Hybrid Models of Traffic Flow: Impacts of Bounded Vehicle Accelerations

by

Jorge Andrés Laval

Ingeniero Civil (Universidad Católica de Chile, Santiago) 1999
M.S. (University of California, Berkeley) 2001

A dissertation submitted in partial satisfaction of the requirements for the degree of Doctor of Philosophy in

Engineering-Civil and Environmental Engineering in the

GRADUATE DIVISION
of the
UNIVERSITY OF CALIFORNIA, BERKELEY

Committee in charge:
Professor Carlos F. Daganzo, Chair
Professor Michael J. Cassidy
Professor John Rice

Fall 2004
The dissertation of Jorge Andrés Laval is approved:

Chair

Date

Date

Date

University of California, Berkeley

Fall 2004
Hybrid Models of Traffic Flow:
Impacts of Bounded Vehicle Accelerations

Copyright 2004
by
Jorge Andrés Laval
Abstract

Hybrid Models of Traffic Flow:
Impacts of Bounded Vehicle Accelerations

by

Jorge Andrés Laval

Doctor of Philosophy in Engineering-Civil and Environmental Engineering

University of California, Berkeley

Professor Carlos F. Daganzo, Chair

This dissertation develops extensions to the kinematic wave theory of moving bottlenecks that allow the formulation of parsimonious traffic flow models that turn out to match field data with unprecedented accuracy. These ”hybrid models” explicitly incorporate the effects of a stream of underperforming vehicles, whose trajectories are computed endogenously using realistic accelerations. Hybrid models exhibit important practical advantages stemming from their simplicity and mathematical stability. They do not amplify errors, they require few easily observable parameters and they can be validated by parts.

Two practical applications are included: the effects of heavy vehicles on uphill bottlenecks, and the effects of lane-changing maneuvers. In the first case, we found
an important disagreement with the Highway Capacity Manual (HCM) formulas, possibly because the HCM formulas are based on microsimulation. Truck management schemes near bottlenecks caused by geometry can now be systematically devised; eg, truck-only lanes or truck metering. An approximate expression for the capacity of an upgrade is also developed, which matches simulation results accurately under all conditions.

The second application shows that the capacity losses imposed by the bounded accelerations of lane-changing vehicles is able to explain the relationship between the discharge rate of a moving bottleneck and its speed, as observed in the field. This is arguably the most important result of this dissertation, since it strongly suggests that lane-changing maneuvers may be the main cause for traffic instabilities on other bottleneck types (eg, merge, diverge, incidents).

The results of this dissertation allows for devising ITS solutions that maximize freeway capacity by controlling lane-changing and the arrival pattern of heavy vehicles to the bottleneck. The simulation of such strategies should be straightforward using hybrid models, and real-time applications should be possible for isolated bottlenecks.

Professor Carlos F. Daganzo
Dissertation Committee Chair
To Viviana and Sofía with all my love.
# Contents

List of Figures iv

List of Tables vi

1 Introduction 1

2 Background 5

2.1 The KW model 5

2.1.1 Fundamental Diagrams 7

2.1.2 Godunov’s method 9

2.2 N-curves 11

2.3 Theory of Moving Bottlenecks 12

2.3.1 Discretization of the KW-MB model 15

2.4 Extensions to the KW model 17

2.4.1 Second-order models 17

2.4.2 Single-pipe multi-class models 18

2.4.3 Multi-lane single-class models 19

2.4.4 Multi-lane multi-class models 20

2.4.5 Bounded accelerations 21

2.5 Lane changing 22

2.6 Free motion models 24

2.6.1 Crawl speed 26

3 Single-Pipe Formulation 28

3.1 Introduction 28

3.2 The CFM rule 30

3.3 Implementation 31

3.4 Numerical demonstration 32

3.4.1 Case A: truck is slowed by a queue 34

3.4.2 Case B: moving bottleneck joins a fast queue without slowing 35
3.4.3 Realistic accelerations ........................................... 36
3.5 Discussion ................................................................... 36

