Available online at www.sciencedirect.com

ScienceDirect

Transportation Research Part C xxx (2007) xxx–xxx

TRANSPORTATION RESEARCH PART C

www.elsevier.com/locate/trc

Behavior of the cell transmission model and effectiveness of ramp metering

Gabriel Gomes [a], Roberto Horowitz [a], Alex A. Kurzhanskiy [a], Pravin Varaiya [a,*], Jaimyoung Kwon [b]

[a] University of California, Berkeley, CA 94720-1770, United States
[b] California State University, East Bay, Hayward, CA 94542, United States

Received 5 June 2007; received in revised form 23 October 2007; accepted 29 October 2007

Abstract

The paper characterizes the behavior of the cell transmission model of a freeway, divided into N sections or cells, each with one on-ramp and one off-ramp. The state of the dynamical system is the N-dimensional vector n of vehicle densities in the N sections. A feasible stationary demand pattern induces a unique equilibrium flow in each section. However, there is an infinite set—in fact a continuum—of equilibrium states, including a unique uncongested equilibrium n^{u} in which free flow speed prevails in all sections, and a unique most congested equilibrium n^{con}. In every other equilibrium n^{e} one or more sections are congested, and $n^{\mathrm{u}} \leqslant n^{\mathrm{e}} \leqslant n^{\mathrm{con}}$. Every equilibrium is stable and every trajectory converges to some equilibrium state.

Two implications for ramp metering are explored. First, if the demand exceeds capacity and the ramps are not metered, every trajectory converges to the most congested equilibrium. Moreover, there is a ramp metering strategy that increases discharge flows and reduces total travel time compared with the no-metering strategy. Second, even when the demand is feasible but the freeway is initially congested, there is a ramp metering strategy that moves the system to the uncongested equilibrium and reduces total travel time. The two conclusions show that congestion invariably indicates wastefulness of freeway resources that ramp metering can eliminate.

Keywords: Cell transmission model; Stability; Macrosimulation; Ramp metering; Fundamental diagram; Congestion; Strictly monotone map; Convergent system

1. Introduction

The paper presents a complete analysis of the qualitative properties of the cell transmission model (CTM). CTM is a first-order discrete Godunov approximation (Godunov, 1959), proposed by Daganzo (1994) and Lebacque (1996), to the kinematic wave partial differential equation of Lighthill and Whitham (1955) and Richards (1956). The popularity of CTM is due to its very low computation requirements compared with

* Corresponding author.
E-mail addresses: gomes@path.berkeley.edu (G. Gomes), horowitz@me.berkeley.edu (R. Horowitz), akurzhan@eecs.berkeley.edu (A.A. Kurzhanskiy), varaiya@eecs.berkeley.edu (P. Varaiya), jaimyoung.kwon@csueastbay.edu (J. Kwon).

0968-090X/$ - see front matter
doi:10.1016/j.trc.2007.10.005

Please cite this article in press as: Gomes, G. et al., Behavior of the cell transmission model and effectiveness ..., Transport. Res. Part C (2007), doi:10.1016/j.trc.2007.10.005

micro-simulation models; the ease with which it can be calibrated using routinely available point detector data (Lin and Ahanotu, 1995; Munoz et al., 2004); its extensibility to networks (Buisson et al., 1996a) and urban roads with signalized intersections (Lo, 2001; Almasri and Friedrich, 2005); and the flexibility with which it can be used to pose questions of traffic assignment (Buisson et al., 1996b; Ziliaskopoulos, 2000) and ramp metering (Daganzo and Lin, 1993; Zhang et al., 1996; Gomes and Horowitz, 2006). These topics are also studied using different discrete models (May, 1981; Papageorgiou et al., 1990; Payne, 1979). CTM is a widely used discrete macroscopic model today.

The objective of this paper, however, is not to relate CTM to the kinematic wave equation nor to investigate its utility for simulation, but rather to study it as a class of nonlinear dynamical systems. From this viewpoint, the interest is to determine the structure of its equilibrium points, their stability, and the qualitative properties of the convergence of its trajectories. Surprisingly, these aspects of the cell transmission model have received no attention in the published literature. The paper fills this gap.

Section 3 presents the model, taken from Gomes (2004) and Gomes and Horowitz (2006), which in turn is based on Daganzo (1994). The freeway is divided into N sections, indexed $0,\ldots,N-1$. Section i is characterized by a single state variable, its density n_i, so the state of the freeway is the N-dimensional vector $n=(n_0,\ldots,n_{N-1})$. Vehicle movement in a section is governed by the familiar triangular 'fundamental diagram', which gives flow as a function of vehicle density. If the density is below critical, vehicles move at free flow speed; if it is above critical, the section is congested, speed is lower, and flow from the immediately upstream section is constrained. Thus the state of a freeway obeys a N-dimensional nonlinear difference equation. When the exogenous demand pattern of on-ramp and off-ramp flows is constant, the difference equation is time-invariant, and it is meaningful to study its equilibrium states.

Theorem 4.1 in Section 4 fully characterizes the structure of the equilibrium flows and states in any CTM model. Each demand pattern induces a unique equilibrium flow vector f, and an *infinite* set of equilibrium states E. Corresponding to f is the set of bottleneck sections at which flow equals capacity. If there are K bottlenecks, the freeway partitions into $1+K$ segments, $S^0,\ldots,S^K$, each of which, except S^0, begins at a bottleneck, and decomposes the set E into the product $E=E^0\times\cdots\times E^K$, with E^k being the equilibrium set for segment S^k. Each equilibrium in E^k determines an integer j so that the most downstream j sections in S^k are congested (density is above critical), and the remaining sections are uncongested. The equilibrium set E forms a topologically closed, connected, K-dimensional surface in the N-dimensional state space. Two equilibria are special: the unique *uncongested* equilibrium n^{u} in which free flow speed prevails in all sections; and the unique *most congested* equilibrium n^{con}. Every other equilibrium $n^{\mathrm{e}}\in E$ is bounded by these, $n^{\mathrm{u}}\leqslant n^{\mathrm{e}}\leqslant n^{\mathrm{con}}$. Theorem 4.1 provides an explicit closed form expression for E. In the special case that the demand is *strictly* feasible, i.e., the equilibrium flow in each section is strictly below capacity, E reduces to the unique uncongested equilibrium n^{u}.

Section 5 studies the qualitative behavior of all trajectories generated by the CTM model. The model induces a strictly monotone map and some of the trajectory behavior is a consequence of the general theory of strictly monotone maps (Hirsch and Smith, 2005). One interesting consequence is that the unique equilibrium n^{u} for a strictly feasible demand pattern is globally asymptotically stable: every trajectory converges to n^{u} (Theorem 5.1). A more surprising result is that every equilibrium is (Lyapunov) *stable*: trajectories starting near an equilibrium $n^{\mathrm{e}}\in E$ remain near it forever (Theorem 5.2). The most remarkable result is that the CTM model is *convergent*: every trajectory converges to some equilibrium state $n^{\mathrm{e}}\in E$ (Theorem 5.3).

Section 6 explores two implications for ramp metering. First, if the demand is infeasible and there is no-metering, every trajectory converges to the most congested equilibrium. However, there is a ramp metering strategy that increases flow in every section and reduces total travel time (Theorem 6.1). Second, even with feasible demand if the freeway is in a congested equilibrium, there is a ramp metering strategy that moves the freeway to the uncongested equilibrium while reducing total travel time (Theorem 6.2).

Section 7 summarizes the most significant results. Some proofs are collected in the Appendix.

2. Background of CTM

The simplest continuous macroscopic model was proposed by Lighthill and Whitham (1955) and Richards (1956), hence called LWR. Based on the conservation of vehicles, LWR is described by a single partial differential equation in conservation form. A significant literature proposes and extends discrete approximations to LWR.

Second order models with two equations were proposed by Payne (1971) and Whitham (1974), which Daganzo (1995) showed to be unsuitable for describing traffic, since in these models vehicles may exhibit negative speeds. Aw and Rascle (2000) proposed an improved second order model, further studied in Haut and Bastin (2005), Herty and Rascle (2006), Piccoli and Garavello (2006). A third order Navier–Stokes type model was introduced in Helbing (1995).

As indicated in Godlewski and Raviart (1996), the best numerical method to solve the partial differential equations along roads is a Godunov scheme (Godunov, 1959), as it is first-order, correctly predicts shock propagations, lacks oscillating behavior and has a physical interpretation. The spatial domain is divided into cells and the state is constant in each of them. For LWR, it leads to piecewise approximation of the state (density) $\rho(x)$ at each time step, whose evolution is computed for small time intervals if we know the solution of initial value problems with Heaviside initial conditions

$$\rho_i(x) = \begin{cases} \rho^-, & x < 0, \\ \rho^+, & x > 0, \end{cases}$$

wherein i is the initial cell. Such initial value or Riemann problems can be solved in closed form for scalar conservation laws. For a system of conservation laws—as occurs in a multi-cell freeway—when no closed form solution is available, an approximate solver such as the Roe average method is used (Godlewski and Raviart, 1996; LeVeque, 1992).

Several numerical methods can be used as a Godunov scheme (LeVeque, 1992; Lebacque, 1996). The CTM (Daganzo, 1994) is a Godunov scheme in which the flow as a function of density (fundamental diagram) has triangular (or trapezoidal) form.

However, the popularity of CTM is not due to its intellectual roots in LWR but from its flexible use in macroscopic simulation.

3. The CTM model

Fig. 1 shows the freeway divided into N sections, each with one on- and one off-ramp. Vehicles move from right to left. Section i is upstream of section $i-1$. There are two boundary conditions. Free flow prevails downstream of section 0; upstream of the freeway is an "on-ramp" with an inflow of r_N. The flow accepted by section $N-1$ is $f_N(k)$ vehicles per period or time step k. The cumulative difference leads to a queue of size $n_N(k)$ in period k.

Table 1 lists the model variables and parameters with plausible values. The length of all sections is normalized to 1 by absorbing differences in length in the speeds v_i, w_i. To be physically meaningful one must have $0 < v_i$, $w_i < 1$. Off-ramp flows are modeled as a portion $\beta_i(k)$ of the total flow leaving the section

$$s_i(k) = \beta_i(k)(s_i(k) + f_i(k)), \quad \text{or} \quad s_i(k) = [\beta_i(k)/(1-\beta_i(k))]f_i(k).$$

We assume constant split ratios β_i ($\beta_N = 0$). With $\bar{\beta}_i = 1 - \beta_i$, the CTM model is, for $k \geqslant 0$,

$$n_i(k+1) = n_i(k) - f_i(k)/\bar{\beta}_i + f_{i+1}(k) + r_i(k), \quad 0 \leqslant i \leqslant N-1, \tag{1}$$

$$f_i(k) = \min\{\bar{\beta}_i v_i[n_i(k) + \gamma_i r_i(k)], w_{i-1}[\bar{n}_{i-1} - n_{i-1}(k) - \gamma_{i-1} r_{i-1}(k)], F_i\}, \quad 1 \leqslant i \leqslant N, \tag{2}$$

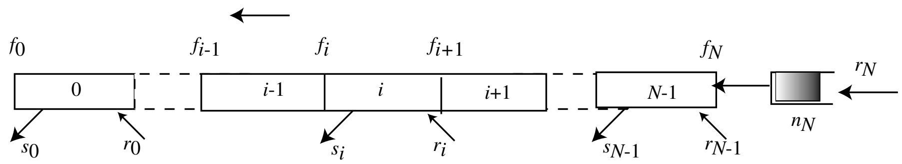

Fig. 1. The freeway has N sections. Each section has one on- and one off-ramp.

Table 1
Model parameters and variables

Symbol	Name	Value	Unit
	Section length	1	miles
	Period (time step)	0.5	min
F_i	Capacity (per lane)	20	veh/period
v_i	Free flow speed	0.5	section/period
w_i	Congestion wave speed	0.5/3	section/period
$\bar{n}_i$	Jam density	160	veh/section
n_i^c	Critical density	40	veh/section
β_i	Split ratio	$\in[0,1)$	dimensionless
$\bar{\beta}_i$	Complementary split ratio $=1-\beta_i$	$\in(0,1]$	dimensionless
γ_i	On-ramp blending factor	$\in[0,1]$	dimensionless
k	Period number	Integer	dimensionless
$f_i(k)$	Flow from section i to $i-1$ in period k	Variable	veh/period
$s_i(k)$, $r_i(k)$	Off-ramp, on-ramp flow in section i in period k	Variable	veh/period
$n_i(k)$	Number of vehicles in section i in period k	Variable	veh/section

$$f_0(k) = \min\{\bar{\beta}_0 v_0[n_0(k) + \gamma_0 r_0(k)], F_0\}, \tag{3}$$
$$n_N(k+1) = n_N(k) - f_N(k) + r_N(k). \tag{4}$$

Flow conservation in section $i \leqslant N-1$ is expressed by

$$n_i(k+1) = n_i(k) - f_i(k) + f_{i+1}(k) + r_i(k) - s_i(k),$$

which is equivalent to (1), using $s_i(k) = \beta_i/\bar{\beta}_i f_i(k)$. Flow conservation at N is expressed by (4). The flow $f_i(k)$ from section i to $i-1$ is governed by the fundamental diagram (2) with this interpretation: $\bar{\beta}_i v_i[n_i(k) + \gamma_i r_i(k)]$ is the number of vehicles that can move from section i to $i-1$ in period k; $w_{i-1}[\bar{n}_{i-1} - n_{i-1}(k) - \gamma_{i-1} r_{i-1}(k)]$ is the number that $i-1$ can accept; and F_i is the capacity or maximum flow from section i to $i-1$. Eq. (3) indicates there is no congestion downstream of section 0. Lastly, it is tacitly assumed that the flows $s_i(k)$ are not constrained by off-ramp capacity.

The parameter values in Table 1 correspond to the fundamental diagram of Fig. 2. Its triangular form incorporates the assumption that is frequently used in our analysis

$$F_i = \bar{\beta}_i v_i n_i^c = w_{i-1}(\bar{n}_{i-1} - n_{i-1}^c). \tag{5}$$

Assumption (5) may appear paradoxical as it relates the capacity F_i to a traffic demand parameter $\bar{\beta}_i$. However, the paradox disappears upon noting that $\bar{\beta}_i v_i n_i(k)$ is the flow from section i to section $i-1$ *downstream* of the off-ramp flow $s_i(k) = \beta_i v_i n_i(k)$.

The *state* of the system is the N-dimensional vector $n(k) = (n_0(k), \ldots, n_{N-1}(k))$. The queue size $n_N(k)$ is *not* included in the state. Section i is said to be *uncongested* or *congested* in period k accordingly as $0 \leqslant n_i(k) \leqslant n_i^c$ or $n_i(k) > n_i^c$.

Remarks. Two reviewers make the following important observations. First, there is empirical evidence of a capacity drop (of about five percent) when the density exceeds its critical value. The triangular fundamental

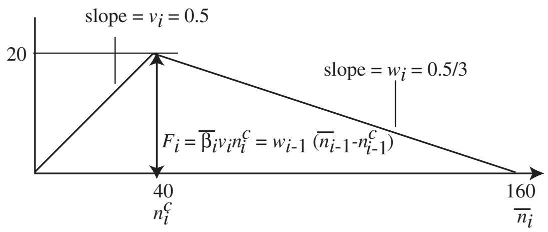

Fig. 2. The fundamental diagram is characterized by the maximum flow F_i and speeds v_i, w_i.

Please cite this article in press as: Gomes, G. et al., Behavior of the cell transmission model and effectiveness ..., Transport. Res. Part C (2007), doi:10.1016/j.trc.2007.10.005

diagram of Fig. 2 then takes the shape of an 'inverse λ'. By neglecting this capacity drop, the benefits of ramp metering evaluated in Section 6 *underestimate* the true benefits. Some emerging research considers the introduction of mainline metering just upstream of each on-ramp, aimed at preserving capacity flow downstream of the on-ramp. Under these conditions, there is no capacity drop (thanks to mainline metering) and the CTM model may be applied for modeling or analysis without change.

Second, taking $v_i < 1$, instead of $v_i = 1$, introduces a numerical error with respect to LWR incurred by the Godunov scheme, as discussed by Leclercq et al. (2005). The choice of $v_i < 1$, $w_i < 1$, needed to accommodate sections of different size, has no impact on the analysis of Sections 4 and 5. The results of Section 6 are also valid, except that the total travel time (due to queuing) would be different from that in LWR.

