**Pergamon**

S0968-090X(96)00002-2

# SOME GENERAL RESULTS ON THE OPTIMAL RAMP CONTROL PROBLEM

H. ZHANG, S. G. RITCHIE and W. W. RECKER

Department of Civil and Environmental Engineering and Institute of Transportation Studies, University of California, Irvine, CA 92717, U.S.A.

**Abstract**—In an effort to relieve peak hour congestion on freeways, various ramp metering algorithms have been employed to regulate the inputs to freeways from entry ramps. In this paper, we consider a freeway system comprised of a freeway section and its entry/exit ramps, and formulate the ramp control problem as a dynamic optimal process to minimize the total time spent in this system. Within this framework, we are able to show when ramp metering is beneficial to the system in terms of total time savings, and when it is not, under the restriction that the controlled freeway has to serve all of its ramp demand, and the traffic flow process follows the rules prescribed by the LWR theory with a triangular flow–density relationship. We also provide solution techniques to the problem and present some preliminary numerical results and empirical validation. Copyright © 1996 Elsevier Science Ltd

## 1. INTRODUCTION

Ramp metering has been recognized as an effective way of relieving freeway congestion, typically the result of either a surge of demand during peak commuting hours or a temporary reduction of freeway capacity. It has been in use for over 30 years and is presently employed in a number of urban areas in North America (Blumentritt, 1989). Theoretically, ramp metering is effective if the traffic volume on the mainline freeway at the section immediately upstream of the ramp is less than the capacity. Under such conditions, metering can be used to ensure that the traffic volume delivered downstream of the ramp does not create a bottleneck situation in which upstream demand exceeds the downstream capacity. This is accomplished by three mechanisms: (1) actual reduction in ramp flow rate, by spreading peak demand over a greater time period, (2) diversion of potential freeway traffic to adjacent surface streets, and (3) breaking of "platooning" of entering vehicles to allow more efficient merging. Under conditions of heavy congestion, mechanisms (1) and (3) are generally ineffective and, theoretically, metering is potentially beneficial only to the extent that freeway traffic can be diverted to surface streets. In actual operation, however, the metering may serve to shorten the duration of congested conditions on the freeway by reducing the magnitude of the associated bottlenecks downstream from the entry ramps. Alternatively, metering under congested conditions may simply exacerbate congestion on surface streets, while providing no significant relief to freeway travel.

Under prolonged periods of heavy congestion, traffic desiring to use the freeway facility can exceed the capacity for an extended period of time. For ramp metering to have a significant effect in shifting demand over time, it would be necessary for vehicles to spend a similar amount of time in the queue. This may be technically possible in some situations, but not acceptable. The need to provide acceptable queue lengths may also negate other benefits provided by ramp metering. To reduce queues to acceptable limits, both in terms of the size of the queue on the ramps and streets and the time spent in the queue, it may be necessary to set the metering rate high enough that a steady stream of vehicles enters the freeway, thus negating the savings from platoon reduction (TRB, 1975). Under conditions of heavy, prolonged demand on the freeway facility, the principal mechanism by which ramp metering can improve the situation is by encouraging diversion to alternate routes.

If enough vehicles divert to alternate facilities, the reduced demand will reduce, or even elimi-nate, freeway congestion. However, since the freeway is part of a larger traffic system, this diversion is truly beneficial only when the alternate routes have sufficient capacity available to carry the diverted traffic (Newman, 1969). Otherwise, the only effect will be to move congestion from the freeway to the surface streets, but not to actually improve overall traffic conditions.

A variety of algorithms have been proposed (and, in some cases, implemented) to control ramp metering; some are largely empirical (ITE, 1985), others are based either on optimization techniques (e.g. Wattleworth and Berry, 1965; Yuan and Kreer, 1971; Wang and May 1973; Chen and Cruz, 1974) or on automatic control theory (e.g. Isaksen and Payne, 1973; Goldstein and Kumar, 1982; Papageorgiou *et al.*, 1990, 1991; Zhang *et al.*, 1994). Popular examples of the latter case are the Time-of-Day control method initiated by Wattleworth and Berry (Wattleworth and Berry, 1965) and the Linear Quadratic feedback control method initiated by Payne *et al.* (Payne *et al.*, 1973). The objective of Time-of-Day control algorithms is to maximize ramp inputs, which is equivalent to minimizing total travel time under non-congested steady-state traffic conditions. The objective of the feedback approach is to minimize deviations from the nominal states, taking into account the traffic evolution, but giving no direct consideration to total travel time, which is a more appealing measure-of-effectiveness to traffic operators.

This paper is concerned with the control of freeway entry ramps during peak periods in order to minimize the total time spent in a freeway system that is defined by the mainline freeway and its entry/exit ramps. It examines the effectiveness of ramp metering under conditions of recurrent congestion and establishes theoretical guidelines for the use of ramp metering under such conditions. The results of this research may assist in the formulation of ramp metering policy in heavily congested urban areas by providing a clear breakdown of conditions and parameters necessary for successful operation of ramp metering under congested conditions. In this exposition, the total time spent in the system is comprised of both travel time on the freeway mainline and also queuing time on the entry ramps. We assume that there is no alternative route to the freeway system (this system has to serve the total traffic demand), and that the traffic flow process is governed by the LWR theory with a triangular flow–density relationship (Daganzo and Lin (1994) considered a similar case using a different approach). Within this framework, we show when ramp metering is beneficial to the system in terms of total time savings, and when it is not. We also provide solution techniques to the problem and present some preliminary numerical results.

## 2. THE RAMP CONTROL OBJECTIVE

The effectiveness of ramp control can be evaluated relative to a number of different performance measures. Some are based on throughput; others are based on travel time or travel delay. In this paper, we choose the total time spent in a freeway system as the objective to be optimized, since it includes both travel times and ramp delays.

Consider a freeway consisting of $n$ sections with n entry ramps, one for each section. Let $L' = [L_1, L_2, ... L_n]$ be the vector of the number of lanes in a freeway section, and $\delta^t = [\Delta_1, \Delta_2, ... \Delta_n]$ be the vector of section lengths, and $a^t = [\Delta_1 L_1, \Delta_2 L_2, ..., \Delta_n L_n]$. Further let $x_1^t(k) = [\rho_1^k, \rho_2^k, ... \rho_n^k]$ be the per lane section densities during time interval $k$, and $x_2^t(k) = [\zeta_1^k, \zeta_2^k, ... \zeta_n^k]$ be the corresponding vector of vehicle queue lengths at entry ramps where the superscript $t$ denotes the transpose of a vector or matrix. Then, the total vehicle time spent on the mainline freeway during some control period $K*T$ is:

$$T_f = T \sum_{k=1}^{K} a^t x_1(k), \qquad (1)$$

and the total time spent in the ramp queues is:

$$T_q = T \sum_{k=1}^{K} 1^t x_2(k), \qquad (2)$$

and the total time spent in the whole system is:

$$TT = T_f + T_q = T \sum_{k=1}^{K} [a'x_1(k) + 1'x_2(k)], \tag{3}$$

where $T$ is the length of each interval, $K$ is the number of such intervals that define the control period, and $1'$ is a row vector of proper dimension, whose elements are all unit values. Then, the ramp control objective is stated simply as

$$\text{minimize } TT = T_f + T_q = T \sum_{k=1}^{K} [a'x_1(k) + 1'x_2(k)]. \tag{4}$$

## 3. THE TRAFFIC FLOW PROCESS

### 3.1. Freeway traffic flow

We assume that, for purposes of this development, the traffic flow process is governed by the macroscopic traffic model developed by Lighthill and Whitham (1955), and Richards (1956):

$$q_x + \rho_t = 0 \tag{5}$$

$$q = f(\rho), \tag{6}$$

and, if the function $f$ is smooth, we have

$$\rho_t + f'(\rho)\rho_x = 0, \tag{7}$$

where $x$ denotes space and $t$ time, $q$ and $\rho$ are flow rate and traffic density, respectively. Here $q$ represents the partial derivative and $f'$ the total derivative.

Equation (7) is the well-known hyperbolic conservation law in one dimension, which can be solved analytically by the method of characteristics. It is known to possess waves that travel at a speed $c = f'(\rho)$. Equation (7) can admit discontinuous solutions, which gives rise to shock waves. The speed of a shock wave is given by $c_s = (f(\rho_1) - f(\rho_2))/(\rho_1 - \rho_2)$, which can be constructed graphically from the $(\rho, f(\rho))$ graph. During the control period considered, we assume that the system demand is such that all sectional speeds are positive for the freeway system, and we call the freeway traffic *uniformly uncongested* if there exist only forward traffic waves for all freeway sections, *uniformly congested* if there exist only backward traffic waves for all freeway sections, and *transient* if forward and backward traffic waves coexist on the freeway.

The flow–density relationship $q = f(\rho)$ is generally referred to as the fundamental diagram of traffic flow. The LWR theory does not specify any form for the speed–density relationship. Rather, this relationship is usually obtained by fitting observed field data. The flow–density relationship must satisfy two boundary conditions: $f(0) = 0 = f(\rho_{jam})$, where $\rho_{jam}$ is the density at which traffic is at a standstill. In addition, there exists a density $\rho_c$ such that $f(\rho_c)$ is maximal, and $f'(\rho) > 0$, if $\rho < \rho_c$ and $f'(\rho) < 0$, if $\rho > \rho_c$. There are many proposals for the form that this relationship should take; for example, Newell (1991) suggests a triangular flow–density relationship:

$$q = \begin{cases} v_f\rho, & \text{if } \rho \le \rho_c \\ c_0(\rho_{jam} - \rho), & \text{if } \rho \ge \rho_c, \end{cases}$$

where $v_f$ is the free flow travel speed and $c_0$ is the wave speed traveling in a direction against the traffic stream.

