AN APPLICATION OF SYSTEM THEORY TO THE OPTIMAL CONTROL OF VEHICULAR TRAFFIC NETWORKS

Thesis for the Degree of Ph. D.

MICHIGAN STATE UNIVERSITY

JEFFREY L. GOODNUFF

1967

[redacted]

This is to certify that the

thesis entitled

An Application of System Theory
to the Optimal Control
of Vehicular Traffic Networks

presented by

Jeffrey L. Goodnuff

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in E.E.

Herman E. Koenig
Major professor

Date May 12, 1967

O-169

ABSTRACT

AN APPLICATION OF SYSTEM THEORY TO THE OPTIMAL CONTROL OF VEHICULAR TRAFFIC NETWORKS

by Jeffrey L. Goodnuff

Vehicular traffic demands are increasing so rapidly that merely increasing the physical size of freeway and street systems is not, in itself, a solution. Present and future traffic networks must be operated at or near their highest efficiency levels. This can only be accomplished through control.

Recognizing the inevitable need for control, this thesis investigates the problem of applying physical system theory to the analysis and control of vehicular traffic systems.

Two complementary variable functions of time, traffic density, $x(t)$, and traffic flow rate, $y(t)$, are defined and used to characterize the dynamics of several traffic system components in the form of mathematical state models. These state models are combined, using the logically consistent procedures of system theory, into state models of traffic systems. As an example of the application of these state models to control, the special case of vehicular traffic control in a high density mode is considered. A state model of such a system is developed and a near time optimal control strategy is derived for a surface street grid of arbitrary size.

AN APPLICATION OF SYSTEM THEORY TO THE OPTIMAL CONTROL OF VEHICULAR TRAFFIC NETWORKS

By

Jeffrey Lynn Goodnuff

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Electrical Engineering

1967

ACKNOWLEDGEMENTS

The author wishes to express his thanks to his major professor, Dr. H. E. Koenig, whose encouragement and constructive suggestions were invaluable. Thanks are also due to Dr. J B. Kreer, whose many thought provoking discussions with the author proved to be most helpful.

The author also wishes to thank the Division of Engineering Research and the Crouse Hinds Company, under whose sponsorship this research was done.

TABLE OF CONTENTS

LIST OF FIGURES

I

INTRODUCTION

Present day vehicular traffic demands are increasing so rapidly that merely increasing the size of physical street and freeway systems is not, in itself, sufficient. If new and existing traffic networks are to operate at their highest efficiency they must be controlled in the same technical sense that the operations of aircraft, powerplants, and other physical systems are controlled. It is this inevitable need for control which motivates the research described in this dissertation.

Before any system can be controlled its dynamics must be characterized by a mathematically tractable state model. Such a model must do more than extrapolate the future from the past. It must go beyond simulation. It must reflect both the mathematical characteristics of each subsystem, or component part, and the interconnection pattern of these subassemblies.

Many of the basic concepts used to characterize physical phenomena can be applied to develop models of the dynamics of vehicular traffic networks. However, quantitative aspects of systems, such as these, in which man is involved are not easily represented. There are a number of reasons for this difficulty.

- Systems involving man as a component part are not as predictable as non-human systems. Certainly no general model characterizing the dynamics of man as a system component exist.

- Systems with men as integral parts are, in general, not easily subjected to experiment. For moral, legal, and political reasons it is sometimes very difficult to perform experiments on these man-machine systems.

- There are no universally agreed upon methods of quantitatively representing non-physical phenomena. For this reason communication between investigators in similar but distinct fields is often difficult and sometimes impossible. Investigators attempting to model non-physical systems are, in a sense, in the same position as the early physical scientists--their first task is to define a set of standard measurements which will characterize the phenomena of interest.

Even in the face of these difficulties considerable progress, at least in the form of qualitative or semiquantitative descriptions and hypothetical models, has been made [GR 1, GR 2, HE 1].

This thesis does not claim to have found an easy answer to the problems of quantitatively describing and controlling non-physical systems. It does, however, offer a logically consistent formulation procedure which, when properly applied to certain types of vehicular traffic systems, results in state models useful in describing, simulating, and controlling these systems.

In order that this thesis may serve as a basis for control, it examines the vehicular traffic phenomena and defines two quantitative measurements, traffic density and traffic flow rate, which are used to characterize vehicular traffic systems. These complementary measurements are used to model some of the basic components of traffic systems. The resulting component models are combined, using the logically consistent procedures of system theory, into a state model of the system.

As an example of the application of mathematical models of this form to the control of traffic systems, the special case of optimal control of vehicular traffic in a high density mode is considered in Chapter III. A state model of such a system is developed and a control scheme which results in a minimum number of control intervals is derived.

II

A CHARACTERIZATION OF VEHICULAR TRAFFIC

A complete analysis of any metropolitan traffic system involves at least three major aspects.

First, there must be an identification or analysis of traffic flow demands, as a function of time, between identified geographic regions as generated by the business, industrial, and other sociological activity of the populace. These flow demands depend on such parameters as the general economic level of a region, the distribution or mix of business and personal property in the region, and many other socio-economic factors. Furthermore, the flow demands are not independent of the traffic system, but are, at least in part, generated by it. This problem of flow demand identification and prediction is generally classified under the heading of origin-destination studies, and receives considerable attention in the literature [CA 1].

The second major aspect of metropolitan traffic analysis is that of determining the distribution of known inter-regional flow demands over alternate paths and modes of transportation. This so called "mode split" problem is not only concerned with predicting the distribution of flow demand over the various transportation modes, but also with predicting the distribution of vehicular flow demand over alternate routes.

The third aspect of vehicular traffic analysis is that of determining the dynamics of flow streams (i.e., the stream flow rates, densities, and delays) from the known inter-regional demands.

While all three aspects of the problem are inter related, the complexity of the problem and the present state of the art in traffic network analysis almost precludes the inclusion of this interdependence in a mathematical analysis or simulation of traffic systems. This chapter, therefore somewhat arbitrarily, assumes that the traffic flow rates over identified inter-regional paths of a traffic network are known as a function of time. It considers the problem of developing a state model characterizing the flow dynamics as an explicit function of these flows, the network structure, and intersection controls.

The following postulates must be satisfied before the methods of physical systems analysis may be applied [KO 1]. Therefore vehicular traffic systems must also satisfy these postulates.

(1) The system must be an interconnection of identifiable subsystems, or components. If this identification is not physically possible, it must be at least conceptually possible.

(2) These subsystems, or components, must be discrete in the sense that they must have a finite number of interfaces with other components (called terminals), and measurements taken at these terminals must completely characterize the component.

(3) All components must be such that they may be characterized by a pair of complementary variables, x satisfying postulate (5), and y satisfying postulate (6) following.

(4) All components with N terminals must be such that they may be characterized with N-1 measurements made at the N terminals.

(5) The variable x must be such that the algebraic sum of all x measurements around a circuit vanish.

(6) The variable y must be such that the algebraic sum of all y measurements corresponding to a cutset vanish.

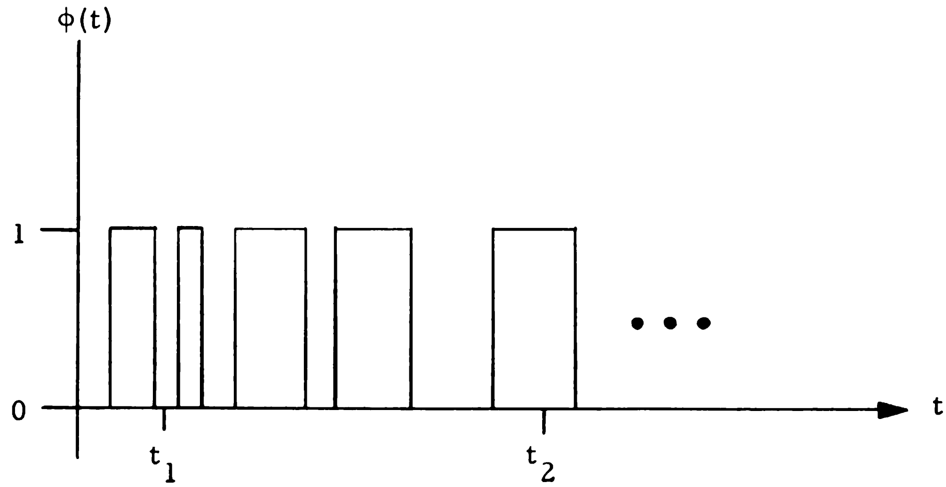
2.1 Fundamental Variables and Parameters

It is convenient to define a pair of complementary variable functions of time, $x(t)$ representing traffic density and $y(t)$ representing the flow rate of the traffic system, which can be used to characterize the traffic phenomena. The units of these variables are chosen as:

x - total vehicle length/total lane length,

y - number of vehicles/minute.

In order to operationally define these variables, consider the output of a standard loop or overhead sonic type detector, $\phi(t)$, shown in Figure 2.1.



2.1 Typical traffic detector output

Let $\phi(t) = 1$ when a vehicle is on the detector and $\phi(t) = 0$ otherwise. Define the integer valued functional $J(\phi(t), t_1, t_2)$ as

$$J(\phi(t), t_1, t_2) = \text{total number of 0 to 1 changes of } \phi(t) \text{ for } t_1 \leq t \leq t_2$$

The flow rate, y, is obtained as

$$y = \frac{J(\phi(t), t_1, t_2)}{(t_2 - t_1)} \tag{2.1}$$

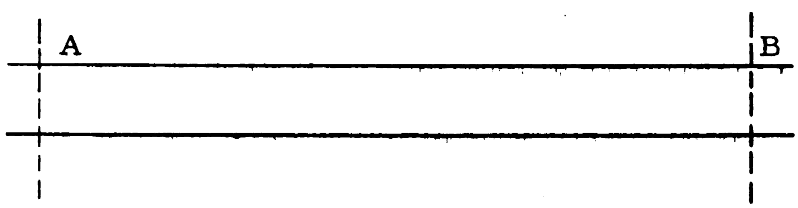
while the traffic density, x, is

$$x = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} \phi(t)\, dt \tag{2.2}$$

From (2.2) it is evident that x is the ratio of two lengths of time and assumes values between zero and one.

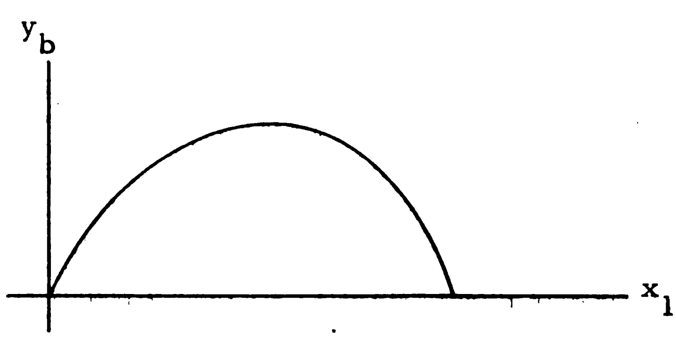
The above definition of traffic density is essentially the same as lane occupancy. If the traffic mix (i.e., average vehicle length) is known, other characteristics such as average stream speed may be computed from these flow and density measurements.

Consider now a uniform section of street as shown in Figure 2.2, and let y_a and y_b represent the flow rates at the points A and B respectively. Furthermore, let x_1 represent the average



2.2 Uniform section of street

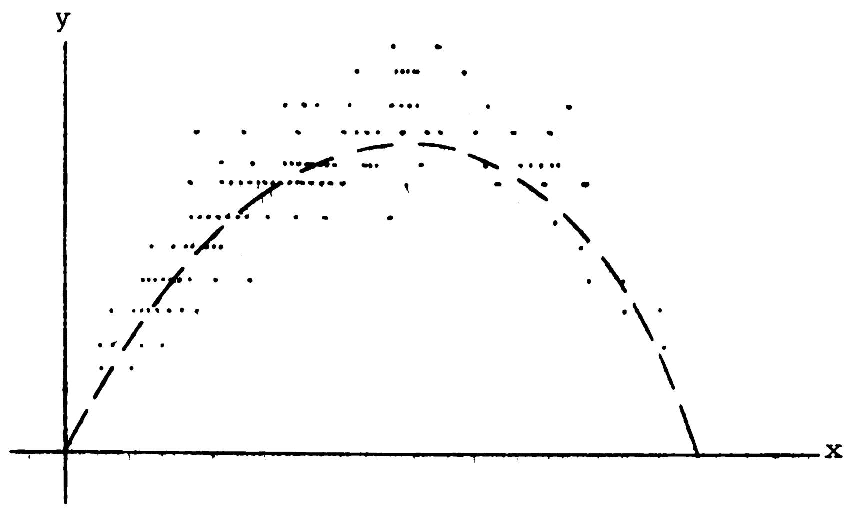
density in that street. In the special case when all densities and flow rates are constant (i.e., $y_a = y_b$ = constant, and x_1 = constant) it is well established that a plot of flow rate vs density is of the form shown in Figure 2.3 [GR 1]. It must be emphasized that the



2.3 Steady state flow density plot

relation shown in Figure 2.3 is valid only under steady-state conditions.

