Pergamon

Transpn Res.-C, Vol. 4, No. 4, pp. 181–210, 1996

0968-090X/96 $15.00+0.00

PII: S0968-090X(96)00011-3

A THEORY OF TRAFFIC FLOW IN AUTOMATED HIGHWAY SYSTEMS[1]

M. BROUCKE* and P. VARAIYA
Department of Electrical Engineering and Computer Science, University of California, Berkeley, CA 94720 mire, USA

(Received 25 August 1995; in revised form 29 June 1996)

Abstract—This paper presents a theory for automated traffic flow, based on an abstraction of vehicle activities like entry, exit and cruising, derived from a vehicle's automatic control laws. An activity is represented in the flow model by the space and time occupied by a vehicle engaged in that activity. The theory formulates Traffic Management Center (TMC) plans as the specification of the activities and velocity of vehicles, and the entry and exit flows for each highway section. We show that flows that achieve capacity can be realized by stationary plans that also minimize travel time. These optimum plans can be calculated by solving a linear programming problem. The theory permits the study of transient phenomena such as congestion, and TMC feedback traffic rules designed to deal with transients. We propose a "greedy" TMC rule that always achieves capacity but does not minimize travel time. We undertake a microscopic study of the "entry" activity, and show how lack of coordination between entering vehicles and vehicles on the main line disrupts traffic flow and increases travel time. We conclude by giving some practical indication of how to obtain the space and time usage of activities from vehicle control laws. Finally, we illustrate the concepts presented in this paper with two examples of how the model is used to calculate the capacities of a one-lane automated highway system. In one example we study market penetration of adaptive cruise control and in the second example we study the effect of platooning maneuvers in a platooning architecture for AHS. Copyright © 1996 Elsevier Science Ltd

1. INTRODUCTION

This paper proposes a theory of automated highway traffic. The theory predicts the performance of an automated highway system (AHS) in terms of achievable (steady state) flows and travel times. The performance predictions can be used to compare alternative AHS designs.

The theory shows how AHS steady state performance is a function of the characteristics of both the control laws that govern the movement of individual vehicles and the traffic management rules that guide the vehicle flow. This functional relationship can be used to suggest changes in vehicle control laws and traffic management rules for improving highway performance.

The theory also explains how the automated highway can become congested, and what sorts of actions need to be taken to prevent congestion from occurring and to eliminate it once it occurs. Thus the theory may be used to design vehicle control and traffic management rules for reducing undesirable transient behavior such as congestion.

Vehicles in an AHS are under automatic control: the distance a vehicle maintains from the vehicle in front, its velocity, and its route from entry into the highway to exit, are all determined by the vehicle's feedback control laws. One may therefore compare the effect on the traffic of changes in vehicle control laws, and seek to calculate the "optimum" control laws. By contrast, in non-automated traffic flow theory, the driver determines a vehicle's headway, its velocity, its movement during a merge, etc. Driver behavior

[1]This research is supported by the California Department of Transportation through the California PATH program and the Federal Highway Administration through the National Automated Highway System Consortium.

is difficult to change significantly. One hypothesizes feedback models of driver behavior and uses real data or experiments to calibrate the model parameters.

Similarly, the Traffic Management Center (TMC) for the AHS can directly influence the flow by issuing orders to vehicles regarding their velocity and route. Those orders will be obeyed because the vehicles are programmed to do so. The TMC for the non-automated highway can also make speed and route suggestions, but drivers may ignore these suggestions or react to them in an unexpected manner. Thus, the influence of TMC policies in the AHS is much stronger and more predictable than its influence on non-automated traffic; and so, one may again seek to determine optimum TMC policies.

Because it is possible to exercise much greater control over the movement of individual vehicles and the traffic as a whole, a theory of AHS traffic flow will tend to be *prescriptive*. Non-automated traffic flow theory is more *descriptive*, by contrast.[2] The following notation will be used in this paper:

Symbol	**Interpretation**
$s(\alpha)$	space used by activity α in m (meters)
$t(\alpha)$	duration of activity α in s (seconds)
α	a vehicle activity
$L(i)$	length of section i in m
t	time index
τ	time period
i	section index
θ	flow type
η	vehicle body type
V	maximum permissible velocity
$v(i,t)$	average velocity of vehicles in section i at time t
$n(i,t,\theta)$	number of vehicles in section i, time t, of flow type θ
$\pi(\alpha,i,t,\theta)$	proportion of vehicles performing activity α in section i at time t of flow type θ
$\lambda(\alpha)$	space–time usage of activity α in m-s
$f(i,t,\theta)$	number of vehicles entering section i at time t of flow type θ
$g(i,t,\theta)$	number of vehicles exiting section i at time t of flow type θ
$\rho(i,t)$	fraction of vehicles in section i at time t that remain in the section at time $t+1$.
$u(t)$	the TMC plan consisting of an activity, velocity, entry and exit plan
$F(T,\theta)$	average input flow over time $t=0,\ldots T$
$G(T,\theta)$	average output flow over time $t=0,\ldots T$
$\underline{f}(i,\theta)$	time-averaged input flow in section i of type θ
$\underline{g}(i,\theta)$	time-averaged output flow in section i of type θ
ϕ	a flow in vehicles/s
$\underline{\phi}(i,\theta)$	average flow of type θ from section $i-1$ to section i
$\underline{v}(i,t)$	stationary speed of section i
$\underline{n}(i,t,\theta)$	stationary number of vehicles in section i at time t of flow type θ
$\underline{\pi}$	stationary activity plan
$\bar{F}(\theta)$	achievable input flows
$G(\theta)$	achievable output flows
$N(i)$	maximum number of vehicles in section i
$\lambda(i)$	average space–time used per vehicle in section i
$n(i,t)$	number of vehicles in section i at time t of all flow types
$\bar{\phi}$	maximum flow out of section i in vehicles/s
ϕ^*	minimum of the maximum flow out of any section
S	gap required for the entry activity in the receiving lane, in m
x_k	kth free space gap, in m

[2]Of course, this descriptive theory is used to design and prescribe ramp metering and other traffic management rules.

D	inter-platoon gap, in m
d	intra-platoon gap, in m
l	vehicle length, in m
n	platoon size
Q	longitudinal sensor range, in m

2. GENERAL ASSUMPTIONS

We introduce the main abstractions and assumptions and the structure of the proposed theory. The theory is based on an *activity* model: the movement of a vehicle is conceptualized as a sequence of activities, such as entry, cruise, and exit, that are realized by vehicle control laws; the highway is viewed as providing the space necessary to carry out each activity; the vehicle control laws and vehicle speed determine the time to complete an activity.[3]

When there is insufficient space in one section of the highway, the rate of activity completion in the section immediately upstream must be reduced. Since the rate of activity completion is proportional to the speed, this causes a reduction in flow.

In this way, the interaction between the demand for space by vehicle activities and the fixed supply of space offered by the highway determines the steady state flows that can be realized, as well as the transient congestion effects that can occur. This interaction is mediated by the vehicle control policies (which determine the space needed for each activity) and the traffic management rules (which determine the activities that are to be carried out in different sections of the highway). That is how the theory relates AHS performance to characteristics of vehicle control and traffic management rules.

We now introduce the main assumption which we call "safety needs space," that binds together activities, vehicles and highway.

To fix ideas, we assume that the AHS has a single lane, with entrances and exits. At each instant of time, every (automated) vehicle is engaged in one of a finite number of activities such as cruising, changing a lane (in case of a multi-lane highway), entering the highway, exiting the highway, etc. If vehicles are organized in closely-spaced platoons, then cruising in a one-vehicle platoon is a different activity from cruising in a two-vehicle platoon, and so on. Cruising in platoons of different sizes are considered different activities because the space needed per vehicle in a cruising platoon decreases with the platoon size. (See Varaiya, 1993.)

The highway is divided into sections, and we will assume that a vehicle executes a single activity in each section through which it travels. Consequently, the passage of a vehicle through the automated highway can be summarized by the sequence of activities that the vehicle executes, starting with the "entry" activity in the section where it enters and terminating with the "exit" activity in the section where it leaves the highway. In this model, vehicles are assumed to travel at a constant average speed within each section, and an assumption of "one activity per section" can be used to tie the spatial discretization of the highway into sections with the temporal discretization of movement into activities. Consequently, variation in speed due to interaction of activities is not captured here. Although not mathematically necessary, we adopt the one-activity-per-section assumption to simplify the model description. (See Daganzo, 1994 for a related modeling move to tie together spatial and temporal discretization.)

While it is engaged in a particular activity, a vehicle's motion is governed by a feedback control law which ensures that this activity is carried out safely. These feedback laws and the resulting vehicle motion can be complicated.[4] But for our purposes we will work with the assumption "safety needs space."

To motivate this assumption, consider the "cruising" activity, in which a vehicle keeps in one lane and its cruise control law guarantees safety by maintaining a minimum

[3]This activity model is inspired by the work in Hall, 1995.

[4]Examples of such feedback laws are given in Sheikholeslam and Desoer, 1990; McMahon *et al.*, 1990; Peng and Tomizuka, 1990; and Frankel *et al.*, 1995.

safe distance between its vehicle and the vehicle in front of it. This distance is an increasing function of vehicle speed.[5] We shall assume a maximum permissible speed and let $s(cruise)$ be the corresponding minimum safe distance between a cruising vehicle and the vehicle in front of it. Thus the safety-needs-space assumption says that its feedback law will guarantee that a cruising vehicle will "occupy" $s(cruise)$ meters of a highway lane for a duration $t(cruise)$.

In general, safety-needs-space says that vehicle control laws cause a vehicle engaged in activity α to occupy a distance $s(\alpha)$ from which, for a specified duration $t(\alpha)$, all other vehicles are excluded. For activities involving vehicles in two lanes, as happens during a lane change and in some implementations of entry/exit, the vehicle occupies a minimum safety distance in both lanes.

The time the vehicle spends in a section is equal to the section length divided by the vehicle velocity. When a vehicle engaged in activity α leaves this section, its $s(\alpha)$ space is available for use by another vehicle from the upstream section. The longer the vehicle stays in its section, the later will its space become available, and this may slow down upstream vehicles. Thus, if the activities that vehicles are executing in different sections are not well coordinated, the speed in some sections may be forced below the maximum or free flow speed, causing congestion. Traffic management rules determine the activities that vehicles undertake and their speed, and thus, ultimately, the AHS steady state performance as well as how well congestion is dissipated.

The remainder of the paper is organized as follows. In section 3 we introduce the formal activity model. This is a system of differential equations, several parameters of which are set by TMC plans, including vehicle speed and activity, and entry and exit flows.

TMC plans and achievable flows are studied in section 4. An achievable flow is any vector of flows (indexed by origin–destination pairs or other characteristics) that can be sustained in the long run. The main result of this section is that the set of achievable flows is convex.

In section 5 we define AHS capacity as the set of undominated achievable flows, and efficient TMC plans as those which minimize travel time. We show that every undominated flow, together with an efficient plan that achieves this flow, can be computed by solving a linear programming problem.

In section 6 we consider transient behavior: how congestion can develop and how TMC feedback rules can mitigate its effects. We exhibit a "greedy" rule that is easy to implement and always achieves capacity, but does not minimize travel time.