For the purpose of exposition, we assume that the freeway is comprised of $n$ sections of equal length and equal number of lanes, and that each section has only one entry and one exit ramp (these assumptions will later be relaxed). We then convert the above equations into difference equations, and add sources/sinks to represent entry/exit ramps in the analysis.

Because of the nature of the discontinuities associated with the traffic flow law, the selection of an appropriate difference scheme and discretization scale is essential (Newell, 1989). Here, we adopt a simple difference scheme that depends on the wave directions. This scheme is known to be stable and capable of capturing traffic shocks.

Let $\rho_i^k$ be the section density of the $i$th section at time step $k$, $r_i^k$, $e_i^k$ the entry and exit rates from ramps in the $i$th section at time step $k$, $s_i^k$ the exit rate/freeway flow rate ratio for exit ramp $i$ at time step $k$, $T$ the length of each time step, $L$ the number of freeway lanes, and $\Delta$ the length of each freeway section. Then, in difference form:

$$\rho_i^{k+1} = \rho_i^k + \frac{T}{\Delta}\left(f_{i-1/2}^k - f_{i+1/2}^k + \frac{r_i^k - e_i^k}{L}\right), \quad i = 1, 2, \dots, n$$

$$f_{i+1/2}^k = \begin{cases} f(\rho_i^k), & \text{if } c_{i+1/2} > 0 \\ f(\rho_{i+1}^k), & \text{if } c_{i+1/2} < 0 \end{cases}, \quad c_{i+1/2} = \begin{cases} \dfrac{f(\rho_i^k) - f(\rho_{i+1}^k)}{\rho_i^k - \rho_{i+1}^k}, & \text{if } \rho_i^k \neq \rho_{i+1}^k \\ f'(\rho_i^k), & \text{if } \rho_i^k = \rho_{i+1}^k \end{cases}$$

$$e_i^k = s_i^k\,(L{\cdot}f(\rho_i^k)), \quad 0 \leq s_i^k \leq 1 \tag{8}$$

This numerical scheme is stable if $(1 + s_i^k)|\frac{T}{\Delta}f'(\rho_i^k)| \leq 1$ for all $\rho_i^k$ and $s_i^k$ (see Appendix). For example, taking the maximum $f'(\rho_i^k) = v_f = 60$ m.p.h., $s_i^k = 0.2$ and $\Delta = 0.5$ mile, then the maximum time step $T = 25$ s. The expression $|\frac{T}{\Delta}f'(\rho_i^k)|$ plays a pivotal rule in the later analysis; we denote it as $\alpha_i^k$, i.e. $\alpha_i^k = |\frac{T}{\Delta}f'(\rho_i^k)|$.

## 3.2. Ramp traffic queues

When a ramp cannot service all of its demand, a queue will build up. For a single ramp $i$, with demand $d_i$, and queue length $\zeta_i$, the queuing process is described by

$$\zeta_i = d_i - r_i \tag{9}$$

or, in difference form,

$$\zeta_i^{k+1} = \zeta_i^k + T(d_i^k - r_i^k). \tag{10}$$

To avoid the perception of signal failure by drivers, the control cycle in most practical applications is restricted from exceeding a certain maximum value (e.g. 20 s), which sets a lower bound $r_{i,\min}^k$ for the metering rate $r_i^k$. Alternatively, the metering rate cannot exceed the virtual demand of a ramp, $d_i^k + \zeta_i^k/T$, i.e.

$$r_{i,\min} \leq r_i^k \leq d_i^k + \zeta_i^k/T. \tag{11}$$

## 3.3. The traffic process in vector notation

To conform to the optimal control formalism, we write the state and constraint equations developed above in vector notation. Let the control vector $u^t(k) = [r_1^k, r_2^k, \dots r_n^k]$ be the ramp metering rate with superscript $t$ denoting the transpose of a vector or matrix, and its lower bound $u_{\min}^t = [r_{1,\min}, r_{2,\min}, \dots r_{n,\min}]$. We further define a disturbance vector $w_2^t(k) = [d_1^k, d_2^k, \dots d_n^k]$ as the vector of on-ramp demand, and another disturbance vector $w_1^t(k) = [q_0^k, s_1^k, s_2^k, \dots q_e^k]$, with $q_0^k$ freeway inflow and $q_e^k$ the freeway outflow and $s_i^k$, $i = 1, \dots n$ defined as in equation (8). Then, eqns (8) and (10) can be written as:

$$x_1(k + 1) = \Phi_1^k(x_1(k), u(k), w_1(k)), \tag{12}$$

$$x_2(k + 1) = x_2(k) + T[w_2(k) - u(k)] = \Phi_2^k(x_2(k), u(k), w_2(k)), \tag{13}$$

or,

$$x_{k+1} = F^k(x_k, u_k, \bar{w}_k), \tag{14}$$

where

$$F_k = \begin{bmatrix} \Phi_1^k \\ \Phi_2^k \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad \overline{w} = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}.$$

Similarly, the constraint represented by eqn (11) becomes:

$$c_1(x_k, \overline{w}_k) \leq u_k \leq c_2(x_k, \overline{w}_k), \tag{15}$$

where

$$c_1(x_k, \overline{w}_k) = u_{\min}, \text{ and } c_2(x_k, \overline{w}_k) = w_2(k) + x_2(k)/T$$

## 4. THE OPTIMAL CONTROL PROBLEM

Optimal control theory is concerned with the control of a dynamic process in the best manner. The dynamic process is generally described by ordinary differential/difference equations, and the objective is to optimize a performance index that consists of states and control variables. In addition, there may be constraints imposed on the controls. This theory has been applied to the ramp control problem (Papageorgiou, 1983), and has recently gained popularity in dynamic network assignment (e.g. Ran et al., 1993). In this section, we present the standard form of the discrete optimal control problem, and state its necessary conditions for optimality, which will be used to analyze the ramp control problem. For more information on optimal control theory, the reader is referred to Bryson and Ho (1975).

Consider the discrete optimal control problem

$$\text{minimize } J = \phi(x_K) + \sum_{k=1}^{K-1} L^k(x_k, u_k), \tag{16}$$

with

$$x_{k+1} = F^k(x_k, u_k), \tag{17}$$

and with initial condition $x_0$ given, and $u_k \in \Omega_k$, where $\Omega_k$ is convex, and where

$$x_k \in R^n, u_k \in R^m, \Omega_k \subset R^m, L^k:R^n \times R^m \to R^1, \phi:R^n \to R^1, F^k:R^n \times R^m \to R^n.$$

We first adjoin state equation (17) with objective function (16) by multiplier (column) vector $\lambda_{k+1}$ to form the Hamiltonian

$$H^k = L^k + (\lambda_{k+1})'F^k. \tag{18}$$

Then, on the optimal state trajectory, the following equation must hold:

$$(\lambda_k)' = H_x^k = L_x^k(x_k, u_k) + (\lambda_{k+1})'F_x^k(x_k, u_k), \tag{19}$$

with end condition given by

$$(\lambda_K)' = \phi_x(x_k).$$

The gradient of the objective function $J$ with respect to control vector $u_k$, denoted here as a row vector $p_k$, is given by

$$p_k = H_u^k = L_x^k(x_k, u_k) + (\lambda_{k+1})'F_x^k(x_k, u_k). \tag{20}$$

The optimal control policy $U^* = [(u_1)',(u_2)',....,(u_{K-1})']'$ must satisfy:

$$u_k^* = \frac{\text{arg min } H^k(x_k, u_k, \lambda_{k+1})}{u_k \in \Omega_k}, \quad k = 1,..., K-1. \tag{21}$$

If the problem is linear in objective and state equation, and the admissible controls are given by box constraints $c_1(x_k) \leq u_k \leq c_2(x_k)$, eqn (21) reduces to:

$$\begin{aligned} u_k^* &= c_1(x_k) && \text{if } p_k > 0 \\ u_k^* &= c_2(x_k) && \text{if } p_k < 0 \\ \text{singular case} && \text{if } p_k \equiv 0 \end{aligned} \tag{22}$$

## 5. ANALYSIS OF OPTIMALITY OF THE RAMP CONTROL PROBLEM

Recall the formulation of the ramp metering objective function in Section 2 above. Let

$$L^k(x_k, u_k) = a'x_1(k) + 1'x_2(k) + 0'u(k),$$

and

$$\phi(x_K) = a'x_1(K) + \beta 1'x_2(K), \beta \in R^1, \beta \geq 1.$$

where $\beta$ is some penalty for the presence of ramp queues at the end of the control period. With this modification, the objective function specified by eqn (4) becomes

$$\text{minimize } TT = T\left[\phi(x_K) + \sum_{k=1}^{K-1} L^k(x_k, u_k)\right].$$

Alternatively, since $T$ is a positive constant, this is equivalent to

$$\text{minimize } J = \phi(x_K) + \sum_{k=1}^{K-1} L^k(x_k, u_k),$$

which is the same form as eqns (16) and (17). In this section, we use the gradient information and the optimality condition set forth in Section 4 above to analyze when ramp metering is beneficial and when it is not.