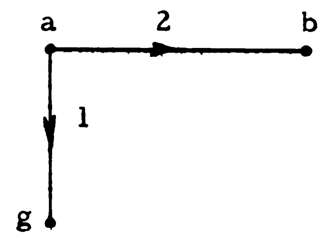
Actual measurements of density-flow rate relationships have, in the past, been taken with very little attention given to the steady state requirement. Such measurements typically result in a clustering of points similar to that shown in Figure 2.4 [JL 1].



2.4 Experimental data from the John C. Lodge freeway

In order to characterize the dynamics of a traffic stream it is clearly necessary to obtain density-flow rate relationships under known dynamic conditions. Until this is done there can be very little correlation between the theoretical analysis of traffic systems and their actual behavior.

For the purpose of applying system theory to traffic analysis, assume that the flow density relationships for the short section of street of Figure 2.2 may be approximated by the terminal equations and graph (2.3). It is important to note that the density, $x_2(t)$, is



$$y_2(t) = f(x_1, x_2)$$
$$y_1(t) = c_1 \frac{dx_1(t)}{dt} \qquad (2.3)$$

a differential density and not an absolute density. The densities $x_1(t)$ and $x_2(t)$ are defined as

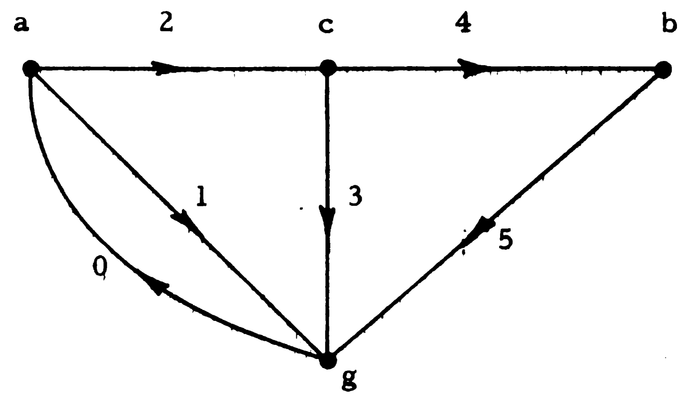
$$x_2(t) = x_a(t) - x_b(t)$$
$$x_1(t) = x_a(t) \qquad (2.4)$$

where $x_a(t)$ and $x_b(t)$ are the absolute densities at points a and b.

The exact form of the function $f(x_1, x_2)$ can not be determined until traffic stream data taken under known dynamic conditions is available. Even though this information must be known before there can be any correlation between the theoretical analysis of traffic systems and their actual behavior, the exact form of $f(x_1, x_2)$ need not be known to demonstrate the application of system theory to vehicular traffic systems. To this end the next section identifies the model of one of the basic subassemblies in vehicular traffic systems.

2.2 A Single Origin-Destination

Consider an origin and destination connected by a single expressway link with no intersections. Assuming the input flow rate, $y_o(t)$, is known, and considering the expressway as a cascaded system of two uniform sections, the system graph is as shown in Figure 2.5.



2.5 System graph for two cascaded sections

The terminal equations corresponding to the system graph are

$$
\begin{aligned}
y_o(t) &\ - \ \text{known} \\
y_i(t) &= c_i \frac{dx_1(t)}{dt} \qquad i = 1, 3 \\
y_j(t) &= f_j(x_{j-1}, x_j) \qquad j = 2, 4 \\
x_5(t) &\ - \ \text{known}
\end{aligned}
\tag{2.5}
$$

The vertices a and b correspond to the ends of the expressway while the vertex c represents the center. If the expressway is uniform throughout its length, the two sections are identical. In addition to y_o, the terminal density, x_5, must be specified.

From the component equations (2.5) and the circuit and cut-set equations of the system graph, the system model is easily derived as [KO 1, WI 1]

$$
\frac{d}{dt}\begin{bmatrix} x_1(t) \\ x_3(t) \end{bmatrix} = \begin{bmatrix} 1/c_1 & 0 \\ 0 & 1/c_3 \end{bmatrix} \begin{bmatrix} y_o(t) - f_2(x_1, x_3 - x_1) \\ f_2(x_1, x_3 - x_1) - f_4(x_3, x_5 - x_3) \end{bmatrix} \tag{2.6}
$$

In the event the expressway is uniform and f_2 and f_4 are given as

$$f_2(\beta, \gamma) = f_4(\beta, \gamma) = (1 - k\gamma)\, g(\beta) \tag{2.7}$$

where the function g is similar in form to Figure 2.3, the state model (2.6) becomes

$$\frac{d}{dt}\begin{bmatrix} x_1(t) \\ \dot{x}_3(t) \end{bmatrix} = \begin{bmatrix} 1/c_1 & 0 \\ 0 & 1/c_3 \end{bmatrix}\begin{bmatrix} y_o(t)\,[1 - k(x_3 - x_1)]\, g(x_1) \\ [1 - k(x_3 - x_1)]\, g(x_1) - [1 - k(x_5 - x_3)\, g(x_3)] \end{bmatrix} \tag{2.8}$$

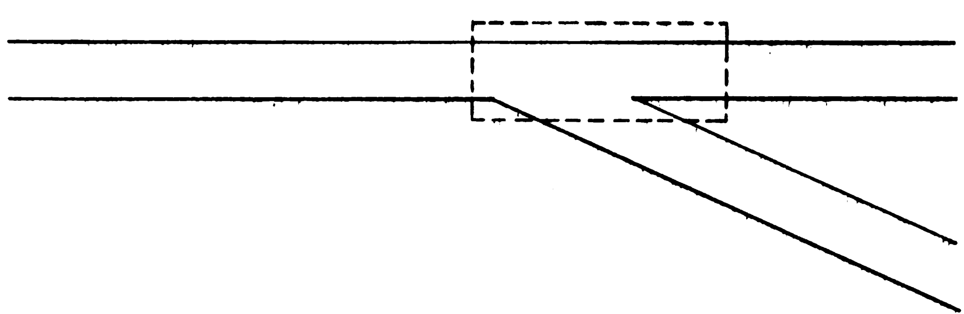
Suppose $x_1(t) = x_3(t) = 0$. These steady state conditions imply

$$y_o(t) = g(x_1) - k\,g(x_1)(x_3 - x_1) = g(x_3) - k\,g(x_3)(x_5 - x_3) \tag{2.9}$$

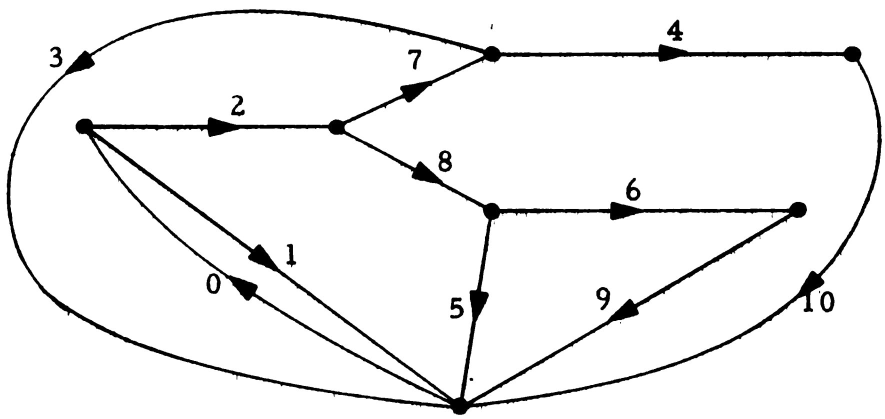
One solution to (2.9), provided the boundary condition $y_o = g(x_5)$ is satisfied, is $x_1(t) = x_3(t) = x_5(t)$. This particular steady state solution corresponds to the situation where the flow rate throughout the entire expressway is constant.

2.3 Multiple Destinations

The single origin, dual destination system shown in Figure 2.6 is composed of four components; three uniform expressway



2.6 Single origin, dual destination system links (characterized in section 2.2) and an exit ramp. The system graph for such a configuration is shown in Figure 2.7. Furthermore, it is assumed that the turn ratio for the exit is known (i.e., the ratio $k_{78} = y_7/y_8$ is known).



2.7 Single origin, dual destination system graph

Let the exit ramp be modeled as a "perfect coupler" with characteristics of the form

$$\begin{bmatrix} y_7 \\ x_8 \end{bmatrix} = \begin{bmatrix} 0 & k_{78} \\ -k_{78} & 0 \end{bmatrix} \begin{bmatrix} x_7 \\ y_8 \end{bmatrix} \tag{2.10}$$

Edges 1 through 6 are used to characterize the three uniform expressway links and have equations of the form

$$\begin{aligned} y_i(t) &= c_i \frac{dx_i(t)}{dt} & i &= 1, 3, 5 \\ y_j(t) &= f_j(x_{j-1}, x_j) & j &= 2, 4, 6 \end{aligned} \tag{2.11}$$

The input flow rate, y_o, and output densities, x_9 and x_{10}, are assumed known.

For a general application the system must be modeled as a four-terminal component. Such a model is easily developed from the component equations, (2.10) and (2.11), and the circuit and cut-set equations of the system graph (Figure 2.7). The result is:

$$\frac{d}{dt}\begin{bmatrix} x_1(t) \\ x_3(t) \\ x_5(t) \end{bmatrix} = \begin{bmatrix} 1/c_1[y_o - f_2(x_1, x_1-x_5-\gamma(x_3-x_5))] \\ 1/c_3[\gamma f_2(x_1, x_1-x_5-\gamma(x_3-x_5)) - f_4(x_3, x_3-x_{10})] \\ 1/c_5[(1-\gamma)f_2(x_1, x_1-x_5-\gamma(x_3-x_5)) - f_6(x_5, x_5-x_9)] \end{bmatrix} \tag{2.12a}$$

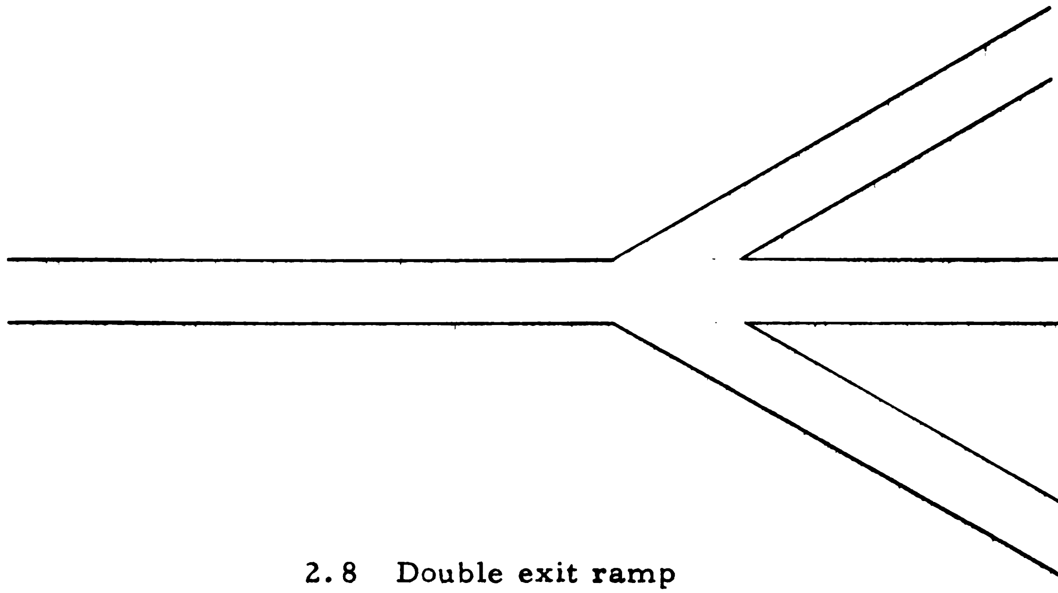
where $\gamma = 1/(1 + k_{78})$ and the terminal variables (identified in the system graph of Figure 2.7) are

$$\begin{bmatrix} x_o(t) \\ y_9(t) \\ y_{10}(t) \end{bmatrix} = \begin{bmatrix} -x_1(t) \\ -f_6(x_5, x_5-x_9) \\ -f_4(x_3, x_3-x_{10}) \end{bmatrix} \tag{2.12b}$$

2.4 The Eight-Way Intersection

When each of the four streets at an intersection carry traffic in opposite directions, the dynamics of all eight streams are "coupled" through interference patterns. Even in the case of a light controlled grade intersection, the oncoming stream interferes with the left turning stream. If the intersection is in the form of a clover leaf with grade separation then, of course, there are no inter-stream interferences and the intersection can be modeled as a combination of four double exit ramps of the form shown in Figure 2.8.

