UC Berkeley
Recent Work

Title

On the Variational Theory of Traffic Flow: Well-Posedness, Duality and Applications

Permalink

https://escholarship.org/uc/item/61v1r1qq

Author

Daganzo, Carlos F.

Publication Date

2006-06-01

**On the Variational Theory of Traffic Flow:
Well-Posedness, Duality and Applications
Carlos F. Daganzo**

WORKING PAPER
UCB-ITS-VWP-2006-2

UC Berkeley Center for Future Urban Transport
A **VOLVO** Center of Excellence

June 2006

ON THE VARIATIONAL THEORY OF TRAFFIC FLOW: WELL-POSEDNESS, DUALITY AND APPLICATIONS[1]

Carlos F. Daganzo, Institute of Transportation Studies, University of California, Berkeley, California, 94720, USA

ABSTRACT

This paper describes some simplifications allowed by the variational theory of traffic flow (VT). It presents general conditions guaranteeing that the solution of a VT problem with bottlenecks exists, is unique and makes physical sense; i.e., that the problem is well-posed. The requirements for well-posedness are mild and met by practical applications. They are consistent with narrower results available for kinematic wave or Hamilton-Jacobi theories. The paper also describes some duality ideas relevant to all these theories. Duality and VT are used to establish the equivalence of eight traffic models. Most of these are not new but VT-duality considerations offer a new insight into their relationship.

1. INTRODUCTION

Consider an infinite one-directional road on which vehicles cannot pass and move in the direction of increasing distance, x. If at some location $x = 0$ we assign consecutive integers to the vehicles we observe as time increases from $-\infty$ to $+\infty$, then the space-time trajectories of all the vehicles are completely defined as proposed in Moskowitz (1965), and further elaborated in Makigami et al (1971), by the integer contours of a surface. This surface is characterized by a continuous function $N(t, x) = n$. The floor $\lfloor n \rfloor$ is the number of the last vehicle to have advanced beyond x by time t. Since passing is not allowed the ordering of the vehicles is preserved everywhere. Therefore we can assume without loss of generality that

---

$N(t, x)$ is non-increasing in t for every x. Moreover, since vehicles move in the direction of increasing $x,$ we can also assume that N is non-decreasing in x for every $t.$

The simplest model of traffic flow further assumes that that N is differentiable almost everywhere (except possibly along some curves that would form ridges in the surface defined by N) and that the first partial derivatives of N are related by a function; i.e.:

$$\partial N / \partial t = Q(-\partial N / \partial x, t, x). \tag{1}$$

This is a Hamilton-Jacobi (HJ) equation with Q as the Hamiltonian. Note that $\partial N/\partial t$ (abbreviated q) is the flow and $-\partial N/\partial x$ (abbreviated k) is the density, and that meaningful solutions require flow and density to be non-negative.

The function Q is called the "fundamental diagram" (FD) by traffic engineers. We assume in rough agreement with experiments that Q is piecewise differentiable and concave in its first argument and returns non-negative values (for every t and x) if the first argument is in an interval $[0, \kappa(t, x)]$ such that $Q(0, t, x) = Q(\kappa, t, x) = 0$, with $\kappa(t, x) < \infty$. The parameter κ is called the "jam density." The maximum of Q, q_{max}, is called the "capacity"; see Fig. 1(a).
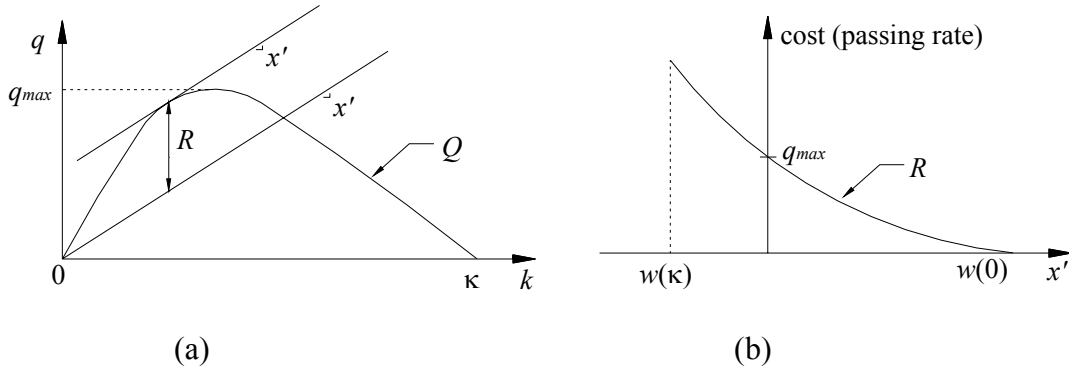

Figure 1. Basic concepts of Variational Theory: (a) fundamental diagram; and (b) cost function.

Note that if (1) is differentiated with respect to x and expressed in terms of density it reduces to the conservation law: $\partial k/\partial t + (\partial Q/\partial k)(\partial k/\partial x) = -\partial Q/\partial x$. This is the classical kinematic wave (KW) formulation of Lighthill and Whitham (1955), and Richards (1956). A simplified solution of some KW problems in terms of the Moskowitz function was later proposed in Newell's seminal trilogy (Newell, 1993). Newell's results have been recently formalized and extended in Daganzo (2003, 2003a and 2005), which proposes a variational theory (VT) able to capture bottlenecks of all types.

1.1 Variational theory

Variational theory also assumes that (1) holds where N is differentiable and that Q is concave, but treats the problem as a capacity-constrained optimization problem. An intuitive explanation is as follows. We know that the flow at any point in space-time is bounded from above by $q_{max.}$, the capacity. A similar capacity constraint should also hold if the road is viewed from a rigid frame of reference that moves with speed x'. In this case the capacity relative to the frame (the "relative capacity") is the maximum rate at which traffic can pass an observer attached to the frame. Since an observer that moves with speed x' next to a traffic stream with density k and flow q is passed by traffic at a rate $q - kx'$, the FD for the moving frame is $Q(k, t, x) - k x'$ and its relative capacity is:

$$R(x',t,x) = \sup_{k}\{Q(k,t,x) - kx'\}. \tag{2}$$

Figure 1(a) shows these relations geometrically; note that the relative capacity R is the intercept of the tangent to curve Q with slope x'. Figure 1(b) shows the relative capacity function (also called the "cost function" in variational theory) with x' as the argument. Note that $-R$ is the Legendre-Fenchel transform of Q, and that as a result, as shown by Fig. 1(b), R is convex and strictly decreasing in the range of "valid" slopes where Q is non-negative; i.e., for $x' \equiv w(k,t,x) \equiv \partial Q / \partial k \in [w(\kappa, t, x), w(0, t, x)]$. Note as well that $R \geq 0$ and $R(u) = 0$, since curve Q touches the origin. Thus, in traffic flow the Legendre-Fenchel transform has an intuitive physical interpretation, which makes its application fairly intuitive as we shall now see.