In section 7 we focus on two particular activities—entry and exit. These activities are likely to be the most important in limiting AHS performance. In section 8 we discuss the substantive modeling questions of how to define an activity and how to compute the amount of space–time an activity needs. In section 9 we compare two alternative AHS designs using the proposed theory. Finally, section 10 collects some concluding remarks.

3. THE ACTIVITY MODEL

We study a one-lane automated highway, divided into sections. Sections are indexed $i = 1,\dots, I$; section i is $L(i)$ m in length. Section $i-1$ is upstream of section i. Time is indexed $t = 0, 1, \dots$. Each time period is τ seconds long.

3.1. Vehicles

Vehicles have types indexed by θ which may stand for their body type (passenger, truck, bus), origin and destination and any other distinguishing characteristics of interest.

[5]This function depends on other parameters such as maximum vehicle braking torque, road surface and tire conditions, etc.

All vehicles in section i at time t have the same velocity, denoted $v(i,t)$, and measured in m/s. It is required that $v(i,t) \leq V$, the maximum permissible or free flow speed. (V, too, may be indexed by i, but we don't do that to ease the notational burden.)

Let $n(i,t,\theta)$ be the number of vehicles of type θ in section i at time t. We adopt the notational convention that $n(i,t)$ is the array indexed by θ, $n(i)$ is the array indexed by (t,θ), and so on.

3.2. *Activity plan*

There are finitely many activities, indexed by α. An *activity plan* is any array of non-negative numbers $\pi = \{\pi(\alpha,i,t,\theta)\}$ such that for every i,t,θ

$$\sum_{\alpha} \pi(\alpha, i, t, \theta) \equiv 1.$$

$\pi(\alpha,i,t,\theta)$ is the fraction of the $n(i,t,\theta)$ vehicles engaged in activity α.

Associated with each activity α is the space–time (in m-s) $\lambda(\alpha) > 0$ of the section occupied by each vehicle engaged in that activity. Thus $n(i,t)$ vehicles engaged in activities $\pi(i,t)$ will occupy

$$\sum_{\alpha} \sum_{\theta} \lambda(\alpha)\pi(\alpha, i, t, \theta)n(i, t, \theta)$$

m-s of section i in a period τ.

Two vehicles with the same (i,t,θ) index and engaged in the same activity cannot be further distinguished within the model. In that sense, this is a theory of vehicle *flow*. The theory aggregates individual vehicle movement through the use of activities.

3.3. *Velocity plan*

A *velocity plan* is an array of nonnegative numbers $v = \{v(i,t)\}$ (in m/s), each less than V. All $n(i,t)$ vehicles move at $v(i,t)$ m/s to conform to the plan. This restriction, in part, is imposed by the single lane highway: since vehicles cannot pass each other, relative velocities cannot be too great. However, the restriction also presupposes that the sections are not so long that vehicles with significantly different speeds can coexist in the same section.

It is possible, at the cost of further notational complexity, to introduce the following features. Suppose the vehicle type θ also signifies vehicle body type: light duty, truck, bus, etc. Then we can insist that the space–time required depends also on vehicle type, i.e. we have $\lambda(\alpha,\theta)$. We can also insist that vehicle maximum velocity is a function of $\theta, V(\theta)$, and require that the velocity $v(i,t)$ be smaller than the maximum permissible velocity, i.e. $n(i,t,\theta) > 0$ implies $v(i,t) \leq V(\theta)$. These features are very useful and easy to introduce in the simulation system, but they would make this paper difficult to read.

3.4. *Highway configuration*

We have already specified parts of the highway configuration. We have a one-lane highway, divided into sections $i = 1,\ldots, I$ of length $L(i)$. Section i is immediately downstream of section $i-1$. Each section has at most one entrance and one exit. Vehicles can make an entry through some dedicated infrastructure that connects a non-automated highway or street to the AHS entrance. Vehicles can exit the AHS through another transitional infrastructure.[6] We can require that an entering vehicle must engage in a distinguished "entry" activity, and an exiting vehicle must engage in "exit". These activities will occupy more space–time than most other activities because they will involve merging from a ramp or a transition lane into the main AHS lane.

[6]See Godbole *et al.*, 1995 for several transitional infrastructure designs, and Rao and Varaiya, 1994 for a similar highway configuration.

In a following paper we will extend the model to a multi-lane AHS. Such an extension then permits one to consider the "lane change" activity. It also permits the possibility of modeling entry and exit as a kind of lane change.

3.5. *Entry and exit plans*

An *entry plan* is an array $f = \{f(i,t,\theta)\}$ of non-negative numbers. $f(i,t,\theta)$ is the number of vehicles of type θ that enter the highway in section i in period t.

An *exit plan* is an array $g = \{g(i,t,\theta)\}$ of non-negative numbers. $g(i,t,\theta)$ is the number of vehicles of type θ that exit the highway in section i in period t.

If entry or exit in a particular section, say j, is forbidden, one merely adds the constraint: $f(j,t,\theta)\equiv 0$ or $g(j,t,\theta)\equiv 0$, for all t,θ. We will shortly impose more complex constraints on all the plans.

3.6. *Dynamics*

The *state* of the system at time t is $n(t) = \{n(i,t,\theta)\}$. Suppose that we are given an activity plan π, a velocity plan v, an entry plan f, and an exit plan g. Let $n(t)$ be the state at time t. Then, for all t and $1 \leq i \leq I$,

$$n(i, t+1, \theta) = \rho(i, t)n(i, t, \theta) + [1 - \rho(i-1, t)]n(i-1, t, \theta) + f(i, t, \theta) - g(i, t, \theta). \quad (1)$$

Since the AHS sections are $i = 1,\ldots, I$, we also have the boundary conditions,

$$n(0, t, \theta) = 0, \text{ for all } t, \theta, \quad (2)$$

$$n(I+1, t, \theta) = 0, \text{ for all } t, \theta. \quad (3)$$

Equation (1) should be interpreted as follows. First, by definition,

$$1 - \rho(i, t) := \frac{v(i, t) \times \tau}{L(i)}. \quad (4)$$

Here $\rho(i,t)$ is the fraction of vehicles in section i at time t that remain in that section for time $t+1$. Thus, $[1-\rho(i,t)]$ is the fraction of vehicles in section i at time t that leave that section at the end of that period. By definition (4), the fraction of vehicles that leave is equal to the fraction of the section length $L(i)$ that is traveled in time τ by vehicles moving at velocity $v(i,t)$. Thus this definition assumes a spatial homogeneity of the disposition of vehicles in each section. Obviously this is not the case at the level of individual vehicles. But in our model, a homogeneity assumption of this kind is necessary since we want the state simply to be the number of vehicles in each section.[7]

Thus, the first term on the right in (1) is the number of vehicles in i at time t that remain in i at time $t+1$, and the second term is the number of vehicles in $i-1$ at time t that move into i at time $t+1$. The last two terms are straightforward: $f(i,t,\theta)$ is the number of vehicles of type θ that enter the AHS according to the entry plan, and $g(i,t,\theta)$ is the number that leave the AHS.

The boundary condition (3) implies that all vehicles in section I leave the AHS:

$$g(I+1, t, \theta) = [1 - \rho(I, t)]n(I, t, \theta), \; f(I+1, t, \theta) = 0. \quad (5)$$

Fact 1. $n(t)$ is indeed a state, i.e. given $n(0)$ and activity, velocity, entry and exit plans $u(t = [\pi(t), v(t), f(t), g(t)]$, $t \geq 0$, there is a unique state trajectory $n(t)$, $t \geq 0$, that satisfies (1)–(4).

[7] Equation (4) also ties together the time and space discretization parameters τ and $L(i)$. Since the maximum velocity is V, the maximum value of the right hand side of (4) is $V\times\tau/L(i)$. This ratio must be less than one. Typical values are $V = 25$ m/s, $\tau = 10$ s, and $L(i) = 500$ m.

4. ACHIEVABLE FLOWS AND A STATIONARY TMC PLAN

In this section we will specify the constraints that a flow must satisfy in order to be a feasible solution of (1). We define what stationary, or time-invariant flows are achievable and construct a TMC plan that can realize the achievable flows.

We will call $u(t) = [\pi(t),v(t),f(t),g(t)]$, $t \geq 0$, a TMC *plan*. By choosing this plan, the TMC controls the traffic flow. In this section we study the flows and throughput that TMC plans can achieve.

4.1. Feasibility constraint

A trajectory-plan $(n(t),u,(t))$ must satisfy two physical constraints

$$n(i, t, \theta) \geq 0, \tag{6}$$

$$\sum_{\alpha}\sum_{\theta} \pi(\alpha, i, t, \theta)n(i, t, \theta)\lambda(\alpha) \leq L(i)\tau. \tag{7}$$

The non-negativity requirement (6) is clear. Constraint (7) expresses the requirement that there is enough space and time in the section over the period τ to safely carry out the activities assigned by the plan.

There are, in addition, three constraints dealing with entry and exit. First, vehicles of certain types may not be allowed to enter or exit from certain sections. This constraint is of the form

$$f(i, t, \theta) \equiv 0, \text{ or } g(i, t, \theta \equiv 0,$$

for all t and for specified values of i,θ.

Second, suppose that a vehicle's body type, entry and exit are encoded in its type, i.e. θ is of the form $\theta = (\eta,j,k)$ where η is the body type, j is the entry section and k is the exit section. Then vehicles of type (η,j,k) can enter only from section j. That is,

$$f(i, t, (\eta, j, k)) \equiv 0, \ i \neq j.$$

Similarly, vehicles of type (η,j,k) exit only from section k. That is,

$$g(k, t, (\eta, j, k)) = [1 - \rho(k-1, t)]n(k-1, t, (\eta, j, k)),$$

or, equivalently,

$$n(k, t, (\eta, j, k)) \equiv 0.$$

Lastly, we may require that when a vehicle of type (η,j,k) enters, it must first carry out an entry activity. If this activity is labeled α_{in}, the requirement may be expressed as $\pi(\alpha_{in},j,t,(\eta,j,k)) = 1$, or $\pi(\alpha,j,t,(\eta,j,k)) = 0$ for $\alpha \neq \alpha_{in}$. Other maneuver restrictions can be expressed in a similar way.[8]

All these constraints can more generally and more uniformly be expressed by specifying three subsets T_f, T_g and T_n of section-type pairs, and one subset T_π of activity-section-type triples, and the requirement that for all t,

$$f(i, t, \theta) = 0, \text{ for all } (i, \theta) \in T_f, \tag{8}$$

$$g(i, t, \theta) = 0, \text{ for all } (i, \theta) \in T_g, \tag{9}$$

[8]For example, one may require that vehicles of a particular type must execute maneuver α_1 in section i_1, α_2 in section i_2, and so on.

$$n(i, t, \theta) = 0, \text{ for all } (i, \theta) \in T_n, \tag{10}$$

$$\pi(\alpha, i, t, \theta) = 0, \text{ for all } (\alpha, i, \theta) \in T_\pi. \tag{11}$$

We will say that a trajectory-plan (n,u) is *feasible* if the constraints (6)–(11) are satisfied. To prevent trivial cases we will not allow $f(i,t,\theta)$ and $g(i,t,\theta)$ both to be positive, by insisting that every (i,θ) is either in T_f or in T_g.