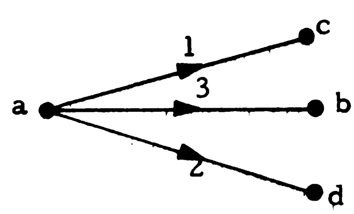
This double exit ramp is a simple generalization of the single exit



2.8 Double exit ramp

ramp of section 2.3. It is modeled in the "perfect coupler" form

$$\begin{bmatrix} y_1 \\ y_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & k_1 \\ 0 & 0 & k_2 \\ -k_1 & -k_2 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ y_3 \end{bmatrix} \qquad (2.13)$$



where constants k_1 and k_2 represent the ratios of left turns to through traffic, and right turns to through traffic respectively. These equations characterize the intersection in terms of the traffic composition. The cloverleaf model is formulated by simply joining four of these double exits.

Although complicated because of the number of equations, the procedure for formulating a model of the intersection from the component models and the constraint equations of the system graph is straight-forward and yields a set of terminal equations for a cloverleaf of the form:

$$\begin{bmatrix} y_1' \\ y_2' \\ y_3' \\ y_4' \\ x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \left[\begin{array}{c|c} 0 & -A^t \\ \hline A & 0 \end{array}\right] \begin{bmatrix} x_1' \\ x_2' \\ x_3' \\ x_4' \\ y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} \tag{2.14a}$$

where

$$A = \begin{bmatrix} \frac{1}{1+k_1+k_2} & 0 & 0 & 0 \\ 0 & \frac{1}{1+k_3+k_4} & 0 & 0 \\ 0 & 0 & \frac{1}{1+k_5+k_6} & 0 \\ 0 & 0 & 0 & \frac{1}{1+k_7+k_8} \end{bmatrix} \begin{bmatrix} 0 & -k_2 & -1 & -k_1 \\ -k_3 & 0 & -k_4 & -1 \\ -1 & -k_5 & 0 & -k_6 \\ -k_8 & -1 & k_7 & 0 \end{bmatrix} \tag{2.14b}$$

The eight constants k_i, $i = 1, 2, \ldots, 8$ are the traffic composition constants which must be known in order to characterize the intersection. It should be noted that the terminal equations retain the skew symmetric form characteristic of "perfect coupler" type components.

At this point it is useful to consider the problem of traffic stream interference in more detail, in particular, concerning the interference to left turn traffic generated by the oncoming traffic stream.

Even if traffic flow rates are known in advance, the particular distribution of vehicles (hence the opportunity to cross a traffic stream) is a random phenomenon. It is therefore necessary to employ statistics to compute the interference characteristics of a traffic stream.

Consider the case where the traffic stream is a single lane and the gaps between vehicles is a random variable, g. Let the probability density function (p.d.f.) for g be denoted by $f_g(g)$. It is known that there exists a gap acceptance probability associated with each car-driver gap combination [HE 1]. Let this probability be denoted by $\alpha(t)$, where

$$\begin{aligned} \alpha(t) &= 0 & t \leq T \\ &= 1 - e^{-\lambda(t-T)} & t > T \end{aligned} \tag{2.15}$$

The conditional probability, $\alpha(t)$, may be interpreted as the probability of turning left through the oncoming stream between t and t + dt seconds (i.e., $\alpha(t) = P(\text{left turn} \mid t \leq g \leq t + dt)$.

Knowing the p.d.f. for the random variable g implies

$$\begin{aligned} P(\text{turning left}) &= \int_{\text{all } t} P(\text{turning left} \mid t \leq g \leq t + dt)\, P(t \leq g \leq t + dt) \\ &= \int_{-\infty}^{\infty} \alpha(t)\, f_g(t)\, dt \end{aligned} \tag{2.16}$$

For the case of Poisson traffic

$$f_g(t) = \sigma e^{-\sigma t} \qquad t > 0$$
$$= 0 \qquad t \leq 0 \tag{2.17}$$

and (2.16) becomes

$$P(\text{turning left}) = \int_{-\infty}^{\infty} \sigma(1 - e^{-\lambda(t-T)}) e^{-\sigma t} dt$$

$$= \frac{\lambda e^{-\sigma T}}{\lambda + \sigma} \tag{2.18}$$

where σ^{-1} is the average gap length between vehicles.

In a similar manner it is easily shown that the probability of completing a left turn across N lanes of traffic is given by

$$P(\text{turning left}) = \frac{\lambda \exp\{-\sum_{\lambda=1}^{N} \sigma_i T\}}{\lambda + \sum_{i=1}^{N} \sigma_i} \tag{2.19}$$

The four grade corner intersection differs from the clover-leaf described above in that the left-turn traffic experiences "interference" passing through the on-coming traffic in the adjacent lane. The total out of the left turn exit of an intersection is, therefore, decreased by an amount that depends on the flow-rate in the adjacent on-coming traffic stream. When the model of the grade separation intersection given in (2.14) is altered to include this interference function, the model takes the form

$$\begin{bmatrix} y_1' \\ y_2' \\ y_3' \\ y_4' \end{bmatrix} = \begin{bmatrix} & & \\ & -A^t & \\ & & \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} - \begin{bmatrix} f_1(y_4, y_2) \\ f_2(y_1, y_3) \\ f_3(y_2, y_4) \\ f_4(y_3, y_1) \end{bmatrix} \tag{2.20a}$$

and

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} & & \\ & A & \\ & & \end{bmatrix} \begin{bmatrix} x_1' \\ x_2' \\ x_3' \\ x_4' \end{bmatrix} + \begin{bmatrix} h_1(y_3, x_4') \\ h_2(y_4, x_1') \\ h_3(y_1, x_2') \\ h_4(y_2, x_3') \end{bmatrix} ; \tag{2.20b}$$

The functions f_i and h_i are defined as

$$f_i = b_{i,i+1}(1 - \tau_{i+3})\, y_{i+1}$$

and (2.21)

$$h_i = a_{i-1,i}(1 - \tau_{i+2})\, x_{i+3} \qquad i = 1, 2, 3, 4$$

where b_{ij} and a_{ij} represent the entries of $-A^t$ and A, respectively, and all subscripts are modulo 4. The factor τ_j is a function of the flow y_j, is called the "highway transparency" and is a monotonic decreasing function of y_j having the following properties

(1) $\tau_j(0) = 1$

(2) $0 \leq \tau_j(y_j) \leq 1$ for all y_j (2.22)

It is to be noted that when the cross-traffic flow is zero (i.e., $\tau = 1$) or when there is no left-turn flow, (2.20) reduces to the interference-free case of the cloverleaf given in (2.14). This is to be expected since there can be no "interference" in the case when there is no cross traffic or in the case where there is no traffic with which to interfere. The transparency function, τ, may be taken as

$$\tau = 1 - P(\text{turning left}) \tag{2.33}$$

or, from (2.19)

$$\tau = \frac{\lambda\left(1 + \exp\left[-\sum_{i=1}^{N} \sigma_i T\right]\right) + \sum_{i=1}^{N} \sigma_i}{\lambda + \sum_{i=1}^{N} \sigma_i} \tag{2.34}$$

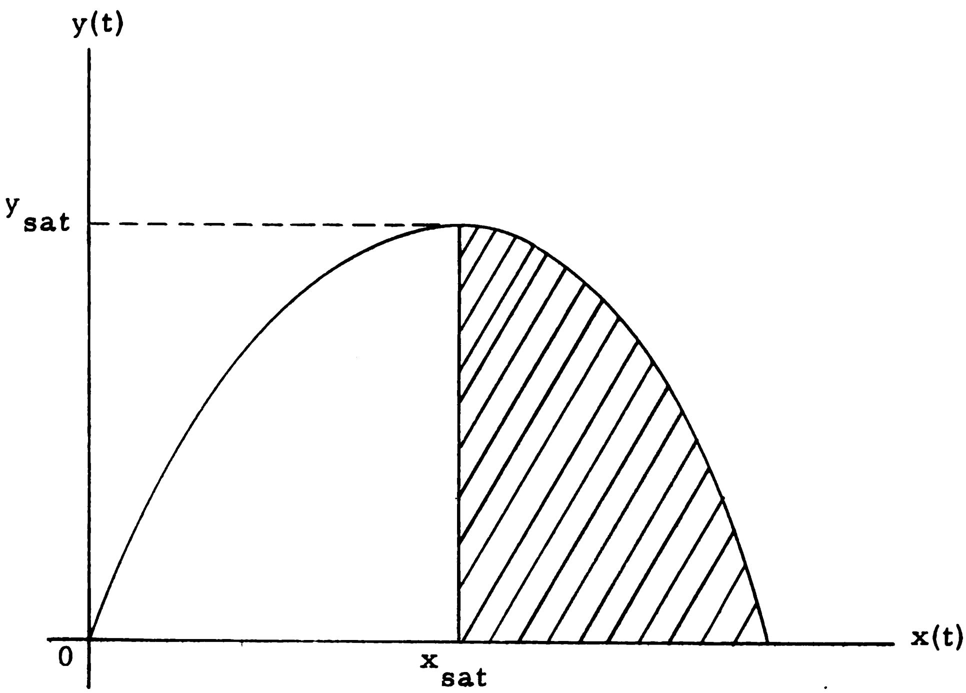
This chapter does not claim that the component and state models introduced above are in a complete or final form. As pointed out in section 2.1, considerable experimental research must be completed before a high degree of correlation between theoretical and actual results can be obtained. This chapter does, however, point the direction for such research and illustrates the potentials of a system theoretic approach.

Using the fundamental variables and methods defined in this chapter, the particular problem of control of vehicular traffic in a high density mode is discussed in Chapter III.

III

OPTIMAL CONTROL IN HIGH DENSITY MODES

It is the objective of this chapter to characterize the vehicular traffic phenomenon with a mathematically tractable model to which some of the procedures of optimal control theory may be applied. Consider the general character of vehicular traffic as characterized by traffic density, x(t), and traffic flow rate, y(t). When one plots traffic flow, y(t), vs traffic density, x(t), curves of the general form shown in Figure 3.1 result [GR 1].



3.1 Flow-density characteristics

It should be noted that the curve plotted in Figure 3.1 is a plot of flow rate and absolute density, not differential density as previously defined. Curves of this form are obtained under free flow steady-state conditions, i.e., the time derivative of density is zero, and there is no interruption of the traffic stream. Again referring to Figure 3.1 observe that as the absolute density increases, the steady-state flow rate also increases. This relationship is nearly linear until y(t) reaches about 80% of the maximum possible flow rate, y_{sat}. At this point the flow rate, y(t), begins to level off until, at $y(t) = y_{sat}$, the slope, $\frac{dy}{dx}$, is zero. At this point any further increase in traffic density clearly decreases the flow rate. Furthermore, it is well known that without external control, recovery from a staurated condition takes a very long time [MA 1]. It is therefore logical to try, by some means of control, to keep the traffic density below this saturation level, x_{sat}.