Clearly, an observer travelling with a valid speed $x'(t) \in [w(\kappa), w(0)]$ along a "valid" space-time path $\mathscr{P}$ from point D to point P cannot see a change in vehicle number greater than the integral with respect to time of its relativistic capacities; i.e., an upper bound to change is:

$$\Delta(\mathscr{P}) = \int_{t_D}^{t_P} R(x',t,x)dt, \tag{3}$$

where t_D and t_P are the times associated with the path endpoints. Therefore, an upper bound to the vehicle number N_P observed at a point P can be written by considering the set $\boldsymbol{P}$ of all valid observer paths to P from the points of a boundary $\boldsymbol{D}$ where the vehicle numbers are known. In other words, if $D(\mathscr{P}) \in \boldsymbol{D}$ is the beginning of a valid path $\mathscr{P}$, and $N_{D(\mathscr{P})}$ is the known vehicle number at $D(\mathscr{P})$, then it must be true that N_P, must satisfy:

$$N_P \leq \inf\{N_{D(\mathscr{P})} + \Delta(\mathscr{P}) : \forall \mathscr{P} \in \boldsymbol{P}\}. \tag{4}$$

Equation (4) is the capacity constraint mentioned at the outset.

In variational theory the solution domain $\boldsymbol{S}$ is the set of points P such that all infinitely long valid paths ending at P intersect the boundary. For example, the solution domain for the initial value problem (IVP) is the half plane, $t \geq 0$. Variational theory assumes that capacity constraint (4) is binding; i.e., that the actual value of N_P for P $\in \boldsymbol{S}$ is the largest possible allowed by (4):

$$N_P = \inf \{N_{D(\mathcal{P})} + \Delta(\mathcal{P}) : \forall \, \mathcal{P} \in \boldsymbol{P} \,\}. \tag{5}$$

This is a calculus of variations problem. It is well known that under some regularity conditions (5) characterizes both the viscosity solution of the HJ-IVP and also the entropy solution of the KW-IVP.[2] A key advantage of VT over the HJ and KW theories is its natural framework for expressing the relatively complicated problems arising in traffic flow applications (including bottlenecks and finite roads), and the convenient way in which the "well-posedness" of such problems can be assessed; see below.

2. SIMPLIFICATIONS

2.1 Homogeneous problems with point bottlenecks: solution existence and uniqueness.

In traffic flow theory it is often necessary to consider "point bottlenecks". These are usually slower vehicles or fixed obstructions that reduce the maximum rate at which traffic can flow past them. A point bottleneck is defined by its space-time trajectory $x_B(t)$, assumed to be a valid path, and by its relative capacity (maximum passing rate) $r_B(t)$. The bottleneck imposes the condition $dN(t, x_B(t))/dt \leq r_B(t)$ in HJ theory. This type of constraint seems not to have received much attention in the mathematics literature. The constraint is even more complicated when expressed in terms of KW theory. But the complication disappears in VT.

In VT a bottleneck reduces the original relative capacity of the road along the bottleneck trajectory. This is recognized by using $r_B(t)$ instead of R as the integrand in (3) for the portion of any path that overlaps $x_B(t)$.[3] Nothing else needs to be changed: (4) and (5) continue to apply. Hence, in VT, point bottlenecks are just shortcuts through space-time, which preserve the shortest-path character of the problem without increasing its complexity. The solution should be equally easy to find. The question is whether the solution with bottlenecks is continuous and varies with t and with x at allowable rates.

We look for solutions that satisfy the following Lipschitz-continuity conditions:

$$(N(t, x_1) - N(t, x_2)) / (x_2 - x_1) \in [0, \kappa] \qquad \text{if } x_1 < x_2 \quad \text{and} \tag{6a}$$

$$(N(t_2, x) - N(t_1, x)) / (t_2 - t_1) \in [0, q_{max}] \qquad \text{if } t_1 < t_2. \tag{6b}$$

A solution satisfying (6) is obviously continuous; thus, vehicles have continuous trajectories. Furthermore, if (6) holds, vehicles can neither reverse direction nor overtake an object moving with speed u; i.e. their average speed is always bounded in a physically meaningful

---

[2] If Q is not concave (4) continues to be true but (5) may not match the HJ and the KW solutions; other constraints come into play.

[3] For a bottleneck to have an effect its relative capacity should be less than the original; i.e., $r_B(t) < R(dx_B/dt, t, x)$.

way.[4] Therefore, solutions satisfying (6) will be called "valid". A VT problem whose solution is valid will be said to be "well-posed".

We examine below whether these conditions are satisfied for "homogeneous" problems in which Q and R are time-independent and space-independent. Therefore, they will be expressed from now on as functions of one argument, $Q(k)$ and $R(x')$; the parameters κ, q_{max}, etc. will become constants. It will be useful to keep in mind that for homogeneous problem without bottlenecks straight lines turn out to be optimum paths and the RHS of (3) reduces to $R(v_{DP})(t_P - t_D)$, where v_{DP} is the slope of segment DP: $v_{DP} = (x_P - x_D)/(t_P - t_D)$; see Daganzo (2005). In this special case, thus, the calculus of variations problem (5) reduces to an ordinary minimization for the point on the boundary ($D \in \boldsymbol{D}$) that produces the minimum cost. We can now state the following.

THEOREM 1: *A VT-IVP with any number of piecewise differentiable bottlenecks "B" is well-posed if the initial data satisfy* (6a) *and the bottlenecks satisfy*: $r_B(t) \geq 0$ *and* $dx_B(t)/dt \geq 0$.

Proof: see Appendix A (Lemmas 5 and 6).

Lemmas 5a and 6 of Appendix A (summarized as Theorem 2, below) prove a similar result for the finite highway problem (FHP) with bottlenecks. The highway extends from $x = 0$ to $x = x^o > 0$. Given are the vehicle numbers along the boundary: $n(0, x) = N(0, x)$ for $0 \leq x \leq x^o$, and $n(t, 0)$ and $n(t, x^o)$ for $t \geq 0$. We look for $N(t, x)$ for $t > 0$ and $x \geq 0$. The solution domain is $\boldsymbol{S} = \{(t, x): t \geq 0 \text{ and } 0 \leq x \leq x^o\}$. Because the boundary at $x = 0$ and $x = x^o$ can be reached by valid paths from the boundary, we add the necessary consistency condition stipulated in Daganzo (2005) for problems with complex boundaries: namely, that the least cost of reaching a boundary point with a path from the boundary be no less than the cost specified for that point; i.e., that $N(t, 0) = n(t, 0)$ and $N(t, x^o) = n(t, x^o)$ for $t \geq 0$. In our case, this means that the boundary lines $x = 0$ and $x = x^o$ must be optimum paths.

THEOREM 2: *A FHP with bottlenecks is well posed if*: (*i*) *the boundary data n(t, x) satisfies* (6); (*ii*) *the bottlenecks satisfy* $r_B(t) \geq 0$, $dx_B(t)/dt \geq 0$ *and* $x_B(t) \geq 0$; *and* (*iii*) the *consistency condition is satisfied*: $N(t, 0) = n(t, 0)$ and $N(t, x^o) = n(t, x^o)$ for $t \geq 0$.