We note some properties of feasible trajectories that will be used to define achievable flows.

Fact 2. There is a uniform bound which applies to all feasible trajectory-plans.

Proof. From (7), $n(i,t,\theta) \leq L(i)\tau/\min_\alpha\lambda(\alpha)$, i.e. all trajectories are uniformly bounded. From (1) it follows that entry and exit plans must be uniformly bounded. □

Let $(n(t),u(t))$, $t=0,1,\ldots$ be a feasible trajectory–plan. Summing (1) over i, and cancelling some terms, gives

$$\sum_{i=1}^{I+1} n(i, t+1, \theta) = \rho(I+1, t)n(I+1, t, \theta) + \sum_{i=1}^{I+1} n(i-1, t, \theta) + \sum_{i=1}^{I+1} f(i, t, \theta) - \sum_{i=1}^{I+1} g(i, t, \theta).$$

Using the boundary conditions (2), (3) gives

$$\sum_{i=1}^{I} [n(i, t+1, \theta) - n(i, t, \theta)] = \sum_{i=1}^{I} [f(i, t, \theta) - g(i, t, \theta)].$$

Summing over $t=0, 1,\ldots, T-1$ and dividing by T gives

$$\frac{1}{T}\sum_{i=1}^{I} [n(i, T, \theta) - n(i, 0, \theta)] = F(T, \theta) - G(T, \theta),$$

where

$$F(T, \theta) := \frac{1}{T}\sum_{t=0}^{T-1}\sum_{i=1}^{I} f(i, t, \theta),\ G(T, \theta) := \frac{1}{T}\sum_{t=0}^{T-1}\sum_{i=1}^{I} g(i, t, \theta),$$

are, respectively, the average number of vehicles of type θ that enter and leave the AHS during $t=0,\ldots, T-1$. It follows from Fact 2 that

$$\lim_{T\to\infty} F(T, \theta) - G(T, \theta) = 0. \tag{12}$$

Definition. A vector $F = \{F(\theta)\}$ of flows is *achievable* if there is a feasible trajectory-plan and a sequence of times $T_k \to \infty$, such that

$$\lim_{k\to\infty} F(T_k, \theta) = \lim_{k\to\infty} G(T_k, \theta) = F(\theta), \text{ for all } \theta. \tag{13}$$

A feasible trajectory-plan $(n(t),u(t))$, $t=0, 1,\ldots$ is *stationary* if the sequence $(n(t),u(t))$ does not depend on t.

Theorem 1. Every achievable flow can be realized by a stationary plan which, moreover, minimizes travel time.

Proof. Let $(n(t),u(t)=[\pi(t),v(t),f(t),g(t)])$ be a feasible pair and $T_k \to \infty$ such that (13) holds, i.e. the flow $F=\{F(\theta)\}$ is realized. We will construct a stationary pair, $(\underline{n}, \underline{u})$ which realizes F.

Because F is achievable we define the limits

$$\underline{f}(i, \theta) = \lim_{k\to\infty} \frac{1}{T_k}\sum_{t=0}^{T_k-1} f(i, t, \theta),$$

$$\underline{g}(i,\theta) = \lim_{k\to\infty} \frac{1}{T_k} \sum_{t=0}^{T_k-1} g(i,t,\theta).$$

Summing over $t=0,1,\ldots, T_k-1$, dividing by T_k, and taking the limit $T_k \to \infty$ of the right- and left-hand sides of (1), one obtains

$$\lim_{k\to\infty} \frac{1}{T_k} \sum_{0}^{T_k-1} (n(i,t+1,\theta) - \rho(i,t)n(i,t,\theta)) = \underline{\phi}(i-1,\theta) + \underline{f}(i,\theta) - \underline{g}(i,\theta). \quad (14)$$

$\underline{\phi}(i-1,\theta)$ is the average flow from section $i-1$ to section i and is defined as

$$\underline{\phi}(i-1,\theta) = \lim_{k\to\infty} \frac{1}{T_k} \sum_{0}^{T_k-1} [1-\rho(i-1,t)]n(i-1,t,\theta).$$

This limit exists by the uniform boundedness of n (Fact 2) and by taking a subsequence of $\{T_k\}$ if necessary. Thus, the limit on the left-hand side of (14) exists and we are interested in the stationary case where $n(i,t+1,\theta) = n(i,t,\theta)$. In other words,

$$\underline{\phi}(i,\theta) = \underline{\phi}(i-1,\theta) + \underline{f}(i,\theta) - \underline{g}(i,\theta), \quad (15)$$

where, as above,

$$\underline{\phi}(i,\theta) = \lim_{k\to\infty} \frac{1}{T_k} \sum_{0}^{T_k-1} [1-\rho(i,t)]n(i,t,\theta).$$

Now we construct a stationary plan $\underline{u} = [\pi, \underline{v}, \underline{f}, \underline{g}]$ and a stationary trajectory as follows. We first define the velocity plan

$$\underline{v}(i,t) \equiv V,$$

where V is the maximum permissible velocity, which we assume is the same for all sections. This gives

$$\underline{\rho}(i) := 1 - \frac{V \times \tau}{L(i)}.$$

Next we define the trajectory

$$\underline{n}(i,t,\theta) = \frac{\underline{\phi}(i,\theta)}{1-\underline{\rho}(i)} = \frac{\underline{\phi}(i,\theta) \times L(i)}{V \times \tau}.$$

$\underline{n}(i,t,\theta)$ is a valid stationary trajectory because $\underline{n}(i,t,\theta) \geq 0$, it satisfies the stationary flow equation (15), and it satisfies constraint (10): if $n(i,t,\theta) = 0$ for all $(i,\theta) \in T_n$, then $\underline{n}(i,t,\theta) = 0$ for all $(i,\theta) \in T_n$.

It remains only to define the activity plan $\underline{\pi}$ and to show that the space–time constraint (7) holds. Define

$$\underline{\pi}(\alpha,i,t,\theta) = \begin{cases} 0 & \text{if } (\alpha,i,\theta) \in T_\pi \\ 1 & \text{if } \alpha = \arg\min \lambda(\alpha) \\ 0 & \text{otherwise.} \end{cases}$$

Thus, while respecting the constraint (11), $\underline{\pi}$ assigns activities that occupy the least space–time. We verify that (7) holds with the following chain of inequalities:

$$L(i)\tau \geq \frac{1}{T_k}\sum_{0}^{T_k-1}\sum_{\alpha}\sum_{\theta}\pi(\alpha, i, t, \theta)n(i, t, \theta)\lambda(\alpha)$$

$$\geq \frac{1}{T_k}\sum_{0}^{T_k-1}\sum_{\alpha}\sum_{\theta}\underline{\pi}(\alpha, i, t, \theta)n(i, t, \theta)\lambda(\alpha)$$

$$\geq \sum_{\alpha}\sum_{\theta}\underline{\pi}(\alpha, i, t, \theta)\underline{n}(i, t, \theta)\lambda(\alpha).$$

The first inequality holds since (n,u) is feasible; the second follows from the fact that $\underline{\pi}$ is the least space-occupying activity plan; the third inequality is a consequence of the definition of $\underline{n}$ and that $[1 - \underline{\rho}(i, t)] \geq [1 - \rho(i, t)]$ for all i, t, so that

$$\underline{n}(i, t, \theta) \leq \lim_{\kappa\to\infty}\frac{1}{T_k}\sum_{0}^{T_k-1} n(i, t, \theta).$$

Since vehicles in $(\underline{n},\underline{u})$ are traveling at the maximum velocity, their travel time is minimized[9] and the assertion is proved.

Theorem 2. The set of achievable flows is convex and compact.

Proof. Let (n^k, u^k) be the stationary trajectory-plan defined in the proof of Theorem 1 that achieves flow $\{F^k(\theta)\}$, $k = 1, 2$. Then

$$n^k(i, \theta) = \rho^k(i)n^k(i, \theta) + [1 - \rho^k(i-1)]n^k(i-1, \theta) + f^k(i, \theta) - g^k(i, \theta);$$

$$1 - \rho^k(i) = \frac{V \times \tau}{L(i)}.$$

Note that $\rho^k = \rho$, independent of k.

Let $\mu^k \geq 0$, $\mu^1 + \mu^2 = 1$. We will find a stationary trajectory-plan (x, u) that realizes the flow $\Sigma\mu^k F^k$. Define $x := \mu^1 n^1 + \mu^2 n^2$. Then,

$$x(i, \theta) = \sum \mu^k\rho^k(i)n^k(i, \theta) + \sum \mu^k[1 - \rho^k(i-1)]n^k(i-1, \theta) + \sum \mu^k[f^k(i, \theta) - g^k(i, \theta)]$$
$$= \rho(i)x(i, \theta) + [1 - \rho(i-1)]x(i-1, \theta) + f(i, \theta) - g(i, \theta),$$

where $f(i, \theta) := \Sigma\mu^k f^k(i, \theta)$ and $g(i, \theta) := \Sigma\mu^k g^k(i, \theta)$.

Finally, we define the activity plan by

$$\pi(\alpha, i, \theta) = \frac{\sum_k \mu^k n^k(i, \theta)\pi^k(\alpha, i, \theta)}{\sum_k \mu^k n^k(i, \theta)}.$$

[9]If vehicles can travel over different routes in an AHS network, it is more complicated to find a plan that minimizes travel time.

It is now straightforward to check that (x, $u=[\pi, v, f, g]$) is a feasible pair. Thus the set of achievable flows is convex.

To show that it is closed, consider a convergent sequence of feasible stationary pairs (n^k, u^k), $k=1, 2$.... It is easy to see that the limiting pair is feasible. Boundedness of achievable flows follows from Fact 2.

The next result is intuitively obvious.

Fact 3. If F is achievable and if $0 \leq H(\theta) \leq F(\theta)$, then H is achievable.

Proof. Let (n, $u=[\pi, v, f, g]$) be a plan that achieves F:

$$n(i, t+1, \theta) = \rho(i, t)n(i, t, \theta) + [1 - \rho(i-1, t)]n(i-1, t, \theta) + f(i, t, \theta) - g(i, t, \theta). \quad (16)$$

Define $0 \leq \gamma(\theta) \leq 1$ by $H(\theta) = \lambda(\theta)F(\theta)$. Then it is easy to check by multiplying (16) by $\gamma(\theta)$ that H is achieved by the trajectory-plan (n', $u'=[\pi', v', f', g']$):

$$\pi' = \pi, \; v' = v, \; f'(i, t, \theta) \equiv \gamma(\theta)f(i, t, \theta), \; g'(i, t, \theta) \equiv \gamma(\theta)g(i, t, \theta), \; n'(i, t, \theta) \equiv \gamma(\theta)n(i, t, \theta).$$

This proves the claim. □

5. CAPACITY AND OPTIMAL PLANS

We first show that the set of achievable flows is a convex polygon.