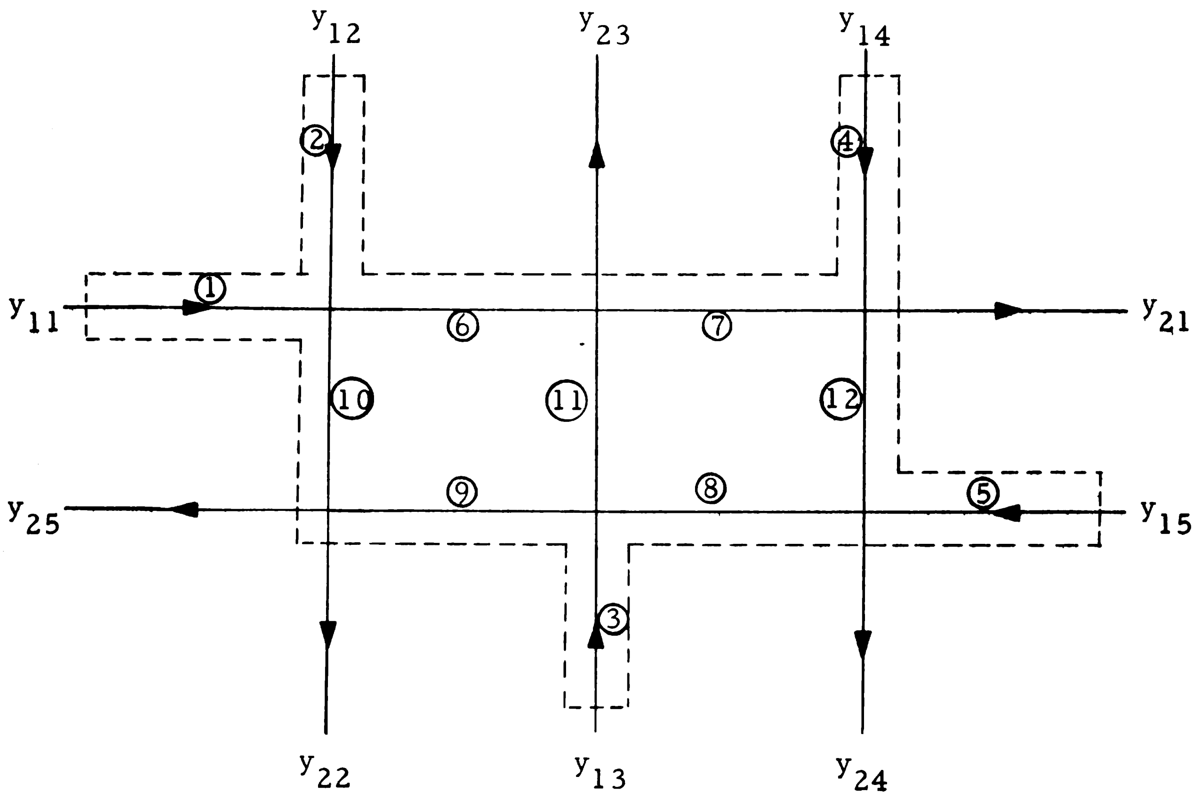
Many attempts have been made in the past to implement control schemes with just this objective in mind [GR 2, MA 2]. In general these control schemes have improved existing traffic flow conditions during peak travel hours on the order of 20%. They are very attractive for particular bottlenecks which cover a small area, such as a tunnel or short section of limited access expressway, as they are easy to implement, realize substantial improvements in a short time, and are very inexpensive when compared to the only other alternative--addition of more road surface.

These volume-control techniques, however, have been limited to a few distinct locations, where conditions are unusually bad. No attempt has been made to implement a general control scheme which could be applied to an arbitrary area.

With this thought in mind let the vehicular traffic flow be divided into three distinct modes; (1) high density mode, (2) low density mode, and (3) transition or medium density mode. The particular mode of a given street is determined by monitoring queue lengths. When the length of the queue is such that each vehicle must wait for one or more complete traffic signal cycles before passing the intersection (i.e., the queue length does not go to zero during any given cycle), that section of street is assumed to be operating in the high density mode. When traffic is free flowing (i.e., each vehicle must wait no longer than one traffic signal cycle), the system will be controlled under the low density mode. In any other situation the system is said to be in the medium density or transition mode. Inasmuch as the greatest cost-benefit returns are potentially realizable at high densities of operation, this thesis places primary emphasis on the high density mode of operation.

3.1 The System Model

Consider the problem of controlling an m by n grid of intersecting surface streets under these conditions. Such a grid is shown schematically in Figure 3.2. The area under control is



3.2 m by n rectangular grid

shown within the dotted boundary, all streets are assumed to be one way, and the small circled numbers on the diagram serve to identify each "stub." It is further assumed that the input flow rates, $y_{1i}(t)$, are known as a function of time. Since the streets are short (i.e., the street transit time is small with respect to the rates of change of the traffic flow rate and density), and the high density conditions exist (i.e., queues are always present), the transit time for each street stub may be neglected and a control scheme based only on queue lengths. Furthermore, since the high density traffic control mode is designed to operate with streets in a saturated condition, it is extremely useful in recovering from catastrophic disturbances.

For the m by n rectangular grid of Figure 3.2, the flow rates at the i th grid input and output are designated as y_{1i} and y_{2i} respectively. The density (or in this case, queue length) in the i th street will be denoted as x_i.

Let the components, x_i, of the state vector, $\vec{X}$, represent the densities at the respective streets in the grid. The state of the grid, as a function of time, and the inputs to the grid can be expressed as a vector difference equation of the form:

$$\vec{X}(k) = F[\vec{X}(k-1), \vec{Y}(k-1)] \tag{3.1}$$

where the vector $\vec{X}(k)$ represents the state vector at the k th control interval, which is of duration T(k).

If the grid is to be controlled, rather than just modeled and simulated, a set of control variables must be identified. Suppose $u_i(k)$ signifies the number of vehicles to be removed from street section i during the k th control interval. Then the detailed form of (3.1) is:

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \\ x_3(k+1) \\ x_4(k+1) \\ x_5(k+1) \\ x_6(k+1) \\ x_7(k+1) \\ x_8(k+1) \\ x_9(k+1) \\ x_{10}(k+1) \\ x_{11}(k+1) \\ x_{12}(k+1) \end{bmatrix} = \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \\ x_4(k) \\ x_5(k) \\ x_6(k) \\ x_7(k) \\ x_8(k) \\ x_9(k) \\ x_{10}(k) \\ x_{11}(k) \\ x_{12}(k) \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{vmatrix} y_{11}(k) \\ y_{12}(k) \\ y_{13}(k) \\ y_{14}(k) \\ y_{15}(k) \end{vmatrix} + \begin{bmatrix} -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ a_1 & b_2 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & a_6 & -1 & 0 & 0 & 0 & b_{11} & 0 \\ 0 & 0 & 0 & 0 & a_5 & 0 & 0 & -1 & 0 & 0 & 0 & b_{12} \\ 0 & 0 & b_3 & 0 & 0 & 0 & 0 & a_8 & -1 & 0 & 0 & 0 \\ b_1 & a_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & a_3 & 0 & 0 & 0 & 0 & b_8 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & a_4 & 0 & 0 & b_7 & 0 & 0 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} u_1(k) \\ u_2(k) \\ u_3(k) \\ u_4(k) \\ u_5(k) \\ u_6(k) \\ u_7(k) \\ u_8(k) \\ u_9(k) \\ u_{10}(k) \\ u_{11}(k) \\ u_{12}(k) \end{bmatrix}$$

(3.2a)

where the output flows are given by:

$$\begin{bmatrix} y_{21}(k+1) \\ y_{22}(k+1) \\ y_{23}(k+1) \\ y_{24}(k+1) \\ y_{25}(k+1) \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & b_4 & 0 & 0 & a_7 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & b_9 & a_{10} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & b_6 & 0 & 0 & 0 & 0 & a_{11} & 0 \\ 0 & 0 & 0 & 0 & b_5 & 0 & 0 & 0 & 0 & 0 & 0 & a_{12} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} u_1(k) \\ u_2(k) \\ u_3(k) \\ u_4(k) \\ \vdots \\ u_{12}(k) \end{bmatrix} \tag{3.2b}$$

and the constants, b_i, are the percent of vehicles in street i which turn.

$$a_i = 1 - b_i \tag{3.2c}$$

When the grid of Figure 3.2 is extended to an arbitrary m x n rectangular grid the state model is of order N = 2(m + n + mn + 1) and of the form:

$$\vec{X}(k+1) = \vec{X}(k) + \begin{bmatrix} U \\ 0 \end{bmatrix} \vec{Y}(k) + \begin{bmatrix} -U & 0 \\ A & B \end{bmatrix} \begin{bmatrix} \vec{\gamma}_1(k) \\ \vec{\gamma}_2(k) \end{bmatrix} \tag{3.3}$$

Indeed, for any complete m by n rectangular lattice there clearly are m(n+1) + n(m+1) = m + n + 2mn edges. For a grid of the form shown in Figure 3.2 there exist, in addition to the closed lattice, m + 1 + n + 1 edges. Thus N = m + n + 2mn + m + 1 + n + 1 = 2(m + n + mn + 1). The N dimensional state model will always be of the form of (3.3) if the stub numbering system is chosen so that the m + n + 2 external stubs are numbered first.

Consider now the general N dimensional discrete state model

$$\vec{X}(k+1) = C\,\vec{X}(k) + D\,\vec{\gamma}(k), \quad k = 0, 1, 2, \ldots \tag{3.4}$$

Under the assumption that the matrix D is N by N, recursive solution of (3.4) gives

$$\vec{X}(k) = C^{k}\,\vec{X}(0) + [H_1, \; D] \begin{bmatrix} \vec{\Gamma}_1 \\ \vec{\Gamma}_2 \end{bmatrix} . \tag{3.5}$$

where the submatrix, H_1, and the subvectors, $\vec{\Gamma}_1$, and $\vec{\Gamma}_2$ are

$$H_1 = [C^{k-1} D, \ldots, D] \; ; \quad \vec{\Gamma}_1 = \begin{bmatrix} \vec{\gamma}(0) \\ \vec{\gamma}(1) \\ \vdots \\ \vec{\gamma}(k-1) \end{bmatrix} ; \quad \vec{\Gamma}_2 = \vec{\gamma}(k) \tag{3.6}$$

Assuming D^{-1} exists we can solve (3.5) for $\vec{\Gamma}_2$ and obtain

$$\vec{\Gamma}_2 = D^{-1}[\vec{X}(k) - C^{k}\vec{X}(0) - H_1\vec{\Gamma}_1] . \tag{3.7}$$

<u>Theorem 3.1</u>: A discrete state model in the form shown in (3.3) is controllable if and only if B^{-1} exists.

<u>Proof</u>: From the definition of controllability [GO 1] it is clear that any discrete state model of the form (3.4) is controllable if and only if it can be solved for $\vec{\Gamma}_2$. Equation (3.7) implies an explicit solution exists for $\vec{\Gamma}_2$ if and only if D^{-1} exists. Observe that the form of D in (3.3) is:

$$D = \begin{bmatrix} -U & 0 \\ A & B \end{bmatrix}$$

Using the fact that the determinant of a product is equal to the product of the determinant yields

$$|D| = \left| \begin{bmatrix} -U & 0 \\ A & B \end{bmatrix} \right| = \left| \begin{bmatrix} U & 0 \\ -A & U \end{bmatrix} \begin{bmatrix} -U & 0 \\ 0 & B \end{bmatrix} \right| = \begin{vmatrix} -U & 0 \\ 0 & B \end{vmatrix} = (-1)^{m+n} |B| \tag{3.8}$$

The strategy used as a basis of control, is to select the control vectors, $\vec{\gamma}_1(k)$ and $\vec{\gamma}_2(k)$, in (3.3) in such a manner that the state is driven to the origin in minimum time subject to certain constraints imposed on the variables in the model by physical considerations. These constraints are considered next.

First of all, since each section of street is of finite length, the state variables must satisfy the contraint

$$0 \leq x_i(k) \leq X_i \text{ for all } k, \text{ and } i, \; i = 1, 2, \ldots, N \tag{3.9}$$

where the constants X_i represent the maximum storage capacity of the i th section of street. Secondly, since more vehicles cannot be removed from a street than exist in that section the controls must be constrained such that

$$0 \leq u_i(k) \leq x_i(k) \text{ for all } k \text{ and } i, \quad i = 1, 2, \ldots, N \tag{3.10}$$

Finally, each intersection can transmit only a finite number of vehicles in a given time, therefore

$$M_i \leq u_i(k) \leq \Gamma_i \text{ for all k and i, } i = 1, 2, \ldots, N \qquad (3.11)$$

The constants M_i and Γ_i, are the minimum and maximum number of vehicles which are permitted to cross through the ith intersection in time T(k). Psychological factors will dictate the values of M and Γ for each intersection. These values, together with such traffic stream parameters as vehicular acceleration time, will in turn determine the time, T(k), associated with the kth control interval. It is very important to note that each control interval is not necessarily of the same duration as the other intervals. As is shown later, each control interval is complete when Γ_i vehicles have been counted through the ith intersection, $i = 1, 2, \ldots, N$.

