On the Numerical Treatment of Moving Bottlenecks

Carlos F. Daganzo and Jorge A. Laval

Institute of Transportation Studies and Department of Civil and Environmental Engineering

University of California, Berkeley CA 94720

(March 18, 2003)

ABSTRACT

This report is part of PATH Task Order 4141 and shows how moving obstructions can be modeled numerically with kinematic wave theory. It shows that if a moving obstruction is replaced by a sequence of fixed obstructions at nearby locations with the same "capacity", then the error in vehicle number converges uniformly to zero as the maximum separation between the moving and fixed bottlenecks is reduced. This result implies that average flows, densities, accumulations and delays can be predicted as accurately as desired with this method. Thus, any convergent finite difference scheme can now be used to model moving bottlenecks. An example is given.

TABLE OF CONTENTS

1. INTRODUCTION

Our experience as drivers tells us that a truck on an uphill, a vehicle changing lanes or a vehicle accelerating while entering the freeway can affect significantly freeway traffic streams. Empirical findings confirm this intuition. It is known, for example, that lane changes are an important contributor to the formation of stop-and-go waves (Mauch and Cassidy, 2002), and that slow trucks contribute to the formation of sag bottlenecks (Koshi et al, 1992). Controlled experiments show that the effects of moving obstructions are indeed significant (Muñoz and Daganzo, 2002). This suggests that "moving bottlenecks" should be incorporated into practical traffic models.

Gazis and Herman (1992) were first to recognize the moving bottleneck problem and to address some of its issues. Later, moving obstructions were introduced into the kinematic wave (KW) theory of Lighthill and Whitham (1955), and Richards (1956), and this led to a complete theory of moving bottlenecks (Newell, 1998). Because this theory did not fully agree with observation, it was revised to improve its predictions (Muñoz and Daganzo, 2002.) With these models it is possible to predict the effects of any set of moving obstructions on a KW traffic stream. Unfortunately, an efficient way of discretizing these models has not yet been found.

The issue is not trivial because in any finite difference scheme the time-space (t, x) trajectory of the moving obstruction $\phi(t)$ is most simply modeled by a step function $\phi^{(a)}(t)$ with steps equal to the lattice spacing, Δx; see Fig. 1. Although it is possible to choose $\phi^{(a)}(t) \to \phi(t)$ by letting $\Delta x \to 0$, it is not generally possible to match the flows and densities of the solutions on both sides of ϕ and $\phi^{(a)}$. To see this note that if a bottleneck ϕ is in an active steady state (i.e., it is holding back a queue) while traveling at a speed $v = d\phi/dt > 0$, then the two flow-density (q, k)

states directly upstream and downstream of the bottleneck must satisfy the "shock condition" $\Delta q/\Delta k = v$. This means that on the (q, k) plane they must be on a "capacity line" with slope v; see points U and D on Fig. 2. Thus, the flow in these states must be different. Yet, for a bottleneck with zero speed, such as any step of $\phi^{(a)}$, the difference in flows $\Delta q^{(a)}$ must be zero, as shown by points $U^{(a)}$ and $D^{(a)}$ of the figure. Thus, conservative finite difference schemes that produce a KW solution by introducing boundary conditions of the form $\phi^{(a)}$, with $\Delta q^{(a)} = 0$ if the bottleneck is holding back a queue, cannot precisely match the true flows on both sides of an active bottleneck, since $\Delta q \neq 0$ when the bottleneck is moving. Fortunately, this difficulty is not fatal. We will show that if the capacity of $\phi^{(a)}$ is carefully chosen, then the Moskowitz function of the approximation tends to the true function, $N^{(a)} \rightarrow N$, as $\Delta x \rightarrow 0$.[1] Thus, practical approximation schemes can be built, and this will be demonstrated.

Section 2 below describes how errors in the input data of KW moving bottleneck problems affect the Moskowitz solutions. Section 3 shows how errors due to an imperfect bottleneck trajectory affect the solution—the errors can be reduced as much as desired by reducing Δx. In Secs. 2 and 3 time is continuous. Section 4 then shows how the results can be used to approximate the Moskowitz function with discrete-time numerical schemes. Section 5 discusses future work.

---

[1] Recall that that the Moskowitz function (see Moskowitz, 1965, and Makigami et al, 1971) is the integral of flow that gives the vehicle number on the (t, x) plane. Moskowitz functions are continuous with a bounded rate of variation, and are related to the density and flow functions by $\partial N/\partial x = -k$ and $\partial N/\partial t = q$.

2. ERROR BOUNDS: SAME BOUNDARY, DIFFERENT DATA

This section compares two KW solutions defined on the same boundary with different input densities along the boundary ("exact" and "approximate"). We start with conventional initial-value problems; then extend the results to more general boundary value problems, and finally to problems including moving bottlenecks. In all cases, it is shown that the solution error at any point $|N^{(a)} - N|$ can never exceed the maximum input error.

2.1 Homogeneous Initial-Value Problems

It is known (Lax, 1973) that the integral $\int_{-\infty}^{\infty} \left| k^{(a)}(t,x) - k(t,x) \right| dx$ corresponding to any two density functions, $k^{(a)}(t, x)$ and $k(t, x)$, arising from the solution of two KW problems with the same equilibrium flow-density relation, $F(k) \in C^1$, but different initial data is non-increasing in t. In other words, the separation of the two density profiles at a given time, as measured by the L_1 norm, cannot increase with t. This result is intimately related to the ideas of uniqueness and stability since it implies that solutions with the same input data must match everywhere (uniqueness), and also implies that small perturbations to initial data cannot grow into the solution (stability).

It turns out that the Moskowitz function satisfies a similar property with respect to the L_∞ norm. It has been shown (Daganzo, 2001 and 2003) that if F is continuous, non-negative and satisfies $q = 0$ for $k \notin [0, k_j]$, where k_j is the "jam density", then $sup_x |N^{(a)}(t, x) - N(t, x)|$ is non-increasing in t.[2] Thus, if we denote by "E" the absolute difference (i.e., the error) between the two Moskowitz functions, we can write:

$$E \le sup_x[E(0, x)]. \tag{1a}$$

[2] Actually, these references show that $sup_t |N^{(a)}(t, x) - N(t, x)|$ is non-decreasing in x for initial boundary problems where flows are given at $x = 0$. The result extends trivially to the initial value problem by reversing the roles of x and t.

The same result can be expressed in a way that will be useful later if we call the curve defined by the Moskowitz function for a fixed value of t an "N-profile" and denote it by M_t, and if we also define a kinematic wave operator $\mathcal{W}_t$ that returns M_t from M_0; i.e., $M_t = \mathcal{W}_t M_0$. This operator is a property of F. If we now use double bars to denote the L_∞ norm then the property in question becomes $\| \mathcal{W}_t M^{(a)} - \mathcal{W}_t M \| \leq \| M^{(a)} - M \|$, for any $M^{(a)}$, M and t. In other words, the kinematic wave operator is a (non-strict) contraction mapping for any t. Operators with this property will be simply called "contractions". Note that for time-independent F-functions $\mathcal{W}_{t+t'} = \mathcal{W}_t \mathcal{W}_{t'}$. Thus, $\mathcal{W}_t$ can be expressed as the composition of many identical "differential" operators, $\mathcal{W}$. Since the composition of contractions is a contraction, an equivalent statement to (1a) is: "$\mathcal{W}$ is a contraction".

2.2 General Boundaries

The contraction property also applies to more general boundary value problems. Let C be a continuous curve on which two sets of well-posed density data have been defined, and let P_0 be a point (t_0, x_0) that is part of the solution obtained from C. See Fig. 3. We now show that (1a) is generalized by:

$$E_0 \leq sup_{R \in C} [E_R], \tag{1b}$$

where $E_0 = | N^{(a)}(t_0, x_0) - N(t_0, x_0) |$ and E_R is the absolute difference in N-values for a point, R, on the curve.