Fact 4. $\{F(\theta)\}$ is achievable if and only if there exist stationary flows $\phi(i,\theta)$, a trajectory $\{n(i,\theta)\}$, and plans $\{f(i,\theta), g(i,\theta), \pi(\alpha,i,\theta)\}$, all of them non-negative, such that the following linear constraints hold:

$$\phi(i, \theta) = \phi(i-1, \theta) + f(i, \theta) - g(i, \theta), \quad (17)$$

$$\phi(0, \theta) = 0, \quad (18)$$

$$\phi(I+1, \theta) = 0, \quad (19)$$

$$n(i, \theta) = \frac{\phi(i, \theta) \times L(i)}{V \times \tau}, \quad (20)$$

$$\sum_{\alpha}\sum_{\theta} \pi(\alpha, i, \theta)n(i, \theta)\lambda(\alpha) \leq L(i)\tau, \quad (21)$$

$$\sum_{\alpha} \pi(\alpha, i, \theta) = 1, \quad (22)$$

$$f(i, t, \theta) = 0, \text{ for all } (i, \theta) \in T_f, \quad (23)$$

$$g(i, t, \theta) = 0, \text{ for all } (i, \theta) \in T_g, \quad (24)$$

$$n(i, t, \theta) = 0, \text{ for all } (i, \theta) \in T_n, \quad (25)$$

$$\pi(\alpha, i, t, \theta) = 0, \text{ for all } (\alpha, i, \theta) \in T_\pi. \quad (26)$$

Constraint (20) is linearized by the nonlinear transformation $p(\alpha,i,\theta) = \pi(\alpha,i,\theta)n(i,\theta)$.

Definitions. An achievable flow $F=\{F(\theta)\}$ is *undominated* if for any achievable flow H with $H(\theta) \geq F(\theta)$, $H(\theta) = F(\theta)$ for all θ. The *capacity* of the AHS is the set of all undominated flows. See Fig. 1. A trajectory-plan is *efficient* if it minimizes travel time.

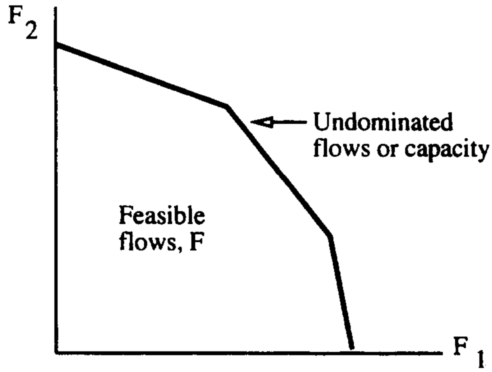

Fig. 1. The convex polygon is the set of feasible flows. The bold boundary is the set of undominated flows.

Theorem 3. A flow F^* is undominated if and only if it is the optimal solution of the linear programming problem:

$$\max \sum w(\theta)F(\theta)$$

subject to constraints (17) – (26)

for some weights $w(\theta)\geq 0$, not all zero. Moreover, the optimal solution yields an efficient pair that achieves F^*.

Proof. This follows from Fact 4 and the theory of linear programming. □

6. TRANSIENT BEHAVIOR AND TMC RULES

A TMC plan specifies activities, velocity, entry and exit flows in each section and for all times. The plan may be specified ahead of time, with no measurement of the traffic state. (In control engineering, this is said to be an "open loop" specification.) Open loop specifications are very useful for analytical study but they should not be implemented in practice. This is because the state equation model (1) is an idealization which ignores the uncertainty in model parameters and the presence of random fluctuations. These departures from idealization cause the actual traffic trajectory to be different from the open loop trajectory predicted by the model.

It is, therefore, preferable to design a TMC plan in the form of a (feedback) rule. The rule gives the plan values at time t as a function of the state $n(t)$ at that time. A rule can be evaluated by its steady state and transient behaviors. A well-designed rule would achieve capacity and minimum travel time in the absence of fluctuations, independent of the initial state; and small fluctuations would cause small departures of the achieved flow from capacity.

Since a rule specifies the plan as a function of the state, implementation of the rule requires sensors that measure the state, and communicating measurements to appropriate locations where the plan is computed. A rule requiring fewer state measurements is, everything else being equal, preferable to one that requires more measurements. A rule in which a plan for section i requires state measurements in sections near i, is preferable to one which requires measurements in sections remote from i, because the former will require less communications facilities.

We illustrate some of the issues using the example of Fig. 2. The figure shows two trajectory-plan pairs. The highway configuration is as follows. Each section is 100 m long. There is only one entry (in section 1) with flow f, and one exit (in section I) with flow g. There are two activities. Activity 1 must be carried out in all sections except I and activity 2 (the exit activity) must be carried out in section I. $\lambda(1)=10\tau$ m-s, $\lambda(2)=20\tau$ m-s. The maximum speed is 20 m/s. Section I is a "capacity bottleneck." At

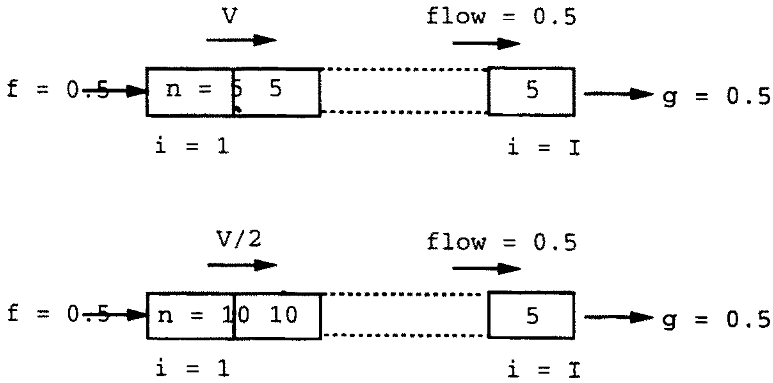

Fig. 2. Both trajectory-plan pairs achieve the maximum flow of 0.5. The upper pair minimizes travel time; the lower pair nearly doubles travel time because vehicles travel at $V/2$ in sections 1 through $I-1$.

most, 5 vehicles can be accommodated in section I, and so the maximum value of g, using the fundamental equation of traffic flow, $g=\frac{5}{100}20=1$. Hence the highway capacity is 1 vps.

Both trajectory-plan pairs in Fig. 2 achieve the capacity. In the upper pair, the velocity is 20 m/s, so the travel time is minimized. In the lower pair, the velocity is 10 m/s, so the travel time is twice the minimum in sections 1 through $I-1$.

A rule must specify the velocity in each section, and f, g in the sections 1 and I respectively. The rule for the last section g is obvious: $v(t)=20$ m/s, and $g(t)=[1-\rho]n(I,t)$. A reasonable velocity rule for all other sections is to have the maximum possible velocity (up to 20 m/s). Of course, what the maximum velocity in section i turns out to be at any time depends on the space available in section $i+1$. If the state n is as shown in the lower part of Fig. 2, the maximum possible speed is 10 m/s; if it is as in the upper part, the maximum speed is 20 m/s.

6.1. *A greedy rule*

The example motivates the need for rules or policies, both for velocity and entry, in order to achieve the maximum achievable flow, while not exceeding the space limit in each section. We will specify "greedy" policies for velocity and the entry flow f and show that they achieve the maximum steady-state flow.

To obtain the velocity policy, consider the space–time freed up by vehicles leaving section i over time t to $t+1$

$$\frac{v(i,t)}{L(i)}\sum_{\alpha}\sum_{\theta}\lambda(\alpha)\pi(\alpha,i,t,\theta)n(i,t,\theta).$$

Thus the free space–time in i is

$$L(i)\tau-[1-\frac{v(i,t)}{L(i)}]\sum_{\alpha}\sum_{\theta}\lambda(\alpha)\pi(\alpha,i,t,\theta)n(i,t,\theta). \tag{27}$$

We will choose $v(i-1,t)$ so that the space needed by vehicles leaving section $i-1$

$$\frac{v(i-1,t)}{L(i-1)}\sum_{\alpha}\sum_{\theta}\lambda(\alpha)\pi(\alpha,i,t,\theta)n(i-1,t,\theta). \tag{28}$$

is exactly the space–time available in section i, as long as the velocity does not exceed V.

Let us simplify notation by eliminating indices for θ and α. Define $n(i,t)$, the total number of vehicles in section i as

$$n(i,t)=\sum_{\theta}n(i,t,\theta)$$

and $\pi(\alpha,i,t)$, the proportion of vehicles performing activity α as

$$\pi(\alpha, i, t) = \frac{\sum_{\theta} \pi(\alpha, i, t, \theta) n(i, t, \theta)}{\sum_{\theta} n(i, t, \theta)}.$$

Then $\lambda(i)$, the average space–time used per vehicle in section i, is

$$\lambda(i) = \sum_{\alpha} \lambda(\alpha) \pi(\alpha, i, t).$$

$\lambda(i)n(i,t)$ is the space–time used by vehicles in section i. Also, the maximum number of vehicles in section i, $N(i)$ is given by

$$N(i) = \frac{L(i)\tau}{\lambda(i)}.$$

Using this notation the appropriate expression for velocity in section $i-1$ is

$$V(i-1, t) = \min\left\{ V, \frac{L(i)L(i-1)\tau}{n(i-1, t)\lambda(i)} - [1 - \frac{v(i, t)}{L(i)}] \frac{n(i, t)L(i-1)}{n(i-1, t)} \right\} \tag{29}$$

We can check that if one applies (29) and $v(i-1,t) < V$ in section $i-1$, then section i achieves its space limit. This can be seen by substituting (29) in the flow equation (1) (after summing over θ)

$$n(i, t+1) = (1 - \frac{v(i, t)}{L(i)}) n(i, t) + \frac{v(i-1, t)\tau}{L(i-1)} n(i-1, t)$$

$$= \frac{L(i)\tau}{\lambda(i)}.$$

Now the flow out of section i is

$$\phi(i, t) = [1 - \rho(i, t)] n(i, t) = \frac{v(i, t)\tau}{L(i)} n(i, t),$$

while the maximum flow $\bar{\phi}(i)$ is

$$\bar{\phi}(i) = \frac{V\tau^2}{\lambda(i)}.$$

We will need the minimum of these flows to prove existence of an equilibrium solution of the flows; therefore, we make the following definition.

Definition. ϕ^* is the minimum of the maximum possible flow out of any section or

$$\phi^* = \min_i \frac{V\tau^2}{\lambda(i)} = \min_i \bar{\phi}(i).$$

We will assume for simplicity that section I is the "bottleneck," i.e. $\phi^* = \bar{\phi}(I)$ and $\phi^* < \bar{\phi}(i)$ for $i \neq I$.

Theorem 5. Assume the velocity policy (29) is applied and $v(I,t) \equiv V$, then for every i and t, either $v(i,t) = V$ or $\phi(i,t) \geq \phi^*$.

Proof. The proof follows by induction. Considering first $i = I$, by assumption we have $v(I,t) = V$. Now assume that the statement of the theorem is true for section i. We will show that it is true for section $i-1$. Fixing t, we must show either

(a) $v(i-1,t) = V$, or

(b) $[1-\rho(i-1,t)]n(i-1,t) \geq \phi^*$.