3.2 The Control Problem

Stated as an optimal control problem, the high density vehicular traffic control problem becomes:

For the system in (3.3) find the control, $\vec{u}(k)$, from the admissable set Ω,

$$\Omega = \{u_i(k) \mid 0 \leq u_i(k) \leq x_i(k) : \text{ for } i = 1, \ldots, N;\ k = 1, 2, \ldots, \ell\} \cap$$
$$\{u_j(k) \mid M_j \leq u_j(k) \leq \Gamma_j : \text{ for } j = 1, \ldots, N;\ k = 1, 2, \ldots, \ell\}.$$

such that:

(1) $\vec{X}(\ell) = 0$ for minimum ℓ, and

(2) $0 \leq x_i(k) \leq X_i$ for all k, and i, $i = 1, 2, \ldots, N$.

At this point it will be useful to examine the Pontryagin maximum principle for discrete systems [PO 1, HA 1, HA 2, HO 1]. Consider systems characterized by difference equations of the following form

$$\vec{X}(k+1) - \vec{X}(k) = \vec{F}[\vec{X}(k), \vec{u}(k)], \ k = 1, 2, \ldots, \ell$$

where the state vector is an element $\vec{X}$ of a Euclidean space E^N, and the control vector is an element $\vec{u}$ of a Euclidean space E^r. A subset $\Omega \subset E^r$, called the admissable set, is specified and all control vectors are required to be members of Ω. For every k, $k = 1, 2, \ldots, \ell$ the vector valued function $\vec{F}[\vec{X}, \vec{u}]$ is assumed to satisfy the following conditions:

(1) $\vec{F}[\vec{X}, \vec{u}]$ is defined for all $(\vec{X}, \vec{u}) \in E^N \times \Omega$,

(2) for every $u \in \Omega$, $\vec{F}[\vec{X}, \vec{u}]$ is twice continuously differentiable with respect to $\vec{X}$,

(3) $\vec{F}[\vec{X}, \vec{u}]$ and all its first and second partial derivatives with respect to $\vec{X}$ are uniformly bounded over $A \times \Omega$ for any bounded set A, $A \subset E^N$,

(4) the matrix $U + \partial \vec{F}[\vec{X}, \vec{u}] / \partial \vec{X}$ is non-singular on $E^N \times \Omega$,

(5) the set $\{\vec{F}[\vec{X}, \vec{u}] \mid \vec{u} \in \Omega\}$ is convex for all $\vec{X} \in E^N$. *

---

*In general condition (5) may be considerably relaxed, but in this context it is not restrictive [HO 1].

It is also necessary to define an initial set

$$\{\vec{X} \mid \alpha_i(\vec{X}) = 0,\ i = 1, 2, \ldots s,\ s \leq N\},$$

a terminal set

$$\{\vec{X} \mid \beta_i(\vec{X}) = 0,\ i = 1, 2, \ldots, t,\ t \leq N\},$$

And an object function $f_o(\vec{X})$. The scalar functions $\alpha_i(\vec{X})$, $\beta_i(\vec{X})$ must be twice continuously differentiable.

The sequences $\vec{\underline{u}}(1)$, $\vec{\underline{u}}(2)$, ..., $\vec{\underline{u}}(\ell)$ and $\vec{\underline{X}}(1)$, $\vec{\underline{X}}(2)$, ..., $\vec{\underline{X}}(\ell)$ are said to be optimal controls and trajectories respectively if they satisfy the conditions:

(1) $\alpha_i(\vec{\underline{X}}(1)) = 0$ for $i = 1, 2, \ldots, s$,

(2) $\vec{\underline{X}}(k+1) - \vec{\underline{X}}(k) = \vec{F}(\vec{\underline{X}}(k), \vec{\underline{u}}(k))$ for all $k = 1, 2, \ldots, \ell$,

(3) $\vec{\underline{u}}(k) \in \Omega$ for all $k = 1, 2, \ldots, \ell$,

(4) $\beta_i(\vec{\underline{X}}(\ell)) = 0$ for $i = 1, 2, \ldots, t$,

and if the functional

$$J = \sum_{k=1}^{\ell} f_o(\vec{X}(k))$$

attains its minimum value, for $\vec{X}(k) = \vec{\underline{X}}(k)$, subject to these constraints.

If the function f_o is unity, the control which minimizes the functional J is that which satisfies constraint (4) with minimum ℓ.

In order to state the Pontryagin maximum principle it is necessary to augment the N state equations with the scalar equation

$$x_o(k+1) - x_o(k) = f_o(\vec{X}(k)),\ k = 1, 2, \ldots, \ell$$

which results in the N+1 equations

$$x_i(k+1) - x_i(k) = f_i(\vec{X}(k), \vec{u}(k))\ k = 1, 2, \ldots, \ell,\ i = 0, 1, \ldots, N$$

The Pontryagin maximum principle for discrete systems states:

If the sequences $\vec{\underline{u}}(1), \vec{\underline{u}}(2), \ldots, \vec{\underline{u}}(\ell)$ and $\vec{\underline{X}}(1), \vec{\underline{X}}(2), \ldots, \vec{\underline{X}}(\ell)$ are optimal then there exists a sequence of non zero vectors $\vec{P}(1), \vec{P}(2), \ldots, \vec{P}(\ell)$ such that:

$$(1)\quad \sum_{j=0}^{N} f_j[\vec{\underline{X}}(k), \vec{\underline{u}}(k)]\ p_j(k+1) \geq \sum_{j=0}^{N} f_j[\vec{\underline{X}}(k), \vec{u}(k)]\ p_j(k+1),$$

for all $k = 1, 2, \ldots, \ell$ and all $u \in \Omega$,

$$(2)\quad \vec{P}(k+1) - \vec{P}(k) = -\left(\frac{\partial}{\partial \vec{X}} \vec{F}[\vec{X}, \vec{\underline{u}}]\Big|_{\vec{X}=\vec{\underline{X}}(k)}\right)^t \vec{P}(k+1),$$

for all $k = 1, 2, \ldots, \ell$.

Condition (1) is, of course, the maximization of Hamiltonian, while the adjoint equations are represented by (2). For a proof of the Maximum Principle in this form see [HA 2].

3.3 Solution of the Optimal Control Problem

The traffic control problem as formulated in Section 3.2 is in a form to which the discrete maximum principle may be applied. Let the state model for the traffic control system of (3.3) be augmented to include the object function f_o. Further, let f_o be unity, then the state model for the minimum interval case becomes:

$$\begin{bmatrix} x_o(k+1) \\ \vec{x}_1(k+1) \\ \vec{x}_2(k+1) \end{bmatrix} - \begin{bmatrix} x_o(k) \\ \vec{x}_1(k) \\ \vec{x}_2(k) \end{bmatrix} = \begin{bmatrix} 1 \\ \vec{\beta} \\ 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ -U & 0 \\ A & B \end{bmatrix} \begin{bmatrix} \gamma_o(k) \\ \vec{\gamma}_1(k) \\ \vec{\gamma}_2(k) \end{bmatrix} \tag{3.12}$$

If the grid is n by m let $N = 2(m + n + mn + 1)$ and $\sigma = m + n + 2$, then the state model in (3.12) is $N + 1$ dimensional. The vectors $\vec{x}_1$, $\vec{\beta}$, and $\vec{\gamma}_1$ have σ components while $\vec{x}_2$ and $\vec{\gamma}_2$ have $N - \sigma$ components. The Hamiltonian is

$$H(\vec{P}, \vec{X}, \vec{u}) = \sum_{j=0}^{N} p_j f_j = p_o + \vec{\beta}\bullet\vec{P}_1 - \vec{\gamma}_1\bullet\vec{P}_1 + (A\vec{\gamma}_1 + B\vec{\gamma}_2)\bullet\vec{P}_2 \tag{3.13}$$

where the vector $\vec{P}_1$ has σ components and $\vec{P}_2$ has $N - \sigma$ components. The adjoint equations are

$$p_i(k+1) - p_i(k) = -\sum_{j=0}^{N} \frac{\partial f_j(x,\gamma)}{\partial x_i} p_i(k+1), \quad i = 0, 1, \ldots, N; \tag{3.14a}$$
$$k = 1, 2, \ldots, \ell .$$

Clearly, since $f_j(\vec{X}, \vec{u})$ is independent of the state variable $\vec{X}$, (3.14a) becomes

$$\begin{bmatrix} p_o(k+1) \\ \vec{P}_1(k+1) \\ \vec{P}_2(k+1) \end{bmatrix} - \begin{bmatrix} p_o(k) \\ \vec{P}_1(k) \\ \vec{P}_2(k) \end{bmatrix} = 0 . \tag{3.14b}$$

The co-state variables, p_i, corresponding to time optimal control are all constants.

It follows that (3.13) can be written as

$$H(\vec{P}, \vec{X}, \vec{u}) = p_o + \vec{\beta}\bullet\vec{P}_1 + \vec{D}^t\vec{\gamma}_1 + \vec{E}^t\vec{\gamma}_2 \tag{3.15}$$

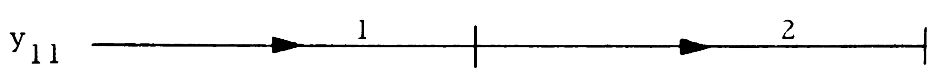
where $\vec{D}^t = (\vec{P}_2{}^t A - \vec{P}_1{}^t)$ and $\vec{E}^t = \vec{P}_2{}^t B$. $H(\vec{P}, \vec{X}, \vec{u})$ is a maximum when the i th components of $\vec{\gamma}_1$ and $\vec{\gamma}_2$ are $\max[\operatorname{sign}(d_i)\gamma_i]$ and $\max[\operatorname{sign}(e_i)\gamma_i]$, respectively. This is simply the familiar result that all time optimal controls for linear systems are on the boundary of the admissable set Ω[PO 1]. This fact allows optimal trajectories for a number of traffic systems to be sketched.

3.4 Cascaded Signals

Let $\sigma = 1$ and $N = 2$. Equation (3.3) then reduces to the two scalar equations

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \end{bmatrix} - \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ a & b \end{bmatrix} \begin{bmatrix} \gamma_1(k) \\ \gamma_2(k) \end{bmatrix} + \begin{bmatrix} y_{11} \\ 0 \end{bmatrix} \tag{3.16}$$

where a is the percent of γ_1 which enters street 2, while b = -1. This corresponds to an array of streets of the form shown in Figure 3.3, that is, a series connection of two streets.



3.3 Series connection of two streets

The admissable set Ω is

$$\Omega = \{\gamma_i \mid 0 \leq \gamma_i \leq \min(\Gamma_i, x_i), \text{ for } i = 1, 2\} \tag{3.17}$$

Consider now the optimal trajectories resulting from the control $\gamma_1 = \Gamma_1$ and $\gamma_2 = \Gamma_2$. Equation (3.16) becomes

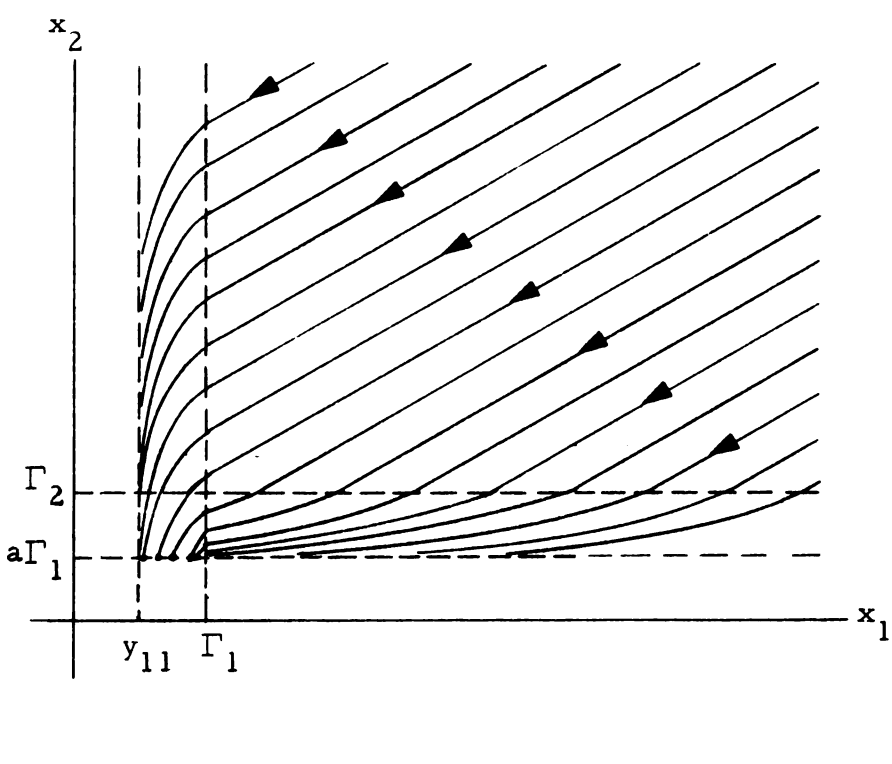
$$x_1(k) = x_1(0) - k(\Gamma_1 - y_{11}) \tag{3.18a}$$