The logic behind (1b) uses the fact that $E = | N^{(a)} - N |$ is a continuous function of x and t with a bounded rate of variation (since N and $N^{(a)}$ are continuous with bounded rates of variation.) Assume that (1b) is false; i.e., that there is an $\varepsilon > 0$ such that $E_0 \geq \varepsilon + E_R$, $\forall\ R \in C$. If we let Δt be a small time increment, we know from (1a) that there must be a point P_1 in the domain of dependence of point P_0 on the line $t = t_0 - \Delta t$ such that $E_1 \geq E_0$. The same happens at P_1. Thus, we can form a

sequence of points $\{P_i\}$, approaching the curve C , for which $E_i \geq E_0$; see Fig. 3. For all these points $E_i \geq E_0 \geq \varepsilon + E_R$; i.e., $E_i - E_R \geq \varepsilon, \forall\ R \in C$. This is independent of Δt. Since a point P_i arbitrarily close to C can always be found by reducing Δt, we see that E cannot have a bounded rate of variation near C. Thus, (1b) must hold.

2.3 Problems with Moving Bottlenecks

First we need some definitions. From now on, flows and densities for the states ("*U*" and "*D*") next to an active moving bottleneck with speed v (see Fig. 2) will be denoted by capitalized letters and capitalized subscripts; e.g., Q_D, K_U. The downstream flow, Q_D, will be called "the capacity" of the moving bottleneck, and its subscript will be omitted. Recall too, from the theory of moving bottlenecks, that the intercept of line $\overline{UD}$ is the maximum passing rate, Q_r.

We consider now moving bottlenecks with exogenous trajectories and capacities, and show that (1a) and (1b) continue to hold. A separate discussion is necessary because the bottleneck trajectory does not define a boundary condition of the type discussed in Sec. 2.2 since the densities along it are now endogenously determined—only its speed and capacity are exogenous. To establish the result we first modify the differential wave operator to include bottlenecks, and then show that it remains a contraction.

If a homogeneous freeway includes a moving bottleneck at $x = x_0$ and time t, then the N-profile upstream of the bottleneck at time $t + dt$ can be obtained by applying the wave operator to the following two profiles at time t (i.e., ignoring the bottleneck) and comparing the results: (*i*) the original N- profile, M_t , and (*ii*) an artificial profile, $M_t^{(U)}$, which is identical to M_t except that it is in state "*U*" downstream of the bottleneck:

$$M_t^{(U)}(x) = M_t(x)\,, \qquad \text{if } x \leq x_0 \tag{2a}$$

$$= M_t(x_0) - K_U(x - x_0) \qquad \text{if } x > x_0\,. \tag{2b}$$

Case (*ii*) gives the correct solution when the bottleneck is active and case (*i*) when it is not. The reader can verify by exhaustively examining all possibilities that the correct solution is always the lower of the two cases. Thus, if we use an operator $\mathcal{U}$ to express (2), $M_t^{(U)} = \mathcal{U}M_t$, the solution upstream of the bottleneck is:

$$M_{t+dt} = inf\{\mathcal{W}M_t\,;\, \mathcal{W}\mathcal{U}M_t\} \qquad \text{for } x \le x_0 + vdt\,, \tag{3}$$

where the *inf* operation denotes the lower envelope of the two curves. (For simplicity, our notation does not explicitly reflect the dependence of $\mathcal{U}$ on x_0.)

For the downstream part we define another location-dependent operator $\mathcal{D}$ that changes the upstream part of the profile:

$$M_t^{(D)}(x) = M_t(x)\,, \qquad \text{if } x \ge x_0 \tag{4a}$$

$$= M_t(x_0) - K_D(x - x_0) \qquad \text{if } x < x_0\,; \tag{4b}$$

and use a similar minimum rule:

$$M_{t+dt} = inf\{\mathcal{W}M_t\,;\, \mathcal{W}\mathcal{D}M_t\} \qquad \text{for } x \ge x_0 + vdt\,. \tag{5}$$

The combination of (3) and (5) is the desired result. The complete operation will be denoted "$\mathcal{B}$", so that $M_{t+dt} = \mathcal{B}M_t$.

Note as an aside that M_{t+dt} must be continuous since it is part of a KW solution. Reassuringly, the vehicle numbers predicted by both (3) and (5) at $x = x_0 + vdt$ always match. As expected, these numbers never exceed $N_t(x_0) + Q_r\,dt$, and they equal this value when the bottleneck is active. We now show that $\mathcal{B}$ is a contraction.

Lemma 1: If (a, A, b, B) are real numbers such that $|A-B| \le \varepsilon$ and $|a-b| \le \varepsilon$, then $|min(A, a) - min(B, b)| \le \varepsilon$. □

Proof: If $(A, B) <$ (or $>$) (a, b) then $|min(A, a) - min(B, b)| = |A-B|$ or $|a-b|$, and the lemma is obviously true. Otherwise, when $(A, b) <$ (or $>$) (a, B), $|min(A, a) - min(B, b)| = |A-b|$ or $|B-a|$. By symmetry, it suffices to pursue the first of these two instances, checking the magnitude of $|A-b|$ when $A < a$ and $b < B$. If $A \geq b$, then $|A-b| \leq |a-b| \leq \varepsilon$. Otherwise, $|A-b| \leq |A-B| \leq \varepsilon$. Thus, $|A-b| \leq \varepsilon$ in both cases, and this concludes the proof. □

Theorem 1: Operation $\mathcal{B}$ is a contraction. □

Proof: For the upstream part (3), we need to show that

$$\left\| M^{(a)} - M \right\| \leq \varepsilon \;\Rightarrow\; \left\| inf\{\mathcal{W}M^{(a)};\, \mathcal{W}\mathcal{U}M^{(a)}\} - inf\{\mathcal{W}M;\, \mathcal{W}\mathcal{U}M\} \right\| \leq \varepsilon. \tag{6a}$$

Note first that (2) is a contraction. Therefore, $\mathcal{U}$ is a contraction. Since $\mathcal{W}$ is a contraction, $\mathcal{W}\mathcal{U}$ is a contraction too. Hence, if $\left\| M^{(a)} - M \right\| \leq \varepsilon$ then $\left\| \mathcal{W}M^{(a)} - \mathcal{W}M \right\| \leq \varepsilon$ and $\left\| \mathcal{W}\mathcal{U}M^{(a)} - \mathcal{W}\mathcal{U}M \right\| \leq \varepsilon$. Lemma 1 can now be invoked with: $A = \mathcal{W}M^{(a)}$, $B = \mathcal{W}M$, $a = \mathcal{W}\mathcal{U}M^{(a)}$, and $b = \mathcal{W}\mathcal{U}M$, and the result is (6a). Similar logic applies to the downstream part (5), and we find that

$$\left\| M^{(a)} - M \right\| \leq \varepsilon \;\Rightarrow\; \left\| inf\{\mathcal{W}M^{(a)};\, \mathcal{W}DM^{(a)}\} - inf\{\mathcal{W}M;\, \mathcal{W}DM\} \right\| \leq \varepsilon \tag{6b}$$

The combination of (6a) and (6b) establishes that $\mathcal{B}$ is a contraction, and this concludes the proof. □

Theorem 1 only establishes the desired result for the initial value problem, but the logic of section 2.2 again reveals that (1b) holds. In fact, one can easily see along similar lines of thought that the theorem and the inequality hold for general boundary problems with multiple bottlenecks.

3. ERROR BOUNDS: DIFFERENT BOUNDARIES

We now use (1) to (6) to see how a moving bottleneck, $\phi(t)$ with capacity $Q(t)$, can be approximated by another one whose trajectory is a step function, $\phi^{(a)}$, defined on a spatial lattice $x_i = i\Delta x$. We assume that $\| \phi^{(a)} - \phi \| \le \Delta x$ and $\phi^{(a)} \le \phi$ as in Fig.1.

It will be shown in this section that if $Q^{(a)}(t) = Q(t)$ for all t then $N^{(a)} \to N$ as $\Delta x \to 0$. We will prove this assertion first for piecewise linear problems, where $F(k)$, $N(0, x)$ and ϕ are piecewise linear, and $Q(t)$ is piecewise constant. Piecewise linear problems are easy to treat because their solutions, $k(t, x)$, are piecewise constant in a polygonal tessellation of the (t, x) plane. The edges of these polygons are "interfaces" that satisfy the "shock condition". The vertices of the polygon can be associated with a sequence of times, $\{\tau_p\}$; see Fig. 4.