4. Structure of equilibria

The parameters $\gamma_i \in [0, 1]$ in (2) and (3) reflect the relative position of the on-ramp in section i (Gomes, 2004; Gomes and Horowitz, 2006). For simplicity we assume $\gamma_i = 0$, indicating that the on-ramp is at the beginning of each section as in Fig. 1. (However, the results below hold for a different choice of γ_i.) With $\gamma_i = 0$, Eqs. (1)–(4) simplify

$$n_i(k+1) = n_i(k) - f_i(k)/\bar{\beta}_i + f_{i+1}(k) + r_i(k), \quad 0 \leqslant i \leqslant N-1, \tag{6}$$

$$f_i(k) = f_i(n(k)) = \min\{\bar{\beta}_i v_i n_i(k), w_{i-1}[\bar{n}_{i-1} - n_{i-1}(k)], F_i\}, \quad 1 \leqslant i \leqslant N, \tag{7}$$

$$f_0(k) = f_0(n(k)) = \min\{\bar{\beta}_0 v_0 n_0(k), F_0\}, \tag{8}$$

$$n_N(k+1) = n_N(k) - f_N(k) + r_N(k). \tag{9}$$

In view of (5) a useful alternative to (7) is

$$f_i(k) = \min\{\bar{\beta}_i v_i n_i(k), F_i - w_{i-1}[n_{i-1}(k) - n^{c}_{i-1}], F_i\}, \tag{10}$$

and if section $i-1$ is uncongested $(n_{i-1}(k) \leqslant n^{c}_{i-1})$, (10) simplifies to

$$f_i(k) = \min\{\bar{\beta}_i v_i n_i(k), F_i\}. \tag{11}$$

Fix the split ratios $\beta_0 \ldots, \beta_{N-1}$. Assume stationary demands $r_i(k) \equiv r_i$. Each on-ramp *demand* vector $r = (r_0, \ldots, r_N)$ induces a unique equilibrium *flow* vector $f(r) = (f_0, \ldots, f_N)$ calculated as

$$f_N = r_N, \tag{12}$$

$$f_i = \bar{\beta}_i(f_{i+1} + r_i), \quad 0 \leqslant i \leqslant N-1. \tag{13}$$

The function $r \mapsto f(r)$ defined by (12) and (13) is one-to-one. A demand r is said to be *feasible* if $0 \leqslant f_i \leqslant F_i$, $0 \leqslant i \leqslant N$; it is *strictly feasible* if $0 \leqslant f_i < F_i$, $0 \leqslant i \leqslant N$. A strictly feasible demand induces an equilibrium flow with excess capacity in every section.

A state $n = (n_0, \ldots, n_{N-1})$ is an *equilibrium* for a feasible demand r with induced flow $f = f(r)$, if the constant trajectory $n(k) \equiv n$ is a solution of (6)–(8)

$$f_i = \min\{\bar{\beta}_i v_i n_i, F_i - w_{i-1}(n_{i-1} - n^{c}_{i-1}), F_i\}, \quad 1 \leqslant i \leqslant N-1, \tag{14}$$

$$f_0 = \min\{\bar{\beta}_0 v_0 n_0, F_0\}. \tag{15}$$

An equilibrium n is said to be *uncongested* if all sections are uncongested; otherwise it is *congested*.

Let $E = E(r)$ be the set of equilibria, i.e., all solutions of the nonlinear system of Eqs. (14) and (15), corresponding to a demand r. This section is devoted to characterizing $E(r)$ and the pattern of congested sections for each $n \in E(r)$. If r is not feasible, there is no solution to (14) and (15), so $E(r) = \emptyset$. Lemma 4.1 implies that $E(r) \neq \emptyset$ if r is feasible.

Lemma 4.1. *A feasible demand r has a unique uncongested equilibrium $n^{u}(r)$.*

Proof. *Existence*: Let $f = f(r)$ be the equilibrium flow. Define

$$n_i^{\mathrm{u}} = (\bar{\beta}_i v_i)^{-1} f_i, \quad 0 \leqslant i \leqslant N-1. \tag{16}$$

Then $n_i(k) \equiv n_i^{\mathrm{u}}$ satisfies (6), because (6) is equivalent to (13). Next, because $0 \leqslant f_i \leqslant F_i$ and $F_i = \bar{\beta}_i v_i n_i^{\mathrm{c}}$ (see (5)), $n_i^{\mathrm{u}} = (\bar{\beta}_i v_i)^{-1} f_i \leqslant (\bar{\beta}_i v_i)^{-1} F_i = n_i^{\mathrm{c}}$. So n^{u} is uncongested. It remains to prove that n^{u} is an equilibrium, i.e., satisfies (14), which simplifies to (11) because n^{u} is uncongested. From (16), $f_i = \bar{\beta}_i v_i n_i^{\mathrm{u}}$, and since r is feasible, $f_i \leqslant F_i$. So (11) holds.

Uniqueness: Suppose $\{0 \leqslant n_i \leqslant n_i^{\mathrm{c}};\ 0 \leqslant i \leqslant N-1\}$ is an equilibrium, i.e., satisfies (14) and (15). Since $n_i \leqslant n_i^{\mathrm{c}}$, $\bar{\beta}_i v_i n_i \leqslant \bar{\beta}_i v_i n_i^{\mathrm{c}} = F_i$, therefore (14) reduces to

$$f_i = \min\{\bar{\beta}_i v_i n_i, w_{i-1}(\bar{n}_{i-1} - n_{i-1})\}.$$

If $f_i \neq \bar{\beta}_i v_i n_i$, it must be that $\bar{\beta}_i v_i n_i > w_{i-1}(\bar{n}_{i-1} - n_{i-1}) \geqslant w_{i-1}(\bar{n}_{i-1} - n_{i-1}^{\mathrm{c}}) = F_i$. This contradicts $\bar{\beta}_i v_i n_i \leqslant F_i$, hence f_i must equal $\bar{\beta}_i v_i n_i$, so $n_i = n_i^{\mathrm{u}}$. □

Proposition 4.1. *Suppose at equilibrium n, section $i-1$ is uncongested and sections $i, \ldots, i+j$ are congested. Then*

$$f_i = F_i, \quad \bar{\beta}_i^{-1} F_i - r_i = f_{i+1} < F_{i+1}, \ldots, f_{i+j+1} < F_{i+j+1}. \tag{17}$$

Proof. Since $n_{i-1} \leqslant n_{i-1}^{\mathrm{c}}$, from (11),

$$f_i = \min\{\bar{\beta}_i v_i n_i, F_i\}.$$

Since $n_i > n_i^{\mathrm{c}}$, one has $\bar{\beta}_i v_i n_i > F_i$ from (5); and since r is feasible, $F_i \geqslant f_i$. Hence $f_i = F_i$ and, by (13),

$$f_{i+1} = \bar{\beta}_i^{-1} f_i - r_i = \bar{\beta}_i^{-1} F_i - r_i.$$

Again, as $n_i > n_i^{\mathrm{c}}$, (14) implies

$$f_{i+1} = \min\{\bar{\beta}_{i+1} v_{i+1} n_{i+1}, F_{i+1} - w_i(n_i - n_i^{\mathrm{c}}), F_{i+1}\} < F_{i+1}.$$

Lastly, if section $i+k$ is congested, $n_{i+k} > n_{i+k}^{\mathrm{c}}$, hence

$$f_{i+k+1} = \min\{\bar{\beta}_{i+k+1} v_{i+k+1} n_{i+k+1}, F_{i+k+1} + w_{i+k}[n_{i+k}^{\mathrm{c}} - n_{i+k}], F_{i+k+1}\} < F_{i+k+1},$$

and the remainder of the assertion follows. □

Corollary 4.1. *A strictly feasible demand r has a unique equilibrium, so $E(r) = \{n^{\mathrm{u}}\}$.*

Proof. If the equilibrium $n = (n_0, \ldots, n_{N-1})$ is uncongested, then $n = n^{\mathrm{u}}$ by Lemma 4.1. So suppose there is at least one congested section. There are two cases to consider. In the first case, section 0 is congested. Since $f_0 = \min\{\bar{\beta}_0 v_0 n_0, F_0\} < F_0$ by strict feasibility, so $n_0 < n_0^{\mathrm{c}}$, which means section 0 is not congested. In the remaining case, there must exist a pair of adjacent sections $i-1$, i with $i-1$ uncongested and i congested. But then by Proposition 4.1, $f_i = F_i$, which contradicts strict feasibility of r. □

The next result is a partial converse to Proposition 4.1.

Proposition 4.2. *Suppose $f_i = F_i, f_{i+1} < F_{i+1}, \ldots, f_{i+j} < F_{i+j}$. Suppose at equilibrium n section $i+j$ is congested. Then sections $i, i+1, \ldots, i+j$ are all congested at n.*

Proof. Because section $i+j$ is congested, i.e., $n_{i+j} > n_{i+j}^{\mathrm{c}}$, and $f_{i+j} < F_{i+j}$,

$$f_{i+j} = \min\{\bar{\beta}_{i+j} v_{i+j} n_{i+j}, F_{i+j} - w_{i+j-1}(n_{i+j-1} - n_{i+j-1}^{\mathrm{c}}), F_{i+j}\} = F_{i+j} - w_{i+j-1}(n_{i+j-1} - n_{i+j-1}^{\mathrm{c}}) < F_{i+j},$$

so $n_{i+j-1} > n_{i+j-1}^{\mathrm{c}}$, i.e., section $i+j-1$ is congested. The result follows by induction. □

We say that i is a *bottleneck* section for demand r (or induced flow f) if $f_i = F_i$. (The reason for the name will become clear in Theorem 4.1.) Suppose there are $K \geqslant 0$ bottleneck sections at $0 \leqslant I_1 < I_2 \cdots < I_K \leqslant N-1$. Partition the freeway into $1+K$ segments $S^0, \ldots, S^K$ comprising contiguous sections as follows:

$$S^0 = \{0, \ldots, I_1 - 1\}, S^1 = \{I_1, \ldots, I_2 - 1\}, \ldots, S^K = \{I_K, \ldots, N - 1\}. \tag{18}$$

If there are no bottleneck sections, $K = 0$, $I_1 = N$, and $S^0 = \{0, \ldots, N - 1\}$ is the entire freeway. On the other hand, if the most downstream section is congested, $I_1 = 0$, and S^0 is the empty segment.

Proposition 4.3. *The sections immediately downstream of the segments $S^0, \ldots, S^K$ are uncongested. Consequently, for $k = 1, \ldots, K$,*

$$f_{I_k} = \min\{\bar{\beta}_{I_k} v_{I_k} n_{I_k}, F_{I_k}\}. \tag{19}$$

Proof. The assertion is true of segment S^0 because downstream of section 0 is free flow by assumption. Consider segment S^k. Since

$$f_{I_k} = \min\{\bar{\beta}_{I_k} v_{I_k} n_{I_k}, F_{I_k} - w_{I_k-1}(n_{I_k-1} - n^{\mathrm{c}}_{I_k-1}), F_{I_k}\} = F_{I_k},$$

we must have $n_{I_k-1} \leqslant n^{\mathrm{c}}_{I_k-1}$, i.e., section $I_k - 1$ is uncongested, and (19) holds. □

Partition the N-dimensional state $n = (n_0, \ldots, n_{N-1})$ into sub-vectors $n = (n^0, \ldots, n^K)$ in conformity with the segments $S^0, \ldots, S^K$, so n^k has components $\{n_i,\ i \in S^k\}$. Since the equilibrium flow immediately upstream of segment S^k is known (it is equal to capacity) and the section immediately downstream of S^k is uncongested, the equilibrium conditions (14) and (15) partition into $1 + K$ *decoupled* conditions, one for each segment. Thus n^0 satisfies

$$f_0 = \min\{\bar{\beta}_0 v_0 n_0, F_0\}, \tag{20}$$
$$f_i = \min\{\bar{\beta}_i v_i n_i, F_i - w_{i-1}(n_{i-1} - n^{\mathrm{c}}_{i-1}), F_i\}, \quad 1 \leqslant i \leqslant I_1 - 1, \tag{21}$$

n^k satisfies for $k = 1, \ldots, K$,

$$f_{I_k} = \min\{\bar{\beta}_{I_k} v_{I_k} n_{I_k}, F_{I_k}\}, \tag{22}$$
$$f_i = \min\{\bar{\beta}_i v_i n_i, F_i - w_{i-1}(n_{i-1} - n^{\mathrm{c}}_{i-1}), F_i\}, \quad I_k + 1 \leqslant i \leqslant I_{k+1} - 1. \tag{23}$$

These decoupled conditions decompose the equilibrium set.

Proposition 4.4. *The set of equilibria $E(r)$ factors into the product set,*

$$E(r) = E^0(r) \times \cdots \times E^K(r), \tag{24}$$

in which $E^0(r)$ is the set of solutions n^0 of (20) and (21) and $E^k(r)$ is the set of solutions n^k of (22) and (23) for $k \geqslant 1$.

We now fully characterize the components $E^0(r), \ldots, E^K(r)$. Recall that the flow in all non-bottleneck sections is strictly below capacity:

$$f_i < F_i, \quad i \notin \{I_1, \ldots, I_K\}. \tag{25}$$

Lemma 4.2. *$E^0(r) = \{n^{\mathrm{u},0}\}$ consists of a single point, the component of the uncongested equilibrium n^{u} corresponding to segment S^0. Hence $n^{\mathrm{u},0}$ is given by*

$$n^{\mathrm{u},0}_i = (\bar{\beta}_i v_i)^{-1} f_i, \quad 0 \leqslant i \leqslant I_1 - 1. \tag{26}$$

Proof. Because of (25) the equilibrium flows $f_0, \ldots, f_{I_1-1}$ are strictly below capacity and so, by Corollary 4.1, $E^0(r) = \{n^{\mathrm{u},0}\}$, and (26) follows from (16). □

The next result gives an explicit expression for the equilibrium set $E^k(r)$.

Lemma 4.3. *$n^k \in E^k(r)$ if and only if either* (i) *there is no congested segment at n^k and $n^k = n^{\mathrm{u},k}$, or* (ii) *there exists $j \in \{I_k, \ldots, I_{k+1} - 1\}$ such that at n^k sections $I_k, \ldots, j$ are congested, sections $j + 1, \ldots, I_{k+1} - 1$ are uncongested and n^k is given by* (*see Fig. 3*)

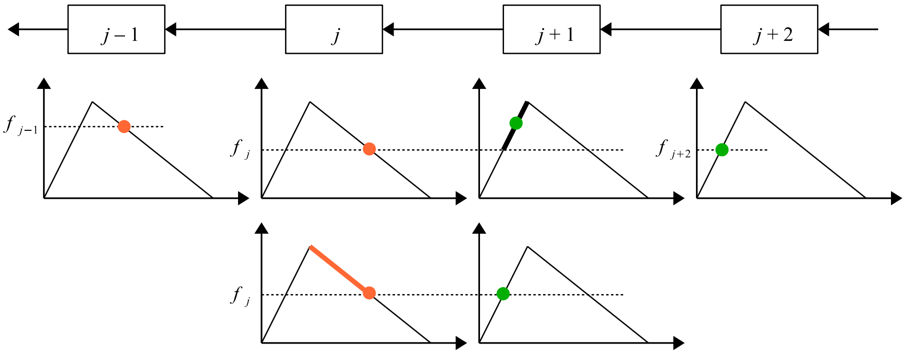

Fig. 3. Equilibrium satisfying (27)–(29) (top) or (30) (bottom).

$$n_i^k = n_i^c + w_i^{-1}(F_{i+1} - f_{i+1}), \quad I_k \leqslant i \leqslant j-1, \tag{27}$$

$$n_i^k = n_i^u = (\bar{\beta}_i v_i)^{-1} f_i, \quad j+2 \leqslant i \leqslant I_{k+1} - 1, \tag{28}$$

$$n_j^k = n_j^c + w_j^{-1}(F_{j+1} - f_{j+1}) \quad and \quad n_{j+1}^k \in [(\bar{\beta}_{j+1} v_{j+1})^{-1} f_{j+1}, n_{j+1}^c], \quad or \tag{29}$$

$$n_j^k \in (n_j^c, n_j^c + w_j^{-1}(F_{j+1} - f_{j+1})] \quad and \quad n_{j+1}^k = n_{j+1}^u = (\bar{\beta}_{j+1} v_{j+1})^{-1} f_{j+1} \tag{30}$$

Proof. Let $n^k \in E^k(r)$ be an equilibrium. Then (i) follows from Lemma 4.1. Next, according to Proposition 4.2 there exists j such that sections $I_k, \ldots, j$ are congested and $j+1, \ldots, I_{k+1} - 1$ are uncongested. Hence for $i \leqslant j - 1$,

$$f_i = \min\{\bar{\beta}_i v_i n_i^k, F_i - w_{i-1}(n_{i-1}^k - n_{i-1}^c), F_i\} = F_i - w_{i-1}(n_{i-1}^k - n_{i-1}^c),$$

because $n_i^k > n_i^c, n_{i-1}^k > n_{i-1}^c$, which proves (27).