Equivalently, we will assume that $v(i-1,t) < V$ and show that $[1-\rho(i-1,t)]n(i-1,t) \geq \phi^*$. The first case is when $v(i,t) = V$. We calculate the flow out of section $i-1$

$$\phi(i-1,t) := [1-\rho(i-1,t)]n(i-1,t)$$
$$= \frac{v(i-1,t)\tau}{L(i-1)} n(i-1,t).$$

Substituting the velocity policy (29) and using $v(i-1,t) < V$

$$\phi(i-1,t) = \frac{L(i)\tau}{\lambda(i)} - [1 - \frac{V\tau}{L(i)}]n(i,t)$$
$$\geq \frac{L(i)\tau}{\lambda(i)} - [1 - \frac{V\tau}{L(i)}]\frac{L(i)\tau}{\lambda(i)}$$
$$= \frac{V\tau^2}{\lambda(i)}$$
$$\geq \phi^*.$$

The second case is when $\phi(i,t) \geq \phi^*$. Then using the fact that $n(i,t)$ never exceeds the space limit $L(i)\tau/\lambda(i)$

$$\phi(i-1,t) = \frac{L(i)\tau)}{\lambda(i)} - [1 - \frac{v(i,t)\tau}{L(i)}]n(i,t)$$
$$\geq \frac{L(i)\tau}{\lambda(i)} - n(i,t) + \phi^*$$
$$\geq \phi^*.$$

Thus, if $v(i-1,t) < V$, then $\phi(i-1,t) \geq \phi^*$ which proves that (a) or (b) is true. This completes the induction and the proof of the theorem. □

It only remains to find a rule for controlling entry, i.e. f. As above, we propose a greedy policy for f that fills the available space in section 1. We assume there is no limit on f so the first section will remain filled after $t=0$. One can easily check that the rule for f is

$$f(t) = \frac{L(1)\tau}{\lambda(1)} - n(1,t) + \frac{v(1,t)\tau}{L(1)} n(1,t). \quad (30)$$

Corollary 1. Using (30) as the rule for f and (29) as the rule for v, $f(t) \geq \phi^*$ for all t.

Proof. Following Theorem 5, there are two cases to examine. First, when $v(1,t) = V$,

$$f(t) = \frac{L(1)\tau}{\lambda(1)} - n(1, t) + \frac{v\tau}{L(1)} n(1, t)$$

$$\geq \frac{L(1)\tau}{\lambda(1)} - (1 - \frac{V\tau}{L(1)}) \frac{L(1)\tau}{\lambda(1)}$$

$$= \frac{V\tau^2}{\lambda(1)}$$

$$\geq \phi^*.$$

The second case is when $\phi(1,t) \geq \phi^*$, so that

$$f(t) = \frac{L(1)\tau}{\lambda(1)} - n(1, t) + \phi(1, t)$$

$$\geq \frac{L(1)\tau}{\lambda(1)} - n(1, t) + \phi^*$$

$$\geq \phi^*. \qquad \square$$

Fact 5. If at time t section i is full, i.e. $n(i,t) = N(i)$, then $\phi(i,t) \geq \phi^*$.

Proof. Suppose $\phi(i, t) < \phi^*$. Then by Theorem 5 $v(i,t) = V$ and

$$\phi(i, t) = \frac{V\tau}{L(i)} N(i) \geq \phi^*$$

which is a contradiction. □

Fact 6. If $n(i,t) = N(i)$ and

$$\phi(i, t) < \overline{\phi}(i)$$

for all t, then $n(i+1,t) = N(i+1)$ and $\phi(i+1,t) \geq \phi^*$ for all t.

Proof. Since $\phi(i, t) < \frac{V\tau}{L(i)} N(i)$ and $n(i,t) = N(i)$ for all t, it must be that $v(i,t) < V$. Hence $v(i,t)$ is space-filling, and so $n(i+1,t) = N(i+1)$. From Fact 5 this implies $\phi(i+1,t) > \phi^*$. □

Theorem 6. Using the greedy policies (29) and (30) for v and f, respectively, and assuming $v(I,t) = V$ for all t, then f, g, n and v converge to a unique equilibrium solution for (1), i.e. as $t \to \infty$

$$f(t) \to \phi^*$$

$$g(t) \to \phi^*$$

$$\phi(i, t) \to \phi^*$$

$$n(i, t) \to N(i)$$

$$v(i, t) \to \frac{\phi^* L(i)}{N(I)\tau}.$$

Proof. From Corollary 1 we know $f(t) \geq \phi^*$ for all t. Also, $g(t) \leq \phi^*$. Since $\Sigma_t f(t) - g(t) < \infty$, we must have $f(t) \to \phi^*$ and $g(t) \to \phi^*$. We must now show that $n(i,t) \to N(i)$. This can be done by induction. Because $f(t)$ is space-filling $n(1,t) \equiv N$ (1). Assume $n(i,t) \equiv N(i)$ for $t > T$. We will show that $n(i+1,t) = N(i+1)$ for $t > T_1$, for some T_1.

We know from Theorem 5 that either $v(i,t) = V$ or $\phi(i,t) \geq \phi^*$. If $\phi(i,t) \geq \phi^*$, section i is space-filling so $n(i+1,t) = N(i+1)$. If $v(i,t) = V$ then $\phi(i,t) = \bar{\phi}(i)$. Since $\Sigma_t \phi(i,t) - g(t)$ is bounded, $\phi(i,t)$ can equal $\bar{\phi}(i)$ only a finite number of times. So there exists T_1 such that $\phi(i,t) < \bar{\phi}(i)$, for $t > T_1$. By Fact 6 $n(i+1,t) = N(i+1)$, $t > T_1$, completing the induction.

Next we will show by induction that $\phi(i,t) \to \phi^*$. Since section I is the bottleneck, $\phi(I,t) = \phi^* = g$. Assume $\phi(i,t) = \phi^*$ for $t > T$. After T, $n(i,t) = N(i)$ so $\phi(i-1,t) = \phi(i,t) = \phi^*$ which completes the induction. Finally, by substituting $\phi(i,t)$ and $n(i,t)$ in an expression for $v(i,t)$ we obtain $v(i,t) \to \frac{\phi^* L(i)}{N(i)\tau}$, which completes the proof. □

As a final note observe that the information needed for the greedy velocity policy can be obtained from vehicle-borne sensors and requires no extra sensor information from the roadside. The policy can be implemented by a vehicle longitudinal control law that tracks velocity V while maintaining a safe distance from the vehicle ahead.

7. ENTRY AND EXIT

An automated highway will make contact with a non-automated highway at points of entry and exit. In current design proposals (Godbole *et al.*, 1995) a "transition area" serves as interface between the two highways where vehicles undergo "check-in" and "check-out" and where vehicle control is transferred from driver to system upon entry to the AHS and from system to driver upon exit. We call these two activities "entry" and "exit."

Automation of these activities is a complex task. A vehicle entering the AHS must negotiate its passage through the transition area and coordinate its entry with vehicles on the automated lane. If this coordination is poor, there will be congestion at the entrance, slowing down upstream vehicles. A vehicle leaving the AHS may similarly disrupt traffic, thereby reducing capacity. By contrast, in between entry and exit, traffic on the automated lane should proceed very smoothly. Thus, it seems that AHS capacity and transient behavior are likely to be limited by the entry and exit activities. In this section we will formulate a micro-level queuing model for entry and show how the space occupied by the entry activity may determine the capacity of the highway. Then we show that the amount of delay incurred by upstream vehicles due to an entering vehicle depends on the sophistication of the feedback control law that implements entry.

Figure 3 shows a long automated lane, with one entrance. Distance along the highway is denoted by d, and the entrance is located at $d = E$. Vehicles are organized in platoons of closely spaced vehicles. (For simplicity assume that platoons have a fixed number of vehicles.) Platoons can engage in two activities: cruise and entry, with $\lambda(cruise) = D$ m (D does not include the platoon length) and $\lambda(entry) = S$ m, with $S > D$. The maximum velocity is V m/s. Let $f(c)$ denote the number of platoons per hour that come cruising from upstream of the entrance; and let $f(e)$ be the flow of entering platoons. An entering platoon must first engage in the entry activity; it then switches to cruise.

We want to compute the achievable throughput vectors $F = (f(c), f(e))$. By Theorem 1, we may assume that a stationary trajectory-plan achieves F, with platoons traveling at maximum velocity V. Let L be the length of the entry section, so a platoon stays in this section for time L/V hours. Hence the number of cruising platoons in this section is $n(c) = f(c) \times L/V$, and the number of entering platoons is $n(e) = f(e) \times L/V$. The space constraint is $D \times n(c) + S \times n(e) \leq L$, or

$$D \times f(c) + S \times f(e) \leq V,$$

so the capacity of this AHS is the set of all vectors $F = (f(c), f(e)) \geq 0$ that satisfy

$$D \times f(c) + S \times f(e) \leq V. \tag{31}$$

This capacity estimate is optimistic. The estimate is based on our model which assumes that the inter-platoon distance among the cruising platoons is distributed in such a way that a gap of size S meters appears every time a platoon is about to enter. This requires perfect coordination between the cruising platoons and the entry platoons. If this perfect coordination is lacking, then the cruising platoons will be forced to slow down in order to create the needed gap of S meters for an entering platoon, resulting in an increase in total travel time. In order to estimate the total delay, we need to know the distribution of inter-platoon distances. We will assume a random distribution.

Suppose that the inter-platoon distances are iid (independent, identically distributed) random variables, denoted z. The cruise control law guarantees that $z \geq D$ (the safe cruising distance) with probability 1, and we assume that $x := z - D$ is an exponentially distributed random variable with mean μ^{-1}, i.e. x has the probability density

$$p(x) = \mu e^{-\mu x}, x \geq 0.$$

For convenience, also denote $p_1(x) \equiv p(x)$.

Suppose that a platoon enters at some time t at distance E. This is platoon #0 in Fig. 3. (Note: in the figure, platoons are indicated by points.) Number the cruising platoons that follow #0 by #1, #2,..., and the distance between the end of platoon #$i-1$ and the beginning of platoon #i by $z_i = D + x_i$. If $x_1 < S$, then platoon #1 will have to slow down until it creates a distance of S; if $x_1 + x_2 < S$, then #2 will have to slow down, too, and so on. This "shock wave" will affect a random number M of platoons, where

$$M = m \Longleftrightarrow \sum_1^m x_i \leq S < \sum_1^{m+1} x_i.$$

We want to calculate the statistics of M, and the amount of slowdown.

It will be convenient to consider the distribution of $\sum_1^n x_i$,

$$p_n(x) := p(\sum_1^n x_i = x) = \mu^n \frac{x^{n-1}}{(n-1)!} e^{-\mu x}, x \geq 0. \qquad (32)$$

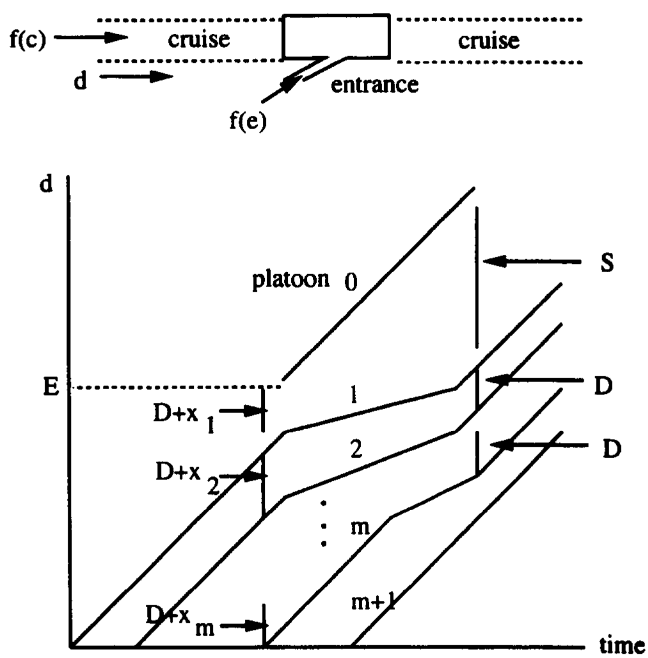

Fig. 3. There is one entry in a long highway. The trajectories show how entry of platoon #0 slows down platoons #1,..., #m.