If the initial data on the line $t = 0$ includes a finite number of constant-density sates, and the moving bottleneck is also composed of a finite number of linear segments with constant Q, then the tessellation has a finite number of vertices and the series $\{\tau_p\}$ where interfaces converge or diverge has a finite number of elements. [3]

[3] This should be intuitive if F is concave and the bottleneck eventually leaves the road because, in the concave case, non-bottleneck vertices eliminate interfaces. Thus, once the bottleneck has left the road the number of interfaces declines at each τ_p. The process stops (after a finite number of steps) when one interface remains, or before if all the interfaces diverge. More elaborate geometrical considerations show that finite stopping also occurs in the non-concave case, even if the bottleneck stays on the road.

Lemmas 2 and 3, below, show that if $|N^{(a)} - N|$ is bounded by ε_p at time $t = \tau_p$ then $|N^{(a)} - N|$ is bounded by $\varepsilon_p + 3k_j\Delta x$ in the interval (τ_p, τ_{p+1}) . Thus, if we choose $\varepsilon_o = 0$, we see by iteration of this result for $p = 0, 1, \ldots P$ that $|N^{(a)} - N| \le 3pk_j\Delta x$ in every interval $[\tau_p, \tau_{p+1}]$, and since the number of intervals is finite, this establishes that $N^{(a)} \to N$ uniformly in (t, x); i.e., that the following is true:

Theorem 2. For piecewise linear problems with P vertices, if we choose $Q^{(a)}(t) = Q(t)$ and $\phi^{(a)}$ is of the form shown in Fig. 1, then $| N^{(a)} - N | \le 3Pk_j\Delta x$. □

We now turn our attention to the Lemmas. Let $\{t_n\}$ be the set of instants where $\phi^{(a)}$ has a step. We examine first the propagation of errors in stationary intervals that begin with a step, and then relax the "step condition".

Lemma 2. If $\tau_p \in \{t_n\}$ and $| N^{(a)} - N | \le \varepsilon_p$ for all x and $t = \tau_p$, then $| N^{(a)} - N | \le \varepsilon_p + k_j\Delta x$ for all x and $t \in [\tau_p, \tau_{p+1}]$. □

Proof: Consider the two bottleneck trajectories of Fig. 5a and note that they intersect at $t = \tau_p$, as required by the lemma. Let N^* be a "test" solution of the approximate bottleneck problem with the same initial data as the exact problem; i.e., such that $N^* = N$ for $t = \tau_p$. We shall compare N, N^*, and $N^{(a)}$.

For $t \in [\tau_p, \tau_{p+1}]$ the bottleneck must be in a steady state with a constant passing rate q_p. This rate must satisfy $q_p \le Q_r$, where Q_r is the maximum possible passing rate. Equality means that the bottleneck is active, but the bottleneck may or may not be active. If the bottleneck is inactive the same traffic state should be observed upstream and downstream of it. Figure 5b shows the two

states (B or B') associated with a specific $q_p < Q_r$. Recall, however, that if the bottleneck is active the steady states will be different on its upstream and downstream sides (U and D).

Consideration of the test solution for all possible steady states reveals that N^* always equals N downstream of ϕ. Differences between the two solutions only arise in the region upstream of ϕ. For example, if the bottleneck were inactive in state B' as in Fig. 5b the exact solution would consist of state B' on both sides of the bottleneck. The test solution, however, would be as in Fig. 5c, which induces states other than B' – but only upstream of ϕ.

If the bottleneck was active then the exact solution would be as in Fig. 5d, but the test solution would be as in Fig. 5e; again, with no changes downstream of ϕ. This is always true. Thus, $N = N^*$ for $x \geq \phi$. An observer moving in front of the bottleneck would never know whether the true or approximate bottleneck was in force.

Consider now the region between ϕ and the shifted curve $\phi^{(s)} = \phi - \Delta x$, which includes $\phi^{(a)}$; see Fig. 5f. Since the width of this region is Δx, and $N = N^*$ along its top edge, it follows that $|N^* - N| \leq k_j \Delta x$ in the region.

This bound also holds beneath $\phi^{(s)}$, because in this region the solution is determined by data on the boundary C (see Fig. 5f), and in these cases (see (1b)) the discrepancy in the two solutions is bounded by the maximum discrepancy along C, which is itself bounded by $k_j \Delta x$. Thus $|N^* - N| \leq k_j \Delta x$ for $t \in [\tau_p, \tau_{p+1}]$.

If Theorem 1 is now applied to $N^{(a)}$ and N^* with the line $t = \tau_p$ as the boundary, where the error is bounded by $|N^{(a)} - N^*| \equiv |N^{(a)} - N| \leq \varepsilon_p$, we find that $|N^{(a)} - N^*| \leq \varepsilon_p$ for $t \in [\tau_p, \tau_{p+1}]$. Thus, $|N^{(a)} - N^*| + |N^* - N| \leq \varepsilon_p + k_j \Delta x$ for $t \in [\tau_p, \tau_{p+1}]$, and the triangle inequality $|N^{(a)} - N| \leq |N^{(a)} - N^*| + |N^* - N|$ ensures that $|N^{(a)} - N| \leq \varepsilon_p + k_j \Delta x$, as claimed. □

This lemma assumed that $\tau_p \in \{t_n\}$ but the result can be generalized.

Lemma 3. If $|N^{(a)} - N| \le \varepsilon_p$ for all x and $t = \tau_p$, then $|N^{(a)} - N| \le \varepsilon_p + 3k_j\Delta x$ for all x and $t \in [\tau_p, \tau_{p+1}]$. □

Proof: In this case there is a discrepancy, $\delta = \phi - \phi^{(a)} > 0$ at $t = \tau_p$; see Fig. 6. We consider the auxiliary shifted solution, $N^{(s)}(x, t) = N(x+\delta, t)$, formed by shifting the exact bottleneck and the exact data by δ space units. Note that the errors induced by the shift are bounded by $|N^{(s)} - N| \le k_j\delta$. We consider this solution in the interval $[\tau_p, \tau_{(p+1)}]$, where $\tau_{(p+1)}$ is the first time the shifted trajectory intersects either a step of $\phi^{(a)}$ (as shown in the figure) or the line $t = \tau_{p+1}$.[4]

We first estimate the error $|N^{(a)} - N^{(s)}|$ in that interval by introducing an auxiliary approximate bottleneck $\phi^{(s)}$ with steps at τ_p and $\tau_{(p+1)}$ that matches $\phi^{(a)}$ in $[\tau_p, \tau_{(p+1)}]$; see the dashed horizontal line in the figure. This bottleneck satisfies the step condition of Lemma 2. Thus, the lemma can be used to compare $N^{(a)}$ and $N^{(s)}$ in $[\tau_p, \tau_{(p+1)}]$. Since the step size of $\phi^{(s)}$ is bounded by $(\Delta x - \delta)$, see figure, Lemma 2 implies:

$$|N^{(a)} - N^{(s)}| \le \varepsilon_{a,s} + k_j(\Delta x - \delta) \text{ for } t \in [\tau_p, \tau_{(p+1)}], \tag{7}$$

where $\varepsilon_{a,s}$ is the maximum error between $N^{(a)}$ and $N^{(s)}$ at time τ_p. Since $|N^{(a)} - N^{(s)}| \le |N^{(a)} - N| + |N - N^{(s)}|$ everywhere, the supremum of the left side of this inequality along the initial boundary, $\varepsilon_{a,s}$, is bounded from above by the supremum of the right side. And since the right side is itself bounded by $\varepsilon_p + k_j\delta$, we see that $\varepsilon_{a,s} \le \varepsilon_p + k_j\delta$. If this inequality is now introduced in the right side of (7), we find:

$$|N^{(a)} - N^{(s)}| \le \varepsilon_p + k_j\Delta x \quad \text{for} \quad t \in [\tau_p, \tau_{(p+1)}]. \tag{8}$$

---

[4] In other words, $\tau_{(p+1)}$ is the first member of $\{t_n\}$ in the interval $(\tau_p, \tau_{p+1}]$, if such a member exists, or else it is τ_{p+1}.