For $i \geqslant j+2$,

$$f_i = \min\{\bar{\beta}_i v_i n_i^k, F_i - w_{i-1}(n_{i-1}^k - n_{i-1}^c), F_i\} = \bar{\beta}_i v_i n_i^k,$$

because $n_i^k \leqslant n_i^c, n_{i-1}^k \leqslant n_{i-1}^c$, which proves (28).

Lastly, because $f_{j+1} < F_{j+1}$,

$$f_{j+1} = \min\{\bar{\beta}_{j+1} v_{j+1} n_{j+1}^k, F_{j+1} - w_j(n_j^k - n_j^c), F_{j+1}\} = \min\{\bar{\beta}_{j+1} v_{j+1} n_{j+1}^k, F_{j+1} - w_j(n_j^k - n_j^c)\}$$

Hence either $f_{j+1} = F_{j+1} - w_j(n_j^k - n_j^c)$ and then (29) holds, or $f_{j+1} = \bar{\beta}_{j+1} v_{j+1} n_{j+1}^k$ and then (30) holds. □

Three densities appear in the expression of $E^k(r)$, namely $n_i^u = (\bar{\beta}_i v_i)^{-1} f_i$, the *uncongested* equilibrium density; n_i^c, the *critical* density; and the *congested* equilibrium density

$$n_i^{con} = n_i^c + w_i^{-1}(F_{i+1} - f_{i+1}). \tag{31}$$

By Lemma 4.3

$$E^k(r) = \{n^{u,k}\} \bigcup_{j \in S^k} E_j^k(r), \tag{32}$$

in which $E_j^k(r)$ is the set of n^k satisfying (27)–(30):

$$E_j^k(r) = \{(n_{I_k}^{con}, \ldots, n_{j-1}^{con}, n_j, n_{j+1}^u, \ldots, n_{I_{k+1}-1}^u) | n_j \in (n_j^c, n_j^{con}]\} \\ \bigcup \{(n_{I_k}^{con}, \ldots, n_j^{con}, n_{j+1}, n_{j+2}^u, \ldots, n_{I_{k+1}-1}^u) | n_{j+1} \in [n_{j+1}^u, n_{j+1}^c]\}. \tag{33}$$

Observe that $f_{I_k} = F_{I_k}$, $n_{I_k}^u = (\bar{\beta}_{I_k} v_{I_k})^{-1} F_i = n_{I_k}^c$. So it follows from (33) that $n^{u,k} \in E_{I_k}^k(r)$. Hence (32) simplifies to

$$E^k(r) = \bigcup_{j \in S^k} E_j^k(r). \tag{34}$$

Observe next that the first set on the right hand side in (33) forms a straight line segment E^k_{j-} connecting the points

$$n^k(j-) = (n^{\rm con}_{I_k}, \ldots, n^{\rm con}_{j-1}, n^{\rm c}_j, n^{\rm u}_{j+1}, \ldots, n^{\rm u}_{I_{k+1}-1}) \tag{35}$$

and

$$n^k(j) = (n^{\rm con}_{I_k}, \ldots, n^{\rm con}_{j}, n^{\rm u}_{j+1}, \ldots, n^{\rm u}_{I_{k+1}-1}). \tag{36}$$

Denote this line segment in terms of its endpoints as

$$E^k_{j-}(r) = (n^k(j-), n^k(j)]. \tag{37}$$

Similarly, the second set on the right hand side in (33) forms a straight line segment connecting the points $n^k(j)$ and

$$n^k(j+) = (n^{\rm con}_{I_k}, \ldots, n^{\rm con}_{j}, n^{\rm c}_{j+1}, n^{\rm u}_{j+2} \ldots, n^{\rm u}_{I_{k+1}-1}), \tag{38}$$

and denoted as

$$E^k_{j+}(r) = [n^k(j), n^k(j+)]. \tag{39}$$

The two line segments have exactly one point, $n^k(j)$, in common. Thus

$$E^k_j(r) = E^k_{j-}(r) \cup E^k_{j+}(r), \tag{40}$$

and, by comparing (35) and (38) one sees that

$$n^k(j+) = n^k(j+1-), \tag{41}$$

so that $E^k_j(r)$ and $E^k_{(j+1)}(r)$ have exactly this point in common. Lastly, since the densities $n^{\rm u}_i \leqslant n^{\rm c}_i \leqslant n^{\rm con}_i$ are ordered, so are the endpoints:

$$\cdots \leqslant n^k(j-) \leqslant n^k(j) \leqslant n^k(j+) = n^k((j+1)-) \leqslant n^k(j+1) \leqslant \cdots \tag{42}$$

(For vectors x and y, $x \leqslant y$ means $x_i \leqslant y_i$ for all components i.)

Fig. 4 depicts the projection of $E^k_j(r) = E^k_{j-}(r) \cup E^k_{j+}(r)$ on the two-dimensional space spanned by n^k_j, n^k_{j+1} and the projection of $E^k_{j+1} = E^k_{(j+1)-}(r) \cup E^k_{(j+1)+}(r)$ on the space spanned by n^k_{j+1}, n^k_{j+2}. According to (41) the two highlighted points in the figure are the same.

Observe lastly that the straight line segments $E_{(j+1)-}$ and E_{j+} are aligned.

Theorem 4.1 follows from Proposition 4.4, and Lemmas 4.2 and 4.3.

Theorem 4.1. *Let r be a feasible demand, f the induced equilibrium flow, and $E(r)$ the equilibrium set. If r is strictly feasible, $E(r)$ consists of the unique uncongested equilibrium $n^{\rm u}$. Otherwise, partition the freeway into*

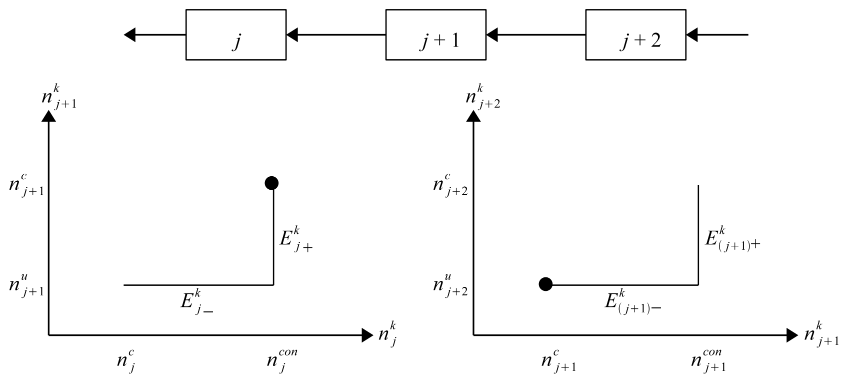

Fig. 4. Projection of $E^k_j(r)$ on coordinates n^k_j, n^k_{j+1} (left) and projection of $E^k_{j+1}(r)$ on coordinates n^k_{j+1}, n^k_{j+2} (right).

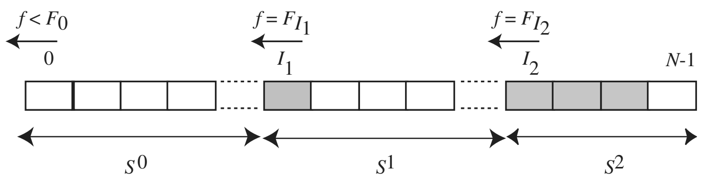

Fig. 5. The demand induces two bottleneck sections and three segments. S^0 is uncongested. In the depicted equilibrium S_1 has one congested section and S^2 has three congested sections.

segments $S^0, \ldots, S^K$ *corresponding to the bottleneck sections* $0 \leqslant I_1 < \cdots < I_K \leqslant N-1$. *Then* $E(r)$ *is the direct product* (24):

$$E(r) = E^0(r) \times \cdots \times E^K(r).$$

Each $E^k(r)$ *decomposes as the union* (34):

$$E^0(r) = \{n^{u,0}\}, \quad E^k(r) = \bigcup_{j \in S^k} E_j^k(r), k \geqslant 1.$$

Each $E_j^k(r)$ *is the union of two connected line segments, given by the 'closed form' expression* (37), (39) *and* (40). E^k (r) *is the union of* $|S^k|$ *connected straight line segments.* ($|S_k| = |I_{k+1} - I_k|$ *is the number of sections in* S^k.)

Consecutive sets $E_j^k(r)$ *and* $E_{j+1}^k(r)$ *have exactly one point in common, and they are ordered: if* $n^k \in E_j^k(r)$ *and* $n'^k \in E_{j+1}^k$, *then* $n^k \leqslant n'^k$. *In particular, the most congested equilibrium in* $E^k(r)$ *is* $n^{\mathrm{con},k}$ *with components* $n_i^{\mathrm{con}}, i \in S^k$, *given by* (31). *Every* $n^k \in E^k(r)$ *lies between the uncongested equilibrium* $n^{\mathrm{u},k}$ *and* $n^{\mathrm{con},k}$, *i.e.,* $n^{\mathrm{u},k} \leqslant n^k \leqslant n^{\mathrm{con},k}$. *Hence for all* $n \in E(r)$,

$$n^{\mathrm{u}} \leqslant n \leqslant n^{\mathrm{con}},$$

in which the most congested equilibrium is $n^{\mathrm{con}} = (n^{\mathrm{u},0}, n^{\mathrm{con},1}, \ldots, n^{\mathrm{con},K})$.

Lastly, $E(r)$ *forms a connected, topologically closed surface of dimension* K *in the* N*-dimensional state space.*

Proof. Only the last assertion needs proof, which follows from the observation that $E(r)$ is the product of $1+K$ sets, $E^0(r), \ldots, E^K(r)$, the first of which being a single point has dimension 0, and each of the rest being a union of connected line segments has dimension 1. ☐

Fig. 5 illustrates the use of Theorem 4.1. The demand induces a flow that gives rise to bottlenecks at I_1, I_2 which partition the freeway into three segments S^0, S^1, S^2. S^0 is uncongested. An equilibrium determines the number of congested sections in the other segments. The figure illustrates an equilibrium in which one section in S^1 and three sections in S^2 are congested (depicted by shaded rectangles); the others are uncongested. The congested sections must lie immediately upstream of the corresponding bottleneck. This simple consequence of the CTM model, which conforms to empirical observations, has surprisingly not been previously noticed.

5. Dynamic behavior

Theorem 4.1 fully characterizes the equilibrium behavior of any CTM model. This section is devoted to the complete description of the qualitative behavior of all trajectories of the N-dimensional difference equation system (6)–(8). We assume a constant feasible demand r and write (6)–(8) as

$$n_i(k+1) = g_i(n(k)), \quad 0 \leqslant i \leqslant N-1. \tag{43}$$

Let $g = (g_0, \ldots, g_{N-1})$. We will consider initial conditions

$$n \in \Sigma = \{n | 0 \leqslant n_i \leqslant \bar{n}_i, \ 0 \leqslant i \leqslant N-1\}. \tag{44}$$

Each initial condition $n(0) \in \Sigma$ generates a *trajectory* $\{n(k), k \geqslant 0\}$ according to $n(k+1) = g(n(k))$.

Please cite this article in press as: Gomes, G. et al., Behavior of the cell transmission model and effectiveness ..., Transport. Res. Part C (2007), doi:10.1016/j.trc.2007.10.005

For two vectors x, y in R^N, write

$$x \leqslant y \Longleftrightarrow x_i \leqslant y_i,$$
$$x < y \Longleftrightarrow x \leqslant y, x \neq y,$$
$$x \ll y \Longleftrightarrow x_i < y_i.$$

Following Hirsch and Smith (2005) say that g is *strictly monotone* if, for x, $y \in \Sigma$,

$$x < y \Rightarrow g(x) < g(y);$$

g is *strongly monotone* if

$$x < y \Rightarrow g(x) \ll g(y).$$

The Proof of Lemma 5.1 can be found in the Appendix.

Lemma 5.1. *The map g is strictly monotone, but it is not strongly monotone.*

Hirsch and Smith (2005) survey the theory of monotone maps. The most powerful results, however, require strong monotonicity, and do not apply to the CTM model.

Let the equilibrium flow induced by the demand r result in bottlenecks at $0 \leqslant I_1 < \cdots < I_K \leqslant N-1$, and let $S^0, \ldots, S^K$ be the corresponding freeway partition. By Theorem 4.1 every equlibrium lies between the uncongested equilibrium n^{u} and the most congested equilibrium n^{con},

$$n^{\mathrm{u}} \leqslant n \leqslant n^{\mathrm{con}}, \quad n \in E(r). \tag{45}$$

Let $\hat{n}(k), k \geqslant 0$ be the trajectory starting with the empty freeway, $\hat{n}(0) = 0$, and let $\bar{n}(k), k \geqslant 0$ be the trajectory starting with the completely jammed freeway, $\bar{n}_i(0) = \bar{n}_i,\ 0 \leqslant i \leqslant N-1$. Let $n(k)$ be a trajectory starting in any state $n(0) \in \Sigma$. The next result shows how much monotonicity of g constrains the trajectories of the CTM model.

Lemma 5.2

(i) *Every trajectory lies between* $\{\hat{n}(k)\}$ *and* $\{\bar{n}(k)\}$:

$$\hat{n}(k) \leqslant n(k) \leqslant \bar{n}(k), \quad k \geqslant 0. \tag{46}$$

(ii) *The trajectory starting with the empty freeway converges to the uncongested equilibrium* n^{u}:

$$\lim_{k\to\infty} \hat{n}(k) = n^{\mathrm{u}}. \tag{47}$$

(iii) *The trajectory starting with the completely jammed freeway converges to the most congested equilibrium* n^{con}:

$$\lim_{k\to\infty} \bar{n}(k) = n^{\mathrm{con}}. \tag{48}$$

Proof

(i) Since $\hat{n}(0) \leqslant n(0) \leqslant \bar{n}(0)$, monotonicity implies $\hat{n}(1) \leqslant n(1) \leqslant \bar{n}(1)$, and then (46) follows by induction.

(ii) Since $\hat{n}(1) \geqslant \hat{n}(0) = 0$, monotonicity implies $\hat{n}(2) = g(\hat{n}(1)) \geqslant g(\hat{n}(0)) = \hat{n}(1)$. By induction, the trajectory is increasing: $\hat{n}(k+1) \geqslant \hat{n}(k)$. Since the trajectory is bounded above by the jam density, it must converge to some equilibrium point, say $\hat{n}^{\mathrm{e}}$. Furthermore, since $n(k) \equiv n^{\mathrm{u}}$ is also a trajectory, by (46) one must have $\hat{n}^{\mathrm{e}} \leqslant n^{\mathrm{u}}$, and so (45) implies $\hat{n}^{\mathrm{e}} = n^{\mathrm{u}}$.

(iii) Since $\bar{n} = \bar{n}(0) \geqslant \bar{n}(1)$, monotonicity implies that the trajectory is decreasing: $\bar{n}(k+1) \leqslant \bar{n}(k)$. Since the trajectory is bounded below by 0, it must converge to an equilibrium, say $\bar{n}^{\mathrm{e}}$. As $n(k) \equiv n^{\mathrm{con}}$ is also a trajectory, by (46) one must have $\bar{n}^{\mathrm{e}} \geqslant n^{\mathrm{con}}$, and so (45) implies $\bar{n}^{\mathrm{e}} = n^{\mathrm{con}}$. □

Lemma 5.2 leads to the first interesting result, Theorem 5.1: If the demand is strictly feasible, then n^{u} is a globally, asymptotically stable equilibrium.

Please cite this article in press as: Gomes, G. et al., Behavior of the cell transmission model and effectiveness ..., Transport. Res. Part C (2007), doi:10.1016/j.trc.2007.10.005

Theorem 5.1. *Suppose r is strictly feasible. Then every trajectory converges to* n^{u}.

Proof. By Lemma 4.1 $E(r) = \{n^{\mathrm{u}}\}$, so $n^{\mathrm{con}} = n^{\mathrm{u}}$. Hence both $\hat{n}(k)$ and $\bar{n}(k)$ converge to n^{u}. By (46), every trajectory $n(k)$ converges to n^{u} as well. □

If r is not strictly feasible, the equilibrium set $E(r)$ is infinite and there is no easy way to analyze how trajectories behave. The main result of this section, Theorem 5.3, is that every trajectory converges to some equilibrium. Before getting into the complexities of the proof, we pause to study three examples.