From equation (14), the Jacobian of $F^k$ with respect to states and controls can be represented as

$$F_x^k = \begin{bmatrix} A^k & 0 \\ 0 & B^k \end{bmatrix}, F_u^k = \begin{bmatrix} C^k \\ D^k \end{bmatrix},$$

$$A^k = \phi_{1,x}^k(x_{1,k}, u_k, w_{1,k}), \quad B^k = \phi_{2,x}^k(x_{2,k}, u_k, w_{2,k})$$

$$C^k = \phi_{1,u}^k(x_{1,k}, u_k, w_{1,k}), \quad D^k = \phi_{2,u}^k(x_{2,k}, u_k, w_{2,k})$$

and from eqns (1)–(4):

$$(L_x^k)' = \begin{bmatrix} a \\ 1 \end{bmatrix}, L_u^k = 0, (\phi_u^K)' = \begin{bmatrix} a \\ \beta 1 \end{bmatrix}$$

and

$$\begin{bmatrix} \lambda_1^k \\ \lambda_2^k \end{bmatrix} = \begin{bmatrix} a \\ 1 \end{bmatrix} + \begin{bmatrix} (A^k)' & 0 \\ 0 & (B^k)' \end{bmatrix} \begin{bmatrix} \lambda_1^{k+1} \\ \lambda_2^{k+1} \end{bmatrix}, \begin{bmatrix} \lambda_1^k \\ \lambda_2^k \end{bmatrix} = \begin{bmatrix} a \\ \beta 1 \end{bmatrix} \tag{23}$$

or,

$$\lambda_1^k = a + (A^k)'\lambda_1^{k+1}, \quad \lambda_1^k = a \tag{24}$$

$$\lambda_2^k = 1 + (B^k)'\lambda_2^{k+1}, \quad \lambda_2^k = \beta 1. \tag{25}$$

As in eqn (20), the gradient of the objective function with respect to control is

$$(p_k)' = [(C^k)'(D^k)'] \begin{bmatrix} \lambda_1^{k+1} \\ \lambda_2^{k+1} \end{bmatrix} = (C^k)'\lambda_1^{k+1} + (D^k)'\lambda_2^{k+1}. \tag{26}$$

With the triangular flow–density relationship, traffic may have two constant wave speeds, one forward, the other backward. Therefore, $\alpha_i^k$ assumes two values:

$$\alpha_i^k = |(T/\Delta)f'(\rho_i^k)| = (T/\Delta)v_f = \alpha \quad \text{if } \rho_i^k < \rho_c.$$

and

$$\alpha_i^k = |(T/\Delta)f'(\rho_i^k)| = (T/\Delta)c_0 = \alpha' \quad \text{if } \rho_i^k < \rho_c.$$

Further, we have

$$a = (\Delta L)1, \quad (B^k)' = I_d, \quad (C^k)' = \frac{T}{\Delta L}I_d, \quad (D^k)' = T \cdot I_d,$$

where $I_d$ is the identity matrix of proper dimension.
Equation (26) can now be simplified as:

$$(p_k)' = \frac{T}{\Delta L}\lambda_1^{k+1} - T\lambda_2^{k+1} = T\left(\frac{\lambda_1^{k+1}}{\Delta L} - \lambda_2^{k+1}\right) = T(\mu_1^{k+1} - \lambda_2^{k+1}) \tag{27}$$

$$\mu_1^{k+1} \equiv \frac{\lambda_1^{k+1}}{\Delta L}.$$

Finally, dividing both sides of eqn (24) by $\Delta L$, we have

$$\mu_1^k = 1 + (A^k)'\mu_1^{k+1}, \; \mu_1^k = 1. \tag{28}$$

We next derive some basic properties from these relationships regarding the effectiveness of ramp metering control for the three regimes of traffic identified earlier.

### 5.1. Case 1: traffic is uniformly congested/uncongested

In this scenario, the Jacobian $A^k$ forms a single Jordan block, with the diagonal elements $1 - (1 + s_i^k)\alpha$, lower off-diagonal elements $\alpha$ if traffic is uniformly uncongested, as shown in eqn (29), or diagonal elements $1 - (1 - s_i^k)\alpha'$, upper off-diagonal elements $\alpha'$ if traffic is uniformly congested, as shown in eqn (30).

$$A^k = \begin{bmatrix} 1 - (1 + s_1^k)\alpha & 0 & 0 & 0 \\ \alpha & 1 - (1 + s_2^k)\alpha & 0 & 0 \\ 0 & \alpha & 1 - (1 + s_3^k)\alpha & 0 \\ 0 & 0 & \alpha & 1 - (1 + s_4^k)\alpha \end{bmatrix} \tag{29}$$

$$A^k = \begin{bmatrix} 1 - (1 - s_1^k)\alpha' & \alpha' & 0 & 0 \\ 0 & 1 - (1 - s_2^k)\alpha' & \alpha' & 0 \\ 0 & 0 & 1 - (1 - s_3^k)\alpha' & \alpha' \\ 0 & 0 & 0 & 1 - (1 - s_4^k)\alpha' \end{bmatrix} \tag{30}$$

Comparing eqns (25) and (28), and keeping in mind that $\beta \geq 1$, we have $p_K \leq 0$. As max $\{(1 + s_i^k)\alpha, (1 - s_i^k)\alpha'\}$; $i = 1,...,n$; $k = 1,...K$ is in the interval $(0, 1]$, both the spectra (the set of eigenvalues) of $A^k$ and its off-diagonal elements lie within the union disk; hence $A^k$ is a contraction mapping, while $B^k = I_d$ is constant with its spectra on the union disk. Thus, $p_k < 0$, for $k = 1, 2,...,K-1$, and is decreasing as $k$ decreases. Because $p_k < 0$, regardless of the controls (metering rate), according to the optimality condition of eqn (22), the performance function obtains its global minimum at the upper bound of controls (because the problem is linear in objective, state equations and constraints when traffic is uniformly congested/uncongested, and hence is convex in its controls). This result means that it is better not to meter when traffic is uniformly congested/uncongested, a result that coincides with a previous result found by Daganzo and Lin in a different analysis (Daganzo and Lin, 1994).

### 5.2. Case 2: traffic is in transition

When traffic is in transition, that is, forward and backward traffic waves co-exist in the freeway section during some periods, then the Jacobian matrix $A^k$ may become tridiagonal, as shown in eqn (31),

$$A^k = \begin{bmatrix} 1 - (1 + s_1^k)\alpha & 0 & 0 & 0 \\ \alpha & 1 + s_2^k\alpha' & \alpha' & 0 \\ 0 & 0 & 1 - (1 - s_3^k)\alpha' & \alpha' \\ 0 & 0 & 0 & 1 - (1 - s_4^k)\alpha' \end{bmatrix} \tag{31}$$

where section 1 has forward waves, while section 2, 3 and 4 have backward waves. In such a situation, some of the spectra of $A^k$ lie outside of the union disk, and the mapping $A^k$ becomes expansive, which may cause some of the $p_k$s to be positive. For such situations, according to eqn (22), the objective function can be decreased by reducing the control variables corresponding to the positive $p_k$s. That is, ramp metering can reduce total time spent in the system under such situations.

### 5.3. Case 3: unequal section lengths and number of lanes

As can be seen from the above analysis, the form of the Jacobian matrix $A^k$ and its spectra play pivotal roles in determining when ramp metering can be beneficial under the present assumptions. If we now relax the assumptions of equal freeway section length and number of freeway lanes, it is clear that the Jacobian matrix $A^k$ assumes the same forms as above, with elements taking values from the set $\{1 - (1 + s_i^k)\alpha_i, 1 - (1 - s_i^k)\alpha_i, \alpha_i\}$, with $\alpha_i = (T/\Delta_i)v_f$, if $\rho_i < \rho_c$, and $\alpha_i = (T/\Delta_i)c_0$ if $\rho_i > \rho_c$. However, since max $\{(1 + s_i^k)\alpha_i, i = 1,..., n, k = 1,..., K\}$ must be less than or equal to unity in order that eqn (8) be stable, the contraction or expansive properties of $A^k$ under these relaxed conditions continue to hold.

Next, we show that the control gradient $p_k$ is negative under uniform congested/uncongested conditions with lane drops and varying section lengths, except in a few time intervals near the end of the time horizon considered, where it may be non-negative. Under the new conditions, the Jacobian of states with respect to the control is

$$C^k = T \begin{bmatrix} 1/\Delta_1 L_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & 1/\Delta_n L_n \end{bmatrix} \leq \frac{T}{\Delta' L'} I_d,$$

$$\Delta' L' = \min\{\Delta_i L_i, i = 1,...,n\},$$

and

$$(D^k)^t = -T * I_d$$

and the control gradient becomes

$$(p^k)^t = (C^k)^t \lambda_1^{k+1} + (D^k)^t \lambda_2^{k+1} \leq T(\mu_1^{k+1} - \lambda_2^{k+1}), \tag{32}$$

where

$$\mu_1^k \equiv \frac{1}{\Delta' L'} \lambda_1^k$$

Now, let $\theta \equiv \frac{1}{\Delta' L'} a$. Instead of eqn (28), we have

$$\mu_1^k = \theta + (A^k)^t \mu_1^{k+1}, \quad \mu_1^K = \theta, \text{ and } \frac{\Delta L}{\Delta' L'} 1 \geq \theta \geq 1, \tag{33}$$

where

$$\Delta L = \max \{\Delta_i L_i, i = 1,...,n\}.$$

Solving eqn (33), we obtain

$$\mu_1^k = (I_d + M^k + M^k M^{k+1} +...+ M^k M^{k+1}...M^{K-k})\theta, \, k = 1,...,K - 1, \, M^k = (A^k)^t.$$