We can now estimate the total error. Since $|N^{(s)} - N| \leq k_j\delta$ for $t \in [\tau_p, \tau_{(p+1)}]$ we can use the triangle inequality and (8) to write:

$$|N^{(a)} - N| \leq |N^{(a)} - N^{(s)}| + |N^{(s)} - N| \leq \varepsilon_p + k_j\Delta x + k_j\delta \leq \varepsilon_p + 2k_j\Delta x \quad \text{for} \quad t \in [\tau_p, \tau_{(p+1)}].$$

Obviously, if $\tau_{(p+1)} = \tau_{p+1}$, the proof is completed. Otherwise, $\phi^{(a)}$ has a step at time $\tau_{(p+1)}$, as in Fig. 6, and Lemma 2 can be used to estimate the error in $[\tau_{(p+1)}, \tau_{p+1}]$. Since the revised input error is $\varepsilon_p + 2k_j\Delta x$ at time $\tau_{(p+1)}$, the resulting bound is $\varepsilon_p + 3k_j\Delta x$. □

This Lemma is the last step in the proof of Theorem 2, which provides an error bound for piecewise linear problems. Bounds can also be obtained for the class of problems whose input data can be approximated arbitrarily well by a sequence of piecewise linear problems. In this case, if we let P_i denote the number of vertices for the i^{th} problem and approximate its solution using a bottleneck step $(\Delta x)_i = (\Delta x)_o/(iP_i)$ we see from the Theorem that the error bound for the i^{th} problem is $3(\Delta x)_o k_j/i$, which tends to zero as $i \to \infty$.

The proposed approach also works for continuous F-functions. To see this, we recall from Daganzo (2001) that the error in the solution introduced by using a piecewise linear approximation $F^{(a)}$ to F, instead of F, is of order Δq, where $\Delta q = sup\,|F^{(a)} - F|$. The rationale for our claim is the following. Since Δq can be made as small as desired by increasing the number of pieces in $F^{(a)}$, we see that the piecewise solution tends to the continuous solution as $\Delta q \to 0$, both with the exact and the step bottlenecks. Hence, since the step solution can be as close as desired to the exact in the piecewise linear case for any Δq, it follows that the same must happen in the continuous case.

The same argument reveals that the method can also be used with inhomogeneous highways (and networks) since they can always be approximated arbitrarily well by combination of homogeneous pieces.

4. DISCRETE-TIME APPROXIMATION: AN EXAMPLE

The above results guarantee that any finite difference (discrete-time) approximation that converges to the true solution of kinematic wave problems with fixed bottlenecks such as the "cell-transmission" (CT) model can also be successfully used with moving bottlenecks. All one has to do is choose a sufficiently small Δx to approximate the bottleneck trajectory, and an even smaller lattice spacing for the finite difference scheme. In practice, however, one can choose both lattice spacings to be equal with acceptable results.

As an illustration we now solve with the CT model a problem where a vehicle embedded in traffic suddenly slows down, and then compare the results with the exact solution of KW theory. We examine a one-mile homogeneous section of a freeway, whose F-function is an isosceles triangle with free flow speed, v_f = 60 mph, jam density, k_j = 300 veh/mi, and wave velocity in congestion, $w = -v_f$. Although the isosceles shape is not realistic, it is useful because the CT model does not introduce any numerical error for isosceles triangles if we choose $\Delta x/\Delta t = v_f$ (Daganzo, 1994). Thus, the discrepancies between exact and approximate solutions about to be presented can be exclusively attributed to the approximation method of this paper.

We assume that the freeway is flowing at capacity when, at t_0 = .3 min and x = .3 mi, a vehicle instantaneously slows down to a constant speed v = 20 mph and then resumes free-flow speed at t = 2.1 min. We assume that the moving restriction halves the capacity; i.e., Q = 4500 vph. Figure 7 shows the solution of this problem on the (t, x) plane. Capital letters indicate the states observed in the shown regions of the (t, x)-plane—"C" stands for freeway capacity and, as before, "D and U", for the states next to the active bottleneck. This part of the figure also shows a set of detectors at locations $\{x_i\}$, spaced 1/5 mi apart. The exact N-values observed at these detectors,

$N(t, x_i)$, will be used for the comparisons, using $N = 0$ for the last vehicle unaffected by the obstruction; see figure.

Four numerical solutions were obtained with the following meshes (Δt sec, Δx mi) = (12, 1/5); (6, 1/10); (3, 1/20) and (1, 1/60). In all cases $\Delta x/\Delta t = v_f$, to ensure that the CT results do not contain numerical errors. Figure 8 shows the four numerical solutions, overlaid over the exact solution. The black dots are the interfaces between traffic states in the exact solution. Note how the discrepancies disappear, as the mesh is refined and the very high resolution that can be achieved. Figure 9 shows the maximum discrepancy for each case by means of a dot, and uses a solid line for the bound of Theorem 2, for $P = 2$, which is amply met. The exact discrepancies are 22.50, 11.25, 5.63 and 1.88 vehicles for Δt = 12, 6, 3 and 1 sec, respectively.

For problems with only one queuing episode and no error in the initial data, such as ours, one can use directly the (tighter) bound of Lemma 2, $k_j \Delta x$. This is also shown in the Figure by means of a dashed line. For these cases, and if the F-function is triangular, it is also possible to derive even tighter bounds that use as an input additional bottleneck data, such as its minimum speed (Laval, 2003). This could be useful if one wished to estimate the exact error for a particular problem.

Note in closing that practical measures of performance for traffic streams are always defined over finite regions of the (t, x)-plane and that they can usually be evaluated if one knows the Moskowitz function in the region of interest. This is certainly the case for average flows, densities, vehicle-miles, vehicle-hours, and any other measures of vehicular accumulation, delay or travel time, as shown by the formulas in Sec. 4.1.1 of Daganzo (1997). Furthermore, practical measures of performance always turn out to be continuous functionals of the Moskowitz function; see for example equation (4.6) in Daganzo (1997). Therefore, an error as small as desired for a measure of

performance can always be obtained by reducing the error in the Moskowitz function; i.e., by reducing Δx and using the procedure proposed in this paper.

5. DISCUSSION

The results in this paper show that it is possible to simulate the effect of under-performing vehicles embedded in a traffic stream, and to estimate their effect in situations such as the beginning of uphill grades, where they may be creating bottlenecks. We believe that an improved quantitative understanding of "sag" bottlenecks and the most important factors that affect their capacities is now within reach. The results in this paper also allows us to model lane-changes on multi-lane freeways as randomly occurring moving bottlenecks in a single-pipe KW stream. The output of such a model should qualitatively match the Mauch and Cassidy (2002), since lane-changes in such a model tend to increase oscillation amplitudes without changing how they propagate—as observed.

It is also possible to extend the hybrid modeling ideas in this paper to multi-pipe models of traffic flow. This has the appeal that an exogenous random capacity does not have to be defined for each lane change, although one needs to formulate instead a model of inter-lane interaction. This possibility is explored in Laval (2003).

REFERENCES

Daganzo, C. F. (1994) “The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory" *Trans. Res.* **28B**, 269-287.

Daganzo, C. F. (1997) *Fundamental of Transportation and Traffic Operations*, Elsevier, New York, N.Y.

Daganzo, C. F. (2001) "A simple traffic analysis procedure" *Nets. Sptl. Econ.* **1**, 77-101.

Daganzo, C.F. (2003) *A theory of supply chains*, Springer, Heidelberg, Germany.

Gazis, D. C. and R. Herman (1992) “The moving and 'phantom' bottlenecks.” *Trans. Sci.* **26**, 223-229.

Koshi, M., M. Kuwahara and H. Akahane (1992) "Capacity of sags and tunnels in Japanese motorways", *ITE Journal*, May issue, pp. 17-22.