5.1. Examples

Example 1 is a freeway with two identical sections, each one mile long. The fundamental diagram, equilibrium flow, and equilibrium set E are shown in Fig. 6. The critical density $n_i^c = 100$ veh/mile; the jam density $\bar{n}_i = 400$ veh/mile; free flow speed $v = 60$ mph and the congestion wave speed $w = 20$ mph.

The demand vector $r = (r_0 = 1200,\ r_1 = 0,\ r_2 = 4800)$, all in vehicles per hour (vph). The upstream flow $r_2 = f_2 = 4800$ vph, and $f_0 = F_0 = 6000$ vph. Thus section 0 is the only bottleneck. The uncongested equilibrium $n^{\mathrm{u}} = (100, 80)$ and the most congested equilibrium $n^{\mathrm{con}} = (160, 160)$. By Theorem 4.1, the equilibrium set E consists of two straight line segments shown in the figure (also see Fig. 4).

The phase portrait of Fig. 7 displays the orbits of the two-dimensional state with initial conditions on the boundary of the square $\Sigma = [0, 400] \times [0, 400]$. An orbit is the set of states $\{n(k) | k \geqslant 0\}$ traversed by a trajectory $k \to n(k)$. We analyze the orbit structure displayed in the figure. The observations made below hold in general.

1. Every trajectory converges to an equilibrium point in E. As a consequence, the state space Σ is partitioned as

$$\Sigma = \bigcup_{n \in E} \Sigma(n),$$

in which $\Sigma(n)$ is the set of all initial states whose trajectories converge to the equilibrium n. By monotonicity $\Sigma(n^{\mathrm{u}})$ includes all initial states $n \leqslant n^{\mathrm{u}}$, and $\Sigma(n^{\mathrm{con}})$ includes all initial states $n \geqslant n^{\mathrm{con}}$. By contrast, for all other equilibrium states $\Sigma(n)$ is simply a one-dimensional manifold. (In the general case, $\Sigma(n)$ is a $(N - K)$-dimensional manifold, Corollary 5.1.)

2. The figure shows four equi-time contour plots, labeled $k = 12$ s, ..., 600 s. For example, the contour plot $k = 60$ s is the set of points reached by all trajectories at $k = 60$ s. As k increases, the contour plots converge towards the equilibrium set E. As might be expected, the contours initially converge rapidly and the convergence slows down as E is approached. More interestingly, consider the orbit going through the state $n = (340, 50)$ on the $k = 60$ contour. In this state section 0 is congested but section 1 is not. However, by time 200 (whose contour plot is not shown) the state has moved to approximately (250, 150), indicating

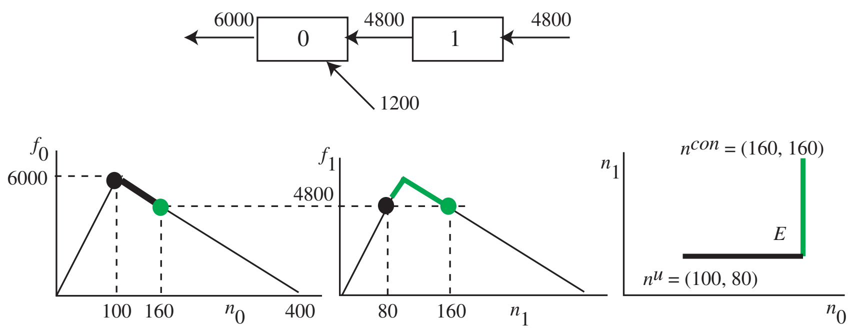

Fig. 6. Freeway, equilibrium flows, fundamental diagram, and equilibrium set E of Example 1.

Please cite this article in press as: Gomes, G. et al., Behavior of the cell transmission model and effectiveness ..., Transport. Res. Part C (2007), doi:10.1016/j.trc.2007.10.005

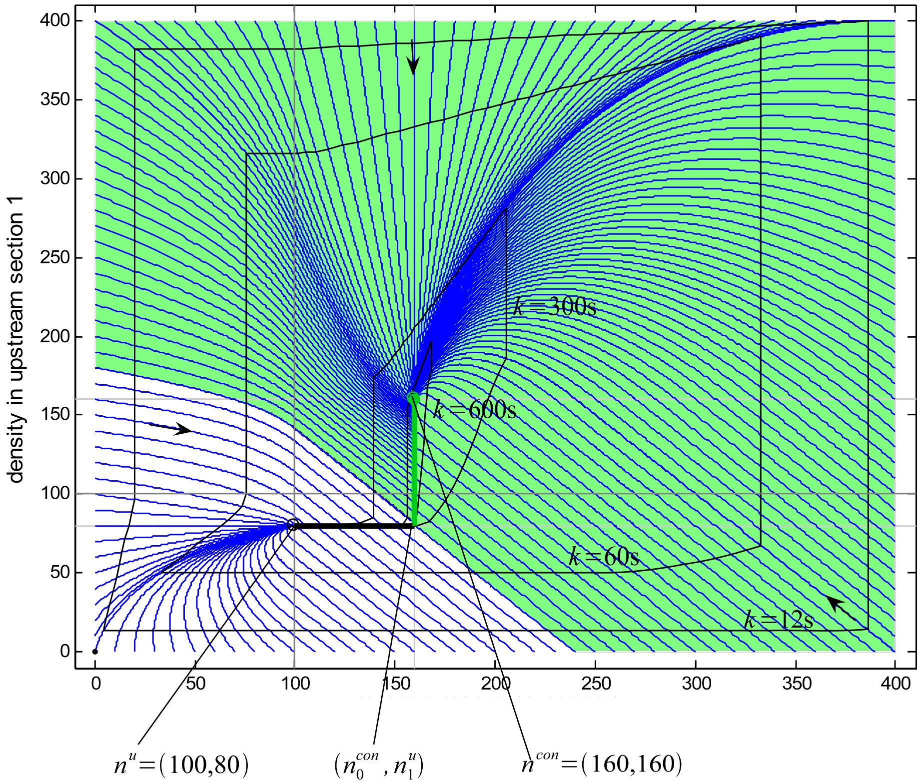

Fig. 7. Equilibrium set and orbits of Example 1.

both sections are congested. The time difference of $200 - 60 = 140$ s is roughly predictable: because the congestion wave speed is 20 mph it takes about 3 min for the congestion wave to travel the one mile-long section.

3. According to Theorem 4.1 the equilibrium set is ordered: if n, n' are two equilibria, either $n \leqslant n'$ or $n' \leqslant n$. Consequently, downstream sections must get congested before an upstream section. As seen in the figure, every trajectory in which section 1 is getting congested also congests section 0. The unshaded area in the figure is the set of all initial states from which trajectories converge to an equilibrium in which the upstream section is *not* congested.
4. All equilibria support the same equilibrium flows. However at equilibrium n^{u} the speed is $v = 60$ mph throughout, whereas in n^{con} the speed is 4800/160 (flow/density) or 30 mph. Thus, although both n^{u} and n^{con} achieve the same throughput, the freeway travel time in n^{u} is one-half of that in n^{con}.

In Example 2 the flow r_2 is slightly reduced from 4800 to 4750 vph, so the demand becomes strictly feasible and the equilibrium set collapses to the single uncongested equilibrium n^{u}. The resulting phase portrait in Fig. 8 can be compared with Fig. 7. The trajectories are nearly identical, except that when they approach the equilibrium set of Example 1 they turn and converge to $n^{\mathrm{u}} \approx (100, 80)$.

Example 3 shown in Fig. 9 is a modification of Example 1 in that there are three identical sections. The fundamental diagram is the same as in Example 1. The demand $r_0 = 1200$, $r_1 = r_2 = 0$, $r_3 = 4800$. Again section 0 is the only bottleneck. The equilibrium set now comprises three straight line segments, connecting the uncongested equilibrium $n^{\mathrm{u}} = (100, 80, 80)$ and the most congested equilibrium $n^{\mathrm{con}} = (160, 160, 160)$. The orbit structure supports the observations made earlier: although it is less apparent in the figure, $\Sigma(n)$ is a two-dimensional manifold if $n \neq n^{\mathrm{u}}, n^{\mathrm{con}}$.

We resume the general discussion. As before let r be a demand vector and ϕ the resulting equilibrium flow vector, i.e. (see (12) and (13))

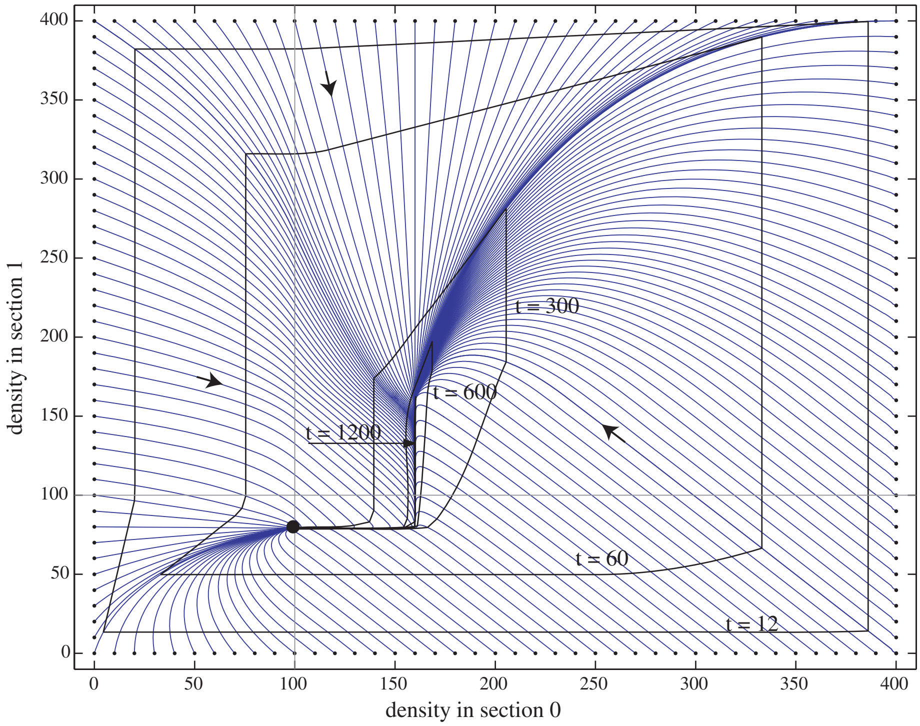

Fig. 8. Equilibrium set and orbits of Example 2.

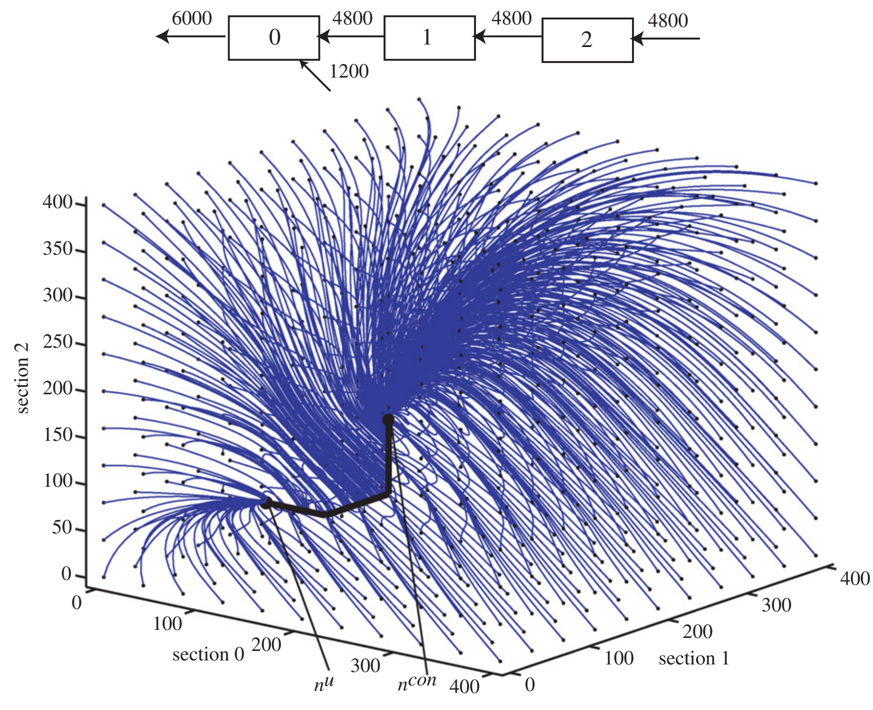

Fig. 9. Freeway, equilibrium set and orbits of Example 3.

$$\phi_N = r_N, \quad \phi_i = \bar{\beta}_i(\phi_{i+1} + r_i), \quad 0 \leqslant i \leqslant N-1. \tag{49}$$

Let $0 \leqslant I_1 < \cdots < I_K \leqslant N-1$ be the bottlenecks, and $S^0, \ldots, S^K$ the corresponding freeway partition. By Theorem 4.1 the equilibrium set decomposes as

$$E(r) = \{n^{u,0}\} \times E^1(r) \times \cdots \times E^K(r). \tag{50}$$

Let $\hat{n}(k), k \geqslant 0$, be the trajectory starting at 0 and converging to n^{u}. Let $\bar{n}(k)$ be the trajectory starting at $\bar{n}$ and converging to n^{con}.

Fix an initial condition $n \in \Sigma$ and let $n(k), k \geqslant 0$, be the trajectory starting at n.

We recall some facts from the general theory of dynamical systems. The ω-*limit* set of n is the set of all limit points of the trajectory $\{n(k)\}$:

$$\omega(n) = \left\{p \in \Sigma | \text{there is a subsequence } k_m \text{ with } \lim_{m\to\infty} n(k_m) = p\right\}.$$

$\omega(n)$ is non-empty, compact, and *invariant*, i.e., if $p \in \omega(n)$ the trajectory starting at p stays within $\omega(n)$. Furthermore the trajectory converges to $\omega(n)$, i.e., $\lim_k d(n(k), \omega(n)) = 0$, with $d(x, \omega(n)) = \min\{|x-p| \,|\, p \in \omega(n)\}$.

Our objective is to prove that the trajectory $\{n(k)\}$ converges to an equilibrium, which is achieved in two steps. The first step shows that $\omega(n)$ always contain an equilibrium (Lemma 5.4). The second step shows that every equilibrium is stable (Theorem 5.2).

We adopt the following notation. For any $p \in \Sigma$,

$$f_i(p) = \begin{cases} \min\{\bar{\beta}_0 v_0 p_0, F_0\}, & i = 0, \\ \min\{\bar{\beta}_i v_i p_i, w_{i-1}[\bar{n}_{i-1} - p_{i-1}], F_i\}, & i \geqslant 1, \end{cases}$$

and

$$f_i(k) = f_i(n(k)).$$

Lemma 5.3

(i) *Suppose* $n^{\mathrm{u}} \leqslant p \leqslant n^{\mathrm{con}}$. *Then*

$$f_i(p) \geqslant \phi_i, \quad \textit{all } i, \quad \textit{and} \quad f_i(p) = F_i, \quad i \in \{I_1, \ldots, I_K\}. \tag{51}$$

(ii) *If* $p \in \omega(n)$, $n^{\mathrm{u}} \leqslant p \leqslant n^{\mathrm{con}}$.

(iii) *Along the trajectory* $\{n(k)\}$

$$\lim_{k\to\infty} \inf f_i(k) \geqslant \phi_i, \quad \textit{all } i, \quad \textit{and} \quad \lim_{k\to\infty} f_i(k) = F_i, \quad i \in \{I_1, \ldots, I_K\}. \tag{52}$$

Proof. Evaluate the three alternatives in $f_i(p) = \min\{\bar{\beta}_i v_i p_i, w_{i-1}[\bar{n}_{i-1} - p_{i-1}], F_i\}$:

$$\begin{aligned} f_i(p) &= \bar{\beta}_i v_i p_i \geqslant \bar{\beta}_i v_i n_i^{\mathrm{u}} = \phi_i, \quad \text{by (16), or} \\ &= w_{i-1}[\bar{n}_{i-1} - p_{i-1}] \geqslant w_{i-1}[\bar{n}_{i-1} - n_{i-1}^{\mathrm{con}}] = \phi_i, \quad \text{by (31), or} \\ &= F_i \geqslant \phi_i, \text{ always.} \end{aligned}$$

Hence $f_i(p) \geqslant \phi_i$ and assertion (i) follows since for bottleneck sections $\phi_i = F_i$. (ii) follows from the observations that $\hat{n}(k) \leqslant n(k) \leqslant \bar{n}(k)$ and $\hat{n}(k) \to n^{\mathrm{u}}, \bar{n}(k) \to n^{\mathrm{con}}$ by Lemma 5.2; hence every limit point p of $\{n(k)\}$ satisfies $n^{\mathrm{u}} \leqslant p \leqslant n^{\mathrm{con}}$. To prove (iii) consider a subsequence $\{k_m\}$ along which $f_i(k_m) \to \lim\inf f_i(k)$ and $n(k_m) \to p \in \omega(n)$. Then $\lim\inf f_i(k) = f_i(p) \geqslant \phi_i$, by (i) and (ii). □

To simplify the discussion we assume that n_N, the upstream ramp queue, is always so large as to maintain

$$f_N(k) \equiv r_N = \phi_N, \quad k \geqslant 0. \tag{53}$$

Lemma 5.4. $\omega(n) \cap E(r) \neq \emptyset$.