In applications, Theorem 2 can be interpreted in terms of a competition between "upstream demand" and "downstream capacity". Let $U(t)$ be a real function satisfying (6b), giving the number of cumulative desired entrances at $x = 0$; and let $N^D(t, 0)$ be the available capacity at $x = 0$ from conditions downstream. We define $N^D(t, 0)$ as the infimum of the costs of reaching point $(t, 0)$ from the boundary with valid paths in the solution domain starting at an earlier time; i.e., from the set of points $\boldsymbol{D}_t = \{(t', x): (t', x) \in \boldsymbol{D} \text{ and } 0 \leq t' < t\}$, with end point $E(\mathscr{P}) = (t, 0)$. Thus, $N^D(t, 0) = \inf\{N_{D(\mathscr{P})} + \Delta(\mathscr{P}); \forall \mathscr{P} \in \boldsymbol{D}_t \text{ and } E(\mathscr{P}) = (t, 0)\}$. Likewise, let $C(t)$ be a real function satisfying (6b), giving a bound on the cumulative exits at x^o; and define the

---

[4] The proof of this statement is so simple we only give the logic: (i) direction reversals cannot occur because they require N to be constant for increasing t and decreasing x, but this is obviously incompatible with (6); (ii) vehicles cannot overtake a valid path with u everywhere because this would imply that N increases along such path, but this is incompatible with the requirement $R(u) = 0$.

upstream demand at $x = x^o$, $N^U(t, x^o)$, as the infimum of the costs of reaching point (t, x^o) from the boundary with valid paths in the solution domain starting from $\boldsymbol{D}_t$; i.e., $N^U(t, x^o) = \inf\{N_{D(\mathscr{P})} + \Delta(\mathscr{P}); \forall \mathscr{P} \in \boldsymbol{D}_t \text{ and } E(\mathscr{P}) = (t, x^o)\}$. Then, we have the following corollary:

COROLLARY 1: *The FHP with bottlenecks is well posed if*: (*i*) *the initial data* $n(0, x)$ *satisfies* (6a); (*ii*) *the bottlenecks satisfy* $r_B(t) \geq 0$, $dx_B(t)/dt \geq 0$ *and* $x_B(t) \geq 0$; *and* (*iii*) *the upstream and downstream data are* $n(t, 0) = \min\{U(t), N^D(t, 0)\}$ *and* $n(t, x^o) = \min\{C(t), N^U(t, x^o)\}$.

The corollary is true because it satisfies conditions (i), (ii) and (iii) of Theorem 2. It satisfies (i) and (iii) because according to (5): $N(t, 0) = \min\{n(t, 0), N^D(t, 0)\} = n(t, 0)$, and $N(t, x^o) = \min\{n(t, x^o), N^U(t, x^o)\} = n(t, x^o)$; and both results obviously satisfy (6b). The functions $U(t)$ and $C(t)$ can be chosen in any way consistent with (6b). For example, if there is a highway with bottlenecks and different Q and R for $x < 0$, we can define $U(t)$ for the downstream problem as the demand at $x = 0$ arising from the upstream problem: $U(t) \equiv N^U(t, 0)$; and choose $C(t)$ for the upstream problem as the available capacity arising from the downstream problem at $x = 0$: $N^D(t, 0)$. To stitch together the two solutions we stipulate $n(t, 0) = \min\{N^U(t, 0), N^D(t, 0)\}$ for both problems. This ensures that both problems are well-posed, and is a natural way to treat inhomogeneous highways.

The Theorems and Corollary are consistent with results available for KW (or LWR) theory for the case without bottlenecks. The results also generalize the demand vs. capacity metaphor of the cell-transmission model (Daganzo, 1993, 1994) and the related formulations in Daganzo (1993a), Lebacque (1993).[5]

The results can also be applied to time-dependent problems. Well-posedness can be checked in this case by slicing the solution space into successive time-independent problems and verifying that each time-independent slice satisfies the conditions of one of the above theorems. Unfortunately, well-posedness cannot always be tested *a priori* (before solving the problem) as in the time-independent case because for the initial data of a slice to be valid (and the theorems to hold) the solution obtained at the end of the previous slice must satisfy (6a) with the jam density κ specified *a priori* for the current slice.

2.2 Linear cost functions

In this subsection Q is triangular in k. Now, the problem simplifies even more because the cost function (3) is linear (Daganzo, 2003a). If we use $u \equiv w(0)$ and $-w \equiv w(\kappa)$ for be the slopes of the rising and dropping branches of Q (in traffic flow lingo u is the "free-flow speed" and w the "backward wave speed"), then (2) becomes:

---

[5] Well-posedness with point bottlenecks has not been systematically studied in KW theory. Thus, we propose using the conditions of Theorems 1 and 2 and Corollary 1 to verify the well-posedness of KW problems with concave Q. This is reasonable because a valid VT solution satisfies the conservation law derived from (1) where the VT solution is differentiable. And, since a valid VT solution is "stable" with respect to the l_∞ norm, one would also expect it to be an entropy solution of the conservation law..

$$R(x') = (1 - x'/u)q_{\max}, \qquad\qquad \text{for } x' \in [-w, u]. \qquad\qquad (7)$$

Note that $R(x')$ decreases. We shall abbreviate its maximum: $R(-w) = (1+w/u)q_{\max}$, by the symbol r. This parameter (the maximum relative capacity) will be of importance later. Experiments show that r is about 15% greater than q_{max}.

This case is so simple because when R is linear the cost of a path (4) is a linear function of the path's duration and distance; i.e., if $\mathcal{P}$ goes from D to P:

$$\Delta(\mathcal{P}) = q_{\max}(t_P - t_D) - (q_{\max}/u)(x_P - x_D). \qquad\qquad (8)$$

Hence, if as is often the case in traffic flow the boundary data (i.e., the coordinates t_D and x_D and the values N_D for all points D$\in$**D**) are given as piecewise linear functions of a parameter, $t_D = t(\alpha)$, $x_D = x(\alpha)$) and $N_D = n(\alpha)$, then (5) becomes:

$$N_P = \inf{}_\alpha \{q_{\max}[t_P - t(\alpha)] - (q_{\max}/u)[x_P - x(\alpha)] + n(\alpha)\}, \qquad\qquad (9)$$

which is just the minimization of a piecewise linear function. Obviously, we can find its minimum by inspecting the corners of the objective function.