Laval, J.A. (2003) “Simulation of Moving Bottlenecks”, Ph.D. dissertation (in progress). Institute of Transportation Studies, Department of Civil and Environmental Engineering, University of California, Berkeley.

Lax, P.D. (1973) *Hyperbolic systems of conservation laws and the mathematical theory of shockwaves*, SIAM Regional Conference Series in Applied Mathematics, J.W. Arrowsmith Ltd., Bristol, U.K.

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.” *Proc. Roy. Soc. A* **229**, 281-345.

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

Mauch, M. and Cassidy, M.J. (2002) "Freeway Traffic Oscillations: Observations and Predictions." Proc. 15th ISTTT, Adelaide, Australia (M.A.P. Taylor, editor), pp. 653-674, Pergamon.

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 Res. Rec. 99, 43-44.

Muñoz, J.C. and C.F. Daganzo (2002). "Moving bottlenecks: A theory grounded on experimental observation", Proc. 15th ISTTT Adelaide, Australia, (M.A.P. Taylor, editor), pp.441-462, Pergamon.

Newell, G.F. (1998) "A moving bottleneck" Trans. Res. 32B, 531-537.

Richards P.I. (1956) "Shockwaves on the highway." Opns. Res., 4, 42-51.

LIST OF FIGURES

1. Exact and approximate moving bottleneck trajectories.

2. Possible states next to an active moving bottleneck.

3. Hypothetical problem with a curved boundary.

4. Typical solution of a piece-wise linear problem.

5. Relevant diagrams when ϕ and $\phi^{(a)}$ intersect at $t = \tau_p$.

6. Shifted solution when ϕ an $\phi^{(a)}$ do not intersect at $t = \tau_p$.

7. Exact (t, x) solution of a moving bottleneck example.

8. Numerical solutions for the example for different mesh sizes (in seconds): (a) $\Delta t = 12$;

(b) $\Delta t = 6$; (c) $\Delta t = 3$; (d) $\Delta t = 1$.

9. Actual errors for the example (dots) with the upper bounds arising from Theorem 2 (solid line) and Lemma 2 (dashed line).

FIGURES

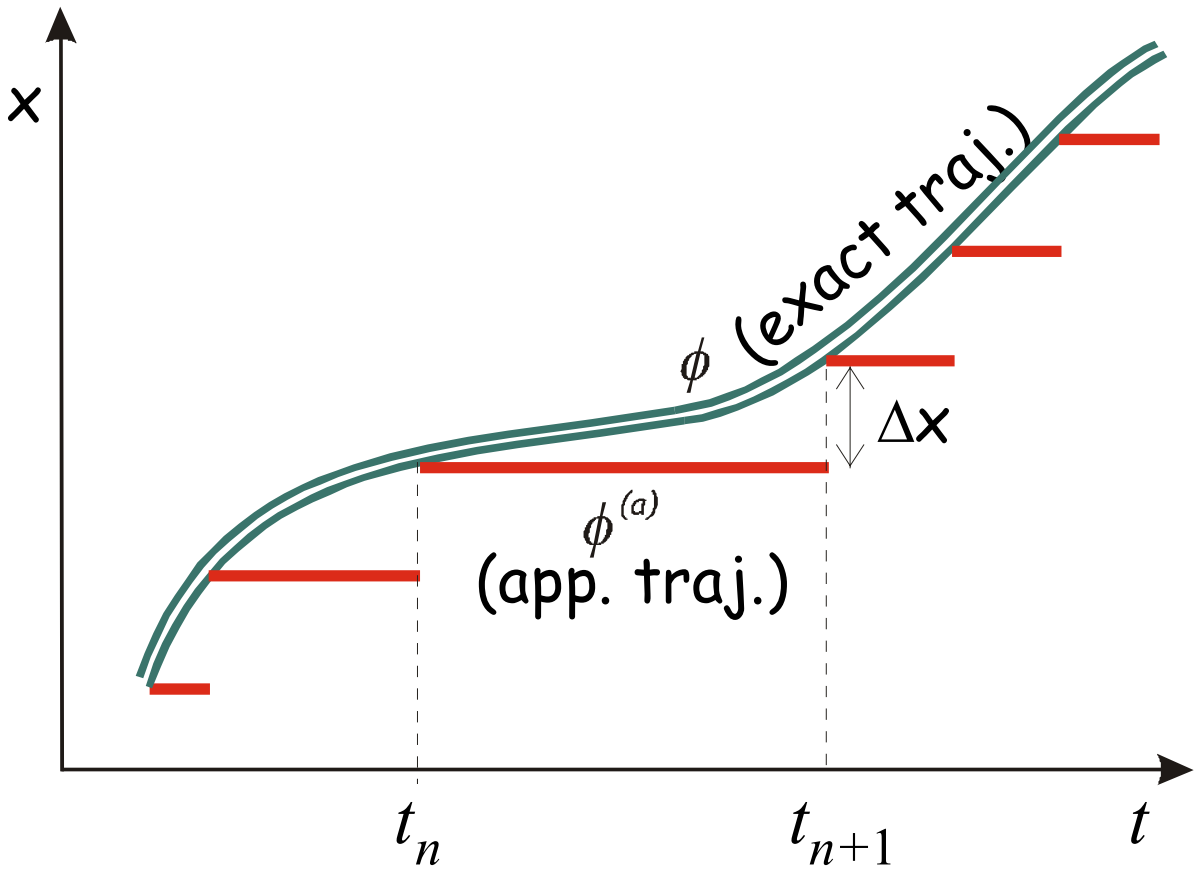

Figure 1: Exact and approximate moving bottleneck trajectories.

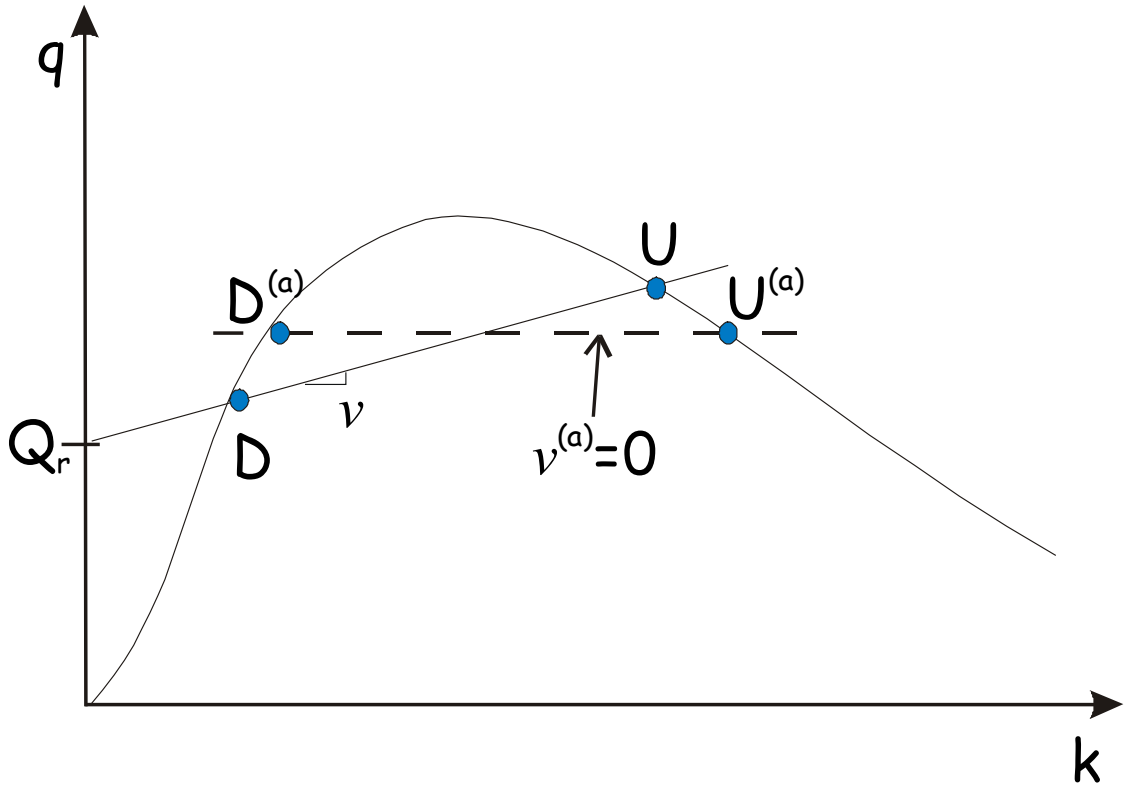

Figure 2: Possible states next to an active moving bottleneck.