So the probability that $M=m$, i.e. m platoons will be disturbed, is given by

$$P_S(m) = \text{Prob}\left\{\sum_1^m x_i \le S < \sum_1^{m+1} x_i\right\}.$$

One can calculate the probabilities $P_S(m)$ from the p_n by observing that

$$P_S(m) = \int_0^S p_1(x_1 \ge S-y) \times p_m(y)dy.$$

A little calculus then gives the following formula:

$$P_S(m) = e^{-\mu S}\frac{(\mu S)^m}{m!} = P_S(m-1) \times \frac{\mu S}{m}, m = 0, 1, \cdots \tag{33}$$

As expected, eq. (33) is the formula for a Poisson distribution. Thus the number M of platoons disturbed by the deviation S has a Poisson distribution. In particular, the mean number of disturbed (or delayed) platoons is $E[M] = \mu S$. If we write the mean inter-platoon distance as $Z := E[z]$, and recall the definition $\mu^{-1} = E[x] = E[z-D]$, we conclude that

$$\text{Average number of delayed platoons } = E[M] = \frac{S}{Z-D}. \tag{34}$$

Observe that the average flow of cruising platoons is $f(c) = V/Z$, whose maximum value is V/D. As expected, (34) implies that as $Z \to D$, $E[M] \to \infty$, i.e. as the flow of cruising platoons increases, the shock wave from each entering platoon passes through an increasing number of platoons, on average. Another interesting point in (34) is that the average number of delayed platoons grows *linearly* with the size of the safe entry distance, S.

We can now calculate the total delay incurred by upstream traffic due to the entering platoon, platoon #0. The entering platoon will require S m; however, if the entering platoon encounters a free space gap, then the actual space B "borrowed" from the upstream cruise platoons will be between 0 and S. We will consider the probability distribution of B after first examining the case of a fixed space S.

In order to create a gap of S m, platoons #1,..., #M are slowed down, where M is the random variable above. Platoon #i is slowed down by a distance

$$S - \sum_{j=1}^{i}(z_j - D) = S - \sum_{j=1}^{i} x_j, i = 1, \cdots, M.$$

So the total slowdown δ (measured in platoon × meters) is the sum of these M numbers,

$$\text{slowdown} := \delta = \sum_{i=1}^{M}[S - \sum_{j=1}^{i} x_j] = MS - \sum_{i=1}^{M}\sum_{j=1}^{i} x_j. \tag{35}$$

We want to calculate $E[\delta]$, the average slowdown.

Introduce the partial sums $y_0=0$, $y_i=\sum_{j=1}^{i} x_j$ for $i>0$, and write $\delta=MS-\sum_1^M y_i$. Then

$$E[\delta] = \sum_{m=0}^{\infty}\left\{mS - \sum_{i=1}^{m} E[y_i|M=m]\right\}P_S(m). \tag{36}$$

Since in (33) we have an expression for $P_S(m)$, the probability that m platoons are delayed given that the space borrowed upstream is S, and we found $E[M]$ in (34), it remains to calculate $E[y_i | M = m]$.

Fact 7. We have

$$p(y_1, \cdots, y_{m+1}) = p(y_1, \cdots, y_m | y_{m+1}) p(y_{m+1})$$

$$= \frac{m!}{(y_{m+1})^m} p_{m+1}(y_{m+1}) 1(y_1 < y_2 < \cdots < y_{m+1})$$

$$= \mu^{m+1} e^{-\mu y_{m+1}} 1(y_1 < y_2 < \cdots < y_{m+1}), \tag{37}$$

where $p_{m+1}(y)$ is given by (32) and $1(\cdot)$ is the indicator function.

Proof. The first equation in (37) is Bayes rule. Since

$$y_{m+1} = \sum_1^{m+1} x_i, \; p(y_{m+1} = y) = p_{m+1}(y)$$

from (32). Second, since $y_i - y_{i-1} = x_i$ are iid and exponential, therefore, given y_{m+1}, the y_i are uniformly and independently distributed over $[0, y_{m+1}]$, constrained to $y_1 < y_2 < \cdots < y_{m+1}$. This gives the second relation. The third relation now follows upon substitution for p_{m+1} from (32).

We now calculate $E[y_i | M = m]$:

$$E[y_i | M = m] = E[y_i | y_m < S \le y_{m+1}]$$

$$= \frac{E[y_i 1(y_m < S \le y_{m+1})]}{E[1(y_m < S \le y_{m+1})]} = \frac{P}{Q}, \text{ say},$$

where

$$P = \int_0^\infty \cdots \int_0^\infty y_i 1(y_m < S \le y_{m+1}) p(y_1, \cdots, y_{m+1}) dy_1 \cdots dy_{m+1}$$

$$Q = \int_0^\infty \cdots \int_0^\infty 1(y_m < S \le y_{m+1}) p(y_1, \cdots, y_{m+1}) dy_1 \cdots dy_{m+1}$$

$$= \int_0^{y_2} dy_1 \int_0^{y_3} dy_2 \cdots \int_0^{y_m} dy_{m-1} \int_0^S dy_m \int_S^\infty p(y_1, \cdots, y_{m+1}) dy_{m+1}$$

$$= \int_S^\infty \int_0^S \frac{y_m^{m-1}}{(m-1)!} dy_m \times \mu^{m+1} e^{-\mu y_{m+1}} dy_{m+1}$$

$$= \frac{S^m \mu^m}{m!} e^{-\mu S}.$$

A slightly more laborious calculation gives

$$P = \frac{iS}{m+1}\frac{S^m\mu^m}{m!}e^{-\mu S},$$

and so

$$E[y_i|M = m] = \frac{P}{Q} = \frac{iS}{m+1}. \quad (38)$$

Substituting this into (36) gives

$$E[\delta] = \sum_{m=0}^{\infty}[mS - \sum_{i=1}^{m}\frac{iS}{m+1}]P_S(m)$$

$$= \frac{S}{2}\sum_{m=0}^{\infty} mP_S(m)$$

$$= \frac{S^2}{2(Z-D)} \text{ platoon} - \text{meters},$$

where we used (34) in the last relation.

Fact 8. Each entering platoon on average disturbs $S/(Z-D)$ platoons and they suffer a total slowdown of $S^2/2(Z-D)$ platoon–meters.[10]

As noted above, if the entering platoon is aligned with a free space gap in the cruise lane the actual space borrowed from upstream B will be between 0 and S. As an example, suppose B is a uniformly distributed random variable with probability density

$$p(B) = \frac{1}{S}, 0 \leq B \leq S.$$

Using the expression for $E[\delta]$ above, the average slowdown now is

$$E[\delta] = \int_0^B \frac{(B)^2}{2(Z-D)}p(B)dB$$

$$= \frac{S^2}{6(Z-D)} \textit{platoon} - \textit{meters}.$$

As expected, the average slowdown is reduced when we account for the borrowed space B.

7.1. *Total time delay constraint*

As we have seen, lack of coordination causes an increase in travel time for cruise vehicles but does not reduce capacity. It is interesting to consider what happens if we

[10]We can compare this slowdown with the case when inter-platoon distance is exactly Z. (This requires a cruising control strategy that achieves equal inter-platoon distance.) In this case platoon #1 is slowed down distance $S-(Z-D)$, #2 is slowed down $S-2(Z-D)$,..., #M by $S-M(Z-D)$ and $M = S/(Z-D)$. (We are neglecting the requirement that M has to be an integer.) The sum of these slowdowns is $S^2/2(Z-D)-S/2$. Thus, exponentially distributed inter-platoon distances cause an extra slowdown of $S/2$, on average, compared with the case of equal intra-platoon distances.

impose the requirement that the total time delay per cruise vehicle from a given entry maneuver does not exceed σ. This requirement introduces an extra constraint on the cruise and entry flows, namely

$$\frac{f(e)}{f(c)}\frac{E[\delta]}{V} \leq \sigma. \tag{39}$$

Substituting for $E[\delta]$ and recalling that $Z = V/f(c)$, (39) can be rewritten as

$$S^2 f(e) + 2V\sigma D f(c) \leq 2V^2\sigma. \tag{40}$$

(40) is an extra linear constraint on $f(c)$ and $f(e)$ which may be appended to the constraints (17)–(26) of the linear programming problem of Theorem 3.

Observe that if $f(e) = 0$ the constraint reduces to

$$f(c) \leq V/D$$

which is equivalent to the space constraint (31), in this case. As expected, there is no additional constraint for total time delay if there are no entering vehicles.

7.2. Entry disturbance length

We have calculated the average slowdown from the entry maneuver. We would like to know how far up the highway the disturbance propagates on average. Ideally, the distance between entrances should be more than the average distance of vehicles delayed upstream. Let's call W the distance the disturbance propagates upstream. W is given by

$$W = \sum_{i=1}^{m}[n_i l + d(n_i - 1) + D + x_i] \tag{41}$$

$$= m(D - d) + y_m + \sum_{i=1}^{m} n_i(l + d), \tag{42}$$

where d is the inter-vehicle spacing within a platoon, l is the vehicle length, and n_i is the size of the ith platoon. We will assume that n_i are independent randomly distributed variables with mean $E[n] = \eta$. We calculate $E[W|S]$,

$$E[W|S] = \int_0^\infty \int_0^\infty \sum_{m=0}^{\infty}[m(D-d) + y_m + \sum_{i=1}^{m} n_i(l+d)]p\{m, y_m, n|S\}dydn$$

$$= (D-d)E[M|S] + E[y_m|S] + (l+d)E[n]E[M|S].$$

Using $E[y_m|S] = E[E[y_m|m,S]|S] = SE[m/(m+1)|S]$, we obtain

$$E[W|S] = (D-d)\mu S + S - \frac{1}{\mu} + \frac{e^{-\mu S}}{\mu} + (l+d)\eta\mu S. \tag{43}$$

For plausible values of $D = 60$ m, $S = 149$ m, $d = 1$ m, $l = 5$ m, $\mu = 1/20$, and $\eta = 15$, we find $E[W|S] = 1239$ m.

It is interesting to consider the effect of the size of free space gaps and the size of platoons on W. First, let us rewrite (43) in terms of $E[x]$ and drop the fourth term which is small

$$E[W|S] = \frac{(D-d)S}{E[x]} + S - E[x] + \frac{(l+d)E[n]S}{E[x]}.$$

We can see that as $E[x]$ increases $E[W|S]$ will decrease. Now suppose we increase the average platoon size while keeping the cruise flow constant. This gives a relation between $E[n]$ and $E[x]$

$$f(c) = \frac{nV}{E[x] + D + E[n]l + (n-1)d} = f.$$

Then

$$E[x] = E[n](\frac{V}{f} - l - d) - D + d.$$

Since $-D+d<0$ and $x \geq 0$ the coefficient multiplying $E[n]$ must be positive. Thus, for a given cruise flow, as the average platoon size increases the distance that the disturbance propagates upstream decreases.