Observe also that

$$\lambda_2^k = (I_d + I_d + I_d^2 +...+ I_d^{K-k-1})1 + \beta I_d^{K-k} 1 \geq (I_d + I_d + I_d^2 +...+ I_d^{K-k-1} + I_d^{K-k})1.$$

Because $M^k$ is contracting and $I_d$ is constant with its eigenvalues equal to unity, $\mu_1^k$ increases more slowly than does $\lambda_2^k$ as $k$ decreases. In most situations of practical importance, the net change in the number of freeway lanes in any particular freeway section is likely to be less than 2, and the section lengths may be in the range of 0.2–0.5 miles, which requires

$T$ to be in the range of 10–30 s. Since the time step $T$ is assumed to be constant, the model's numerical stability is satisfied if the section lengths are taken to be roughly the same. Then, for most situations of practical interest, we have the following estimates: $5 \geq \theta \geq 1$, $1 \geq \alpha_i \geq 0.5$; $0 \leq s_j^k \leq 0.25$; the magnitude of the spectra of $M^k$ then generally lies in the interval $[0, 0.5]$ when traffic is either uniformly congested or uniformly uncongested, which makes $\lambda_2^k$ increase much faster than $\mu_1^k$ as $k$ decreases. Also, since $K >> n$, the slower increase of $\mu_1^k$ thus leads to negative $p_k$, except for a few time intervals (roughly in the order of $n$) near the end of the time horizon.

The conclusion that ramp metering does not improve total travel time for a uniformly congested freeway system may be counterintuitive, and worthy of more comment. For a controlled freeway section, uniform congestion occurs in three possible scenarios: (1) ramp demand in this section is high enough to congest the whole freeway section during the whole control period, as is typical during a commuting peak; (2) the freeway section is short, and mainline flow dominates ramp flow, and congestion is caused by high mainline demand; (3) the controlled section is immediately upstream of a bottleneck, created by, for example, a severe incident. Under all three possible scenarios, temporarily holding vehicles on entry ramps will not bring the freeway system to congestion-free conditions, but rather will only decrease freeway travel time marginally. At the same time, it inevitably increases the delay time at ramps, which more than offsets the time savings on the freeway system (this is clearly illustrated by the magnitude of $\lambda_1^k$ and $\lambda_2^k$) since vehicles held at ramps will wait until they are being serviced, and any waiting time in the queue is wasted time. Therefore, as long as the freeway can service traffic, the system is always better off by servicing as many vehicles as possible, and as early as possible (this can be seen from the increasing magnitude of negative $p_k$ as $k$ decreases).

On the other hand, when traffic is in transition, holding vehicles at certain ramps temporarily may prevent freeway congestion from building up, thereby reducing freeway travel time to an extent that offsets the delay of queued vehicles. Thus, redistributing demand over time by ramp metering can reduce the total time spent in the system when traffic is in transition.

It is important to emphasize that these conclusions may not hold if vehicles can divert from ramp queues to alternative routes. Under such situations, ramp metering can always reduce the total time spent in the freeway system (mainline and ramps) if the diverted vehicles do not rejoin the controlled freeway section, but it may increase or decrease the total time spent in the whole system (freeways and alternative routes), depending on the traffic impact made by diverted vehicles on the alternative routes. If diverted vehicles rejoin the controlled section, total system time savings, if any, would be marginal because traffic is uniformly congested, that is, redistributing ramp demands over space for a uniformly congested freeway can, at best, reduce total system marginally.

## 6. SOLVING THE OPTIMAL CONTROL PROBLEM NUMERICALLY

In this section, we present numerical methods for solving the optimal ramp control problem. Before proceeding further, we simplify the problem by introducing a new control vector $v_k \in R^n$, such that $0 \leq v_{i,k} \leq 1$, and

$$u_{i,k} = u_{i,k} c_{1,i}(x_k, \bar{w}_k) + (1 - v_{i,k})c_{2,i}(x_k, \bar{w}_k), \quad i = 1, \dots, n. \tag{34}$$

Replacing $u_{i,k}$ by $v_{i,k}$, $i = 1, \dots, n$, $k = 1, \dots, K$, in eqns (14), (15) and (16), and introducing

$$V = [v_1, \dots, v_k, \dots, v_k] \in R^{nK},$$

we have

$$\text{minimize } J(V, W) = \phi(x_K(V, W)) + \sum_{k=1}^{K-1} \hat{L}^k(x_k(V, W), v_k), \tag{35}$$

with constraint

$$0 \leq v_k \leq 1$$

where the vectors $x_k$ $(V,W)$ are obtained by solving the equation

$$x_{k+1} = \hat{F}^k(x_k, v_k, \bar{w}_k),\tag{36}$$

with initial condition $x_0$ and $W = [\bar{w}_1, \bar{w}_2, ..., \bar{w}_k]$.

The problem in this form is a constrained non-linear programming problem in terms of variable $V$, with $W$ viewed as a parameter vector. In light of the special structure of this problem (i.e. the box constraints), numerical methods for unconstrained non-linear programming problems, such as steepest decent, conjugate gradient, or Quasi-Newton methods such as the BFGS algorithm (Luenberger, 1984), can be applied with the special consideration that once a constraint is reached, the moving direction is projected to that constraint.

The gradient information $\hat{p}_k = d\hat{J}/dv_k$ is required to determine a moving direction. From the results presented in Section 4 above, we have $\hat{p}_k = \hat{H}^k_v$, with $\hat{H}^k = \hat{L}^k + (\hat{\lambda}_{k+1})'\hat{F}_k$ and $\hat{\lambda}_k = \hat{H}^k_x$, $\hat{\lambda}_K = \phi_k(x_K)$. This provides the following update for the $i$th iteration.

$$V^{i+1} = V^i + \sigma^i d^i\tag{37}$$

with

$$d^i_k = 0 \text{ if } \hat{p}_k > 0 \text{ and } v_k = 0 \text{ or } \hat{p}_k < 0 \text{ and } v_k = 1$$

$$d^i_k = \hat{p}_k G^i_k \text{ otherwise,}$$

where $G^i_k$ is a linear operator, assuming different forms depending on the direction finding algorithm used. For example, $G^i_k = -I_d$, when the steepest decent is the choice.

The moving size $\sigma^i$ is determined by a line search. Since the gradient information is available, the bisection line search method is recommended.

## 7. TEST EXAMPLES

The optimal control algorithm developed above was tested on a four-lane freeway section of approximately 5 miles in length, with 10 subsections (each section is approximately half a mile long), with seven entry ramps and five exit ramps. Each exit ramp and entry ramp, except for the third entry ramp, has only one lane (Fig. 1). The third entry ramp has two lanes. The time step is $T = 30$ s, and the flow–density relationship used is as for Greenshields (1934) with $v_f = 60$, $\rho_{\text{jam}} = 120$.

We consider two cases; one for transient traffic, and the other for uniformly congested traffic. We assume that the average OD demand pattern is known for the whole control period of 3 h, and that traffic density falls within the range of zero and jam density under such demand patterns. The freeway boundary influx for both cases is shown in Fig. 2.
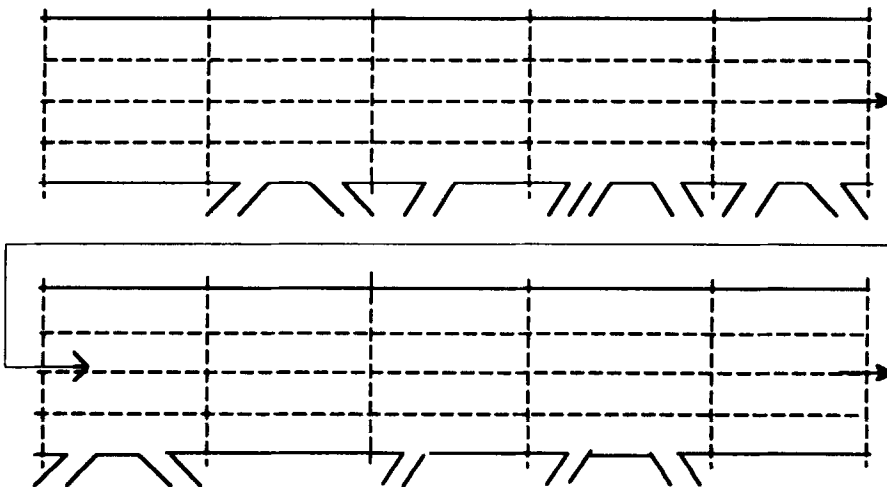


Fig. 1. A test case freeway section.

The ramp demand used in the transient traffic case is shown in Fig. 3(a). The gradient information $\hat{p}_k$ with no metering is plotted in Fig. 3(b). It is clearly seen that metering ramps 4–7 would increase system time, but metering ramps 1–3 during time interval about [80,140] would reduce system time. The optimal metering rate is plotted in Fig. 4(a), and ramp queues in Fig. 4(b). Figures 5(a) and (b) show the section densities before and after ramp metering. The total time saving in this case is 61 vehicle hours.