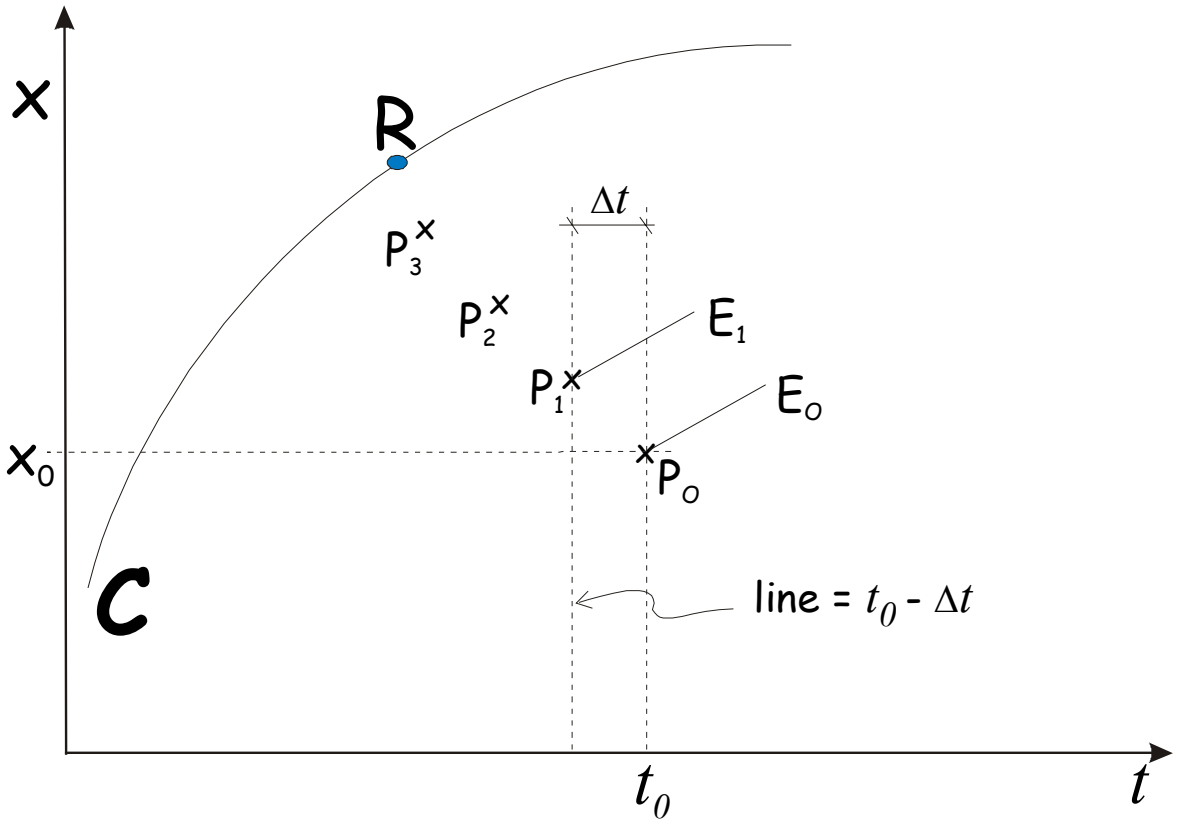

Figure 3: Hypothetical problem with a curved boundary.

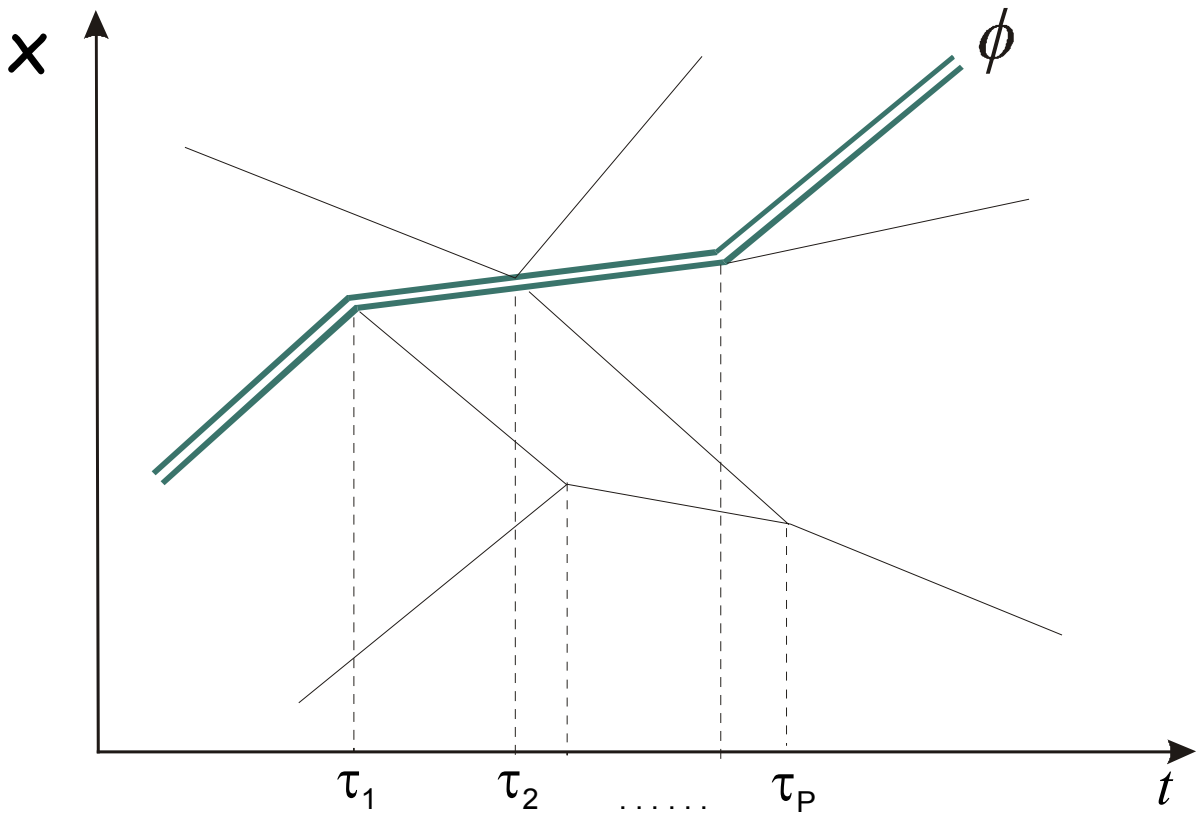

Figure 4: Typical solution of a piece-wise linear problem.

