETURN
      END
```

```
      SUBROUTINE COST(L,MM)
       DOUBLE PRECISION L(10,10)
      DIMENSION NRAMP(10),NCLASS(10),LIGHT(10)
      DOUBLE PRECISION G(10),PHI(10),XLONG(10),XLANE(10)
      DOUBLE PRECISION PHIMAX(10),RHOMAX(10)
      DOUBLE PRECISION KA,FJ,A,B,R,S,U,V,W,GE,GA
      READ,NSTOP
      READ,(NRAMP(I),I=1,MM)
      READ, (NCLASS(I),I=1,MM)
      READ,(LIGHT(I),I=1,MM)
       IF(NSTOP.EQ.0) GO TO 222
      READ, (G(I),I=1,NSTOP)
  222 READ, (PHI(I),I=1,MM)
      READ,(XLONG(I),I=1,MM)
      READ, (XLANE(I),I=1,MM)
      DO 203 I=1,MM
      PHIMAX(I)=2000.*XLANE(I)
  203 RHOMAX(I)=225.*XLANE(I)
      DO 204 I=1,MM
      DO 205 J=1,MM
  205 L(I,J)=0
  204 CONTINUE
      DO 206 I=1,MM
      IF(XLANE(I).EQ.0) GO TO 227
      KA=(.3*RHOMAX(I)-PHIMAX(I)/55.)/(PHIMAX(I)**7.)
  206 L(I,I)=L(I,I)+(42.*KA)*(XLONG(I)*(PHI(I)**5.))
  227 CONTINUE
      DO 207 I=1,MM
      IF(NCLASS(I).EQ.1) GO TO 207
      IF(NCLASS(I).GT.2) GO TO 208
      KK=NRAMP(I)
      FJ=PHI(I)+PHI(KK)
           A=PHIMAX(I)
      B=PHIMAX(KK)
      X=A*(1-(FJ/B))
      R=(X**3.)*((X-PHI(I))**3.)
        R=((3.*X)*PHI(I)-3.*(X**2.)-(PHI(I)**2.))/R
      W=((2.*A)/B)*(PHI(I)**2.)
      W=W*R
      U=(2.*X)/((X-PHI(I))**3.)
      S=(4.*(X**2.)-(3.*PHI(I))*X+(PHI(I)**2.))/(X**2.)
      V=((A/B)*PHI(I))/((X-PHI(I))**3.)
      V=V*S
      L(I,I)=L(I,I)+U+2*V+W
      L(I,KK)=L(I,KK)+V+W
      L(KK,I)=L(KK,I)+V+W
      L(KK,KK)=L(KK,KK)+W
      GO TO 207
  208 KI=LIGHT(I)
      KK=MM-NSTOP+KI
      GE=G(KI)
      IF(NCLASS(I).GT.3) GO TO 209
      U=FUNCT1(GE,PHI(I),PHIMAX(I))
      V=FUNCT2(GE,PHI(I),PHIMAX(I))
      W=FUNCT3(GE,PHI(I),PHIMAX(I))
```

```
      L(KK,KK)=L(KK,KK)+U
      L(I,KK)=L(I,KK)+V
      L(KK,I)=L(KK,I)+V
      L(I,I)=L(I,I)+W
      GO TO 207
  209 IF(NCLASS(I).EQ.5) GO TO 210
      GA=1-GE
      U=FUNCT1(GA,PHI(I),PHIMAX(I))
      V=-FUNCT2(GA,PHI(I),PHIMAX(I))
      W=FUNCT3(GA,PHI(I),PHIMAX(I))
      L(KK,KK)=L(KK,KK)+U
      L(I,KK)=L(I,KK)+V
      L(KK,I)=L(KK,I)+V
      L(I,I)=L(I,I)+W
      GO TO 207
  210 IF(XLANE(I).EQ.0) GO TO 207
      GA=1-GE
      U=FUNCT1(GA,PHI(I),PHIMAX(I))
      L(I,I)=L(I,I)+U
  207 CONTINUE
      PRINT 2000
 2000 FORMAT(' COST MATRIX L:')
      DO 211 I=1,MM
  211 PRINT,(L(I,J),J=1,MM)
      RETURN
      END

      FUNCTION FUNCT1(GI,P,A)
      DOUBLE PRECISION GI,P,A
      FUNCT1=(A*GI)**2.+(A*GI-P)*((2.*A)*GI-P)
      FUNCT1=(2.*((A*P)**2.))*FUNCT1
      IF((A*GI-P).LT.0) GO TO 1
      FUNCT1=FUNCT1/(((A*GI)*(A*GI-P))**3.)
      GO TO 2
1     FUNCT1=FUNCT1/(((A*GI)*(P-A*GI))**3.)
      FUNCT1=-FUNCT1
2     CONTINUE
      RETURN
      END

      FUNCTION FUNCT2(GI,P,A)
      DOUBLE PRECISION GI,P,A
      X=A*GI
      FUNCT2=P**2.-(3.*P)*X+4.*(X**2.)
      FUNCT2=((A*P)*FUNCT2)/(X**2.)
      IF((P-X).LT.0) GO TO 1
       FUNCT2=FUNCT2/((P-X)**3.)
      GO TO 2
1     FUNCT2=-FUNCT2/((X-P)**3.)
2     CONTINUE
      RETURN
      END
```

```
      FUNCTION FUNCT3(GI,P,A)
      DOUBLE PRECISION GI,P,A
      IF(((A*GI)-P).LT.0) GO TO 1
      FUNCT3=((A*GI)-P)**3.
      GO TO 2
    1 FUNCT3=-((P-(A*GI))**3.)
    2 CONTINUE
      FUNCT3=1./FUNCT3
      FUNCT3=(2.*A)*(GI*FUNCT3)
      RETURN
      END

$ENTRY
```

APPENDIX G

REPORT OF INVENTIONS

Although there are no inventions, the work is an advance in that a full sensitivity analysis has been performed on the response of the steady-state (in a freeway corridor network) traffic assignment to external changes including variations in incoming traffic and accidents.

Copies

☆ U. S. GOVERNMENT PRINTING OFFICE: 1977--701-777--206

U. S. DEPARTMENT OF TRANSPORTATION
TRANSPORTATION SYSTEMS CENTER
KENDALL SQUARE, CAMBRIDGE, MA. 02142

OFFICIAL BUSINESS
PENALTY FOR PRIVATE USE, $300

U.S. MAIL

POSTAGE AND FEES PAID
U. S. DEPARTMENT OF TRANSPORTATION
518