Proof. Let $p^0 \in \omega(n)$ and $p(k)$, $k \geqslant 0$, the trajectory starting at p^0. Rewrite (6) in terms of this trajectory, and (49) as

$$p_i(k+1) = p_i(k) - \bar{\beta}_i^{-1} f_i(p(k)) + f_{i+1}(p(k)) + r_i$$
$$r_i = \bar{\beta}_i^{-1}\phi_i - \phi_{i+1}.$$

Adding these together gives

$$p_i(k+1) = p_i(k) + \bar{\beta}_i^{-1}[\phi_i - f_i(p(k))] - [\phi_{i+1} - f_{i+1}(p(k))].$$

Summing this equation for $i = j, \ldots, N-1$ and using (53) leads to

$$\sum_{i=j}^{N-1} p_i(k+1) = \sum_{i=j}^{N-1} p_i(k) + \sum_{i=j}^{N-1} \bar{\beta}_i^{-1}[\phi_i - f_i(p(k))] - \sum_{i=j}^{N-1}[\phi_{i+1} - f_{i+1}(p(k))]$$
$$= \sum_{i=j}^{N-1} p_i(k) + [\phi_j - f_j(p(k))] + \sum_{i=j}^{N-1} \beta_i \bar{\beta}_i^{-1}[\phi_i - f_i(p(k))].$$

By Lemma 5.3, and taking $j = 1$, shows that $\sum_1^{N-1} p_i(k)$ is decreasing, and since it is positive, it converges. Hence $f_i(p(k)) \to \phi_i$ for each i. So if $p^1 \in \omega(p^0)$,

$$f(p^1) = \phi, \quad \text{i.e.,} \quad p^1 \in E(r),$$

from which the assertion follows since $p^1 \in \omega(p^0) \subset \omega(n)$, because $\omega(n)$ is invariant. □

Recall the definition of (Lyapunov) stability: An equilibrium n^e is *stable* if for every $\epsilon > 0$ there is $\delta > 0$ such that $|n - n^e| < \delta$ implies $|n(k) - n^e| < \epsilon$ for all k, in which $\{n(k)\}$ is the trajectory starting at n.

Fix an equilibrium n^e. By Theorem 4.1 n^e has the form

$$n^e = (n^{e,0}, n^{e,1}, \ldots, n^{e,K}),$$

with $n^{e,0} = n^{u,0}$, $n^{e,m} \in E_j^m(r)$ for some $j \in S^m$, $1 \leqslant m \leqslant K$.

Lemmas 5.5 and 5.6 will prove that if $|n - n^e| < \delta$, and $n(k) = (n^0(k), n^1(k), \ldots, n^K(k))$, $k \geqslant 0$ is the trajectory starting at n, then there exists an equilibrium $\tilde{n}^e$, possibly different from n^e, such that

$$|\tilde{n}^{e,m} - n^{e,m}| < \epsilon, \quad \lim_{k\to\infty} n^m(k) = \tilde{n}^{e,m}, \quad 0 \leqslant m \leqslant K, \tag{54}$$

Lemma 5.5. *(54) holds for $m = 0$. In fact*

$$\lim_{k\to\infty} n^0(k) = n^{u,0} = n^{e,0}. \tag{55}$$

Proof. By Lemma 5.2, $\hat{n}(k) \leqslant n(k) \leqslant \bar{n}(k)$, $\hat{n}(k) \to n^u$, $\bar{n}(k) \to n^{con}$. Then (55) follows because, by (50), $n^{con,0} = n^{u,0}$. □

Lemma 5.6. *Suppose (54) holds for $m - 1 \geqslant 0$. Then it holds for m.*

Proof. Consider (54) for $m \geqslant 1$. By Theorem 4.1 $n^{e,m} \in E_{I_m+j}^m(r)$ for some $j \geqslant 0$, so that sections $I_m, \ldots, I_m + j$ are congested and $I_m + j + 1, \ldots, I_{m+1} - 1$ are uncongested as indicated in Fig. 10.

Fig. 10. In equilibrium $n^{e,m}$ sections $I_m, \ldots, I_m + j$ of S^m are congested.

Please cite this article in press as: Gomes, G. et al., Behavior of the cell transmission model and effectiveness ..., Transport. Res. Part C (2007), doi:10.1016/j.trc.2007.10.005

The proof, which separately analyzes the three cases, $j = I_{m+1} - 1$, $j = I_m$, and $I_m < j < I_{m+1} - 1$, is long and placed in the Appendix. □

Theorem 5.2. *Every equilibrium* $n^e \in E(r)$ *is stable. In fact for* $\epsilon > 0$ *there is* $\delta > 0$ *such that if* $|n - n^e| < \delta$, *the trajectory* $\{n(k)\}$ *starting at* n *converges to an equilibrium* $\tilde{n}^e$ *with* $|\tilde{n}^e - n^e| < \epsilon$.

Proof. Lemmas 5.5 and 5.6 prove the second part of the assertion which implies stability. □

Fig. 7 illustrates Theorem 5.1. Trajectories starting close to an equilibrium all converge to some nearby equilibrium.

Theorem 5.3. *The CTM model is a convergent system, i.e. every trajectory converges to some equilibrium in* $E(r)$.

Proof. Consider any trajectory $\{n(k)\}$. By Lemma 5.4 there is an equilibrium n^e and a subsequence $\{k_m\}$ along which $n(k_m) \to n^e$ as $m \to \infty$. By Theorem 5.2 the trajectory must converge to this equilibrium. □

Recall that the stable manifold $\Sigma(n^e)$ of an equilibrium $n^e \in E(r)$ comprises all $n \in \Sigma$ whose trajectories converge to n^e. The next result characterizes the orbit structure.

Corollary 5.1. *If* r *is strictly feasible,* $E(r) = \{n^u\}$ *and* $\Sigma\ (n^u) = \Sigma$. *If* r *is not strictly feasible,* $E(r)$ *is a* K-*dimensional manifold and* $\Sigma(n^e)$ *is a* $(N - K)$-*dimensional manifold for* $n^e \neq n^u$, n^{con}, *whereas* $\Sigma(n^u)$, $\Sigma(n^{con})$ *are* N-*dimensional manifolds with boundary.*

Proof. By Theorem 4.1 $E(r)$ is a K-dimensional manifold. By Theorem 5.3

$$\Sigma = \bigcup_{n^e \in E(r)} \Sigma(n).$$

By Lemma 5.2 every trajectory starting at $n \leqslant n^u$ converges to n^u and every trajectory starting at $n \geqslant n^{con}$ converges to n^{con}. Because $E(r)$ is ordered, it is not very difficult to show, using monotonicity, that the stable manifolds of all equilibria $n^e \neq n^u$, n^{con} are diffeomorphic. The assertion then follows. □

6. Implications for ramp metering

We explore two implications for ramp metering. The first considers the case when the demand vector r is *infeasible*, i.e., the associated equilibrium flow ϕ given by (49) is such that it exceeds the capacity in some section. Peak hour demand may be infeasible in this sense. We begin with an example to illustrate the issues.

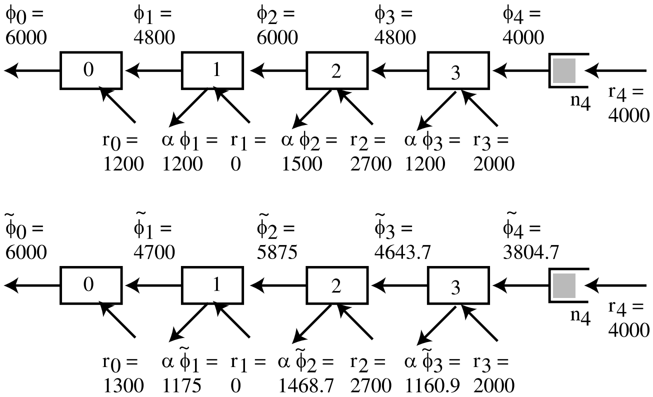

Fig. 11. Freeway, on-ramp and off-ramp flows of example: feasible demand (top); excess demand (bottom).

Please cite this article in press as: Gomes, G. et al., Behavior of the cell transmission model and effectiveness ..., Transport. Res. Part C (2007), doi:10.1016/j.trc.2007.10.005

Example. The upper part of Fig. 11 displays a freeway with four identical sections, each with capacity 6000 vph. The demand vector $r = (r_0 = 1200,\ r_1 = 0, r_2 = 2700,\ r_3 = 2000,\ r_4 = 4000)$. All split ratios are the same: $\beta_i = \beta = 0.2$, so $\bar{\beta} = 0.8$ and $\alpha = \beta[\bar{\beta}]^{-1} = 0.25$. The demand r is feasible and the equilibrium flow $\phi = (\phi_0 = 6000,\ \phi_1 = 4800,\ \phi_2 = 6000,\ \phi_3 = 4800,\ \phi_4 = 4000)$. The off-ramp flow in section i is $\alpha\phi_i$. Sections 0 and 2 are bottleneck sections, with equilibrium flows equal to capacity.

Now consider the demand $\tilde{r}$ in which $\tilde{r}_0 = 1300 > r_1$ and $\tilde{r}_i = r_i, i \geqslant 0$. This demand is not feasible because it would increase ϕ_0 to $\phi_1 + \tilde{r}_0 = 6100$, which exceeds capacity. Evidently, the increased on-ramp flow in section 0 will create congestion in section 0 and force a reduction in the flow out of section 1 from $\phi_1 = 4800$ to $\tilde{\phi}_1 = 4700$. This reduction from ϕ_1 to $\tilde{\phi}_1$ is achieved by a reduction in the flow from section 2 from $\phi_2 = 6000$ to $\tilde{\phi}_2 = 5875$, which in turn reduces the flow from section 3 from $\phi_3 = 4800$ to $\tilde{\phi}_3 = 4643.75$, and ultimately the flow from section 4 from $\phi_4 = 4000$ to $\tilde{\phi}_4 = 3804.6875$. As a result the on-ramp queue n_4 will grow at the rate of $4000 - 3804.6875 = 195.3125$ vph. All sections will become congested.

The reductions in the equilibrium flow from ϕ to $\tilde{\phi}$ will proportionately reduce the discharge at the off-ramps from $\alpha\phi_i$ to $\alpha\tilde{\phi}_i$. The new equilibrium flows are displayed in the lower part of the figure.

The example suggests some observations.

1. The infeasible demand $\tilde{r}$ leads to a unique equilibrium flow $\tilde{\phi}$. This is the flow corresponding to the feasible demand $\tilde{r}^f$, which is the same as $\tilde{r}$, except that the upstream flow is reduced from $\phi_4 = 4000$ to $\tilde{\phi}_4 \approx 3804$. The system converges to the (unique) most congested equilibrium corresponding to $\tilde{r}^f$.
2. The reduction in the flow at the upstream ramp of about $196 = 4000 - 3804$ vph is nearly *twice* the 'excess' demand of $1300 - 1200 = 100$ vph at the ramp in section 0. Suppose that we meter the on-ramp in section 0 and admit only 1200 vph. The queue at this ramp will now grow at 100 vph, but the resulting equilibrium flow and the off-ramp discharges will be the same as in the top of the figure; hence the total discharge will be *higher* by $196 - 100 = 96$ vph.

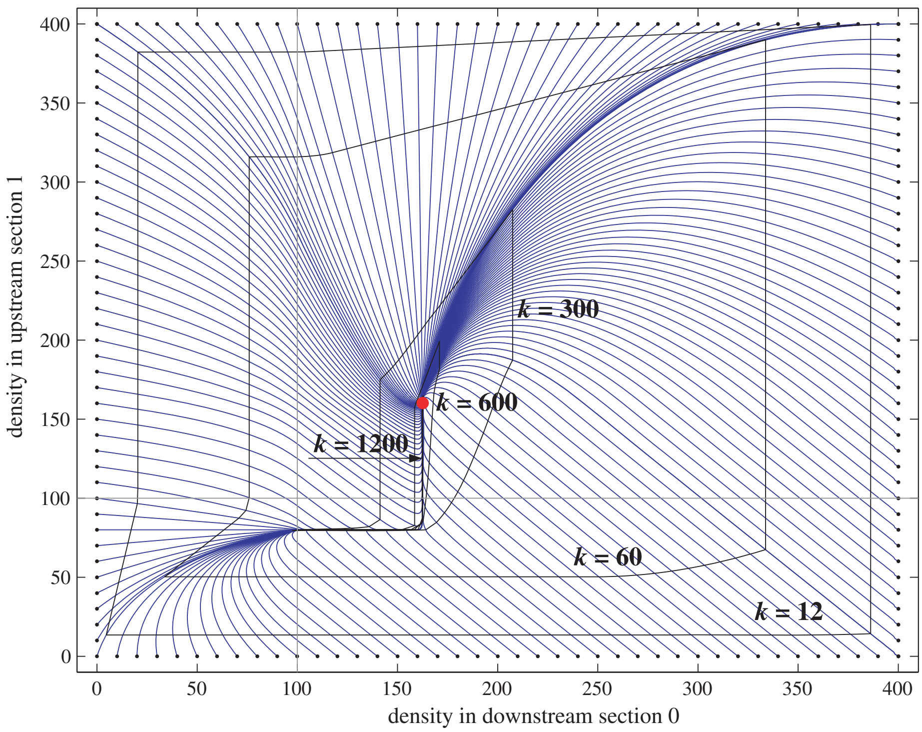

Fig. 12. Orbits of the infeasible demand example.

Please cite this article in press as: Gomes, G. et al., Behavior of the cell transmission model and effectiveness ..., Transport. Res. Part C (2007), doi:10.1016/j.trc.2007.10.005

3. Fig. 12 shows the phase portrait of the freeway considered in Fig. 6 except that the on-ramp flow in section 0 is increased to 1250 so that the demand becomes infeasible. The figure also displays the equilibrium set $E(\tilde{r}^f)$. All of the trajectories converge to the most congested equilibrium in $E(\tilde{r}^f)$. There is a pleasing symmetry with the case of strictly feasible demand, in which every trajectory converges to the uncongested equilibrium as in Fig. 8.

The next result places the example above in a general setting. The freeway structure is the same as in Sections 3–5. Let $r = (r_0, \ldots, r_N)$ be a demand vector. Let ϕ be the solution of (49):

$$\phi_N = r_N, \quad \phi_i = \bar{\beta}_i(\phi_{i+1} + r_i), \;\; 0 \leqslant i \leqslant N-1.$$

Suppose that r is *infeasible*, so that $\phi_i > F_i$ for some i.

To simplify the notation we make two assumptions. First $\phi_0 > F_0$, and if $r_0 = 0$ the demand becomes feasible. Second, if $r_N = 0$ (zero inflow from the upstream ramp) the demand again becomes feasible.

Since $\phi_0 > F_0$, under demand r the entire freeway will become congested as in the example. Since with $r_N = 0$ the demand is feasible,

$$\tilde{r}_N = \max\{\rho \geqslant 0 | \text{ the demand } (r_0, \ldots, r_{N-1}, \rho) \text{ is feasible}\} \tag{56}$$

is well-defined, i.e., $\tilde{r}_N \geqslant 0$. Since with $r_0 = 0$ the demand is feasible,

$$\hat{r}_0 = \max\{\rho \geqslant 0 | \text{ the demand } (\rho, r_1, \ldots, r_N) \text{ is feasible}\} \tag{57}$$

is similarly well-defined.