$$x_2(k) = x_2(0) - ka\Gamma_1 - k\Gamma_2 \tag{3.18b}$$

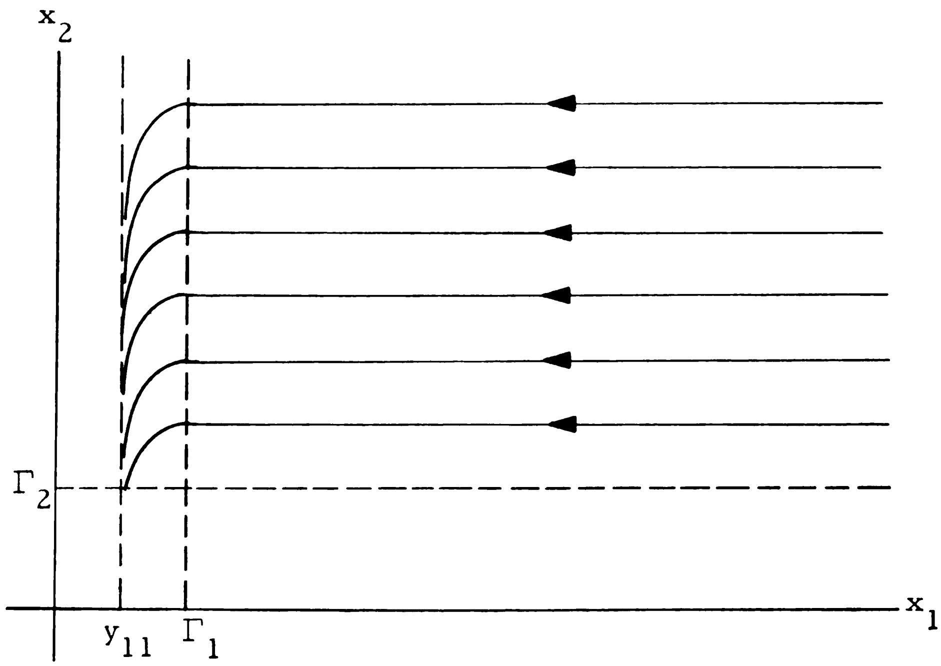
Substitution of (3.18a) into (3.18b) yields (provided $\Gamma_1 - y_{11} \neq 0$):

$$x_2(k) = \frac{(\Gamma_2 - a\Gamma_1)}{(\Gamma_2 - y_{11})} x_1(k) - x_2(0) - \frac{(\Gamma_2 - a\Gamma_1)}{(\Gamma_1 - y_{11})} x_1(0) \tag{3.19}$$

Under the assumption that y_{11} is constant, the phase plane plot is as shown in Figures 3.4a and 3.4b.



3.4a Phase plane for $\Gamma_2 > a\Gamma_1$

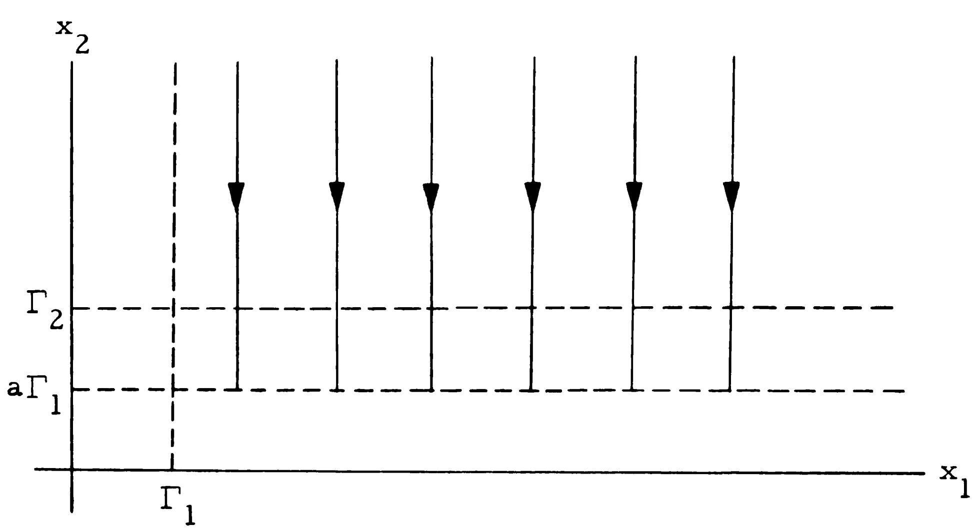


3.4b Phase plane for $\Gamma_2 = a\Gamma_1$

When $y_{11} = \Gamma_1$ (3.18a) reduces to:

$$x_1(k) = x_1(0) = \text{const.} \tag{3.20}$$

Therefore, for this case the phase plane plot takes the form shown in Figure 3.4c.



3.4c Phase plane for $\Gamma_1 = y_{11}$

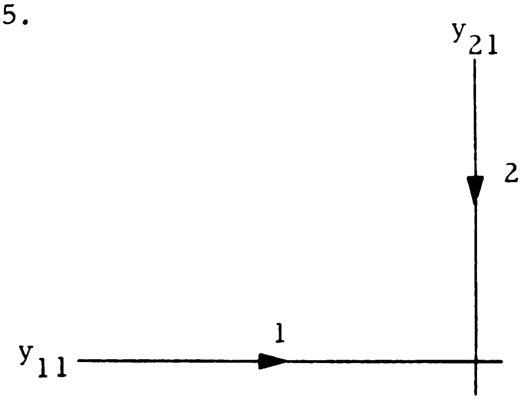
The arrows on the optimal trajectories indicate the direction of increasing k (time).

Clearly when the input demand to the system is so great that $y_{11} > \Gamma_1$ there exists no control, γ_1, in the admissable set Ω, which will reduce the initial value of x_1. For this reason, and because queue lengths are not restricted outside of the control region, y_{11} will always be selected so that $(\Gamma_1 - y_{11}) \geq 0$. Similarly, if $\Gamma_2 < a\Gamma_1$ there exists no admissable control, γ_2, which will reduce the initial value of x_2. Furthermore, in the event $\Gamma_2 < a\Gamma_1$ the maximum capacity of the system is determined by Γ_2. Therefore, in the event $\Gamma_2 < a\Gamma_1$, the value of Γ_1 will be decreased so that $\Gamma_2 \geq a\Gamma_1$. Under these assumptions an optimal control may always be selected which will monotonically decrease any initial state toward the origin.

Referring to Figures 3.4 note that as the optimal trajectories cross the lines $x_1(k) = \Gamma_1$ or $x_2(k) = \Gamma_2$, their slope is no longer constant. This is a consequence of the fact that the admissable set, Ω, is a function of the state of the system. Furthermore, if the state of the system is such that, for at least some i, $x_i \leq \Gamma_i$, the i th queue may be dissipated in one light cycle. If this is the case the density obviously is not high enough for transit time to be neglected (a basic assumption for the high density mode of operation). Notice, however, that application of the proper optimal control monotonically decreases any initial state to the point (Γ_1, Γ_2), at which time the control mode will be transferred to one of low or medium density.

3.5 The Single Intersection

Consider the case of the single intersection as shown in Figure 3.5.



3.5 Single intersection

The state model for this arrangement of streets is:

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \end{bmatrix} - \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix} = \begin{bmatrix} y_{11} \\ y_{12} \end{bmatrix} + \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \gamma_1(k) \\ \gamma_2(k) \end{bmatrix} \tag{3.21}$$

Solving (3.21) recursively yields:

$$x_1(k) = k(y_{11} - \gamma_1) + x_1(0) \tag{3.22a}$$

$$x_2(k) = k(y_{12} - \gamma_2) + x_2(0) \tag{3.22b}$$

In the event that, for some i, $y_{1i} \geq \Gamma_i$, it is clear that no admissable control exists which will decrease the initial state of x_i.

For this reason y_{1i} is restricted so that

$$y_{1i} < \Gamma_i \quad \text{for all } i \tag{3.23}$$

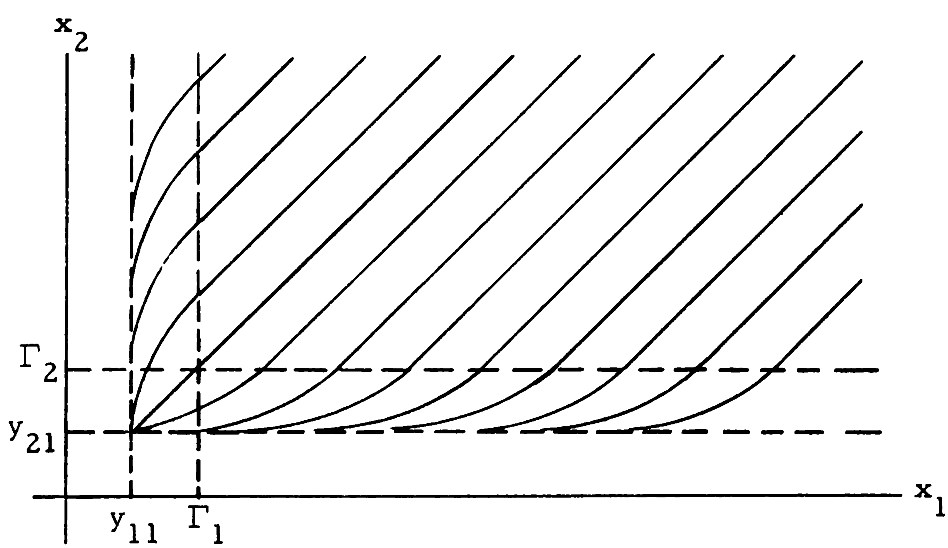
Substituting (3.22a) into (3.22b) yields, under the control $(\gamma_1, \gamma_2) = (\Gamma_1, \Gamma_2)$,

$$x_2(k) = \alpha x_1(k) + x_2(0) - \alpha x_1(0) \tag{3.24}$$

where α, the optimal trajectory slope, is given as:

$$\alpha = \frac{(\Gamma_2 - y_{12})}{(\Gamma_1 - y_{11})} \tag{3.25}$$

The phase plane plot of the optimal trajectories is shown in figure 3.6.

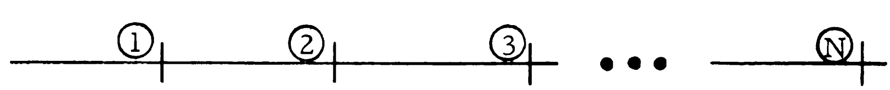


3.6 Single intersection optimal trajectories

In this case, as in the case of two series connected streets, the optimal trajectories are straight lines only for $x_1 \geq \Gamma_1$ and $x_2 \geq \Gamma_2$. Notice that as the phase plane point, (x_1, x_2) approaches the point (y_{11}, y_{12}) the optimal trajectories asymtotically approach the lines $x_1 = y_{11}$ and $x_2 = y_{12}$. These trajectories are clearly all monotonic decreasing, and all pass through the point (y_{11}, y_{12}). Furthermore, note that the point (y_{11}, y_{12}) is always "inside the point (Γ_1, Γ_2), that is $y_{11} < \Gamma_1$ and $y_{12} < \Gamma_2$. This assures that an admissable optimal control always exists which will drive the system, in a monotonic decreasing manner, to a point where the control may be changed to a low or medium density mode.

3.6 Arterials of Arbitrary Length

Consider now the case of a long arterial. That is, assume that the area to be put under control consists of a series connection of N one way street sections, as shown in figure 3.7. Let y_{11} be the input to the end of the arterial, and assume that there is no input to the arterial except at that end. All of the results of the following development can easily be extended to the case where known inputs are allowed at points other than at the end, however, for notational convenience only the single input case will be considered.



3.7 Single arterial

Equation (3.3) for this street arrangement becomes

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \\ x_3(k+1) \\ \vdots \\ x_N(k+1) \end{bmatrix} - \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \\ \vdots \\ x_N(k) \end{bmatrix} = \begin{bmatrix} y_{11} \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} -1 & 0 & 0 & \dots & 0 & 0 \\ a_2 & -1 & 0 & \dots & 0 & 0 \\ 0 & a_3 & -1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & & 0 & \dots & a_N & -1 \end{bmatrix} \begin{bmatrix} \gamma_1(k) \\ \gamma_2(k) \\ \gamma_3(k) \\ \vdots \\ \gamma_N(k) \end{bmatrix} \tag{3.26}$$

Notice that the control matrix is triangular (in fact bidiagonal) with determinant $(-1)^N$.