The solution can also be found with network algorithms; see e.g., Daganzo (2003a). These methods are advantageous when the solution is sought at many points in the solution domain. The networks in question are digraphs with nodes L embedded in space-time, with directed arcs LL'. Arcs are defined only for node pairs that can be connected by a valid path. We call these "valid node pairs." Each arc is assigned a cost, $c_{LL'}$, equal to that of an optimum continuum path between its end nodes; e.g., as given by (8) when Q is triangular. Of interest are networks whose shortest "walks" (network paths) between all valid node pairs are shortest continuum paths. These networks are said to be "sufficient" because by solving the shortest path problem on the network one solves the continuum problem exactly for all its valid node pairs. This is useful because if one puts nodes of a sufficient network on the corners of a piecewise linear boundary, then the network solution identifies the exact N at every node. The solution can be found with the usual dynamic programming recursion:

$$c_{L'} = \min_{L \in F(L')} \{c_L + c_{LL'}\}, \qquad\qquad (10)$$

where $F(L')$ is the set of "from" nodes of L'.

For problems with linear R sparse sufficient networks with as few as 2 links per node can be constructed; thus (10) can be computed fast. The rest of this paper does not consider bottlenecks and uses sufficient networks of the "lopsided" type defined in Daganzo (2003a).[6] A lopsided network (see Fig. 2) is a network with the following properties: (i) its nodes are on a rectangular lattice with space separation δ and time step ε, (ii) the set of links pointing to

---

[6] Lopsided neworks can only be used if there are no bottlenecks; otherwise they need to be modified. This is done by overlying a discrete shortcut with appropriately reduced link costs over the network; see Menendez and Daganzo (2005).

any node is translationally symmetric, (iii) two of the links in this set have slopes u and $-w$, and (iv) none of the links spans a distance greater than δ. Note: since the nodes are on a rectangular lattice, δ/u and $\delta/(-w)$ must be integer multiples of ε, assuming $u, w \neq 0$. These networks will help us compare different ways of finding N. But before this is done we introduce some duality ideas, which will allow us to double the number of models covered under the same umbrella.
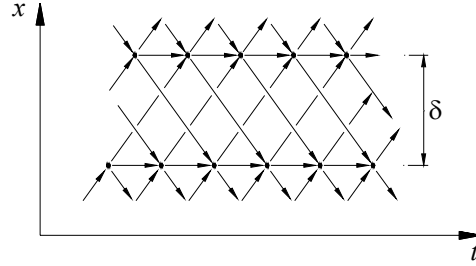

Figure 2: A lopsided network.

3. DUALITY

In this section $Q(k)$ is general—not necessarily triangular. The results apply to problems where $N(t, x)$ strictly decreases with x for every t in the relevant solution domain. Lemma 5b of Appendix A shows that an IVP with strictly decreasing $n(0, x)$ satisfies this condition if it has no bottlenecks with zero relative capacity; and also if there are bottlenecks with zero relative capacity but the solution is only sought upstream of them.[7]

Since N is continuous and declines with x, the relation $N(t, x) = n$ defines an implicit function for x in terms of t and n, $x = X(t, n)$. This function gives the position of vehicle n at time t. It is also continuous and declines with n. Both functions describe the same Moskowitz surface. The two functions are connected by the relation:

$$X(t, N(t, x)) = x, \tag{11a}$$

which merely expresses that the position at time t of the vehicle that was at x at time t must be x. Conversely, we can also write:

$$N(t, X(t, n)) = n, \tag{11b}$$

since the vehicle number found at time t at the position of vehicle n at time t is n. Note that (11b) is obtained from (11a), and vice versa, by interchanging (x, X) with (n, N). Since the (primal) results of the previous section were derived with N as the unknown, this suggests that

---

[7] Note that the initial vehicle positions of an IVP can always be described by a strictly decreasing $N(0, x)$, if the road contains at least one vehicle. Therefore, no generality is lost by assuming that $N(0, x)$ is monotonic.

similar (dual) results can be derived with X as the unknown after swapping the variables and functions for position and vehicle number.

Differentiation of (11a) with respect to t and x yields the following relation among the partial derivatives of the primal and dual functions:

$$\partial N/\partial x = 1/\partial X/\partial n < 0 \tag{12a}$$

and

$$\partial N/\partial t = -\partial X/\partial t/\partial X/\partial n \geq 0. \tag{12b}$$

The same expressions with (x, X) and (n, N) interchanged are obtained if one differentiates (11b). The quantity $v = \partial X/\partial t$ is the vehicle speed, and the quantity $s = -\partial X/\partial n$ the reciprocal of density; i.e., the continuum version of vehicular spacing. If we now insert (12) into (1) we find:

$$\partial X/\partial t = V(-\partial X/\partial n), \tag{13}$$

where V is related to Q by the following transformation:[8]

$$V(s) = Q(1/s)/s, \quad \text{where } s = -\partial X/\partial n. \tag{14}$$

Equation (13), like (1), is an HJ equation. Since we have not reversed the direction of time the solution of (13), which is obtained by transforming with (11) the stable (viscosity) solution of (1), is also a stable solution. Thus, for any given set of boundary conditions $[t_D = t(\alpha), x_D = x(\alpha))$ and $N_D = n(\alpha)]$ the stable solutions of (13) and (1) describe the same Moskowitz surface. Since V, like Q, is concave in the relevant range of its argument, $s \in [1/\kappa, \infty)$, X, like N, can be found with VT. Thus, the Moskowitz surface can be found by solving with the same methods either a primal problem (1) or a dual problem (13).[9]

The dual cost function R^d is given by (2) with V substituted for Q. We find that R^d is the inverse of R, with the roles of speed x' and passing rate n' reversed, and that it still is convex and decreasing in the relevant range of passing rates. In the triangular case the dual cost function is the inverse of (7), and is still linear:

$$R^d(n') = (1 - n'/q_{max})\, u, \qquad\qquad \text{for } n' \in [0, r]. \tag{15}$$

---

[8] Note that the transformation $Q \rightarrow V$ is an involution, which should not be surprising since the swap of x and n is a reflection.

[9] Note too that one can define by differentiating (13) with respect to n a conservation law, $\partial s/\partial t + (\partial V(s)/\partial s)\, \partial s/\partial n = 0$, which is the dual of $\partial k/\partial t + (\partial Q(k)/\partial k)\partial k/\partial x = 0$. The analyses and methods relevant to the primal conservation law also apply to the dual.

Therefore, the sufficient lopsided networks that one could use with (10) now have slopes equal to the cost rates of the primal (0, and r) and cost rates equal to the slopes of the primal (u and $-w$, respectively). Variational theory in its primal and dual forms is used in the next section to examine the connection between eight different traffic models.

4. APPLICATION: EIGHT MODELS OF A TRAFFIC STREAM

In this section the highway is homogeneous and the FD is triangular. We classify traffic models into 4 categories distinguished by the number of variables that are treated discretely: 0-*models* treat all variables continuously, as in the discussion up to this point; 1-*models* treat one independent variable (x or n) discretely; 2-*models* treat both independent variables (t and x; or t and n) discretely; and 3-*models* treat all variables (t, x and n) discretely. Here we shall present primal and dual VT models for each category (eight models in total) and see how they relate to existing ones.