Theorem 6.1

(i) $\tilde{r}_N < r_N$ *is the largest upstream flow for which the demand* $\tilde{r} = (r_0, \ldots, r_{N-1}, \tilde{r}_N)$ *is feasible. The corresponding equilibrium flow* $\tilde{\phi}$ *is*

$$\tilde{\phi}_N = \tilde{r}_N, \quad \phi_i = \bar{\beta}_i(\tilde{\phi}_{i+1} + r_i), \;\; 0 \leqslant i \leqslant N-1.$$

(ii) *With demand* r, *under the no-metering strategy every trajectory converges to the* (*unique*) *most congested equilibrium* $n^{con} \in E(\tilde{r})$ *corresponding to demand* $\tilde{r}$. *Moreover, the queue* $n_N(k)$ *at the upstream ramp grows indefinitely at the rate of* $(r_N - \tilde{r}_N)$ *vehicles per period.*

(iii) $\hat{r}_0 < r_0$ *is the largest flow for which the demand* $\hat{r} = (\hat{r}_0, r_1, \ldots, r_N)$ *is feasible. The corresponding equilibrium flow* $\hat{\phi}$ *is*

$$\hat{\phi}_N = r_N, \quad \hat{\phi}_i = \bar{\beta}_i(\hat{\phi}_{i+1} + r_i), \;\; 1 \leqslant i \leqslant N-1, \quad \hat{\phi}_0 = \bar{\beta}_0(\hat{\phi}_1 + \hat{r}_0).$$

Under the ramp metering strategy that reduces the on-ramp flow in section 0 *from* r_0 *to* $\hat{r}_0$, *every trajectory converges to some equilibrium in* $E(\hat{r})$. *The queue at the on-ramp in section* 0 *grows indefinitely at the rate of* $(r_0 - \hat{r}_0)$ *vehicles per period.*

(iv) *Flows under the ramp metering strategy are larger throughout the freeway*:

$$\tilde{\phi}_i < \hat{\phi}_i, \;\; 1 \leqslant i \leqslant N \quad \text{and} \quad \tilde{\phi}_0 = \hat{\phi}_0 = F_0.$$

Suppose $\beta_i > 0$ *for some* $i \geqslant 1$, *so that there is non-zero off-ramp flow in at least one section. Then the total discharge under the ramp metering strategy is strictly larger than under the no-metering strategy. Moreover,*

$$\mu = \frac{r_N - \tilde{r}_N}{r_0 - \hat{r}_0} = (\bar{\beta}_1, \ldots, \bar{\beta}_N)^{-1} > 1. \tag{58}$$

Proof. (i) follows from (56) and (49). Since the entire freeway becomes congested under r, every trajectory converges to $n^{\text{con}}(\tilde{r})$ by (48) and, by (i), vehicles accumulate at the upstream ramp at the rate of $(r_N - \tilde{r}_N)$ per period. This proves (ii).

To prove (iii) we solve (49) recursively for $\tilde{r}$ and $\hat{r}$, setting $\bar{\beta}_N = 1$, to get

$$\tilde{\phi}_i = \sum_{j=i}^{N-1} (\bar{\beta}_i, \ldots, \bar{\beta}_j) r_j + (\bar{\beta}_i, \ldots, \bar{\beta}_N) \tilde{r}_N, \quad 0 \leqslant i \leqslant N-1, \tag{59}$$

$$\hat{\phi}_i = \begin{cases} \sum_{j=i}^{N-1} (\bar{\beta}_i, \ldots, \bar{\beta}_j) r_j + (\bar{\beta}_i, \ldots, \bar{\beta}_N) r_N, & 1 \leqslant i \leqslant N-1, \\ \bar{\beta}_0 \hat{r}_0 + \sum_{j=1}^{N-1} (\bar{\beta}_0, \ldots, \bar{\beta}_j) r_j + (\bar{\beta}_0, \ldots, \bar{\beta}_N) r_N, & i = 0. \end{cases} \tag{60}$$

Since $\tilde{r}_N < r_N$ it follows from (59) and (60) that $\tilde{\phi}_i < \hat{\phi}_i, 1 \leqslant i \leqslant N$. Also, since $\hat{r}_0$ is the largest flow that keeps $\hat{\phi}_0 \leqslant F_0$, it must be that $\hat{\phi}_0 = F_0$. Similarly $\tilde{\phi}_0 = F_0$. Hence if $\beta_i > 0$ for some $i \geqslant 1$, then $\beta_i \hat{\phi}_i > \beta_i \tilde{\phi}_i$, i.e., the off-ramp discharge under ramp metering is strictly larger in at least one section. Lastly, from (59) and (60),

$$F_0 = \tilde{\phi}_0 = \sum_{j=0}^{N-1} (\bar{\beta}_0, \ldots, \bar{\beta}_j) r_j + (\bar{\beta}_0, \ldots, \bar{\beta}_N) \tilde{r}_N,$$

$$F_0 = \hat{\phi}_0 = \bar{\beta}_0 \hat{r}_0 + \sum_{j=1}^{N-1} (\bar{\beta}_0, \ldots, \bar{\beta}_j) r_j + (\bar{\beta}_0, \ldots, \bar{\beta}_N) r_N,$$

which, upon subtraction, gives

$$\bar{\beta}_0 (r_0 - \hat{r}_0) = (\bar{\beta}_0, \ldots, \bar{\beta}_N)(r_N - \tilde{r}_N),$$

and so

$$r_0 - \hat{r}_0 = (\bar{\beta}_1, \ldots, \bar{\beta}_N)(r_N - \tilde{r}_N),$$

which implies (58) because $\bar{\beta}_i < 1$ for at least one i. □

Theorem 6.1 prompts some observations.

1. The discussion of infeasible demand above assumes that the on-ramp flow in a section takes priority over the flow from the upstream section: the latter cannot block an on-ramp flow, even if the section is congested. This priority is implicit in the treatment of $r_i(k)$ in (1).
2. The unserved demand under the ramp metering strategy is $(r_0 - \hat{r}_0)$ vehicles per period; the unserved demand under the no-metering strategy is $(r_N - \tilde{r}_N)$. By (58), the no-metering strategy *magnifies* the unserved demand under the ramp strategy by $\mu = (\bar{\beta}_1, \ldots, \bar{\beta}_N)^{-1}$. The larger are the split ratios, the larger is the 'multiplier' μ, and worse is the no-metering strategy. (In the example of Fig. 11 $\mu = (0.8 \times 0.8 \times 0.8)^{-1} \approx 2$).
3. The ramp metering strategy increases speed in every section i (hence reduces travel time). Because $\hat{\phi}_i > \tilde{\phi}_i$ and the freeway is congested under the no-metering strategy, the fundamental diagram implies that the density $\hat{n}_i < \tilde{n}_i$ which, in turn, implies that the speed (=flow/density) under ramp metering is higher: $\hat{\phi}_i/\hat{n}_i > \tilde{\phi}_i/\tilde{n}_i$.
4. Because of (58) the total travel time under the no-metering strategy grows arbitrarily larger than under the no-metering strategy. This contradicts the conclusion of Zhang et al. (1996) that the no-metering strategy minimizes total travel time "if traffic is uniformly congested" as is the case when the demand is infeasible. There is no contradiction with a similar conclusion of Daganzo and Lin (1993) which considers the very special case with no off-ramps, so $\bar{\beta}_i = 1$ for all i, and hence $\mu = 1$ in (58). The special case with one off-ramp is graphically analyzed in Yperman et al. (2004) and Lago and Daganzo (2007).
5. An intuitive explanation of the increased off-ramp discharge under the ramp metering strategy might be that the no-metering strategy creates a congestion "queue" that blocks the off-ramps. This explanation is too crude. Note that under the ramp metering strategy, the system can converge to any equilibrium in $E(\hat{r})$, including the most congested equilibrium $n^{\text{con}}(\hat{r})$, and under the no-metering strategy it converges to the most congested equilibrium $n^{\text{con}}(\tilde{r})$. Thus the entire freeway may be congested under both strategies. Nevertheless, the flows in every section, and hence the off-ramp flows, are larger under the ramp metering strategy. Thus a more accurate (but less intuitive) explanation is that the congestion queue under ramp metering "moves faster" than the queue under the no-metering strategy.

Please cite this article in press as: Gomes, G. et al., Behavior of the cell transmission model and effectiveness ..., Transport. Res. Part C (2007), doi:10.1016/j.trc.2007.10.005

While Theorem 6.1 is intuitively evident, the second implication of the theory *is* surprising: Theorem 6.2 says that ramp metering can reduce total travel time even when the demand is feasible.

Fix a feasible (but not strictly feasible demand) r; let ϕ be its equilibrium flow given by (49) and $E(r)$ its equilibrium set. Recall that $f(n^e) = \phi$ for all $n^e \in E(r)$.

To simplify the notation we assume that under r the only bottleneck is section 0; hence $\phi_0 = F_0$, $\phi_i < F_i$, $1 \leqslant i \leqslant N-1$, $\phi_N = r_N \leqslant F_N$. Suppose the freeway is initially in a congested equilibrium $n(0) = n^e$ in which sections $0,\ldots,j$ are congested for some $j \geqslant 1$, with $n_i^e(0) = n_i^{con}, 0 \leqslant i \leqslant j$, and sections $j+1,\ldots,N-1$ are uncongested with $n_i^e = n_i^u, i \geqslant j+1$. For any $p \in \Sigma$ write $n^u \prec p \prec n^e$ if

$$n_i^u < p_i < n_i^e, \;\; 0 \leqslant i \leqslant j, \quad \text{and} \quad n_i^u = p_i = n_i^e, \;\; i > j. \tag{61}$$

Lemma 6.1 refines Lemma 5.3(i).

Lemma 6.1. *If* $n^u \prec p \prec n^e$,

$$f_0(p) = \phi_0 = F_0, \;\; f_i(p) > \phi_i, \;\; 0 < i \leqslant j, \quad \textit{and} \quad f_i(p) = \phi_i, \;\; i > j. \tag{62}$$

Proof. First,

$$f_0(p) = \min\{\bar{\beta}_0 v_0 p_0, F_0\} \geqslant \min\{\bar{\beta}_0 v_0 n_0^u, F_0\} = \phi_0 = F_0.$$

Next, for $0 < i \leqslant j$ evaluate the three terms in $f_i(p) = \min\{\bar{\beta}_i v_i p_i, F_i - w_{i-1}[\bar{n}_{i-1} - p_{i-1}], F_i\}$ gives

$$\begin{aligned} f_i(p) &= \bar{\beta}_i v_i p_i > \bar{\beta}_i v_i n_i^u = \phi_i, \;\; \text{or} \\ &= F_i - w_{i-1}[\bar{n}_{i-1} - p_{i-1}] > F_i - w_{i-1}[\bar{n}_{i-1} - n_{i-1}^{con}] = \phi_i, \;\; \text{or} \\ &= F_i > \phi_i, \end{aligned}$$

so $f_i(p) > \phi_i$. The last clause in (62) follows from $n_i^u = p_i, i > j$. □

We assume strictly positive demand in the congested sections, so

$$\rho = \min\{r_i, 0 \leqslant i \leqslant j\} > 0.$$

We construct a ramp metering strategy that selects the on-ramp flow values as follows:

$$r_i(k) = \begin{cases} r_i - \mu_i(k), & 0 \leqslant i \leqslant j, \\ r_i, & j < i \leqslant N-1. \end{cases} \tag{63}$$

(The $\{\mu_i\}$ are specified below in (67).) Denote by $p(k)$, $k \geqslant 0$, the trajectory starting at $p(0) = n^e$ under the metering strategy (63).

Lemma 6.2. *There is a finite time horizon* K *and a metering strategy* $\{\mu_i(k), k = 0,\ldots,K\}$ *such that the resulting (controlled) trajectory* $p(k), k = 0,\ldots,K$, *satisfies*

$$n^u \prec p(k) \prec n^e, \quad k = 1,\ldots,K-1, \tag{64}$$

and

$$p(K) = n^u. \tag{65}$$

In particular, the ramp metering strategy steers the state from the initial congested equilibrium n^e *to the uncongested equilibrium* n^u.

Proof. Set $\mu_i(k) \equiv 0$, $i > j$. Following (6), the controlled trajectory is given by

$$p_i(k+1) = p_i(k) - \bar{\beta}_i^{-1} f_i(p(k) + f_{i+1}(p(k)) + r_i - \mu_i(k), \quad 0 \leqslant i \leqslant N-1, \;\; k \geqslant 0. \tag{66}$$

Observe that for $i > j$, $p_i(0) = n_i^e = n_i^u$ and $r_i(k) = r_i$. Hence under any metering strategy of the form (63), $p_i(k) \equiv n_i^e, i > j$. Thus the metering strategy affects the densities only in sections $0,\ldots,j$.

Rewrite (66) as

$$p_i(k+1) = g_i(p(k)) - \mu_i(k), \quad 0 \leqslant i \leqslant N-1, \;\; k \geqslant 0,$$

and define the metering strategy by

$$\mu_i(k) = \begin{cases} \rho, & \text{if } g_i(p(k)) \geqslant n_i^{\mathrm{u}} + \rho, \\ g_i(p(k)) - n_i^{\mathrm{u}}, & \text{if } n_i^{\mathrm{u}} \leqslant g_i(p(k)) < n_i^{\mathrm{u}} + \rho. \end{cases} \quad (67)$$

Since $r_i - \mu_i(k) \geqslant r_i(k) - \rho \geqslant 0$, the metering strategy is feasible (on-ramp flows are non-negative). By construction of μ, $n^{\mathrm{u}} \prec p(k)$. By monotonicity, if $p(k) \prec n^{\mathrm{e}}$ then $p(k+1) = g(p(k)) - \mu(k) \prec g(n^{\mathrm{e}}) - \mu(k)$, so (64) holds by induction.

We now prove (65). Recall that

$$r_i = \bar{\beta}_i^{-1}\phi_i - \phi_{i+1},$$

and substitute for r_i in (66) to get

$$p_i(k+1) = p_i(k) + \bar{\beta}_i^{-1}[\phi_i - f_i(p(k)] - [\phi_{i+1} - f_{i+1}(p(k))] - \mu_i(k), \quad 0 \leqslant i \leqslant j.$$

Adding these equations gives

$$\begin{aligned} \sum_{i=0}^{j} p_i(k+1) &= \sum_{i=0}^{j} p_i(k) + \sum_{i=0}^{j} \bar{\beta}_i^{-1}[\phi_i - f_i(p(k)] - \sum_{i=0}^{j}[\phi_{i+1} - f_{i+1}(p(k))] - \sum_{i=0}^{j}\mu_i(k) \\ &= \sum_{i=0}^{j} p_i(k) + \sum_{i=0}^{j}(\bar{\beta}_i^{-1} - 1)[\phi_i - f_i(p(k)] - \sum_{i=0}^{j}\mu_i(k) + \bar{\beta}_0^{-1}[\phi_0 - f_0(p(k)] \\ &\quad - [\phi_{j+1} - f_{j+1}(p(k))] \\ &= \sum_{i=0}^{j} p_i(k) + \sum_{i=0}^{j}(\bar{\beta}_i^{-1} - 1)[\phi_i - f_i(p(k)] - \sum_{i=0}^{j}\mu_i(k) < -\sum_{i=0}^{j}\mu_i(k), \end{aligned}$$

because $f_0(p(k)) = \phi_0 = F_0, f_{j+1}(p(k)) = \phi_{j+1}$ and $\phi_i - f_i(p(k)) < 0$ by (62). Moreover, from (67), $\sum_{i=0}^{j}\mu_i(k) \geqslant \rho$ for each k for which $g_i(p(k) \geqslant n_i^{\mathrm{u}} + \rho$ for some i. It follows that $p(K) = n^{\mathrm{u}}$ for some $K \leqslant [\sum_i (n_i^{\mathrm{e}} - n_i^{\mathrm{u}})]/\rho$. □

Theorem 6.2. *Suppose the freeway begins in a congested equilibrium n^{e} in which sections $0,\ldots,j$ are congested and sections $j+1,\ldots,N-1$ are uncongested. Then there exists a ramp metering strategy over a finite horizon K at the end of which the freeway is in the uncongested equilibrium n^{u}. Furthermore, the flows during $k = 0,\ldots,K$ are larger than the equilibrium flows. Finally, if the split ratio $\beta_i > 0$ for some $0 \leqslant i \leqslant j$, then the total discharge flow is strictly larger and the total travel time is strictly smaller than in the no-metering strategy.*

Proof. By Lemma 6.2 in each section i the flow $f_i(p(k)) > \phi_i$ for at least one k. Hence the difference in the total discharge

$$\sum_{k=0}^{K} \beta_i\bar{\beta}_i^{-1}[f_i(p) - \phi_i] > 0,$$

from which the assertion follows. □

Two observations are worth making.