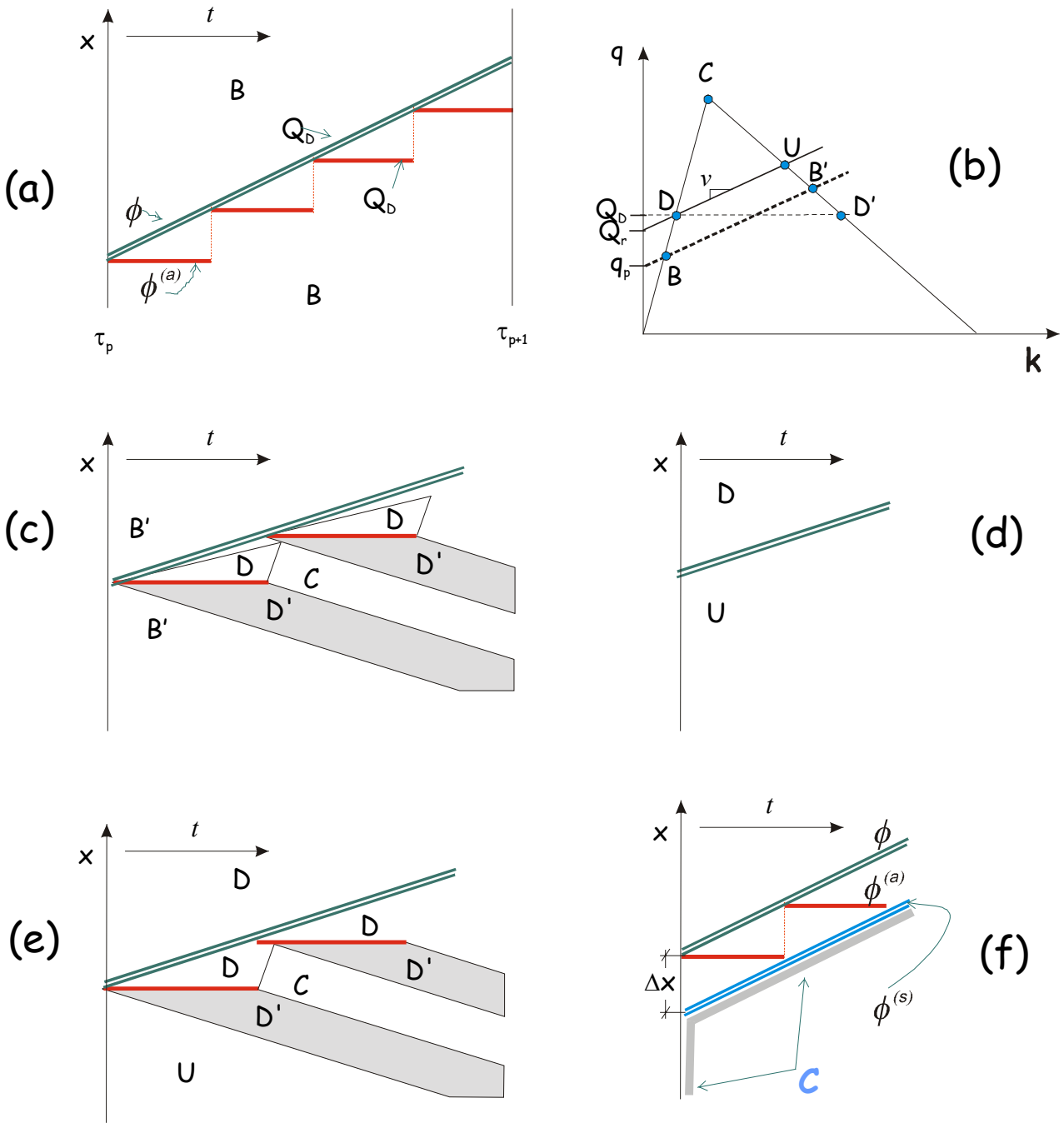

Figure 5: Relevant diagrams when ϕ and $\phi^{(a)}$ intersect at $t = \tau_p$.

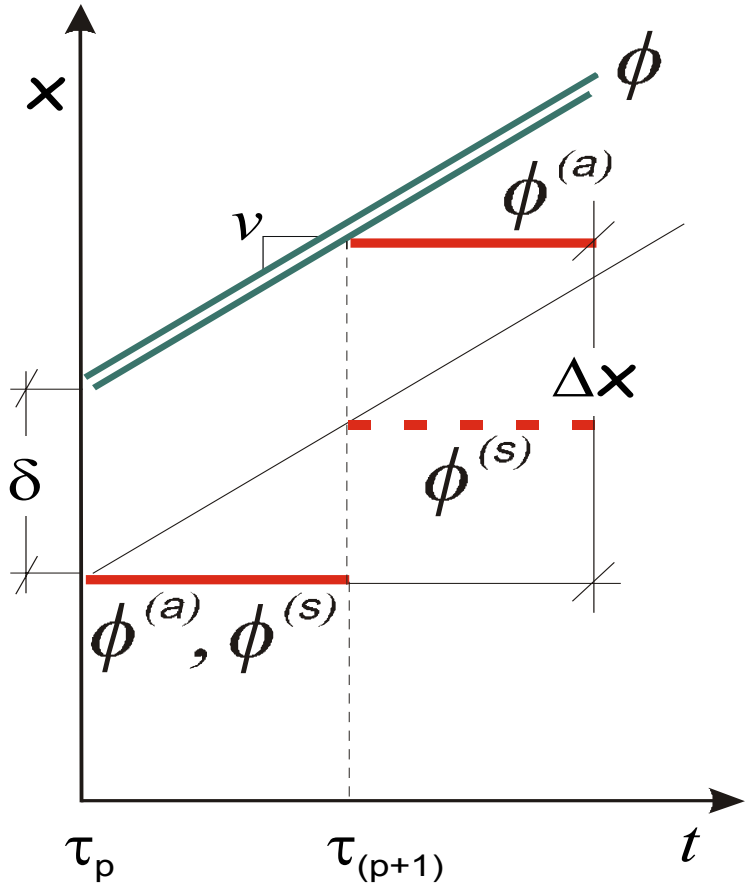

Figure 6: Shifted solution when ϕ an $\phi^{(a)}$ do not intersect at $t = \tau_p$.

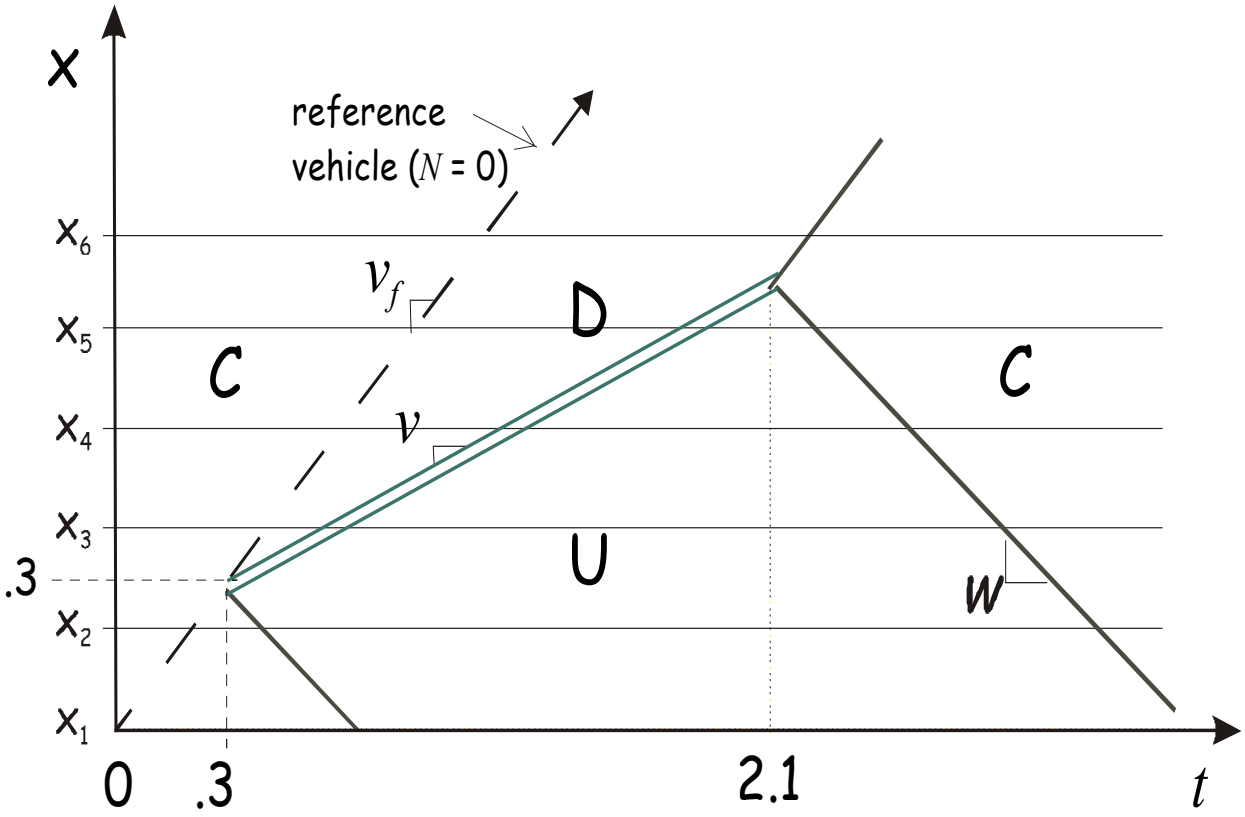

Figure 7: Exact (t, x) solution of a moving bottleneck example.