0-models: These are fluid models. Our primal 0-model is (3, 5). We have already seen that it has the following dual 0-model, where $\mathscr{P}$ is a dual path $n(t)$ from (t_D, n_D) to (t_P, n_P):

$$X_P = \sup \{\Delta(\mathscr{P}) + X_{D(\mathscr{P})}\}, \text{ where } \Delta(\mathscr{P}) = \int_{t_D}^{t_P} R^d(n')dt = u(t_P - t_D) - (u/q_{max})(n_P - n_D). \quad (16)$$

The last equality follows from (15). Note the similarity of $\Delta(\mathscr{P})$ in (16) and (8).

1-models: These are queuing and car-following models. An example of a primal 1-model is Newell's queuing formula (Newell, 1993) which gives the cumulative curve at some point of a highway $N(t, x_M)$ from the vehicle number curves observed at its upstream and downstream ends: $N_U(t)$ and $N_D(t)$. The formula is:

$$N(t, x_M) = \min\{N_U(t - (x_M - x_U)/u), N_D(t - (x_D - x_M)/w) + (x_D - x_M)r\}. \quad (17)$$

The reader can verify that (17) is the result of applying (9) to our boundary data.[10]

We now apply (16) to a "lead vehicle problem". This is a dual problem with boundary conditions: $X(t, 0) = x_0(t)$ for $t \geq 0$ and $X(0, n) = x_n(0)$ for $n \geq 0$. Assume the $x_n(0)$ is linear in n (vehicles are uniformly spaced) and that $dx_0(t)/dt \leq u$. Then, an optimum path to reach point (t, n) for some integer n must begin at one of the two extreme points of the relevant part of the boundary for point (t, n): either point $(0, n)$ or point $(t - n/r, 0)$. The result is:

$$X(t, n) = \min\{x_n(0) + ut, x_0(t - n/r) - ns_j\},$$

---

[10] Since N_U and N_D cannot increase at a rate that exceeds q_{max}, an optimum path to point "P" must emanate from a point on the (upstream or downstream) boundary with the highest possible t. Only two such points generate valid paths. They correspond to the two arguments of (13).

which is the trajectory of vehicle *n*. The parameter $1/r$ (comparable with a second) has the interpretation of a reaction time and we denote it by τ. In practice we are usually interested in the 1-model that seeks the values of *X* for all integer *n*. A recursive expression is obtained by setting $n = 1$ in the above and applying the same recipe to all consecutive vehicle pairs. The result is the following car-following law (Newell, 2002):

$$X(t, n) = \min\{ x_n(0) + ut, x_{n-1}(t-\tau) - s_j \}. \tag{18}$$

2-models: These are numerical versions of fluid models. For the primal we use (10) with a lopsided network with two links per node. We choose $\delta = u\varepsilon$. Therefore, the links with slope *u* (and zero cost) span one time step. The links with slope $-w$ (and cost rate *r*) span $\theta = u/w$ time steps. Hence their cost is $c = r\theta\varepsilon = r\delta/w$. Note: since θ must be integer we are assuming that u/w is an integer—this ratio is comparable with 6 in practice. If we now use sub-indices *l* and *m* to identify the time and distance steps, i.e., so that $N_{lm} \equiv N(l\varepsilon, m\delta)$, (10) becomes:

$$N_{lm} = \min\{N_{l-1,m-1} , N_{l-\theta, m+1} + c \}. \tag{19}$$

Equation (19) expresses the ACT (asynchronous cell transmission) model for cells of size δ; see Appendix B.[11]
A dual 2-model is obtained by applying (10) to a lopsided network on the (t, n) plane as described above with arc slopes $(0, r)$ and arc cost rates $(u, -w)$. We choose the step for variable *n* to be 1 and the time step, $\varepsilon = 1/r = \tau$. This achieves a rectangular lattice, since $\varepsilon r = 1$. The link costs become as a result: $u\tau$ and $-w\tau = -s_j$. Therefore, with the convention: $X_{lm} \equiv X(l\tau, m)$, recursion (10) reduces to:

$$X_{lm} = \min\{X_{l-1,m} + u\tau, X_{l-1,m-1} - s_j \}. \tag{20}$$

This is the CF(L) model (Daganzo, 2006), which merely expresses (18) on a lattice.

3-models: Examples of 3-models are cellular automata (CA) models, where cars are assumed to jump on a lattice. Most CA models are described in dual space, but as we now show primal models can also be derived. Simply, use $\delta = s_j \equiv w/r$ in (19), which yields $c = 1$, and therefore:

$$N_{lm} = \min\{N_{l-1,m-1} , N_{l-\theta, m+1} + 1\}. \tag{21}$$

This expression returns an integer if the input vehicle numbers are integer. Therefore it is a CA model. The expression indicates that the vehicle count at a point in space increases by one if and only if vehicle number at the downstream lattice point had reached the target number θ time steps ago; i.e., if the previous vehicle had jumped from *m* and left it vacant for at least θ time steps. This is the CA(M) rule described in Daganzo (2006).
Consider now the dual formula (20) and express it in dimensionless distance, $Z = X/s_j$. It becomes:

---

[11] The middle term of (B3) turns out to be redundant for the homogeneous highway problem. But if we had used a lopsided network with one horizontal link of cost $q_{max}\varepsilon$, (21) would have included the middle term of (B3).

$$Z_{lm} = \min\{Z_{l-1,m} + u\tau/s_j, \ Z_{l-1,m-1} -1\} = \min\{Z_{l-1,m} + \theta, \ Z_{l-1,m-1} -1 \} \quad\quad (22)$$

We see that if vehicles are initially on the lattice (the Z's are integer) and if θ is an integer, then (22) keeps vehicles on the lattice; i.e., it is a CA model. Equation (22) is the unbounded acceleration model of Nagel and Schreckenberg (1992), called the CA(L) model in Daganzo (2006).

This concludes our review. Duality and variational theory provided a framework that clearly established the equivalence of our models. The best model for any given application depends on the form of the data and the requirements of the output.

5. COMPOSITION INTO NETWORKS AND DISCUSSION

Primal analysis looks for the flow of vehicles from the perspective of the road; and dual analysis the "flow of road" from the perspective of the vehicles.[12] Fixed bottlenecks such as merges and lane-drops are understood by scientists in primal space, from the perspective of the road, since this is the form in which data are available. Moving bottlenecks; e.g., those caused by slow-moving obstructions are understood from the perspective of the moving bottleneck, since data from this perspective is available. The moving-bottleneck effects of lane-changing are most easily expressed in dual space; those of fixed bottlenecks in primal space. The ideas in this paper allow us to combine the effects of fixed and moving bottlenecks, including lane-changing, consistently in whatever framework is most useful (primal or dual) for a practical application.

Since lane changes to a faster stream act as moving bottlenecks on the target lane, and lane changes to a slower stream act as moving bottlenecks on the source lane, the ability to treat moving bottlenecks allows us to compose the very basic component described in this paper—a single lane of traffic—into complex multi-lane streams quite realistically; this approach was explained, proposed and tested with encouraging results in Laval and Daganzo (2005). It has proven to be parsimonious and surprisingly accurate for lane drops, moving bottlenecks and merges; see also Laval et al, (2005). A variant of it has also been applied to HOV lanes, with considerable success (Menendez and Daganzo, 2006).