Recall from the previous analysis of two series connected streets that, although the optimal <u>trajectory</u> is a function of the input, y_{11}, the optimal <u>control</u> is independent of any input. Input levels are only bounded by the requirement that the optimal trajectory be monotonic decreasing. With this in mind, consider the N section series connected street.

Using a technique very similar to dynamic programming [DR 1] each subsection of the arterial will be independently optimized starting at the "output" end (i.e., at $i = N$) and working back until bounds on the input, y_{11}, are obtained.

Since, by the definition of the control region, there is assumed to be an infinite sink at the end of the arterial

$$\gamma_N = \min[\Gamma_N, x_N]. \tag{3.27}$$

Furthermore, the state variable, x_{N-1}, is monotonically decreasing when γ_{N-1} satisfies

$$(\gamma_N - a_N\gamma_{N-1}) \geq 0. \tag{3.28}$$

Equation (3.28) together with the fact that γ_{N-1} must come from the admissable set, Ω, requires

$$\gamma_{N-1} = \min[\min(\Gamma_{N-1}, x_{N-1}), \gamma_N/a_N] \tag{3.29}$$

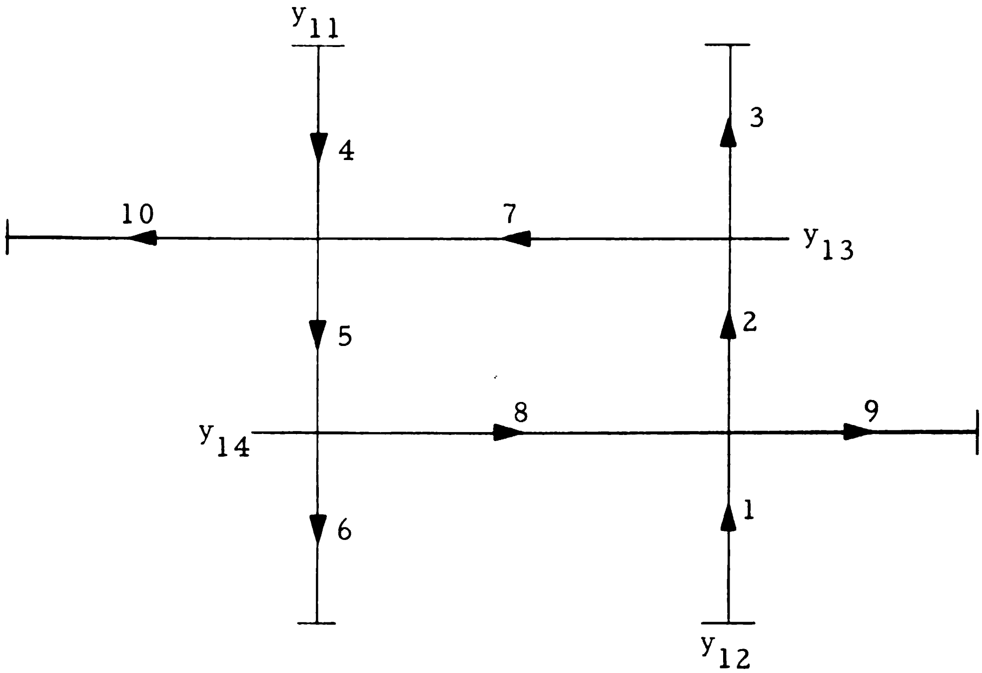
Applying the above argument to the ith section of street yields

$$\gamma_i = \min[\min(\Gamma_i, x_i), \gamma_{i+1}/a_{i+1}] \text{ for } i = 1, 2, \ldots, N-1 \tag{3.30}$$

where γ_N is specified by (3.27). An optimal control recursively derived by (3.27) and (3.30) is clearly one which results in monotonically decreasing optimal trajectories and satisfies the necessary conditions of the Pontryagin Maximum Principle.

3.7 Rectangular Grids

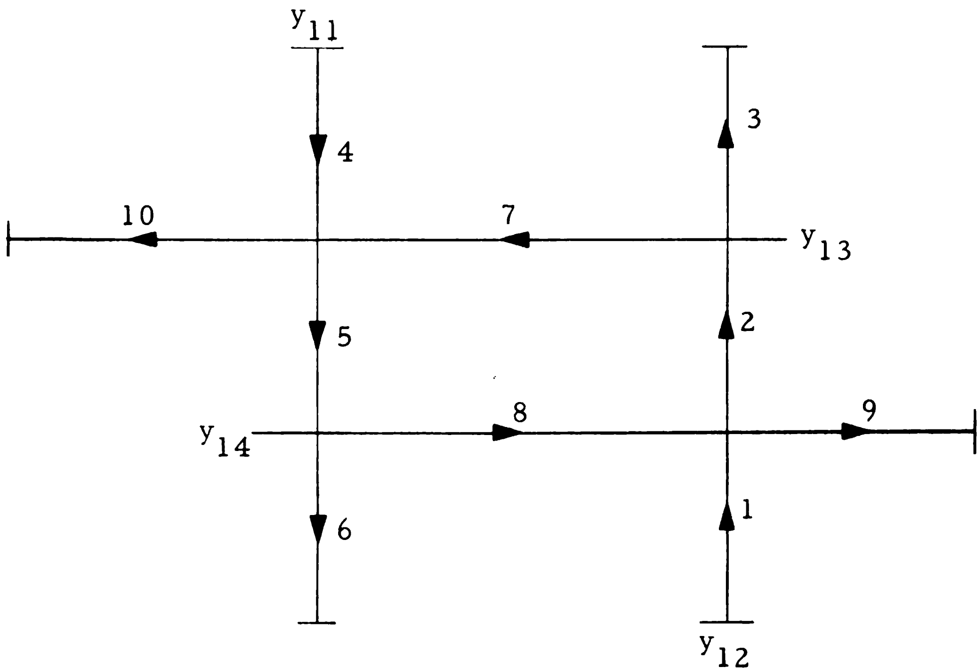
Consider now the most interesting, and probably most useful, grid configuration, a simple grid of intersecting arterials. An example of such a grid is shown schematically in Figure 3.8.



3.8 Schematic diagram of arbitrary grid

The detailed form of the state model for the system of figure 3.7 is

$$
\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \\ x_3(k+1) \\ x_4(k+1) \\ x_5(k+1) \\ x_6(k+1) \\ x_7(k+1) \\ x_8(k+1) \\ x_9(k+1) \\ x_{10}(k+1) \end{bmatrix}
-
\begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \\ x_4(k) \\ x_5(k) \\ x_6(k) \\ x_7(k) \\ x_8(k) \\ x_9(k) \\ x_{10}(k) \end{bmatrix}
=
\begin{bmatrix} y_{12} \\ 0 \\ b_{13}y_{13} \\ y_{11} \\ 0 \\ b_{14}y_{14} \\ a_{13}y_{13} \\ a_{14}y_{14} \\ 0 \\ 0 \end{bmatrix}
+
\begin{bmatrix}
-1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
a_1 & -1 & 0 & 0 & 0 & 0 & 0 & b_8 & 0 & 0 \\
0 & a_2 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & a_4 & -1 & 0 & b_7 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & a_5 & -1 & 0 & 0 & 0 & 0 \\
0 & b_2 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & b_5 & 0 & 0 & -1 & 0 & 0 \\
b_1 & 0 & 0 & 0 & 0 & 0 & 0 & a_8 & -1 & 0 \\
0 & 0 & 0 & b_4 & 0 & 0 & a_7 & 0 & 0 & -1
\end{bmatrix}
\begin{bmatrix} \gamma_1(k) \\ \gamma_2(k) \\ \gamma_3(k) \\ \gamma_4(k) \\ \gamma_5(k) \\ \gamma_6(k) \\ \gamma_7(k) \\ \gamma_8(k) \\ \gamma_9(k) \\ \gamma_{10}(k) \end{bmatrix}
\tag{3.31}
$$



3.8 Schematic diagram of arbitrary grid

The detailed form of the state model for the system of figure 3.7 is

$$
\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \\ x_3(k+1) \\ x_4(k+1) \\ x_5(k+1) \\ x_6(k+1) \\ x_7(k+1) \\ x_8(k+1) \\ x_9(k+1) \\ x_{10}(k+1) \end{bmatrix}
-
\begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \\ x_4(k) \\ x_5(k) \\ x_6(k) \\ x_7(k) \\ x_8(k) \\ x_9(k) \\ x_{10}(k) \end{bmatrix}
=
\begin{bmatrix} y_{12} \\ 0 \\ b_{13}y_{13} \\ y_{11} \\ 0 \\ b_{14}y_{14} \\ a_{13}y_{13} \\ a_{14}y_{14} \\ 0 \\ 0 \end{bmatrix}
+
\begin{bmatrix}
-1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
a_1 & -1 & 0 & 0 & 0 & 0 & 0 & b_8 & 0 & 0 \\
0 & a_2 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & a_4 & -1 & 0 & b_7 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & a_5 & -1 & 0 & 0 & 0 & 0 \\
0 & b_2 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & b_5 & 0 & 0 & -1 & 0 & 0 \\
b_1 & 0 & 0 & 0 & 0 & 0 & 0 & a_8 & -1 & 0 \\
0 & 0 & 0 & b_4 & 0 & 0 & a_7 & 0 & 0 & -1
\end{bmatrix}
\begin{bmatrix} \gamma_1(k) \\ \gamma_2(k) \\ \gamma_3(k) \\ \gamma_4(k) \\ \gamma_5(k) \\ \gamma_6(k) \\ \gamma_7(k) \\ \gamma_8(k) \\ \gamma_9(k) \\ \gamma_{10}(k) \end{bmatrix}
\tag{3.31}
$$

An argument very similar to the one used on the long arterial implies

$$\gamma_3 = \min(\Gamma_3, x_3) \tag{3.32}$$

In this case the condition that x_2 be monotonic decreasing is shown as

$$(\gamma_3 - a_2\gamma_2 - b_{13}y_{13}) \geq 0 \tag{3.33}$$

The inequality (3.33) forces γ_2 to be

$$\gamma_2 = \min[\min(\Gamma_2, x_2), (\gamma_3 - b_{13}y_{13})/a_2] \tag{3.34}$$

In a similar manner

$$\gamma_1 = \min[\min(\Gamma_1, x_1), (\gamma_2 - b_8\gamma_8)/a_1] \tag{3.35}$$

The equations for γ_4, γ_5, γ_6 are:

$$\gamma_6 = \min(\Gamma_6, x_6) \tag{3.36}$$

$$\gamma_5 = \min[\min(\Gamma_5, x_5), (\gamma_6 - b_{14}y_{14})/a_5] \tag{3.37}$$

$$\gamma_4 = \min[\min(\Gamma_4, x_4), (\gamma_5 - b_7\gamma_7)/a_4] \tag{3.38}$$

Equations (3.32) through (3.38) present no problem until attempting to solve for γ_1 and γ_4. In order to solve (3.35) and (3.38) knowledge of γ_8 and γ_7 is needed. Therefore consider the remaining two arterials.

Immediately we see

$$\gamma_9 = \min(\Gamma_9, x_9) \tag{3.39}$$

$$\gamma_8 = \min[\min(\Gamma_8, x_8), (\gamma_9 - b_1\gamma_1)/a_8] \tag{3.40}$$

and

$$\gamma_{10} = \min(\Gamma_{10}, x_{10}) \tag{3.41}$$

$$\gamma_7 = \min[\min(\Gamma_7, x_7), (\gamma_{10} - b_4\gamma_4)/a_7] \tag{3.42}$$

As would be expected, it is necessary to solve (3.35) and (3.40) as a pair of simultaneous inequalities in order to obtain γ_1 and γ_8. Similarly (3.38) and (3.42) must be solved simultaneously to obtain γ_4 and γ_7. Therefore examine (3.35) and (3.40).