The second example illustrates the case where ramp metering does not improve total time spent in the system under uniform congested conditions. Figure 6(a) shows the ramp demand for this case. The control gradient $\hat{p}_k$ with no metering is shown in Fig. 6(b). Because $\hat{p}_k$ is negative over the whole control period, it would not be beneficial to meter any
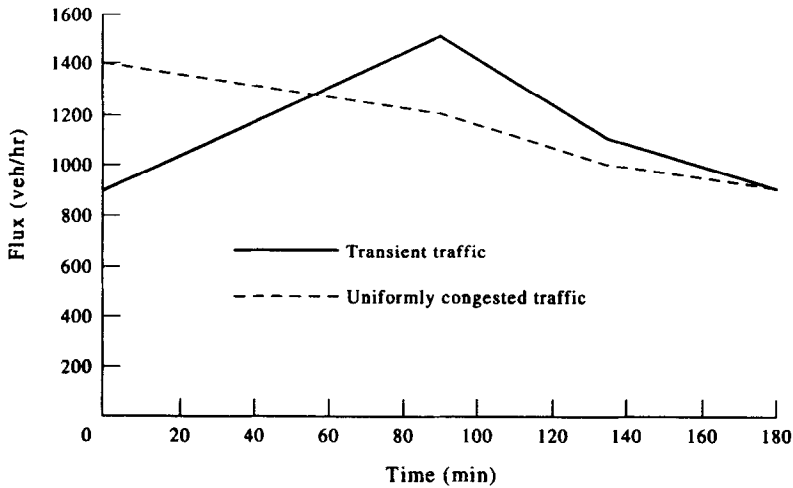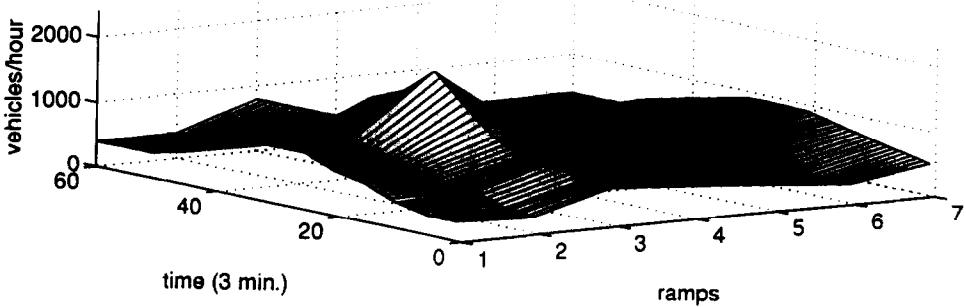
Fig. 2. Freeway demand profiles.

(a) ramp demand profile
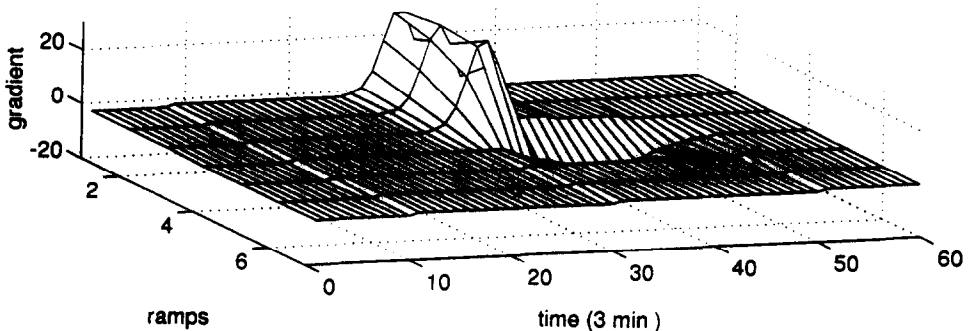
(b) gradient profile without control

Fig. 3. Ramp demand profile and control gradient profile, transient traffic.

ramps. This is confirmed by starting metering rates as half of the ramp demand, and the optimal solution by the control algorithm produces a set of metering rates the same as
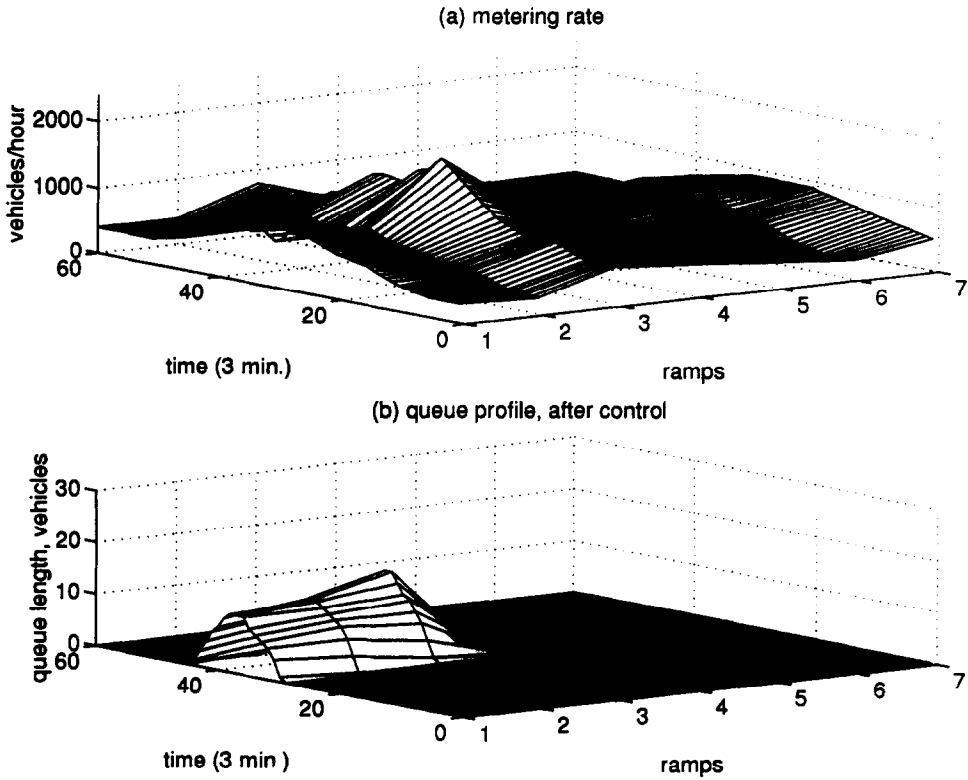
**(a) metering rate**



**(b) queue profile, after control**



Fig. 4. Ramp metering rates and ramp queues, transient traffic.

**(a) density profile, before control**



**(b) density profile, after control**
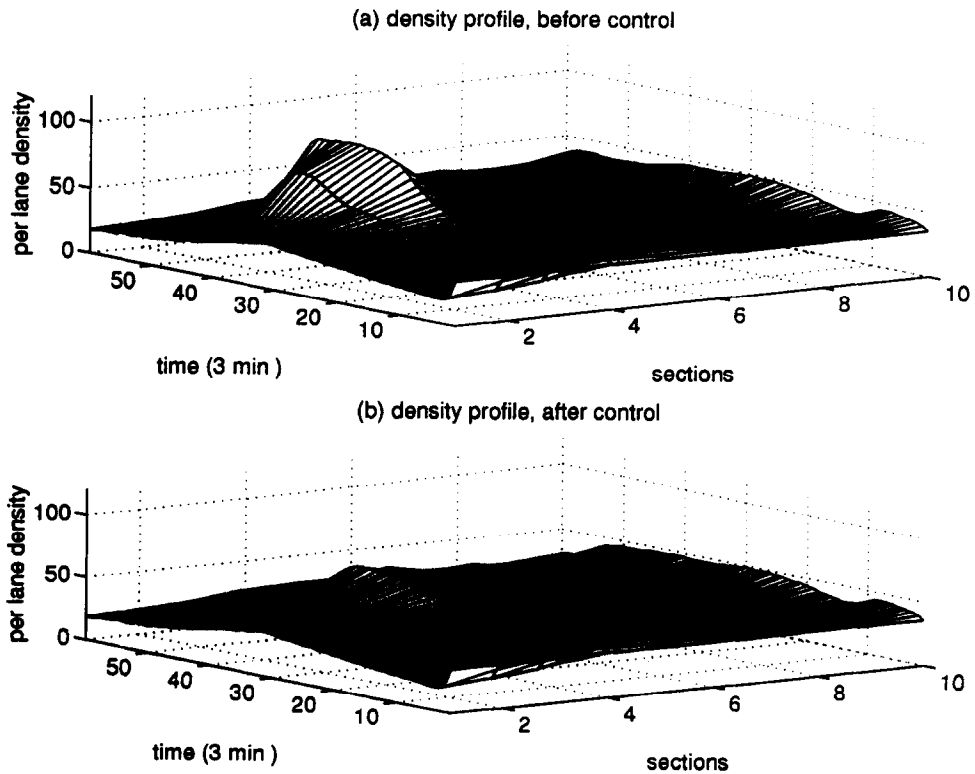


Fig. 5. Traffic density, before and after control, transient traffic.