7.3. Free space distance

We will show that a good entry metering policy is one that uses upstream free space when it is "closer" in a sense to be elaborated below. At time t the *free space* is the distance in meters in a section not reserved by an activity and equals

$$L(i)\tau - \sum_{\alpha}\sum_{\theta}\lambda(\alpha)\pi(\alpha, i, t, \theta)n(i, t, \theta).$$

We adopt the convention, as before, that free space appears immediately downstream from the safety gap in front of a vehicle or a platoon. We will index the free space by k if it is the kth free space gap from a point y on the highway (see Fig. 4).

Definition. The *free space distance* of a segment of free space k of length x_k from a point y along the highway is

$$x_k \sum_{i=1}^{k-1} n_i,$$

where n_i is the number of vehicles in platoon i, n_1 is the number of vehicles in the first platoon upstream from y, and n_k is the number of vehicles in the platoon directly downstream from the free space k.

Thus the distance of the free space is merely the number of vehicles between a reference point on the highway and the location of the free space. Distance does not depend on the Euclidean distance but the density of the flow. Distance is indirectly a function of inter-platoon distance, intra-platoon distance, and safety gaps. We find a simple relation between free-space distance and total time delay caused by the entry maneuver.

Fact 9. The total time delay due to the entry maneuver increases as the total distance of free space from the maneuver increases.

Proof. The total delay is given by

$$\delta = \sum_{i=1}^{m} n_i [\frac{S}{V} - \sum_{j=1}^{i} \frac{x_j}{V}]$$

$$= \sum_{i=1}^{m} \frac{n_i S}{V} - \sum_{i=1}^{m} n_i \sum_{j=1}^{i} \frac{x_j}{V}$$

$$= \sum_{i=1}^{m} \frac{n_i S}{V} - \frac{1}{V}\sum_{i=1}^{m} x_i \sum_{j=i}^{m} n_j.$$

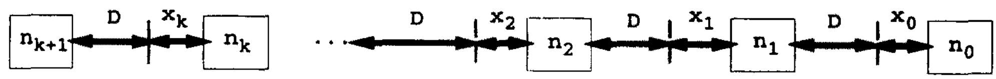

Fig. 4. The arrangement of vehicles, safety gaps and free space in an automated lane.

Now the total distance of free space κ is

$$\kappa = \sum_{k=1}^{m} x_k \sum_{i=1}^{k-1} n_i.$$

Also, let's call

$$N = \sum_{j=1}^{m} n_j \, ; \; X = \sum_{i=1}^{m} x_i \, ; \; \Delta S = S - X.$$

Then

$$\delta = \sum_{i=1}^{m} \frac{n_i S}{V} - \frac{1}{V}(XN - \kappa)$$

$$= \frac{1}{V}[\Delta SN + \kappa].$$

The first term is an additional delay because slightly more free space is needed than that provided by $x_1,..., x_m$. The second term shows that as κ increases the slowdown δ increases, which completes the proof.□

8. ACTIVITY MODEL AND VEHICLE CONTROL

With the exception of the discussion on entry and exit, our treatment of "activity" has been formal: an activity consumes space and time, and the movement of a vehicle through the AHS can be described by a finite activity sequence. In this section we address two pragmatic questions: How should one define an activity in practical terms? How should one determine the space that it consumes? It will turn out that these are questions of AHS design, more particularly, the design of the feedback laws that control vehicle maneuvers, and the TMC rules that govern the flow of traffic. Different AHS designs yield different activities. The designs can then be compared in terms of their steady state capacity and transient behavior using the theory proposed above.

In our theory of automated traffic flow we introduced the notion of a vehicle activity in order to account for the differing amounts of space vehicles take up when they are engaged in maneuvers. Maneuvers are realized by control laws, in automated traffic, and by driver actions, in manual traffic. Thus, it makes sense for activities to be defined in terms of one or a sequence of maneuvers and to examine the control laws that realize maneuvers to characterize the activity.

Since we are dealing with a one-lane AHS, it is necessary only to examine longitudinal control laws. A simplified vehicle model for longitudinal control in the form of a third-order nonlinear differential equation was obtained in Sheikholeslam and Desoer, 1990

$$\dddot{x}_i = b_i(\dot{x}_i, \ddot{x}_i) + a_i(\dot{x}_i)u_i',$$

where the subscript i is an index for the ith vehicle, and x_i is its distance along the highway. This model is linearized using the feedback law

$$u'_i = \frac{1}{a(\dot{x}_i)}[-b(\dot{x}_i, \ddot{x}_i) + u_i]$$

to obtain

$$\ddot{x}_i = u_i.$$

Vehicle maneuvers are specified through u_i. Generally, u_i will consist of a sum of two terms: one term for the desired open loop behavior and a feedback term for tracking the desired open loop behavior. The control objective typically is tracking a velocity profile as a function of time or maintaining a time or distance headway from the vehicle ahead. For example, the control law for the leader of a platoon tracks the velocity of the vehicle ahead (with index $i-1$) and maintains a safe distance by specifying the desired velocity, $\dot{x}_{di} = \dot{x}_{i-1}$ and a desired spacing, $x_{di} = x_{i-1} - l - (a_v\dot{x}_i + a_p)$. l is the vehicle length and a_v and a_p are constants. Then, u_i is given by (see Godbole and Lygeros, 1994)

$$u_i = -3\dot{x}_i - 3(\dot{x}_i - \dot{x}_{di}) - (x_i - x_{di}).$$

We will assume that the time constants for closed-loop tracking of vehicle maneuvers are much faster than the traffic flow time scale, so perturbations due to inexact tracking are ignored and we restrict our attention to the open loop behavior. Then we may define an activity as one or more consecutive vehicle maneuvers characterized by a sequence of desired open loop behaviors.

In this manner, activities are derived from vehicle control laws, and the space used by an activity $\lambda(\alpha)$ is the abstraction that brings activities into our traffic flow theory.

Calling $s(t)$ the space reserved by an activity at time t, and T the duration of the activity, the space–time used by activity α is computed as

$$\lambda(\alpha) = \int_0^T s(t)dt.$$

$s(t)$ can be extracted either directly from the control specification or after some manipulation of the expression for the open loop behavior. $s(t)$ may be parametrized by the vehicle velocity and the initial distance between the given vehicle and the vehicle ahead. We make these points clear by some examples.

Going back to the example of the leader of a platoon, the space reserved by the control law is evident from the expression for desired spacing between platoons,

$$s(t) = l + a_v\dot{x}_i(t) + a_p.$$

Plausible values (see Ioannou and Xu, 1994; Godbole and Lygeros, 1994) are $l = 5$ m, $a_v = 1$ s, $a_p = 10$ m, and $\dot{x} = 25$ m/s, so that $s = 40$ m. Note that if there is no vehicle ahead, the control law will track a desired velocity and the space is effectively reserved.

Vehicles in a platoon with inter-platoon distance d use a velocity-independent spacing

$$s = l + d.$$

For manual driving, we suppose that the driver's control objective is to track a time headway of two seconds to the vehicle ahead. This objective is independent of the relative position or velocities of the vehicles, but depends on the vehicle's own velocity. In this case

$$s(t) = 2\dot{x}_i(t).$$

These examples do not require examination of the vehicle control laws as the space requirement is implicitly expressed by the control objective.

More complicated maneuvers including lane change, platoon merge, platoon split, etc. may be specified as a desired velocity profile $\dot{x}_i(t)$ (assuming the vehicle ahead maintains constant velocity) and an initial relative distance $\Delta x_i = x_{i-1} - x_i$ (which may be fixed by the longitudinal sensor range). From this specification, one can extract the space requirement by numerical integration. These aspects are addressed in Haddon, 1996 and Broucke and Varaiya, 1996.

9. STEADY-STATE CAPACITIES

We consider two alternative designs. We call one design the platoon organization or PO design (Varaiya, 1993). We call the second the adaptive cruise control or ACC design (Ioannou and Xu, 1994).

9.1. *PO design*

There are five activities in the PO design: merge, split, 15 vehicle platoon, entry, and exit. We will determine the steady-state capacity of an automated lane with these activities. We first specify the lane configuration. The lane consists of sections of equal length L. There are three types of sections. In entry sections *entry* and *platoon*15 are allowed; in exit sections *exit* and *platoon*15 are allowed; in all other sections, called cruise sections, either *platoon*15, *merge*, or *split* are allowed. (In a merge maneuver, one platoon first accelerates and then decelerates to join the platoon in front of it; in split, the rear of one platoon first decelerates and then accelerates to form two platoons.)

In order to calculate steady-state capacities, it is necessary to determine the space requirement for each activity, to specify the composition of activities in each section, and to find the section with the strictest space limit which determines the maximum flow.

We specify some physical and design parameters. D is the safety distance maintained by the leaders of platoons, d is the inter-vehicle spacing within a platoon, l is the vehicle length, V is the maximum velocity, n is the platoon size, Q is the range of the longitudinal sensor, a_{min} is the maximum vehicle deceleration, a_{max} is the maximum vehicle acceleration. Representative values used in the PO design are $L=500$ m, $\tau=20$ s, $D=60$ m, $d=1$ m, $l=5$ m, $V=25$ m/s, $n=15$, $Q=60$ m, $a_{max}=2$ m/s^{-1}, and $a_{min}=-2$ m/s^{-1}.

The space requirement for *entry* is $s(t)=D+l=65$ m, so $\lambda(entry) = 65\tau$ m-s; also, $\lambda(exit)=65\tau$ m-s. The space requirement for *platoon*15 is

$$s(t) = \frac{d(n-1)+nl+D}{n}$$

or 10 m, so $\lambda(platoon15)=10\tau$ m-s.

The space requirement for *merge* requires some calculation. We assume that the merge is initiated by one vehicle when the platoon ahead is within the vehicle's sensor range Q. The relative velocity and acceleration between the two cars is initially zero. The merging vehicle accelerates up to a_{max} while keeping a safe relative distance and velocity from the car ahead. The maneuver ends when the vehicle is within distance d m of the platoon ahead. If the activity lasts for less than the time period τ, some extra space must be allotted. Two maneuvers constitute this activity:

$$s(t) = f(\Delta v, \Delta a, Q); \; t_0 \le t \le t_1$$

$$s(t) = \frac{d(n-1)+nl+D}{n}; \; t_1 \le t <= \tau.$$

Δv is the relative velocity of the two vehicles at the beginning of the merge, and Δa is the relative acceleration of the two vehicles at the beginning of the merge. t_1 is the time when the merging vehicle is within d m of the vehicle ahead.

Using the equations for the safe merge developed in Frankel *et al.*, 1995 we obtain a space requirement of 27 m with a duration of $t_1 = 16$ s. To this we add the length of the vehicle l. For the remaining time from t_1 to τ the vehicle requires 10 m. Thus, $\lambda(merge) = 32 \cdot 16 + 10 \cdot 4 = 28\tau$ m-s. A similar exercise for *split*, which takes a vehicle from d m to D m from the platoon ahead and uses a_{min} for deceleration yields $\lambda(split) = 28\tau$ m-s.