Composition of links into networks is possible along traditional lines, e.g., as in the CTM (Daganzo, 2004). But the ideas of this paper allow us to treat turning movements as lane-changes, and junctions as complex multi-lane links. Therefore, they allow us to compose multi-lane links into networks in quite a bit of detail without introducing extra parameters.

---

[12] A possible interpretation of dual VT and its constraints is as follows. Imagine uniformly spaced parked (dual) vehicles by the side of the road. Then, dual VT describes the flow of these vehicles from the perspective of a flexible frame of reference attached to the moving (primal) vehicles; i.e., where (dual) distance increases by a unit with each (primal) moving vehicle. From this frame of reference, the (dual) flow is the rate at which dual vehicles (i.e., units of primal distance) flow past fixed positions in the dual frame (i.e., moving-primal vehicles). Thus, dual flow = primal speed. Conversely, the rate at which a dual vehicle overcomes dual distance (i.e., moving vehicles) is both the dual speed and the primal flow. And the number of parked vehicles between two consecutive moving vehicles is both the dual density and, the primal spacing. Thus, dual-VT can also be interpreted in terms of flows and densities, and its constraints described in terms of relative capacities, but all from the perspective of the flexible frame of reference. Thus, the dual relative capacity is the maximum flow of parked vehicles that can be seen by an observer jumping from primal vehicle to primal vehicle with a fixed jump frequency.

The composition rules, however, require additional data; most notably, the destinations of the vehicles making up the stream. This make-up strongly affects the discharge rates of diverge bottlenecks (see Munoz and Daganzo, 2002) and the performance of intersections controlled by traffic signals. Unfortunately, as a network grows in size, the number of possible destinations grows and the availability of the required input data diminishes. Thus, the practical limit to composition is not theoretical (we could model relatively well almost anything if we knew where vehicles were going) but informational.

We believe that the results in this paper can be of use for the design of small networks such as complex interchanges, but other approaches should be sought for very large networks. See Daganzo (2006a) for some ideas in this direction

REFERENCES

Daganzo, C.F. (1993), "The cell-transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory", Institute of Transportation Studies, Research Report, UCB-ITS-RR-93-7, Univ. of California, Berkeley, CA; abridged in *Transportation Research Part B* **28** (4), 269-287.

Daganzo, C.F. (1993a), "A finite difference approximation for the kinematic wave model", Institute of Transportation Studies, Research Report, UCB-ITS-RR-93-11, Univ. of California, Berkeley, CA; abridged in *Transportation Research Part B* **29** (4) 261-276.

Daganzo, C.F. (1994), "The cell-transmission model. Part II: Network traffic", PATH working paper UCB-ITS-PWP-94-12, Univ. of California, Berkeley, CA; ; abridged in *Transportation Research Part B* **29**(2), 79-94.

Daganzo, C. F. (2003). A Variational Formulation for a Class of First Order PDEs. Institute of Transportation Studies Report UCB-ITS-RR-2003-03, U. of California, Berkeley, CA.

Daganzo, C. F. (2003a). A Variational Formulation of Kinematic Waves: Solution Methods. Institute of Transportation Studies Report UCB-ITS-RR-2003-07, U. of California, Berkeley, CA. *Transportation Research Part B* **39**(2) 934-950.

Daganzo, C.F. (2003b) *A Theory of Supply Chain.* Springer, Heidelberg, Germany.

Daganzo, C. F. (2005). A Variational Formulation of Kinematic Waves: Basic Theory and Complex Boundary Conditions. *Transportation Research Part B* **39**(2), 187-196.

Daganzo, C.F. (2006) In traffic flow cellular automata = kinematic waves *Transportation Research Part B* **40**(5) 396-403.

Daganzo, C.F. (2006a) The nature or urban gridlock U.C. Berkeley Centre for Future Urban Transport Working paper VWP-2006-2, *Transportation Research Part B*, in press.

Laval, J. A. and C. F. Daganzo (2005) Lane-changing in traffic streams. *Transportation Research Part B* **40**(3) 251-264.

Laval, J.A., M.J. Cassidy and C.F. Daganzo (2005) "Impacts of lane changes at on-ramp bottlenecks: a theory and strategies to maximize capacity" Presented at Traffic and Granular Flow05, Berlin, Germany, October 10-12, 2005.

Lebacque, J. P. (1993), "Les modèles macroscopiques de trafic," Annales de Ponts, (3rd trim) **67**, 28-45.

14

Lighthill, M. J. and G. B. Whitham (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 A* **229**, 281-345.

Makigami, Y., G.F. Newell, and R. Rothery (1971), "Three-dimensional representation of traffic flow", *Transportation Science* **5**, 302-313.

Menendez, M. and Daganzo, C.F. (2005) A variational theory of kinematic waves: bottleneck properties and examples in Traffic and Transportation Theory (H. Mahmassani, editor), Proc. ISTTT, College Park, Maryland.

Menendez, M. and Daganzo, C.F. (2006) High occupancy lanes: their effect on traffic flow, draft.

Moskowitz, K. (1965) Discussion of 'freeway level of service as influenced by volume and capacity characteristics' by D.R. Drew and C. J. Keese, *Highway Research Record* **99**, 43-44.

Muñoz, J. C. and C. F. Daganzo (2002) The bottleneck mechanism of a freeway diverge *Transportation Research Part A* **36**, 483-505

Nagel, K and M. Schreckenberg (1992) A cellular automaton model for freeway traffic. *Journal de Physique I* **2**, 2221-2229

Newell, G. F. (1993a) A Simplified Theory of Kinematic Waves in Highway Traffic: (i) General Theory; (ii) Queuing at Freeway Bottlenecks; (iii) Multi-Dimensional Flows. *Transportation Research Part B* **27**, 281-313.

Newell, G. F. (2002) A simplified car-following theory: A lower order model. *Transportation Research Part B* **36**, 195-205.

Richards, P. I. (1956) Shockwaves on the highway *Operations Research* **4**, 42-51.

APPENDIX A: PROOFS

<u>Definition 1.</u> Valid path: A continuous piecewise differentiable function, $x(t)$, such that:

$$(x(t_2) - x(t_1)) / (t_2 - t_1) \in [-w, u] \qquad \text{if} \quad t_1 < t_2 .$$

<u>Definition 2.</u> Cost function with bottlenecks:

$$R_B(x', t, x) = \min\{r_B(t), R(x')\} \quad \text{if} \quad (t, x) \in S, x = x_B(t) \text{ and } x' = x'_B(t) \text{ for some } B.$$
$$= R(x'), \quad \text{otherwise.}$$

<u>Definition 2.</u> Auxiliary cost function:

$$R_A(x') = \max\{0, -x'r/w\} = \max\{0, -x'\kappa\}.$$

<u>Definition 3.</u> Auxiliary path costs, $\Delta_A(\mathscr{P})$, are costs obtained with the auxiliary cost function.