1. If the split ratios in the congested sections are all zero, $\beta_i = 0$, $i = 0,\ldots,j$, the ramp metering strategy does not increase the total discharge, but it moves the system to the uncongested equilibrium n^{u} by 'moving' the 'excess' vehicles $\sum_{i=0}^{j}[n_i^{\mathrm{e}} - n_i^{\mathrm{u}}]$ from the congested sections to their on-ramps. The resulting total travel time is unchanged but traffic in the freeway moves at free flow speeds. If some of the traffic in the queues is diverted to alternative routes, perhaps along arterials, there will be a decline in total travel time just as with non-zero split ratios.
2. There is another compelling reason for maintaining the freeway in free flow. The example of Fig. 6 illustrates a common situation in which the congestion density of 160 vehicles/mile (and speed of 30 mph) compares with the uncongested density of 80 vehicles/mile (and speed of 60 mph) for a three-lane freeway.

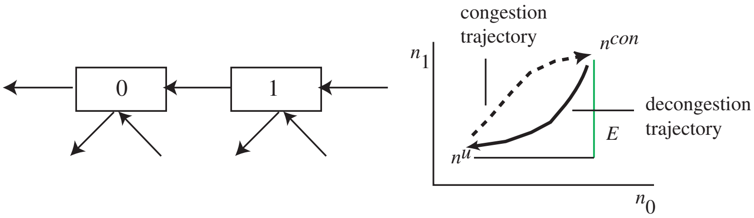

Fig. 13. By creating the cycle from $n^{con} \rightarrow n^{u} \rightarrow n^{con}$ a ramp metering strategy can increase off-ramp discharge.

Storing the 80 additional vehicles would require a 3/4 mile-long one-lane on-ramp (at 50 feet vehicle spacing). Clearly congestion causes the freeway to be used as a very expensive parking place.

3. A '*free lunch*' result lurks behind Theorem 6.2. The result can be understood with the help of Fig. 13 of a two-section freeway whose equilibrium set E is shown on the right. By Lemma 6.2 the flow in section 1 is larger than the equilibrium flow in the rectangle $\{n^{u} < p < n^{con}\}$. The 'decongestion' trajectory constructed in Lemma 6.2 moves the system from n^{con} to n^{u} and causes some additional vehicles to leave the freeway from the off-ramp in section 1. The remainder of the $[(n_1^{con} - n_1^{u}) + (n_0^{con} - n_0^{u})]$ vehicles causing the initial congestion are 'stored' on the on-ramps in sections 0 and 1. Once the sections become uncongested, the ramp metering strategy can now be changed to release the stored vehicles onto the freeway, thereby creating the congestion and moving the state from n^{u} to n^{con} as indicated by the 'congestion' trajectory in the figure. Since this trajectory is inside $\{n^{u} < p < n^{con}\}$ there will again be an additional off-ramp flow. Repeating the two-phase decongestion–congestion cycle provides a free lunch.

7. Conclusions

Despite the widespread use of the CTM model for simulation and analysis of problems in freeway planning and operations, theoretical understanding of the model is spotty. This paper fills this gap by providing a complete analysis of the behavior of the CTM model of a freeway with stationary demand.

The key to the behavior is the location of bottlenecks—sections where flow equals capacity. The bottlenecks partition the freeway into 'decoupled' segments. Each segment starts with a bottleneck and ends just before the next upstream bottleneck. Each segment determines its own equilibrium set; and each equilibrium in the set determines the number of congested sections in the segment.

It is characteristic of the CTM model that within each segment congested segments must lie immediately upstream of the bottleneck. Another characteristic is that congestion must propagate upstream from the bottleneck. Both characteristics are readily observed empirically (Chen et al., 2004; PeMS website, 2007). Of course, as a consequence, as a segment becomes uncongested, upstream sections are relieved before downstream sections.

Bottlenecks may be caused by physical features that reduce capacity. More often, however, bottlenecks depend on the pattern of demand. As the pattern of demand changes by time of day and day of week, bottleneck locations change as well. This characteristic, too, is widely observed.

One surprising conclusion of the analysis is that depending on initial conditions, the *same* demand may leave a segment uncongested or it may congest one or more sections, or even the entire segment. Thus congestion is *not* a sign that capacity exceeds demand, as is commonly believed.

If, however, the demand does exceed capacity in a segment the entire segment will become congested. If ramps are not metered when there is excess demand, freeway utilization will drop and congestion deepen, in comparison with conditions that would prevail under an appropriate ramp metering strategy. When there is excess demand, which typically occurs in every rush hour, proper ramp metering always reduces total travel time because of larger discharge flows, and increases flow and freeway speed. It is a false belief that metering merely transfers delay from a congested freeway to queuing delay on the ramp (except in the singular case of a freeway in which everyone gets off at the same exit).

Please cite this article in press as: Gomes, G. et al., Behavior of the cell transmission model and effectiveness ..., Transport. Res. Part C (2007), doi:10.1016/j.trc.2007.10.005

Surprisingly, even when there is no excess demand, ramp metering can eliminate congestion and reduce total travel time.

Thus the most important practical conclusion of the analysis is that the presence of sustained congestion is a sure indication of the wastefulness of freeway capacity and of travelers' time. Appropriate ramp metering can eliminate this waste. The benefit of ramp metering reported here may underestimate the true benefit if there is a capacity drop in a section when density exceeds its critical value.

Acknowledgements

The authors thank Tarek Hatata, Alex Skabardonis, Fred Dial, Pat Weston and John Wolf for their critical encouragement. This paper is a report of the TOPL (Tools for Operations Planning) Development Group, whose work is supported by the California Department of Transportation through the California PATH Program.

Appendix

Proof of Lemma 5.1. Suppose $x \leqslant y$. We must show

$$g_i(x_{i-1}, x_i, x_{i+1}) \leqslant g_i(y_{i-1}, y_i, y_{i+1}). \tag{68}$$

We verify the inequality one coordinate at a time. Suppose first that $x_{i-1} \leqslant y_{i-1}$ but $x_i = y_i, x_{i+1} = y_{i+1}$. Then from (6) and (7)

$$\begin{aligned} &g_i(x_{i-1}, x_i, x_{i+1}) - g_i(y_{i-1}, y_i, y_{i+1}) \\ &\quad = -\bar{\beta}_i^{-1} \min\{\bar{\beta}_i v_i x_i, w_{i-1}[\bar{n}_{i-1} - x_{i-1}], F_i\} \\ &\quad + \bar{\beta}_i^{-1} \min\{\bar{\beta}_i v_i y_i, w_{i-1}[\bar{n}_{i-1} - x_{i-1}], F_i\} \leqslant 0. \end{aligned}$$

Suppose next that $x_{i+1} \leqslant y_{i+1}$ but $x_{i-1} = y_{i-1}, x_i = y_i$. Then from (6) and (7)

$$\begin{aligned} &g_i(x_{i-1}, x_i, x_{i+1}) - g_i(y_{i-1}, y_i, y_{i+1}) \\ &\quad = \bar{\beta}_{i+1}^{-1} \min\{\bar{\beta}_{i+1} v_{i+1} x_{i+1}, w_i[\bar{n}_i - x_i], F_{i+1}\} \\ &\quad - \bar{\beta}_{i+1}^{-1} \min\{\bar{\beta}_{i+1} v_{i+1} y_{i+1}, w_i[\bar{n}_i - x_i], F_{i+1}\} \leqslant 0. \end{aligned}$$

Lastly suppose $x_i \leqslant y_i$ but $x_{i-1} = y_{i-1}$, $x_{i+1} = y_{i+1}$. To show (68) consider three separate cases.

Case 1:
$x_i < y_i \leqslant n_i^c$. Then from (6), (7) and (11)

$$\begin{aligned} &g_i(x_{i-1}, x_i, x_{i+1}) - g_i(y_{i-1}, y_i, y_{i+1}) \\ &\quad = x_i - \bar{\beta}_i^{-1} \min\{\bar{\beta}_i v_i x_i, w_{i-1}[\bar{n}_{i-1} - x_{i-1}], F_i\} - y_i + \bar{\beta}_i^{-1} \min\{\bar{\beta}_i v_i y_i, w_{i-1}[\bar{n}_{i-1} - x_{i-1}], F_i\} \\ &\quad \begin{cases} = x_i - y_i & \text{if } \bar{\beta}_i v_i x_i \geqslant \min\{\bar{\beta}_i v_i x_i, w_{i-1}[\bar{n}_{i-1} - x_{i-1}], F_i\} \\ \leqslant (1 - v_i) x_i - (1 - v_i) y_i & \text{if } \bar{\beta}_i v_i x_i = \min\{\bar{\beta}_i v_i x_i, w_{i-1}[\bar{n}_{i-1} - x_{i-1}], F_i\} \end{cases} \\ &\quad \leqslant 0, \text{ because } 0 < v_i < 1. \end{aligned}$$

Case 2:
$x_i \leqslant n_i^c < y_i$. Then from (6), (7) and (11)

$$\begin{aligned} &g_i(x_{i-1}, x_i, x_{i+1}) - g_i(y_{i-1}, y_i, y_{i+1}) \\ &\quad = x_i - \bar{\beta}_i^{-1} \min\{\bar{\beta}_i v_i x_i, w_{i-1}[\bar{n}_{i-1} - x_{i-1}], F_i\} + \min\{\bar{\beta}_{i+1} v_{i+1} x_{i+1}, w_i[\bar{n}_i - x_i], F_i\} \\ &\quad - y_i + \bar{\beta}_i^{-1} \min\{\bar{\beta}_i v_i y_i, w_{i-1}[\bar{n}_{i-1} - x_{i-1}], F_i\} - \min\{\bar{\beta}_{i+1} v_{i+1} x_{i+1}, w_i[\bar{n}_i - y_i], F_i\} \end{aligned}$$

If $\bar{\beta}_i v_i x_i \geqslant \min\{\bar{\beta}_i v_i x_i, w_{i-1}[\bar{n}_{i-1} - x_{i-1}], F_i\}$,

$$g_i(x_{i-1},x_i,x_{i+1}) - g_i(y_{i-1},y_i,y_{i+1})$$
$$= x_i - y_i + \min\{\bar\beta_{i+1}v_{i+1}x_{i+1}, w_i[\bar n_i - x_i], F_i\} - \min\{\bar\beta_{i+1}v_{i+1}x_{i+1}, w_i[\bar n_i - y_i], F_i\}$$
$$\begin{cases} \leqslant x_i - y_i \leqslant 0 & \text{if } w_i[\bar n_i - x_i] \leqslant \min\{\bar\beta_{i+1}v_{i+1}x_{i+1}, F_i\} \\ = (1-w_i)(x_i - y_i) \leqslant 0, & \text{if } w_i[\bar n_i - x_i] = \min\{\bar\beta_{i+1}v_{i+1}x_{i+1}, F_i\}, \text{ because } 0 < v_i < 1, \end{cases}$$

and if $\bar\beta_i v_i x_i < \min\{\bar\beta_i v_i x_i, w_{i-1}[\bar n_{i-1} - x_{i-1}], F_i\}$,

$$g_i(x_{i-1},x_i,x_{i+1}) - g_i(y_{i-1},y_i,y_{i+1})$$
$$\leqslant x_i - y_i + \min\{\bar\beta_{i+1}v_{i+1}x_{i+1}, w_i[\bar n_i - x_i], F_i\} - \min\{\bar\beta_{i+1}v_{i+1}x_{i+1}, w_i[\bar n_i - y_i], F_i\} \leqslant 0, \text{ as before.}$$

Case 3:
$n_i^c \leqslant x_i < y_i \leqslant n_i^c$. Then from (6), (7) and (11)

$$g_i(x_{i-1},x_i,x_{i+1}) - g_i(y_{i-1},y_i,y_{i+1})$$
$$= x_i - y_i + \min\{\bar\beta_{i+1}v_{i+1}x_{i+1}, w_i[\bar n_i - x_i]\} - \min\{\bar\beta_{i+1}v_{i+1}x_{i+1}, w_i[\bar n_i - y_i]\}$$
$$\begin{cases} = x_i - y_i \leqslant 0, & \text{if } w_i[\bar n_i - y_i] > \bar\beta_{i+1}v_{i+1}x_{i+1} \\ \leqslant (1-w_i)(x_i - y_i) \leqslant 0, & \text{if } w_i[\bar n_i - y_i] \leqslant \bar\beta_{i+1}v_{i+1}x_{i+1}, \text{ because } 0 < v_i < 1. \end{cases}$$

Thus g is strictly monotone, because if $x \leqslant y$,

$$g_i(x_{i-1},x_i,x_{i+1}) \leqslant g_i(y_{i-1},x_i,x_{i+1}) \leqslant g_i(y_{i-1},y_i,x_{i+1}) \leqslant g_i(y_{i-1},y_i,y_{i+1});$$

moreover, it is trivial to check that if $x \neq y$ then $g(x) \neq g(y)$.

Lastly g is not strongly monotone, because if $x < y$ but $x_{i-1} = y_{i-1}$, $x_i = y_i$, $x_{i+1} = y_{i+1}$, then $g_i(x) = g_i(y)$. □

Proof of Lemma 5.6. Suppose, as induction hypothesis, that (54) holds for $m-1 \geqslant 0$. Fix $m \geqslant 1$. By Theorem 4.1 $n^{e,m} \in E^m_{I_m+j}(r)$ for some $j \geqslant 0$, so that at $n^{e,m}$ sections $I_m, \ldots, I_m + j$ are congested and sections $I_m + j + 1, \ldots, I_{m+1} - 1$ are not congested as in Fig. 10.
We will prove (54) for m, separately analyzing the three cases: $j = I_{m+1} - 1$, $j = I_m$, and $I_m < j < I_{m+1} - 1$.

Case (i): $j = I_{m+1} - 1$. In this case at $n^{e,m}$ the entire segment S^m is congested. The induction hypothesis is not used for this case.By Theorem 4.1 $n^{e,m} = n^{con,m}$ and, by (31),

$$n_i^{e,m} = n_i^c + w_i^{-1}(F_{i+1} - \phi_{i+1}) > n_i^c, \quad i \in S^m. \tag{69}$$

By (52) for $\eta > 0$ we can select $\delta > 0$ so that

$$|n^m - n^{e,m}| < \delta \Rightarrow 0 \leqslant F_{I_m} - f_{I_m}(k) = \eta(k) < \eta. \tag{70}$$

Assume for now that

$$n_i^m(k) > n_i^c, \quad k \geqslant 0, \ i \in S^m, \tag{71}$$

so that

$$\begin{aligned} f_i(k) &= \min\{\bar\beta_i v_i n_i^m(k), F_i - w_{i-1}[n_{i-1}^m(k) - n_{i-1}^c], F_i\} \\ &= F_i - w_{i-1}[n_{i-1}^m(k) - n_{i-1}^c], \quad i = I_m + 1, \ldots, I_{m+1}. \end{aligned} \tag{72}$$

Substituting (69), (70), (72) and (49) in

$$n_i^m(k+1) = n_i^m(k) - \bar\beta_i^{-1} f_i(k) + f_{i+1}(k) + r_i = n_i^m(k) - \bar\beta_i^{-1}[f_i(k) - \phi_i] + f_{i+1}(k) - \phi_{i+1},$$

gives, for $i = I_m$,

$$n_i^m(k+1) = n_i^m(k) - \bar{\beta}_i^{-1}[F_{I_m} - \eta(k) - \phi_i] + F_{i+1} - w_i[n_i^m(k) - n_i^c] - \phi_{i+1}$$
$$= n_i^m(k) - w_i[n_i^m(k) - n_i^c - w_i^{-1}(F_{i+1} - \phi_{i+1})] + \bar{\beta}_i^{-1}\eta(k), \text{ as } F_{I_m} = \phi_{I_m}$$
$$= n_i^m(k) - w_i[n_i^m(k) - n_i^{e,m}] + \bar{\beta}_i^{-1}\eta(k); \quad (73)$$

and, for $i = I_m + 1, \ldots, I_{m+1} - 1$,

$$n_i^m(k+1) = n_i^m(k) - \bar{\beta}_i^{-1}[F_i - w_{i-1}(n_{i-1}^m(k) - n_{i-1}^c) - \phi_i] + F_{i+1} - w_i[n_i^m(k) - n_i^c] - \phi_{i+1}$$
$$= n_i^m(k) + \bar{\beta}_i^{-1} w_i[n_{i-1}^m(k) - n_{i-1}^c - w_i^{-1}(F_i - \phi_i)] - w_i[n_i^m(k) - n_i^c - w_i^{-1}(F_{i+1} - \phi_{i+1})]$$
$$= n_i^m(k) + \bar{\beta}_i^{-1} w_{i-1}[n_i^m(k) - n_i^{e,m}] - w_i[n_i^m(k) - n_i^{e,m}]. \quad (74)$$

Define the vectors $x^m(k)$ with components $x_i^m(k) = n_i^m(k) - n_i^{e,m}, i \in S^m$. In terms of $x^m(k)$ the difference Eqs. (73) and (74) can be written as

$$x^m(k+1) = \begin{bmatrix} 1 - w_{I_m} & 0 & 0 & \ldots & 0 \\ \beta_{I_m+1}^{-1} w_{I_m} & 1 - w_{I_m+1} & 0 & \ldots & 0 \\ . & . & . & . & . \\ 0 & . & . & \beta_{I_{m+1}}^{-1} w_{I_{m+1}-2} & 1 - w_{I_{m+1}-1} \end{bmatrix} x^m(k) + \begin{bmatrix} \beta_{I_m}^{-1}\eta(k) \\ 0 \\ . \\ 0 \end{bmatrix}. \quad (75)$$

The difference Eq. (75) is of the form

$$x^m(k+1) = Ax^m(k) + u(k), \quad x^m(0) = n^m - n^{e,m},$$

and has the solution

$$x^m(k) = A^k(n^m - n^{e,m}) + \sum_{l=0}^{k-1} A^{k-1-l} u(l).$$

The eigenvalues of A are $(1 - w_{I_m}), \ldots, (1 - w_{I_{m+1}-1})$, all of which lie in (0, 1), since $0 < w_i < 1$. Hence $\|A^k\| \leqslant M\lambda^k$ for some $M < \infty$ and $0 < \lambda < 1$. Also $| u(l) | \leqslant (\bar{\beta}_{I_m})^{-1}\eta$. It follows that if $|n^m - n^{e,m}| < \delta$, sufficiently small, then (71) holds and $|x^m(k)| \leqslant \epsilon$ for all $k \geqslant 0$.