We must define the proportion of activities in each section. π_e (π_x) is the proportion of vehicles doing *entry* (*exit*) in an entry (exit) section, π_m (π_s) is the proportion of vehicles doing *merge* (*split*) in a cruise section, π_c is the proportion of vehicles doing *platoon*15 in a cruise section, and π_p is the proportion of vehicles doing *platoon*15 in an entry or exit section. There are some constraints on the proportions:

$$\pi_e = \pi_x,$$

$$\pi_m = \pi_s,$$

$$\pi_p + \pi_e = 1,$$

$$\pi_c + 2\pi_s = 1.$$

The constraints capture that the proportion of vehicles exiting equals the proportion entering, the proportion of vehicles merging equals the proportion splitting, and the sum of proportions of activities in each section type must equal one.

Using these constraints, calling the flow f, and substituting values for $\lambda(\alpha)$, the space constraint for entry/exit sections is

$$[65\pi_e + 10(1 - \pi_e)]f = 25.$$

The space constraint for cruise sections is

$$[10(1 - 2\pi_s) + 28\pi_s + 28\pi_s]f = 25.$$

If we set $\pi_e = .1$ and $\pi_s = .1$, the limiting section is the entry or exit section, and the maximum flow is $f = 5{,}806$ vph.

Suppose we keep π_e fixed but vary π_s between 0 and 0.5. The constraint on the flow due to the entry (exit) sections is 5,806 vph. The constraint due to the cruise section as π_s is varied is shown in Fig. 5.

9.2. *ACC design*

In this design, some of the vehicles are manually driven, and the rest are under adaptive cruise control. So there are four activities: automatic cruise, manual cruise, manual entry, and manual exit. The lane consists of entry, exit and cruise sections. In entry (exit) sections, automatic cruise, manual cruise and entry (exit) are allowed. In cruise sections, automatic and manual cruise are allowed.

The space requirement for manual entry is $\lambda(entry) = D + l = 65\tau$ m. The requirement for manual exit is $\lambda(exit) = D + l = 65\tau$ m. The requirement for manual cruise is $\lambda(mc) = 2V\tau = 50\tau$ m. The requirement for automatic cruise is $\lambda(ac) = V + l + 10 = 40\tau$ m. The only constraint on activity proportions is $\pi_e = \pi_x$, and we set $\pi_e = .1$. Now we write the space constraint for the three types of sections. For entry (exit) sections

$$[(1 - \pi_{ac} - .1)50 + \pi_{ac}40 + 6.5]f = 25.$$

For cruise sections

$$[(1 - \pi_{ac} - .1)50 + \pi_{ac}40]f = 25.$$

We can now compute the capacity. If for example, the proportion of automated vehicles is 0.5, then the maximum flow in an entry (exit) section is 1935.5 vph. Figure 6

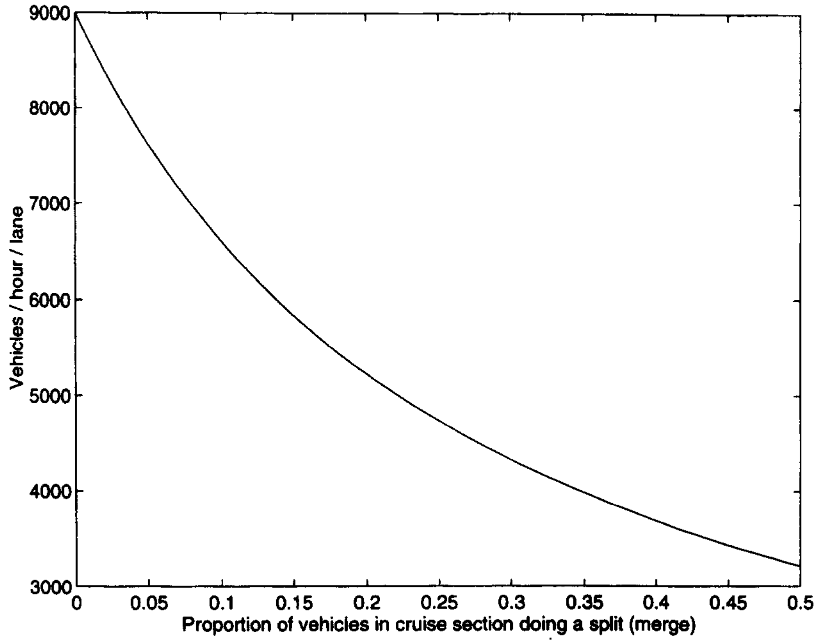

Fig. 5. Maximum flow in cruise sections as a function of the proportion of vehicles doing splits (merges).

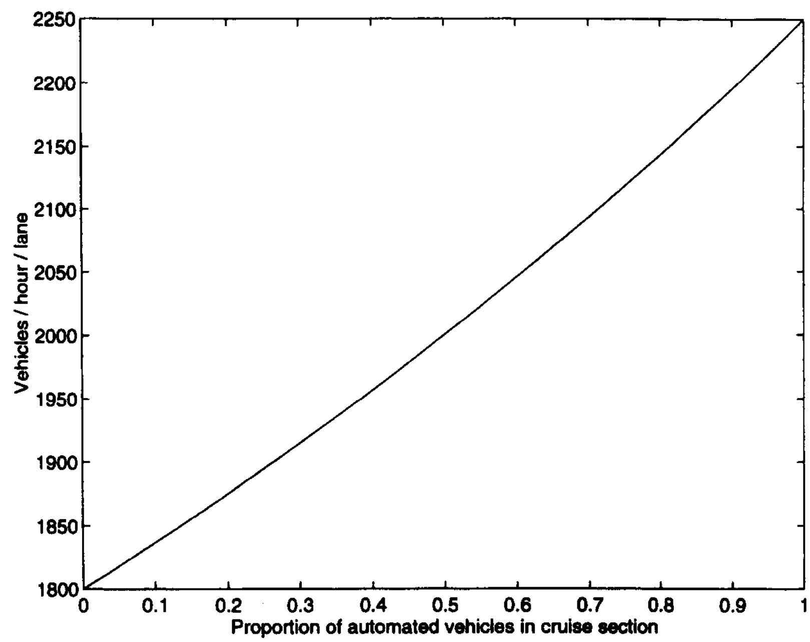

Fig. 6. Maximum flow in cruise sections as a function of the proportion of automated vehicles.

shows the increase in capacity as the proportion of automated vehicles in a cruise section increases.

10. CONCLUSIONS

We have presented a theory for automated traffic flow, based on the notion of vehicle activities. An activity is a sequence of vehicle maneuvers executed by vehicle control laws. The space that it takes up is the abstraction used to represent an activity in the traffic flow model. A plan is defined as the proportion of activities, velocity, entry flow and exit flow in each section. The TMC controls the flow by selection of this plan. We showed that achievable flows can be realized by stationary plans, and maximal achievable flows are obtained by solving a linear programming problem.

These are results about steady-state conditions. However, since conditions may vary over time, perhaps because of incidents, one should use adaptive policies for the entry flow and velocity. We proposed one such policy: the greedy policy attempts to fill up the free space in the next section as quickly as possible. We showed that the greedy policy maximizes steady-state flow, although it does not minimize travel time.

Next we studied entry and exit, which are likely to be the capacity-limiting activities because of the large space they require. We studied the effect of lack of coordination at the entry and found that, although it does not affect capacity, it does increase the travel time of the upstream vehicles. We estimated the upstream distance traveled by the disturbance created by entry and determined that a good metering policy is to carry out the entry maneuver when free space is nearby.

The proposed theory can be compared with the theory of manual traffic flow. The safety-needs-space assumption makes space the crucial resource in our model, and in a one-lane highway, the maximum flow is determined by the most space-constraining section. This insight holds for a network of highways, and the Ford–Fulkerson theorem can be used to relate the maximal or undominated flows with the most constraining sections. The insight is equally valuable in manual traffic. Perhaps the only important distinction is that in manual traffic the "consumption" of space by vehicles has a negative externality, because drivers interact. This interaction between vehicles is absent in our model.

The model has some obvious limitations. The one-activity-per-section assumption means that activities are of roughly the same length and there is one control command per section. We may wish to allow for sequences of activities to be performed in a section and to make sections and activities independent. This can be accommodated at the cost of greater notational burden.

Abstracting activities using space requires care in its application. The space usage is averaged over the duration of the activity, i.e. the time it takes to traverse the section. If the section length is increased, the space usage will change because the activity may require extra space only for a short interval. The selection of section length therefore also affects the space abstraction and should be chosen at a scale where the extra space usage from activities is significant. The space requirement may not be numerically easy to extract from the vehicle control laws which are defined in terms of velocity and relative position of the vehicle ahead.

The usefulness of the proposed theory must be judged by its ability to open up for investigation related questions and in application. Our future work will take steps in both directions, by developing the multi-lane, dynamic case (Broucke and Varaiya, 1996) and by demonstrating the application of the theory to manual driving.

Acknowledgements—The authors are grateful to Datta Godbole, Roberto Horowitz, and John Lygeros for helpful discussions. We thank one of the reviewers for a careful commentary.

REFERENCES

Broucke, M. and Varaiya, P. (1996) Dynamics and control of automated traffic. In preparation.

Daganzo, C. (1994) The cell transmission model: a dynamic representation of highway traffic consistent with the hydrodynamic theory. *Transportation Research-B*, **28** (4), 269–287.

Frankel, J., Alvarez, L., Horowitz, R. and Li, P. (1995) Safety-oriented maneuvers for IVHS. In *Proceedings American Control Conference*, pp. 668–672.

Godbole, D., Eskafi, F., Singh, E. and Varaiya, P. (1995) Design of entry and exit maneuvers for IVHS. In *Proceedings American Control Conference*, pp. 3576–3580.

Godbole, D. and Lygeros, J. (1994) Longitudinal control of the lead car of a platoon. *IEEE Transactions on Vehicular Technology*, **43**(4), 1125–1135.

Haddon, J. (1996) Master's thesis. Electrical Engineering and Computer Science, University of California, Berkeley.

Hall, R. (1995) Longitudinal and lateral throughput on an idealized highway. *Transportation Science*, **29**, 3499–3505.

Ioannou, P. and Xu, Z. (1994) Throttle and brake control systems for automatic vehicle following. *IVHS Journal*, **1**(4), 345–377.

McMahon, D., Hedrick, J. and Shladover, S. (1990) Vehicle modeling and control for automated highway systems. In *Proceedings American Control Conference*, pp. 297–303.

Peng, H. and Tomizuka, M. (1990) Lateral control of front-wheel-steering rubber-tire vehicles, Technical Report UCB-ITS-PRR-90-5, University of California, Institute of Transportation, Berkeley, CA.

Rao, B. S. Y. and Varaiya, P. (1994) Roadside intelligence for flow control in an IVHS. *Transportation Research-C*, **2**(1), 49–72.

Sheikholeslam, S. and Desoer, C. A. (1990) Longitudinal control of platoon of vehicles. In *Proceedings of American Control Conference*, pp. 291–296.

Varaiya, P. (1993) Smart cars on smart roads: problems of control. *IEEE Transactions on Automatic Control*, **38**(2), 195–207.