LEMMA 1: $R_A(x') \le R(x')$ *for* $x' \ge -w.$

Proof: The lemma holds for $x' \geq 0$ since in this case $R_A(x') = 0 \leq R(x')$. For $x' < 0$, $R_A(x')$
$= -x'r/w = -x'R(-w)/w = -x'\sup_k\{Q(k) + kw\}/w = -x'\sup_k\{Q(k)/w + k\}$. Since $0 < (-x') \leq w$,

$$-x'\sup_k\{Q(k)/w + k\} \leq -x'\sup_k\{Q(k)/(-x') + k\} = \sup_k\{Q(k) - x'k\} = R(x').$$

We assume for the rest of this appendix that $x'_B(t) \geq 0$ and $r_B(t) \geq 0$.

LEMMA 2: $0 \leq R_A(x') \leq R_B(x', t, x)$.
  Proof: In view of Lemma 1, we only need to prove Lemma 2 for the first case of definition 2; i.e., it suffices to show that $R_A(x') \leq R_B(x', t, x) = r_B(t)$ when $x = x_B(t)$, $x' = x'_B(t) \geq 0$ and $0 \leq r_B(t) < R(x')$. This is obviously true since for $x' \geq 0$, $R_A(x') = 0$ and $R_B(x', t, x) \geq 0$.

LEMMA 3: If valid path $\mathcal{P}$ goes from point D to point P, $\Delta(\mathcal{P}) \geq \Delta_A(\mathcal{P}) \geq \max\{0, \kappa(x_D - x_P)\}$.
  Proof: The first inequality follows from Lemma 2, since R_B is the cost used to calculate $\Delta(\mathcal{P})$ and R_A is the cost used to calculate $\Delta_A(\mathcal{P})$. The second inequality holds because $\max\{0, \kappa(x_D - x_P)\}$ is the auxiliary cost of the linear path from D to P, which is the least possible because $R_A(x')$ is time- and space-independent.

LEMMA 4 (Existence): *If $0 \leq u$, $w < \infty$, and the initial data satisfy (6) then there is a $\mathcal{P}^*$ which achieves (5) for both the IVP and the FHP.*
  Proof: The set of feasible paths $\mathcal{P}$ from the boundary to P $\in$ **S** is a non-empty, closed and bounded subset of the linear normed space of continuous functions, $x(t)$, with respect to the l_∞ norm. The RHS of (5) is bounded from below because (6) applies to $N_{D(\mathcal{P})}$, and $\Delta(\mathcal{P}) \geq 0$. Since the set of feasible paths is closed, $\mathcal{P}^*$ exists.

LEMMA 5: *If conditions (i), (ii) and (iii) of Theorem 1 in the text apply for an IVP with bottlenecks, then (6a) holds.*
  Proof: Consider two points A and B in the solution domain with coordinates $t_A = t_B = t$ and $x_B > x_A$.
  We first prove that $N_B \leq N_A$. Consider the maximal path, $\mathcal{U}_B$, which reaches B from the boundary with $x' = u$ from a source B'$\in$ **D**, and an optimum path from the boundary to A, $\mathcal{P}^*_A$ which emanates from a point A'$\in$ **D**. If the paths do not intersect (i.e., $x_{B'} > x_{A'}$) then $N_{B'} \leq N_{A'}$ and we can write: $N_B \leq N_{B'} \leq N_{A'} = N_A - \Delta(\mathcal{P}^*_A) \leq N_A$. (The first equality holds because the maximal path imposes a capacity constraint with zero cost, the equality because $\mathcal{P}^*_A$ is a solution of (5), and the last inequality because $\Delta(\mathcal{P}^*_A)$ is non-negative.) Thus, $N_B \leq N_A$ if the paths do not intersect. If the paths intersect, there is a common point C. Clearly, $N_B \leq N_C$ since C is on the maximal path to B, and $N_C \leq N_A$ since C is on the optimum path from the boundary (and the optimum path from C to A has non-negative cost). Thus, $N_B \leq N_A$ if the paths do not intersect.
  To conclude the proof we now show that $N_A \leq N_B + \kappa(x_B - x_A)$ using similar logic. Consider the minimal path, $\mathcal{W}^o_A$, which reaches A from the boundary with $x' = -w$ from a source A'', and an optimum path from the boundary to B, $\mathcal{P}^*_B$, which emanates from a point B''. If the paths do not intersect then $N_{B''} \leq N_{A''}$ and we can write: $N_A \leq N_{A''} + \kappa(x_{A''} - x_A) \leq N_{B''} + \kappa(x_{B''} - x_{A''}) + \kappa(x_{A''} - x_A) = N_{B''} + \kappa(x_{B''} - x_A) = N_B - \Delta(\mathcal{P}^*_B) + \kappa(x_{B''} - x_A) \leq N_B - \Delta_A(\mathcal{P}^*_B) + \kappa(x_{B''}$

$- x_A) \le N_B - \kappa(x_{B''} - x_B) + \kappa(x_{B''} - x_A) = N_B + \kappa(x_B - x_A)$. The first inequality holds because $\kappa(x_{A''} - x_A)$ is an upper bound to the cost of the minimal path from A''; the second because $N_{A''} \le N_{B''} + \kappa(x_{B''} - x_{A''})$ as per (6a); the first equality is algebraic; the second one holds because $\mathscr{P}^*_B$ is an optimum path from the boundary emanating from B''; the third and fourth inequalities hold because of Lemma 3; and the fifth one is algebraic. Thus, $N_A \le N_B + \kappa(x_B - x_A)$ if the paths do not intersect. If the paths intersect, then again there is a common point C and a common part of the optimum path extending from C to B, which we denote $\mathscr{P}^*_{CB}$. And we see using the same logic that: $N_A \le N_C + \kappa(x_C - x_A) = N_B - \Delta(\mathscr{P}^*_{CB}) + \kappa(x_C - x_A) \le N_B - \Delta_A(\mathscr{P}^*_{CB}) + \kappa(x_C - x_A) \le N_B - \kappa(x_C - x_B) + \kappa(x_C - x_A) = N_B + \kappa(x_B - x_A)$. Thus, $N_A \le N_B + \kappa(x_B - x_A)$ whether or not the paths intersect.

LEMMA 5a: *If conditions (i), (ii) and (iii) of Theorem* 2 *in the text apply for a FHP with bottlenecks, then (6a) holds.*

  Proof: The proof of Lemma 1 can be repeated word for word with only two changes. First, for the proof that $N_B \le N_A$ when the maximal and optimal paths $\mathcal{U}_B$ and $\mathscr{P}^*_A$ do not intersect, we need to recognize that one or both paths may emanate from the line $x = 0$, and that it is still true that $N_{B'} \le N_{A'}$ therefore we can still write $N_B \le N_{B'} \le N_{A'} = N_A - \Delta(\mathscr{P}^*_A) \le N_A$. Second, for the proof that $N_A \le N_B + \kappa(x_B - x_A)$ when the minimal and optimal paths $\mathcal{W}^o_A$ and $\mathscr{P}^*_B$ do not intersect, we need to we need to recognize that one or both paths may emanate from the line $x = x^o$. To avoid this problem, extend both paths to the line $t = 0$ with slope $-w$ and define A'' and B'' as the intersection of the extended paths with the line $t = 0$. We imagine that the highway extends to $x = \infty$, is filled with jam density for $x > x^o$, and define $N_{A''}$ and $N_{B''}$ accordingly. Then the extended path $\mathscr{P}^*_B$ continues to be optimal, the inequality $N_{A''} \le N_{B''} + \kappa(x_{B''} - x_{A''})$ is again implied by (6a), and the proof of this case goes through verbatim.