Recall the inequalities which resulted in the right hand terms of (3.35) and (3.40). Written as a matrix inequality they become

$$\begin{bmatrix} a_8 & b_1 \\ b_8 & a_1 \end{bmatrix} \begin{bmatrix} \gamma_8 \\ \gamma_1 \end{bmatrix} \leq \begin{bmatrix} \gamma_9 \\ \gamma_2 \end{bmatrix} \tag{3.43}$$

The assumption that the square matrix in (3.43) is nonsingular (i.e., $a_1 + a_8 - 1 \neq 0$) yields

$$\begin{bmatrix} \gamma_8 \\ \gamma_1 \end{bmatrix} \leq \frac{1}{(a_1 + a_8 - 1)} \begin{bmatrix} a_1 & -b_1 \\ -b_8 & a_8 \end{bmatrix} \begin{bmatrix} \gamma_9 \\ \gamma_2 \end{bmatrix} \tag{3.44}$$

The above inequality together with the left hand terms in (3.35) and (3.40) implies

$$\gamma_8 = \min[\min(\Gamma_8, x_8), (a_1\gamma_9 - b_1\gamma_2)/(a_1 + a_8 - 1)] \tag{3.45}$$

$$\gamma_1 = \min[\min(\Gamma_1, x_1), (a_8\gamma_2 - b_8\gamma_9)/(a_1 + a_8 - 1)] \tag{3.46}$$

In an exactly similar manner γ_7 and γ_4 are given by

$$\gamma_7 = \min[\min(\Gamma_7, x_7), (a_4\gamma_{10} - b_4\gamma_5)(a_4 + a_7 - 1)] \qquad (3.47)$$

$$\gamma_4 = \min[\min(\Gamma_4, x_4), (a_7\gamma_5 - b_7\gamma_{10})(a_4 + a_7 - 1)] \qquad (3.48)$$

If $(a_1 + a_8 - 1) \neq 0$ and if $(a_4 + a_7 - 1) \neq 0$ then (3.32), (3.34), (3.36), (3.37), (3.39), (3.41), (3.45), (3.46), (3.47), and (3.48) yield the required solutions for γ_1 through γ_{10}. In the event that $(a_1 + a_8 - 1) = 0$ the inequality constraints (3.43) on γ_8 and γ_1 are no longer independent, and (3.45) and (3.46) are replaced by (3.49).

$$\gamma_8 \leq \min(\Gamma_8, x_8) \qquad (3.49a)$$

$$\gamma_1 \leq \min(\Gamma_1, x_1) \qquad (3.49b)$$

$$\gamma_8 + \gamma_1 \leq \min[\gamma_9/(1-a_1), \gamma_2/a_1] \qquad (3.49c)$$

That is, the values of γ_8 and γ_1 are no longer unique, but may be chosen arbitrarily provided (3.49) is satisfied.

In a similar manner γ_7 and γ_4 must satisfy (3.50).

$$\gamma_7 \leq \min(\Gamma_7, x_7) \qquad (3.50a)$$

$$\gamma_4 \leq \min(\Gamma_4, x_4) \qquad (3.50b)$$

$$\gamma_7 + \gamma_4 \leq \min[\gamma_{10}/(1-a_4), \gamma_5/a_4] \qquad (3.50c)$$

Therefore the general control strategy for an arbitrary rectangular interconnection of street stubs will be quite simple. Optimal controls for all arterials are independently solved recursively, starting at the boundary of the control area and working back toward the traffic source (on another boundary). At each point where the arterials intersect, a system of two simultaneous equations results. Once the optimal controls for this intersection have been determined the controls for the remainder of the arterials may again be independently determined.

The only information needed to control an arbitrary grid is: (1) the input demands, y_{1i}, (2) the turn coefficients, a_i, and (3) the present state of the system, i.e., the queue lengths. The control scheme is very simple to implement, requiring a minimum of logic for each control area.

Since any controller must keep a record of the densities of all internal grid streets, it would be a relatively simple procedure to program the controller to continuously compute the turn coefficients, a_i. In this sense the system may be made self adaptive.

IV

OBSERVATIONS AND CONCLUSIONS

4.1 Summary

Modern mathematics and control theory are rich with techniques and results which are useful in the analysis and control of complex systems. In order to take advantage of this situation one must be able to characterize the dynamics of the particular phenomenon of interest in the form of a state model to which these techniques may be applied. Furthermore, operational procedures, which are independent of the particular characteristics of the system, exist which generate state models from the component models and their interconnection pattern [KO 1]. These operational procedures require only that the components be characterized by two complementary variables satisfying the postulates of system theory as given in Chapter II.

In Chapter II a pair of complementary variables, traffic density, x, and traffic flow rate, y, which can be used to quantitatively characterize the vehicular traffic phenomenon are defined. Using these complementary variables some useful subsystems, or "components" of traffic systems are, conceptually at least, identified and mathematically characterized. The well defined, operational procedures of system theory are used to combine these component models and the mathematical expression of their interconnection pattern into a state model. The resulting state model is in exactly the form required to apply modern control theory.

A specific problem in traffic control is considered in Chapter III--that of surface street control under a high density mode of operation. A state model, based on traffic queue lengths, is derived and its special properties is studied. It is found that the problem can always be formulated in such a way that the state model, for an arbitary m by n rectangular grid, is in the form of a system of 2(m+n+mn+1) linear, first order, difference equations. An optimal control strategy is derived, using the Pontryagin Maximum Principle, which, for an arbitrary set of initial conditions, reduces the state (queue lengths) to zero in a minimum number of control intervals.

4.2 Some Observations

It is appropriate, at this time, to look at the complete problem of traffic control. In Chapter III a control scheme, useful when queue lengths are so long that they cannot be dissipated in one complete traffic signal cycle, is outlined. The assumptions that allow the vehicular traffic system to be characterized with a set of simultaneous difference equations also provides a natural method of mode control. Whenever queue lengths are reduced to the point that they may be dissipated in one traffic signal cycle, a transition out of the high density mode is in order.

Although a specific low density mode of traffic control is not outlined in this thesis, a few of its properties are clear. It will operate when traffic conditions are such that queues at traffic signals do not accumulate from one cycle to the next. As in the high density mode, the accumulation of queues provides a natural criterion for determining when to enter the low density mode.

Since mode determination is based only on queue length the boundary between low density and high density control may be defined in a dynamic manner. Suppose, for example, a long arterial is to be controlled. It is quite likely that there will be portions of the arterial (near the traffic source) which will remain in the high density mode for quite some time, while other portions (near the traffic sink)may remain in the low density mode. As time progresses (and the vehicular traffic demand lessens) the boundary between the high and low density control modes will move from sink to source.

It must be observed that the possibility of instability of mode selection (hunting) exists. Depending upon the specific nature of the low density control mode, it may be necessary to base mode transition decisions on factors other than queue length alone. Clearly, if a low density control scheme can be developed which will result, as in the high density case, in monotonic decreasing queues there will be no mode selection instability. In such a case the entire traffic control system will remain in the low density mode at nearly all times. The high density mode will be used only for recovery from catastrophic occurrences.

Consider briefly the control philosophy implied by the previously described control schemes. A typical example of a vehicular traffic control system would be a relatively small, highly congested area (probably the central business district), together with a few key arterials feeding that area. The control philosophy described in Chapter III allows the possibility of unlimited queues accumulating outside the control area. If these queues cause undesirable congestion, the boundary of the control area must be extended until the demand is no longer great enough to produce excessively long queues. The control philosophy essentially says "don't use streets as vehicle storage areas as that reduces their efficiency." By carrying this philosophy to its logical conclusion it is clear that inputs to the traffic system must be controlled at the source--parking lots, ramps, and residential areas.

There are some other important features of this control philosophy. Notice that in the strict sense the optimal control scheme of Chapter III is not time optimal, but interval optimal. That is, the state is reduced in a minimum number of control intervals, not necessarily in a minimum time. The control interval length is the amount of time required to count the maximum number of vehicles allowed through the intersection. However, for a given set of pavement conditions, and driver-vehicle traffic mixes (this essentially fixes acceleration time) the control interval lengths are constant and the control is time optimal.

This time independence feature is desirable in that it makes the control strategy self-adaptive to changes in pavement conditions. As would be expected under wet and slippery pavement conditions vehicle acceleration time is increased, resulting in a corresponding increase in control interval length over the dry pavement conditions. Even though the total time to reduce an initial state to an acceptable density level will be lengthened, for these conditions the control is still time optimal. Note that no adjustment in system parameters need be performed to achieve this result.

4.3 Implementation Features

The information needed as input data to the control scheme of Chapter III is restricted to the queue length and turn coefficient values. Queue lengths may be measured by a simple, readily available, presence detector of the loop or overhead sonic type, while values for turn coefficients may be obtained (for various periods of the day) from origin-destination type surveys. In an alternate method of determining turn coefficients the behavior of the traffic could be continuously monitored and these coefficients updated. Although this method is more complicated, it has some outstanding advantages in that unforeseen traffic patterns can be handled with a minimum of delay.

Computation requirements for such a system are not extensive. The only arithmetic operation required is addition, as the controllers' function is to keep a continuous count of queue lengths. Computation speed requirements are sufficiently low that a sizeable system could be handled by a small desk size digital computer in a time sharing mode, or by the relatively low bandwidth available with present fluid logic modules.

This high density mode control scheme has some very useful features. They are: (1) Essentially identical hardware modules can be used to control a traffic area of arbitrary size; (2) since control functions are computed on the basis of local information, only a minimum of information need be transferred between local controllers; and (3) control can be installed on an "as needed" basis, i.e., particularly congested areas of a city may be controlled without the necessity of controlling the entire city.

4.4 Future Research

The research leading to this thesis has created many more questions than it has provided answers. The original objective of this research was to characterize the traffic phenomena in the form of an optimal control problem so that the results of modern control theory could be applied to it. Although this was accomplished, it did create some new and unique optimal control problems.

For example, the optimal control problem associated with the vehicular traffic state model of Chapter III is one where the admissable set of controls, Ω, is a function of the state vector $\vec{X}$. Furthermore, this function is such that the area of Ω is directly proportional of the norm of $\vec{X}$; Ω vanishes when $\vec{X} = 0$. Under these conditions the time optimal control scheme, in the linear continuous state model case, is clearly no longer bang bang. The author is unaware of any treatment of the problem of non-constant admissable sets in the literature.

A second--forever unsolved--problem is that of effective methods of computation of optimal controls in the bounded phase plane case. The physical constraints associated with a general vehicular traffic control problem imply a bounded phase plane. However, in the specific instance of high density control, optimal controls may be chosen in such a way that each state variable is monotonic decreasing.

This author feels that this thesis has merely scratched the surface of a problem which is most interesting and very fruitful for future research efforts. It is hoped that the guidelines and basic results presented in this thesis will stimulate further much needed research in this most rewarding area.

REFERENCES

[CA 1] "Chicago Area Transportation Study," U.S. Department of Commerce, Bureau of Public Roads, 1962.

[DR 1] Dreyfus, S. E., Dynamic Programming and the Calculus of Variations, Academic Press, New York, 1965.

[GR 1] Greenberg, H., "An Analysis of Traffic Flow," Operations Research, Vol. 7, 1959, pp. 79-85.

[GR 2] Greenberg, H., Daou, A., "The Control of Traffic Flow to Increase the Flow," Operations Research, Vol. 8, 1960, pp. 524-532.

[GO 1] Goodnuff, J. L.,"On the Controllability of Discrete State Systems," M.S. Thesis, Michigan State University, 1963.

[HA 1] Halkin, Hubert, "Optimal Control for Systems Described by Difference Equations," Advances in Control Systems: Theory and Applications, Academic Press, New York, 1964, pp. 173-196.

[HA 2] Halkin, Hubert, "A Maximum Principle of the Pontryagin Type for Systems Described by Nonlinear Difference Equations," J. SIAM on Control, Vol. 4, No. 1, 1966, pp. 90-111.

[HE 1] Herman, Robert, and Weiss, George, "Comments on the Highway-Crossing Problem," Operations Research, Vol. 9, No. 6, 1961.

[HO 1] Holtzman, J. M. and Halkin, H., "Directional Convexity and the Maximum Principle for Discrete Systems," J. SIAM on Control, Vol. 4, No. 2, 1966, pp. 263-275.

[JL 1] Actual measurements taken on the John C. Lodge freeway in Detroit, Michigan, 1966.

[KO 1] Koenig, H. E., Tokad, Y., and Kesavan, H. R., Analysis of Discrete Physical Systems, McGraw-Hill, New York, 1967.

[MA 1] May, A. D., "Traffic Characteristics and Phenomena on High Density Controlled Access Facilities," Traffic Engineering, Vol. 31, 1961, pp. 11-16.

[MA 2] May, A. D., "Improving Network Operations With Freeway Ramp Control," 43rd Annual Meeting of Highway Research Board, Washington, D.C., January 13-17, 1964.

[PO 1] Pontryagin, L. S., et. al., The Mathematical Theory of Optimal Processes, Wiley and Sons, Inc., New York, 1962.

[WI 1] Wirth, J. L., "Time Domain Models of Physical Systems and Existence of Solutions," Ph.D. Thesis, Michigan State University, 1962.