the demand. As a result, we obtained identical section density profiles with and without ramp metering, which are shown in Fig. 7(a) and (b).
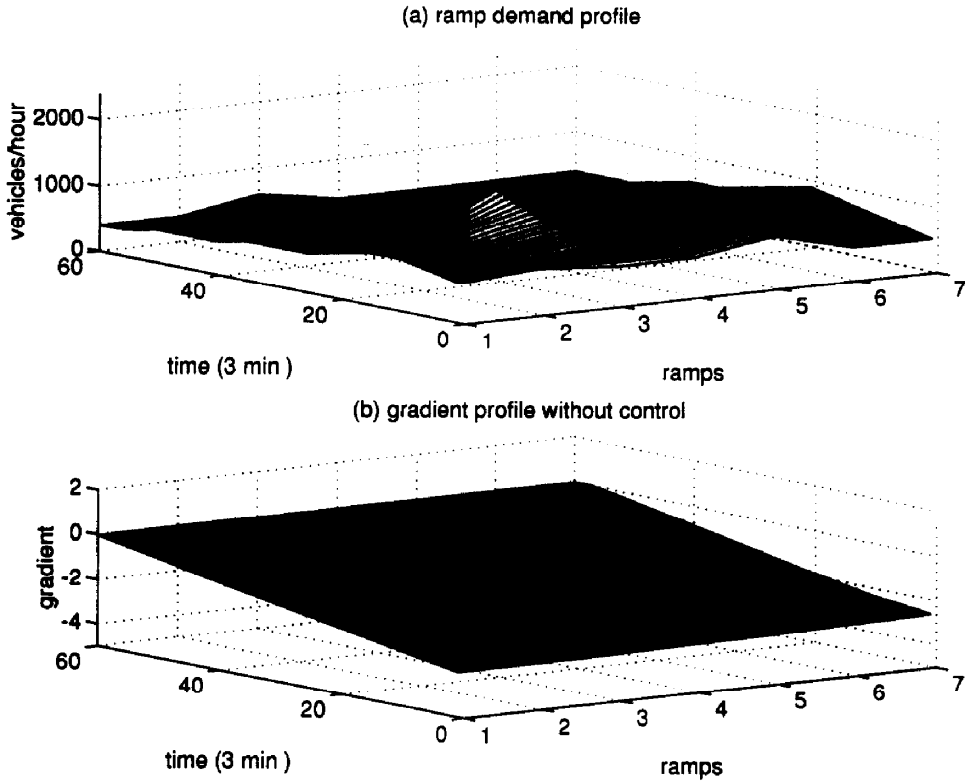
**(a) ramp demand profile**



**(b) gradient profile without control**



Fig. 6. Ramp demand and control gradient profiles, uniformly congested traffic.

**(a) density profile, before control**



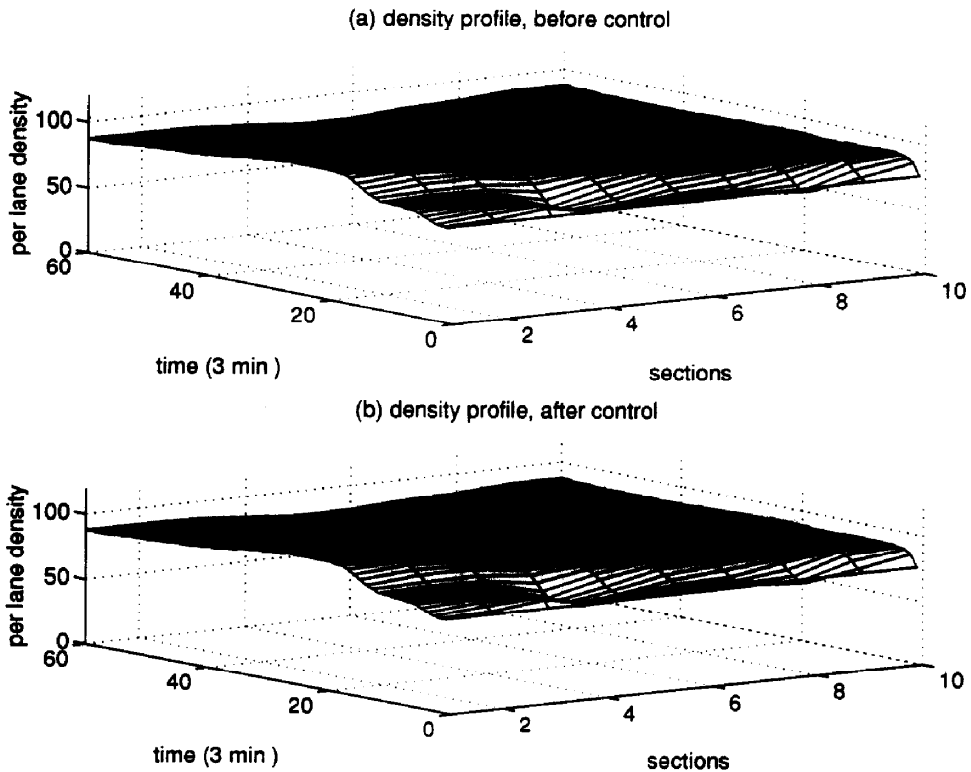**(b) density profile, after control**



Fig. 7. Traffic density, before and after control, uniformly congested traffic.

## 8. SOME EMPIRICAL VALIDATION OF THE THEORETICAL RESULTS

The results presented in the previous sections are based on the regulation of traffic on idealized, hypothetical freeway networks governed by a simplified traffic flow relationship; only superficial consideration was given to the interaction between the freeway and surface street systems and the potential impact of ramp metering strategies on that interaction. In this section, a case study based on existing traffic conditions on a section of a heavily-traveled freeway (and surrounding surface street network) in the Los Angeles area is presented.

Because of the infeasibility of field experimentation, traffic flow is modeled using the computer simulation program INTRAS (INtegrated TRAffic Simulation), developed for the Federal Highway Administration. This program is a microscopic, car-following model which accurately simulates traffic flows while giving the experimenter far more control over the independent variables than would be possible in a real world situation. The model was used to study the behavior of a freeway-arterial system under a variety of parameters including varied traffic volumes, capacity bottlenecks, and different ramp metering strategies. For each case, the effect of the ramp metering on congestion was evaluated.

Selection of the case study site was constrained by the availability of concurrent traffic and control information for both the freeway section and adjacent surface street, including: hourly freeway and ramp volumes, and turn movement counts and associated signal timing information for adjacent intersections. The case study site selected for evaluations consisted of the 3000 ft section of the westbound I-210 freeway between Hill Street and the entry ramp west of Fair Oaks (mileposts 25.140L to 26.822L) and the surface street network bounded by these two streets on the west and east, respectively, and by Villa and Walnut on the north and south, respectively, all within the City of Pasadena, California (Fig. 8). This section of the westbound mainline freeway has five lanes, three entry ramps off of Maple (R1 west of Hill, R2 east of El Molino, and R3 west of Fair Oaks), and two exit ramps to Maple (one near Wilson, the other near Euclid). Peak hour (6.30–7.30 a.m.) volumes for the westbound section of I-210, as determined by Caltrans District 7 for 8 January 1989, are shown in Table 1, for the four 15-min periods comprising the peak hour. The associated entry ramp volumes for the peak hour are shown in Table 2.

These latter figures were derived from data collected by the City of Pasadena, as were similar data for the exit ramps and surface streets (not shown for brevity). Based on discussions with the local Caltrans District, the range of ramp metering headways considered was from 4 to 24 s. Simulations were prepared and run for the various combinations of possible metering headways on the three entry ramps; to keep the number of runs manageable, the range of headways was transformed into five discrete values (4, 8, 12, 20 and 24 s) plus a base case representing no metering. Simulations were run for the four 15-min periods comprising the peak hour, and results were obtained both for the full hour as well as for
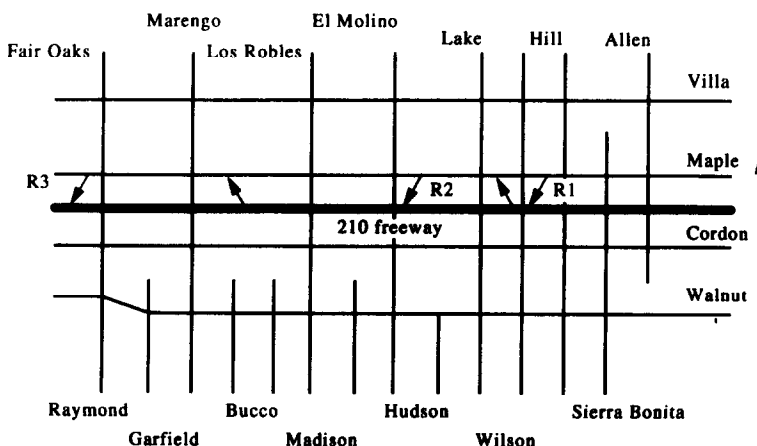


Fig. 8. Case study site.

| Table 1. Peak hour mainline volumes | | |
| --- | --- | --- |
| Period (VPH) | Volume | Nominal D/C |
| 1 | 8064 | 0.81 |
| 2 | 8860 | 0.89 |
| 3 | 9684 | 0.97 |
| 4 | 10,416 | 1.04 |

| Table 2. Peak hour entry ramp volumes | |
| --- | --- |
| Ramp | Volume (VPH) |
| R1 | 408 |
| R2 | 464 |
| R3 | 220 |

each 15-min period; these latter breakdowns were used to estimate the effects of various levels of mainline demand on the efficacy of ramp metering.

Results on the freeway delay for the peak hour simulations are summarized in Fig. 9, which presents the ratio of mainline freeway delay for a given combination of metering headways on R1 and R2 to the delay experienced when the meters are not operating (base case). (For simplicity, results are presented only for a case in which the metering headway on ramp R3 is held constant at 4 s; because of the relatively low volume on ramp R3, the results are generally not sensitive to metering headways on this ramp.) The results demonstrate the general insensitivity of freeway delay to the presence and operation of the ramp meters. Although a portion of the variation evidenced in Fig. 9 can be attributed to stochastic elements within the simulation model, there appears, none the less, to be a trend toward slight improvement with certain combinations of ramp metering rates on R1 and R2. The apparent anomaly of higher delays with certain metering strategies than what with the unmetered base case can be attributed principally to stochastic "noise" and the necessity of metered vehicles to accelerate to freeway speed. Ostensibly, the combination of light metering rates on both R1 and R2 is not sufficient to produce measurable gains in performance of the mainline freeway, although it does result in measurable acceleration delay; anomalies at heavy rates are probably more indicative of stochastic effects in the model.