LEMMA 5b (Strictly monotone problems without bottlenecks): *If for an IVP without bottlenecks the conditions of Lemma 5 are satisfied and n(0, x) is strictly monotone, then N(t, x) is strictly monotone. This is also true for problems with bottlenecks in the part of the solution domain upstream of the bottlenecks.*

  Proof: Strict monotonicity of the data allows us to replace the inequality $N_{B'} \le N_{A'}$ used in the proofs of Lemmas 5 and 5a as part of the relation $N_B \le N_{B'} \le N_{A'} = N_A - \Delta(\mathscr{P}^*_A) \le N_A$ by a strict inequality. This establishes that $N_B < N_A$ in the case where the paths used in the proofs do not intersect. Lack of bottlenecks in the relevant part of the solution domain (upstream of the bottlenecks) allows us to state for the case where the paths intersect at C that the part of the optimum path from C to A has positive cost—since the average slope of this path must then be strictly less than u. Hence the inequality $N_C \le N_A$ is strict and this establishes that $N_B < N_A$ when the paths do intersect.

LEMMA 6 (Bounded flows): *If conditions (i), (ii) and (iii) of Theorem* 1 (*or Theorem* 2) *apply for an IVP (or FHP) with bottlenecks, then (6b) holds.*

  Proof: Since paths with $x' = 0$ are valid and satisfy $R_B \le R = q_{max}$, it follows that $N(t_2, x) - N(t_1, x) \le (t_2 - t_1)q_{max}$ if $t_1 < t_2$. Thus, the upper bound part of (6b) holds. To prove the lower bound part, consider the maximal path from (t_1, x) to the line $t = t_2 > t_1$, reaching it at (t_2, x_2). Assume first that $x_2 \le x^o$. Clearly, Lemma 5 guarantees $N(t_2, x) \ge N(t_2, x_2)$, and since the maximal path has zero cost $N(t_2, x_2) \ge N(t_1, x)$. Thus, $N(t_2, x) \ge N(t_2, x_2) \ge N(t_1, x)$. If $x_2 >$

x^o, there is a time $t_1' = t_1 + (x^o - x)/u < t_2$ at which the maximal path intersects the downstream boundary. The same logic now implies $N(t_1 + (x^o - x)/u, x) = N(t_1', x) \geq N(t_1', x^o) \geq N(t_1, x)$. We can now try a maximal path from (t_1', x) and repeat the argument until we reach the line $t = t_2$ without intersecting the downstream boundary. We find: $N(t_1, x) \leq N(t_1 + (x^o - x)/u, x) \leq N(t_1 + 2(x^o - x)/u, x) \leq \ldots \leq N(t_2, x)$.

COMMENT: For the FHP (and other problems with complex boundaries) valid paths from the boundary to a point P in the solution domain may leave S and return to it. The consistency condition (iii) implies that such paths cannot be unique optima. Hence, they can be ignored when solving (5). This is the recommended option for application because discrete networks for numerical solution then only have to be defined in S. (The reader may want to verify that the proofs of Lemmas 5a and 6 still hold if paths are not allowed to leave S.)

APPENDIX B: ASYNCHRONOUS CELL TRANSMISSION MODEL

In this appendix primes do not denote derivatives. The cell transmission model (CTM) with time step ε, cells of size $\delta = u\varepsilon$, and a triangular FD is:

$$N_{lm} - N_{l-1,m} = \min\{N_{l-1,m-1} - N_{l-1,m}, q_{max}\varepsilon, (\kappa - (N_{l-1,m} - N_{l-1,m+1})/\delta)w\varepsilon\}. \tag{B1}$$

The LHS of (B1) is the flow advancing in one time step across the mth intercell boundary. The RHS is a function of the vehicles currently in the upstream and downstream cells. It is well known (Daganzo, 1994, 1995) that the last term of the CTM formula, which expresses the available capacity of the downstream cell, introduces a first order numerical error when $w \neq \delta/\varepsilon$, and that the error vanishes if $w = \delta/\varepsilon$. These errors can be eliminated by changing the time variable to asynchronous time—as proposed in Sec. 5.2.2 of Daganzo (2003b). The result was called the asynchronous cell transmission model (ACTM).

To summarize, imagine that clocks at each location have been synchronized with the passage of a reference vehicle with negative speed, s, such that: $1/s = 1/u - 1/w < 0$. Thus, the new (asynchronous) time is $t' = t + x/s$, and the new lattice instants at $x = x_m$ are related to the old by: $l' = t_l'/\varepsilon = t_l/\varepsilon + x_m/s\varepsilon = t_l/\varepsilon + m\delta/s\varepsilon = t_l/\varepsilon + m(u\varepsilon)(1/u - 1/w)/\varepsilon = l + m(1 - \theta)$. If θ is an integer then the lattice remains the same, since the lattice instants are displaced from the old by an integer multiple of the time step. This leaves invariant the jam density but changes speed as per: $1/v' = 1/v + 1/w - 1/u$. The advantage of the new coordinate system is that the speed of the backward wave adopts the value $-w'$ such that: $1/(-w') = 1/(-w) + 1/w - 1/u$; i.e., $w' = u$, and therefore $w' = \delta/\varepsilon$. Thus, the formula for available capacity in the new coordinate system, which is $[\kappa - (N_{l'-1,m} - N_{l'-1,m+1})/\delta]w'\varepsilon$, is exact. In terms of the old variables, l and w', this expression becomes $[\kappa - (N_{l'-1,m} - N_{l'-1,m+1})/\delta]w'\varepsilon = [\kappa - (N_{l-1,m} - N_{l-\theta,m+1})/\delta]u\varepsilon = \kappa\delta - (N_{l-1,m} - N_{l-\theta,m+1})$. Substituting this expression for the last term of (B1) we obtain the exact ACTM recipe:

$$N_{lm} - N_{l-1,m} = \min\{N_{l-1,m-1} - N_{l-1,m}, q_{max}\varepsilon, \kappa\delta - (N_{l-1,m} - N_{l-\theta,m+1})\}, \text{ or} \tag{B2}$$

$$N_{lm} = \min\{N_{l-1,m-1}, N_{l-1,m} + q_{max}\varepsilon, \kappa\delta + N_{l-\theta,m+1}\}. \tag{B3}$$