Case (ii): $j = I_m$. In this case sections $j + 1 = I_m + 1, \ldots, I_{m+1} - 1$ are not congested; so $\phi_j = F_j$, $\phi_i < F_i$, $i > j$, $i \in S^m$.By the induction hypothesis, for $\epsilon > 0$ there is $\delta > 0$ such that for $|n - n^e| < \delta$, there is an equilibrium $\tilde{n}^e$ such that

$$|n^{m-1}(k) - \tilde{n}^{e,m-1}| < \epsilon, \quad k \geqslant 0.$$

By Proposition 4.3, $\tilde{n}_{j-1}^{e,m-1} < n_{j-1}^c$; hence, for $\epsilon > 0$ small

$$n_{j-1}^{m-1}(k) < \tilde{n}_{j-1}^{e,m-1} + \epsilon < n_{j-1}^c. \quad (76)$$

Next, by (52) we can select $\delta > 0$ so that

$$|n - n^e| < \delta \Rightarrow 0 \leqslant F_{I_m+1} - f_{I_m+1}(k) = \eta(k) \to 0. \quad (77)$$

By Lemma 4.3, $n^{e,m}$ has the form (see bottom part of Fig. 3)

$$n_i^{e,m} = \begin{cases} n_j^c + (1 - 2\psi)w_j^{-1}(F_{j+1} - \phi_{j+1}), & i = j, \text{ for some } 0 \leqslant \psi \leqslant 1/2, \\ n_i^u = (\bar{\beta}_i v_i)^{-1}\phi_i < n_i^c, & i \geqslant j+1, \ i \in S^m. \end{cases}$$

We now examine the trajectory $\{n(k)\}$ starting at n. Assume for now that

$$n_j^m(k) < n_j^{e,m} + \psi w_j^{-1}(F_{j+1} - \phi_{j+1}) = n_j^c + (1 - \psi)w_j^{-1}(F_{j+1} - \phi_{j+1}), \quad (78)$$
$$n_i^m(k) < n_i^{e,m} + \psi(\bar{\beta}_i v_i)^{-1}(n_i^c - n_i^{e,m}), \quad i \geqslant j+1, \ i \in S^m. \quad (79)$$

From (76) and (78)

$$
\begin{aligned}
f_j(k) &= \min\{\bar{\beta}_j v_j n_j^m(k), F_j - w_{j-1}(n_{j-1}^{m-1}(k) - n_{j-1}^{\mathrm{c}}), F_j\} \\
&= \min\{\bar{\beta}_j v_j n_j^m(k), F_j\} = \begin{cases} F_j, & \text{if } n_j^m(k) > n_j^{\mathrm{c}} \\ \bar{\beta}_j v_j n_j^m(k), & \text{if } n_j^m(k) \leqslant n_j^{\mathrm{c}} \end{cases}.
\end{aligned} \tag{80}
$$

Next,

$$
f_{j+1}(k) = \min\{\bar{\beta}_{j+1} v_{j+1} n_{j+1}^m(k), F_{j+1} - w_j(n_j^m(k) - n_j^{\mathrm{c}}), F_{j+1}\}. \tag{81}
$$

From (79),

$$
\bar{\beta}_{j+1} v_{j+1} n_{j+1}^m(k) < \bar{\beta}_{j+1} v_{j+1}[n_{j+1}^{\mathrm{e},m} + \psi(\bar{\beta}_{j+1} v_{j+1})^{-1}(n_{j+1}^{\mathrm{c}} - n_{j+1}^{\mathrm{e},m})] = \phi_{j+1} + \psi(F_{j+1} - \phi_{j+1}),
$$

and from (78)

$$
F_{j+1} - w_j(n_j^m(k) - n_j^{\mathrm{c}}) \geqslant F_{j+1} - w_j[(1-\psi)w_j^{-1}(F_{j+1} - \phi_{j+1})] = \psi F_{j+1} + (1-\psi)\phi_{j+1}.
$$

Substituting the preceding two inequalities into (81) gives

$$
f_{j+1}(k) = \bar{\beta}_{j+1} v_{j+1} n_{j+1}^m(k). \tag{82}
$$

Lastly, for $i \geqslant j+2$, $i \in S^m$, from (79)

$$
f_i(m) = \min\{\bar{\beta}_i v_i n_i^m(k), F_i - w_{i-1}(n_{i-1}^m(k) - n_{i-1}^{\mathrm{c}}, F_i\} = \bar{\beta}_i v_i n_i^m(k). \tag{83}
$$

Substituting (80), (82), (83) and (49) in

$$
n_i^m(k+1) = n_i^m(k) - \bar{\beta}_i^{-1} f_i(k) + f_{i+1}(k) + r_i = n_i^m(k) - \bar{\beta}_i^{-1}[f_i(k) - \phi_i] + [f_{i+1}(k) - \phi_{i+1}],
$$

gives the difference equation system for $\{n^m(k)\}$:

$$
\begin{aligned}
n_j^m(k+1) &= n_j^m(k) - \bar{\beta}_j^{-1}[\min\{\bar{\beta}_j v_j n_j^m(k), F_j\} - \bar{\beta}_j v_j n_j^{\mathrm{e},m}] + \bar{\beta}_{j+1} v_{j+1}[n_{j+1}^m(k) - n_{j+1}^{\mathrm{e},m}(k)] \\
n_i^m(k+1) &= n_i^m(k) - v_i[n_i^m(k) - n_i^{\mathrm{e},m}(k)] + \bar{\beta}_{i+1} v_{i+1}[n_{i+1}^m - n_{i+1}^{\mathrm{e},m}], \quad i = j+1, \ldots, I_{m+1} - 2 \\
n_{I_{m+1}-1}^m(k+1) &= n_{I_{m+1}-1}^m(k) - v_{I_{m+1}-1}(k)[n_{I_{m+1}-1}^m(k) - n_{I_{m+1}-1}^{\mathrm{e},m}(k)] - \eta(k).
\end{aligned}
$$

In terms of the variables $x_i^m(k) = n_i^m(k) - n_i^{\mathrm{e},m}, i \geqslant j, i \in S^m$, this system can be rewritten as

$$
x_j^m(k+1) = x_j^m(k) - \bar{\beta}_j^{-1}[\min\{\bar{\beta}_j v_j n_j^m(k), F_j\} - \bar{\beta}_j v_j n_j^{\mathrm{e},m}] + \bar{\beta}_{j+1} v_{j+1} x_{j+1}^m(k) \tag{84}
$$

and

$$
\begin{bmatrix} x_{j+1}(k+1) \\ x_{j+2}(k+1) \\ \cdot \\ x_{I_{m+1}-1}(k+1) \end{bmatrix} = \begin{bmatrix} 1 - v_{j+1} & \bar{\beta}_{j+2} v_{j+2} & 0 & \ldots & 0 \\ 0 & 1 - v_{j+2} & \bar{\beta}_{j+3} v_{j+3} & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & 0 & 1 - v_{I_{m+1}-1} \end{bmatrix} \begin{bmatrix} x_{j+1}(k) \\ x_{j+2}(k) \\ \cdot \\ x_{I_{m+1}-1}(k) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \cdot \\ -\eta(k) \end{bmatrix}. \tag{85}
$$

The difference Eq. (85) is of the form

$$
z(k+1) = Az(k) - b\eta(k),
$$

and has the solution

$$
z(k) = A^k z(0) - \sum_{l=0}^{k-1} A^{k-1-l} b\eta(l).
$$

The eigenvalues of A are $(1 - v_{j+1}), \ldots, (1 - v_{I_{m+1}-1})$, all of which lie in (0, 1), since $0 < v_i < 1$. By (77), $\eta(l) \to 0$, $\eta(l) \geqslant 0$. Hence $z(k) \to z^* \leqslant 0$, and (79) is assured by induction. Furthermore, because $\eta(l) \geqslant 0$,

$$
x_{j+1}^m(k) \leqslant M\lambda^k, \tag{86}
$$

for some $M < \infty$ and $0 < \lambda < 1$.Lastly, rewrite (84) as

$$x_j^m(k+1) = \begin{cases} (1-v_j)x_j^m + \bar{\beta}_{j+1}v_{j+1}x_{j+1}^m(k), & \text{if } x_j^m \leqslant -\Delta, \\ x_j^m(k) + \Delta + \bar{\beta}_{j+1}v_{j+1}x_{j+1}^m(k), & \text{if } x_j^m \geqslant -\Delta, \end{cases} \tag{87}$$

in which $\Delta = (1-2\psi)w_j^{-1}(F_{j+1} - \phi_{j+1})$.Because $\Delta > 0$ and $x_{j+1}^m \to 0$, the second alternative in (87) cannot hold for $k \geqslant K$, for some finite K, and so

$$x_j^m(k) = (1-v_j)^{k-K}x_j^m(K) + \sum_{l=K}^{k-1}(1-v_j)^{k-1-l}\bar{\beta}_{j+1}v_{j+1}x_{j+1}^m(l),$$

can be made arbitrarily small, proving (78).

Case (iii): $I_m < j < I_{m+1} - 1$. In this case at $n^{e,m}$ sections $I_m, \ldots, j$ are congested and sections $j+1, \ldots, I_{m+1} - 1$ are uncongested. The proof for this case combines the argument in Case (i) for the congested sections and the argument in Case (ii) for the uncongested sections. The details are omitted. □

References

Almasri, E., Friedrich, B., 2005. Online Offset Optimisation in Urban Networks Based on Cell Transmission Model. ITS Hanover.

Aw, A., Rascle, M., 2000. Resurrection of ‘second order’ models of traffic flow. SIAM Journal on Applied Mathematics 63 (1), 916–938.

Buisson, C., Lebacque, J.-P., Lesort, J.B., 1996a. STRADA, a discretized macroscopic model of vehicular traffic flow in complex networks based on the Godunov scheme. In: CESA’96 IMACS Multiconference.

Buisson, C., Lebacque, J.-P., Lesort, J.B., Mongeot, H., 1996b. The STRADA dynamic assignment model. In: Proceedings of the 1996 ITS Conference.

Chen, C., Skabardonis, A., Varaiya, P., 2004. Systematic identification of freeway bottlenecks. Transportation Research Record 1867, 46–52.

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

Daganzo, C., 1995. Requiem for second-order fluid approximations of traffic flow. Transportation Research Part B 29, 277–286.

Daganzo, C.F., Lin, W.-H., 1993. The spatial evolution of queues during the morning commute in a single corridor. Working Paper UCB-ITS-PWP-93-7, California PATH, University of California, Berkeley, CA 94720.

Godlewski, E., Raviart, P.-A., 1996. Numerical Approximation of Hyperbolic Systems of Conservation Laws. Springer-Verlag.

Godunov, S., 1959. A difference method for numerical calculation of discontinuous solutions of hydrodynamic equations. Matematicheskii Sbornik 47, 271–290.

Gomes, G., 2004. Optimization and microsimulation of on-ramp metering for congested freeways. Ph.D. Thesis, University of California, Department of Mechanical Engineering, Berkeley, CA.

Gomes, G., Horowitz, R., 2006. Optimal freeway ramp metering using the asymmetric cell transmission model. Transportation Research, Part C 14 (4), 244–262.

Haut, B., Bastin, G., 2005. A second order model for road traffic networks. In: Proceedings of the IEEE Intelligent Transportation Systems Conference, pp. 178–184.

Helbing, D., 1995. Improved fluid dynamic model for vehicular traffic. Physical Review E, 3164–3169.

Herty, M., Rascle, M., 2006. Coupling conditions for a class of second order models for traffic flow. SIAM Journal on Mathematical Analysis 38 (2), 595–616.

Hirsch, M.W., Smith, H., 2005. Monotone maps: a review. Journal of Difference Equations and Applications 11 (4–5), 379–398. Also <http://repositories.cdlib.org/postprints/977/>.

Lago, A., Daganzo, C.F., 2007. Spillovers, merging traffic and the morning commute. Transportation Research Part B 41, 670–683.

Lebacque, J.P., 1996. The Godunov scheme and what it means for first order traffic flow models. In: Lesort, J.B. (Ed.), Proceedings of the 13th International Symposium on Transportation and Traffic Theory. Pergamon, pp. 647–677.

Leclercq, L., Laval, J., Chevallier, E., 2005. The Lagrangian coordinates and what it means for first order traffic flow models. In: Allsop, B.R., Bell, M.G.H., Heydecker (Eds.), Proceedings of the 17th International Symposium on Transportation and Traffic Theory. Elsevier.

LeVeque, R.J., 1992. Numerical Methods for Conservation Laws. Birkhäuser.

Lighthill, M., Whitham, G., 1955. On kinematic waves I: Flow movement in long rivers. II: A theory of traffic flow on long crowded roads. Proceedings of the Royal Society of London, Part A 229 (1178), 281–345.

Lin, W.-H., Ahanotu, D., 1995. Validating the Basic Cell Transmission Model on a Single Freeway Link. Technical Note 95-03, California PATH, University of California, Berkeley, CA 94720.

Lo, H.K., 2001. A cell-based traffic control formulation: strategies and benefits of dynamic timing plans. Transportation Science 35 (2), 148–164.

May, A.D., 1981. Models for freeway corridor analysis. Technical report. Transportation Research Special Report 194.

Munoz, L., Sun, X., Sun, D., Gomes, G., Horowitz, R., 2004. Methodological calibration of the cell transmission model. In: Proceedings of American Control Conference, pp. 798–803.

Papageorgiou, M., Blosseville, J.-M., Hadj-Salem, H., 1990. Modelling and real-time control of traffic flow on the southern part of Boulevard Peripherique in Paris. I: Modelling. II: Coordinated on-ramp metering. Transportation Research, Part A 24A (5), 345–370.

Payne, H., 1971. Models of freeway traffic and control. Mathematical Models of Public Systems 28 (1), 51–61.

Payne, H., 1979. FREEFLO: A macroscopic simulation model of freeway traffic. Transportation Research Record 722, 68–77.

PeMS. PeMS website, 2007. <http://pems.eecs.berkeley.edu> (accessed 8/28/2007).

Piccoli, B., Garavello, M., 2006. Traffic Flow on Networks. American Institute of Mathematical Sciences.

Richards, P., 1956. Shock waves on the highway. Operations Research 4 (1), 42–51.

Whitham, G.B., 1974. Linear and Nonlinear Waves. Wiley-Interscience.

Yperman, I., Logghe, S., Immers, B., 2004. Dynamic congestion pricing in a network with queue spillover. <http://www.kuleuven.be/traffic/stats/>.

Zhang, M., Ritchie, S.G., Recker, W.W., 1996. Some general results on the optimal ramp control problem. Transportation Research, Part C 4 (2), 51–69.

Ziliaskopoulos, A.K., 2000. A linear programming model for the single destination system optimum dynamic traffic assignment problem. Transportation Science 34 (1), 37–49.

Please cite this article in press as: Gomes, G. et al., Behavior of the cell transmission model and effectiveness ..., Transport. Res. Part C (2007), doi:10.1016/j.trc.2007.10.005