Although freeway conditions in the case study have been shown to be relatively unaffected by the range of ramp metering strategies examined, the same cannot be said for the impact of metering strategies on the entry ramps and on the adjacent surface street network. A picture of potential adverse effects of heavy metering on the total network (i.e. freeway and adjacent surface streets) is given in Fig. 10. These results displayed in this figure indicate that, for virtually all combinations of metering strategies in the case study, the tradeoff between any gains in freeway performance and adverse effects on the surface street network results in an overall degradation of network performance. There are, however, a few strategies that result in relative improvement to the total network, notably the metering combinations of R1 = 12 s with R2 = 8 s, R1 = 16 s with R2 = 12 s, and R1 = 20 s with R2 = 4 s. These results are based on the aggregate peak hour simulation;
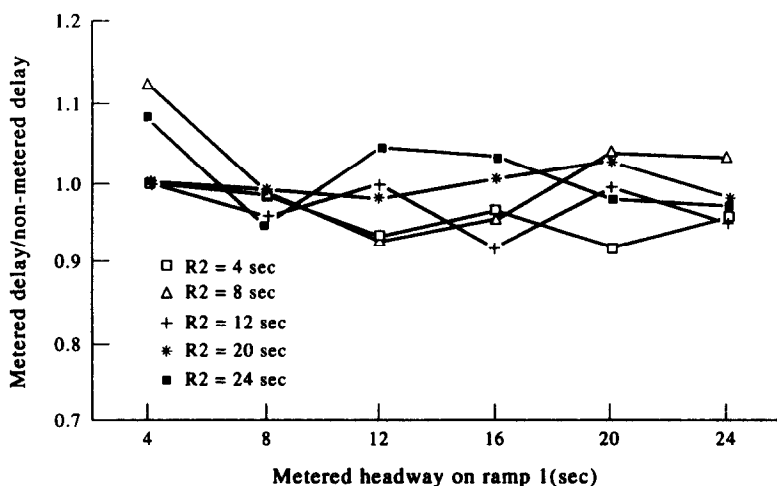


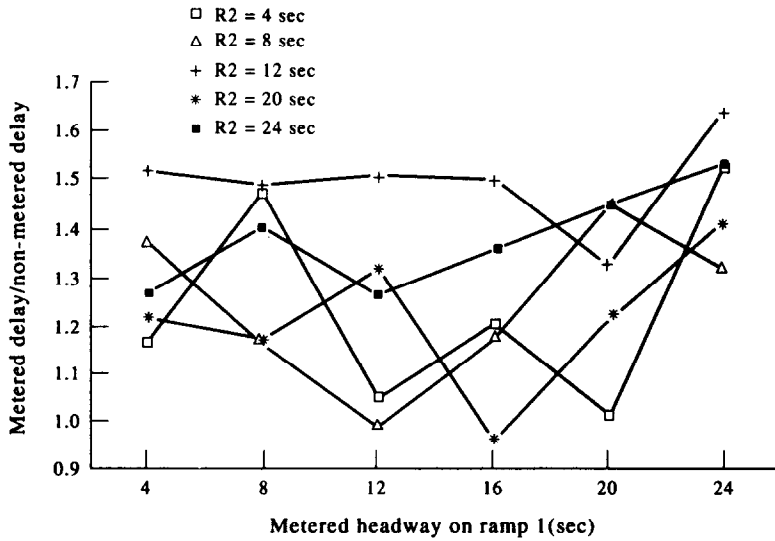Fig. 9. Normalized freeway delay relative to unmetered case.

Fig. 10. Normalized total system delay relative to unmetered case.

the sensitivities of total network delay to the full range of metering strategies for two 15-min periods within the peak hour (the first and third) are shown in Figs 11 and 12. Under the relatively light traffic conditions of the first 15-min period (Fig. 11), virtually all metering strategies result in total network improvement. Conversely, for the third period in which mainline freeway demand approaches the nominal capacity, metering generally results in a substantial degradation in total network performance (Fig. 12).

The results of the case study analysis generally support the conclusions reached in the analysis of the idealized optimal control cases. For the existing ramp volumes in the case study, ramp metering strategies apparently have little impact on freeway performance for conditions in which the nominal mainline D/C is greater than about 0.80. (Conclusions for D/C less than 0.80 are not advanced since the data did not support such analysis.) Such strategies, however, generally can have a deleterious impact on the surface street network performance for conditions in which the mainline freeway demand approaches nominal capacity, as a result of ramp queue buildup and spillback. It is emphasized that these conclusions are restricted
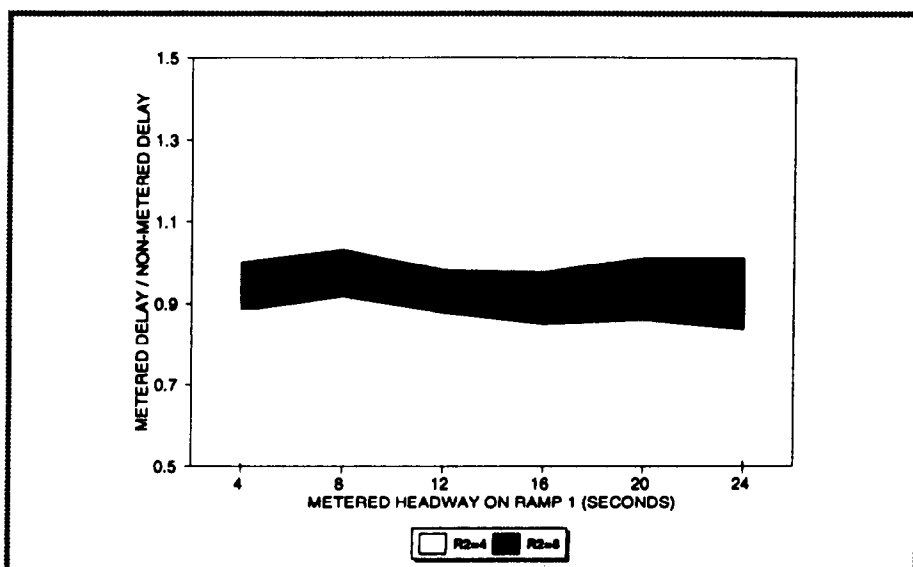


Fig. 11. Sensitivity of total network delay to ramp metering rate on Ramp 1 for D/C = 0.81.
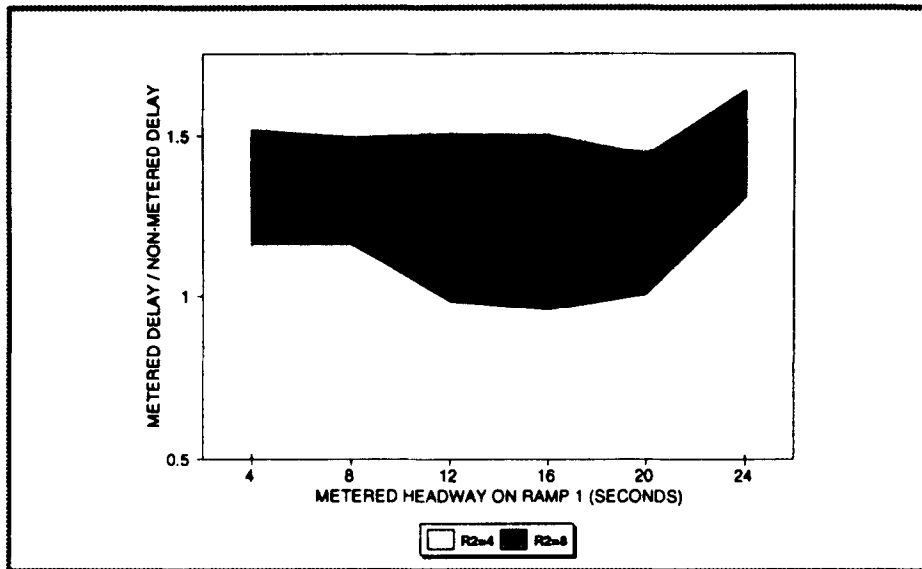
Fig. 12. Sensitivity of total network delay to ramp metering rate on Ramp 1 for D/C = 0.97.

by the assumption of ramp demand being invariant with metering rates; the impact of metering rates on diversion was not considered in this case study.

## 9. CONCLUSIONS

In this paper, we considered a freeway system that consists of a freeway section and its entry/exit ramps, and formulated the ramp control problem as a dynamic optimal process to minimize the total time spent in this system. Under the assumption that the controlled freeway has to serve all its demand, and the traffic flow process follows the rules prescribed by the LWR theory, we are able to show that when traffic is uniformly congested, or uniformly uncongested, ramp metering leads to inferior solutions to the problem. However, it is effective in reducing the total time spent in the system when traffic conditions have the potential to switch between congested and non-congested situations. Numerical methods have been provided to solve the optimal ramp control problem when ramp metering is beneficial. Results of a case study involving an actual freeway system, using a detailed microscopic simulation model, support the theoretical conclusions.

## REFERENCES

Blumentritt C. W., Pinnell C., McCasland W. R., Ross D. W. and Glazer J. (1981) Guidelines for Selection of Ramp Control Systems. NCHRP Report 232, Transportation Research Board, Washington, DC.

Bryson A. E. and Ho Yu-Chi (1975) *Applied Optimal Control*. John Wiley & Sons, New York.