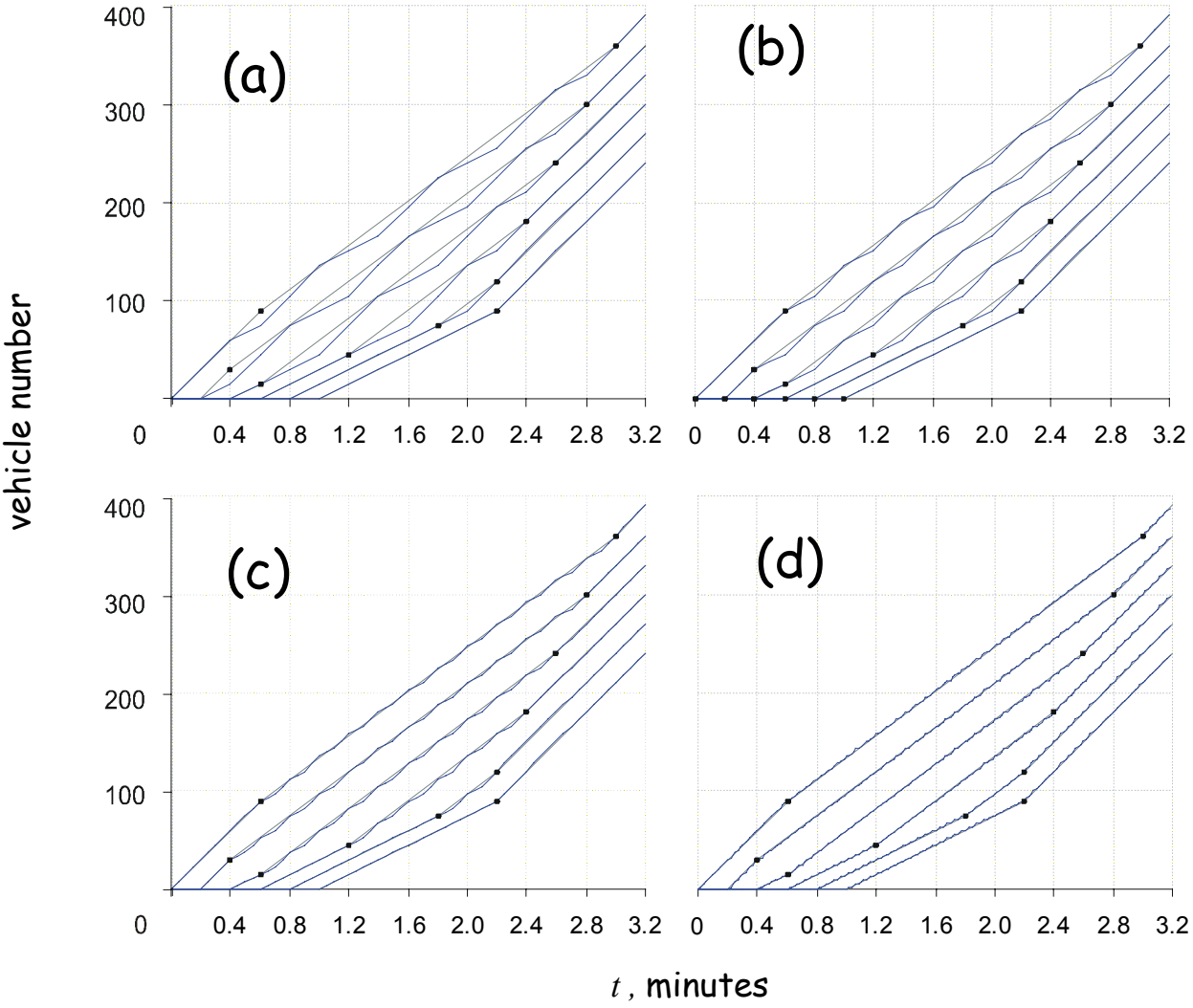

Figure 8: Numerical solutions for the example for different mesh sizes (in seconds):
(a) $\Delta t = 12$; (b) $\Delta t = 6$; (c) $\Delta t = 3$; (d) $\Delta t = 1$.

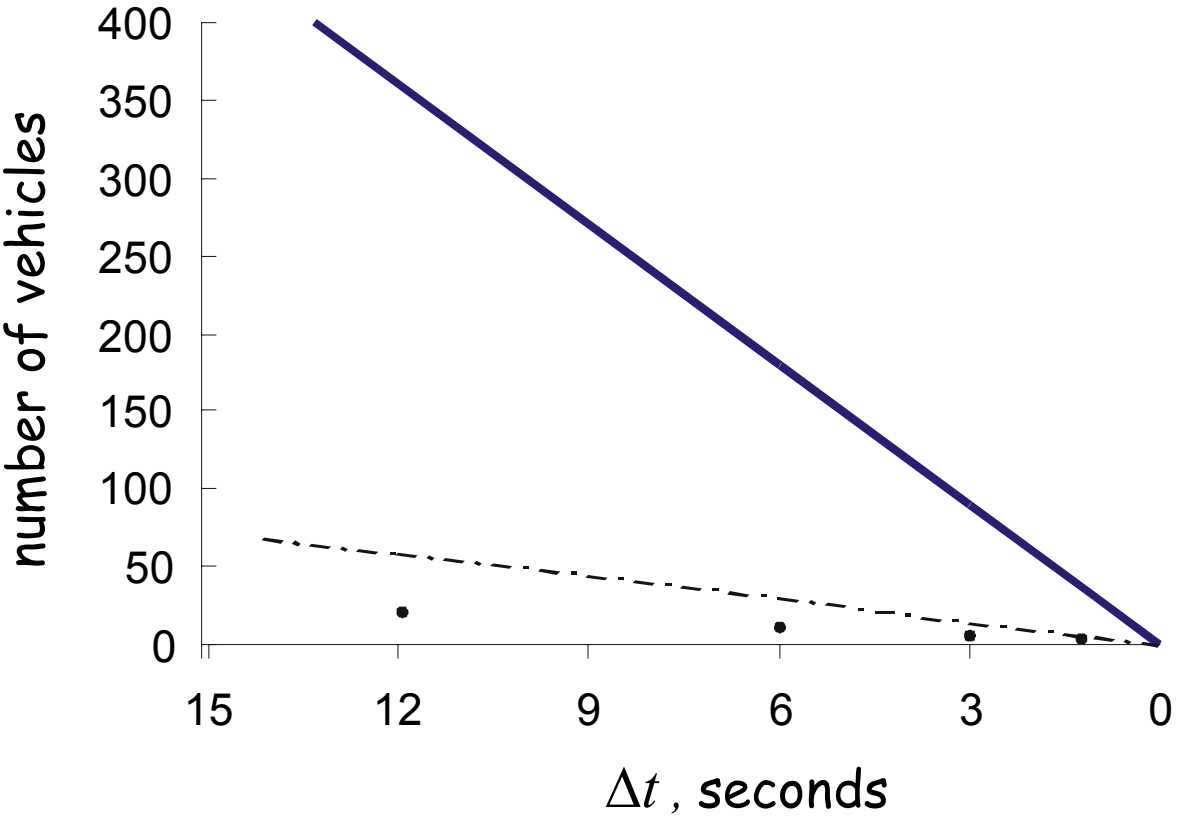

Figure 9: Actual errors for the example (dots) with the upper bounds arising from Theorem 2 (solid line) and Lemma 2 (dashed line).