Chen C.-I., Cruz J. B. Jr and Paquet J. G. (1974) Entrance Ramp Control For Travel Rate Maximization in Expressways. *Transpn. Res.* 8, 503–508.

Daganzo C. F. and Lin Wei-Hua (1994) The Spatial Evolution of Queues During the Morning Commute in a Single Corridor. Preprints, No. 940115, Transportation Research Board, Washington, DC.

Goldstein N. B. and Kumar K. S. P. (1982) A Decantralized Control Strategy for Freeway Regulation. *Transpn Res.-B* 16, 279–290.

Greenshields B. D. (1934) A Study of Traffic Capacity. *Proc. Highway Res. Bd.* 14, 448–477.

Isaksen L. and Payne H. J. (1973) Suboptimal Control of Linear Systems by Augmentation with Application to Freeway Traffic Regulation. *IEEE Trans. Automatic Control* 18, 210–219.

*ITE Traffic Control Systems Handbook ITE* (1985) (Edited by Wilshire R., Black R., Grochoske R. and Higinbotham J.). Institute of Transportation Engineers, pp. 4.18–4.23.

Lighthill M. J. and Whitham G. B. (1955) On Kinematic Waves: II. A Theory of Traffic Flow on Long Crowded Roads. *Proc. R. Soc. Lond. Ser. A.* 229, 317–345.

Luenberger D. (1984) *Linear and Nonlinear Programming*, Second Edition. Addison-Wesley, Reading, Massachusetts.

Lu, Jinfu and Guan, Zhi (1989) *Numerical Methods for Partial Differential Equations* (in Chinese). Qinghua University Press, Beijing.

Newell G. F. (1989) Comments on Traffic Dynamics. *Transpn Res.-B* **23**, 386–389.

Newell G. F. (1991) A Simplified Theory of Kinematic Waves: I General Theory; II Queuing at Freeway Bottlenecks; III Multi-destination Flows. *Transpn Res.-B* **27**, 281–313.

Newman L. *et al.* (1969) Freeway Ramp Control — What It Can and Cannot Do. *Traffic Engng* **39**, 14–21.

Papageorgiou M. (1983) Applications of Automatic Control Concepts to Traffic Flow Modeling and Control, Lecture Notes in Control and Information Sciences, (Balakrishnan A. V. and Thoma M., Eds). Spring, Berlin, 94–100.

Papageorgiou M., Blosseville J.-M. and Habib H.-S. (1990) Modeling and Real-Time Control of Traffic Flow on the Southern Part of Boulevard Peripherique in Paris: Part II: Coordinated On-Ramp Metering. *Transpn Res.-A* **24**, 361–370.

Papageorgiou M., Habib H. S. and Blosseville J. M. (1991) ALINEA: A Local Feedback Control Law for On-ramp Metering. *Transpn Res. Rec.* **1320**, 58–64.

Payne H. J., Thompson W. A. and Isaksen L. (1973) Design of Traffic Responsive Control System for a Los Angeles Freeway, *IEEE Trans. Systems, Man, and Cybernetics* **smc-3**, 213–224.

Ran B., Boyce D. E. and LeBlanc L. J. (1993) A New Class of Instantaneous Dynamic User-Optimal Traffic Assignment Models. *Operat. Res.* **41**, 192–202.

Richards P. I. (1956) Shockwaves on the highway. *Operat. Res.* **4**, 42–51.

Transportation Research Board (1975) Better Use of Existing Transportation Facilities. TRB Special Report 153, Washington, DC.

Wang J. J. and May A. D. (1973) Computer Model for Optimal Freeway On-Ramp Control. *Highway Res. Rec.* **469**, 16–25.

Wattleworth J. A. and Berry D. S. (1965) Peak-period Control of a Freeway System — Some Theoretical Investigations. *Highway Res. Rec.* **89**, 1–25.

Yuan L. S. and Kreer J. B. (1971) Adjustment of Freeway Ramp Metering Rates to Balance Entrance Ramp Queues. *Transpn Res.* **5**, 127–133.

Zhang H., Ritchie S. and Lo Z. (1994) A Local Neural Network Controller for Freeway Ramp Metering. *Int. Fed. Automatic Control, Proc. 7th Symp. Transpn Sys. Theory and Applications Adv. Technol.* August 1994, Tianjin, China.

## APPENDIX

In this appendix, we derive the stability requirement for the finite difference eqn (8) with the triangular fundamental diagram. The concept of stability for a numerical scheme is concerned with the growth or decay of errors introduced at any stage of the computation, and there exists a number of methods to study the stability of various types of numerical schemes (Lu and Guan, 1989). We adopt the Fourier method in this analysis.

Notations:

  $i$:      freeway section index,

  $k$:      time index,

  $j$:      the complex number $\sqrt{1}$,

  $\Delta$:      length of each freeway section,

  $T$:      length of discrete time interval,

  $vf$:     free-flow speed,

  $c_0$:     constant wave speed when traffic density is greater than critical density,

  $s_i^k$:     the ratio of exit ramp flow rate to freeway flow rate of section $i$ at time $k$, $0 \le s_i^k \le 1$,

  $S$:      shift operator, e.g. $S\rho_i^k = \rho_{i+1}^k$,

  $\rho_i^k$:     the density of section $i$ at time $k$ with computation error,

  $*\rho_i^k$:    the density of section $i$ at time $k$ without computation error,

  $\epsilon_i^k$:     computation error at time $k$. $\epsilon_i^k = \rho_i^k - *\rho_i^k$,

  $\hat{\epsilon}_i^k$:     the Fourier transform of $\epsilon_i^k$, defined as $\hat{\epsilon}_i^k = \int_{-\infty}^{\infty} \epsilon_i^k e^{-j\omega(i\Delta)} d\omega$.

Examine the following difference equation:

$$\rho_i^{k+1} = [b_{-1}S^{-1} + b_0 + b_1 S]\rho_i^k, \tag{a1}$$

where $b_{-1}$, $b_0$, $b_1$ are constants.

It is clear that $\epsilon_i^k$ also satisfies eqn (a1), i.e.

$$\epsilon_i^{k+1} = [b_{-1}S^{-1} + b_0 + b_1 S]\epsilon_i^k. \tag{a2}$$

Taking the Fourier transform on both sides of (a2), we have

$$\hat{\epsilon}_i^{k+1} = (b_{-1}e^{-j\omega\Delta} + b_0 + b_1 e^{j\omega\Delta})\hat{\epsilon}_i^k. \tag{a3}$$

Now, let $G(\Delta,\omega) = b_{-1}e^{-j\omega\Delta} + b_0 + b_1 e^{j\omega\Delta}$, we have

$$\hat{\epsilon}_i^{k+1} = G(\Delta,\omega)\hat{\epsilon}_i^k. \tag{a4}$$

Since $G(\Delta,\omega)$ is independent of time step $k$, we can write

$$\hat{\epsilon}_i^k = [G(\Delta,\omega)]^k \hat{\epsilon}_i^0. \tag{a5}$$

For the error to decay over time $kT$, it is necessary and sufficient to require

$$|G(\Delta, \omega)| \le 1, \text{ for all } \omega. \tag{a6}$$

We now examine eqn (8) when traffic is uniformly uncongested, where we have

$$b_1 = \frac{\Delta}{T} v_f \equiv \alpha \tag{a7}$$

$$b_0 = 1 - (1 + s_i^k)\frac{\Delta}{T} v_f = 1 - (1 + s_i^k)\alpha \tag{a8}$$

$$b_{-1} \equiv 0. \tag{a9}$$

Substituting these expressions into $G(\Delta, \omega)$, we have

$$G(\Delta, \omega) = 1 - (1 + s_i^k)\alpha + \alpha e^{-j\omega\Delta}$$
$$= 1 - (1 + s_i^k)\alpha + \alpha \cos(\omega\Delta) - j\alpha \sin(\omega\Delta),$$

and

$$|G(\Delta, \omega)|^2 = [1 - (1 + s_i^k)\alpha + \alpha \cos(\omega\Delta)]^2 + [\alpha \sin(\omega\Delta)]^2$$
$$= [(1 + s_i^k\alpha) - \alpha(1 - \cos(\omega\Delta))]^2 + [\alpha \sin(\omega\Delta)]^2$$
$$= \left[(1 + s_i^k\alpha) - 2\alpha \sin^2\left(\frac{\omega\Delta}{2}\right)\right]^2 + 4\alpha^2 \sin^2\left(\frac{\omega\Delta}{2}\right)\cos^2\left(\frac{\omega\Delta}{2}\right)$$
$$= (1 + s_i^k\alpha) - 4\alpha \sin^2\left(\frac{\omega\Delta}{2}\right)[1 - (1 + s_i^k)\alpha].$$

Clearly, if $1 - (1 + s_i^k)\alpha \geq 0$, i.e. $(1 + s_i^k)\alpha \leq 1$ for all $\alpha$ and $s_i^k$, then $|G(\Delta, \omega)| \leq 1$, and (8) is stable when traffic is uniformly uncongested.
Similarly, when traffic is uniformly congested, one can show that if $(1 - s_i^k)\alpha' \leq 1$, for all $\alpha'$ and $s_i^k$, where $\alpha' = \frac{T}{\Delta} c_0$, then (8) is stable.
In choosing a more stringent condition, we conclude that if $(1 + s_i^k)$ mad $\{\alpha, \alpha'\} \leq 1$ for all $\alpha'$ and $s_i^k$, then (8) is stable under all traffic conditions.