4 Multi-Pipe Formulation ................................................. 40
  4.1 Introduction ................................................................. 40
  4.2 The continuum formulation ......................................... 41
    4.2.1 Solution method ..................................................... 42
  4.3 The discrete formulation ............................................. 43
  4.4 Convergence ............................................................... 46
    4.4.1 The numerical experiment ...................................... 46
    4.4.2 Stability ............................................................... 48
    4.4.3 Consistency .......................................................... 48
  4.5 Moving bottlenecks in multi-lane models ...................... 49

5 Applications ................................................................. 54
  5.1 Upgrade bottleneck .................................................. 55
    5.1.1 Simulation ........................................................... 55
    5.1.2 Analytical solution ............................................... 57
    5.1.3 Comparison with HCM .......................................... 58
  5.2 Disruptive lane-changing (DLC) ................................... 59
    5.2.1 Formulation .......................................................... 60
    5.2.2 Flow downstream of a moving obstruction ............... 61

6 Conclusions and extensions ........................................... 68

Bibliography ................................................................. 71

A Complements for the multi-lane model ............................. 88
  A.1 The IT principle ......................................................... 88
  A.2 Examples of moving bottlenecks in the multi-lane model .......... 90
    A.2.1 A single moving bottleneck join a queue .................... 90
    A.2.2 Passing ............................................................. 91

B Analytical Solution of the Upgrade Bottleneck .................. 95
List of Figures

2.1 (a) Demand and supply functions; (b) triangular FD. Dots depict empirical data for (a) 30-sec and (b) 3-min intervals for loop-detector #7 on I-80 @ University Av. (Berkeley, CA) on 09/28/02 from 6 am to 9 pm, westbound. ......................................................... 8

2.2 (a) Numerical domain of dependence and stencil for Godunov’s method with triangular FD; (b) exact solution of a Riemann problem when \( k_{i+1} \geq k_i \) are congested. ....................................................... 10

2.3 (a) Representation of a moving bottleneck in a triangular FD; (b) solution of a simple example in the time-space plane. ................................................................. 13

2.4 (a) Approximate trajectories; (b) appearance of the solution. ................. 16

2.5 (a) Accelerations predicted by (2.16-2.17); (b) crawl speed predicted by (2.19). ................................................................. 26

3.1 Schematic representation of the hybrid model. ........................................... 29

3.2 \((k, q)\)-diagram showing Cases A and B. ........................................... 33

3.3 (a) \((t, x)\) diagram of the exact solution for example A; (b) numerical surface \((N, t, x)\) using \(\Delta t = 1\) sec. (c) \(k\)-map of numerical solution with \(\Delta t = 2\) sec; (d) with \(\Delta t = .5\) sec; (e) numerical \(N\)-curves along \(\phi\) for different mesh-sizes; (f) numerical \(N\)-curves at 5 locations. ................. 38

3.4 (a) \((t, x)\) density map of numerical solution for example B with \(\Delta t = .5\) sec; (b) numerical \(N_{\phi}\). ......................................................... 39

3.5 A realistic example of a heavy truck with \(v(0) = 5\) mph under the conditions of example A. ................................................................. 39

4.1 (a) Lane-changes in the \(t-x\) plane; (b) discretized freeway representation. 43

4.2 Schematic representations of the IT principle. ........................................ 45

4.3 (a) A 2x1 lane-drop; (b) exact solution in the FD; (c) numerical \(N\)-curves; (d) numerical \(k\)-map with \(\Delta t = 0.2\) sec. ........................................... 51

4.4 (a) Error propagation showing stability of \(M\); (b) total number of lane-changes as a function of \(\Delta t\). ......................................................... 52
4.5  $k$-maps of the lane-drop example for $\Delta t = 2$ sec (top), $\Delta t = 1$ sec (middle) and $\Delta t = 0.5$ sec (bottom). ........................................ 53

5.1  (a) $(t, x)$ density map of a sample simulation; (b) approximate trajectory for analytical solution. .................................................. 56

5.2  Capacity plots showing agreement between simulation and approximation. ............................................................... 64

5.3  Comparison of HCM values and eq. (5.2) for two typical cases. .... 65

5.4  (a) Recoverable flow gap; (b) disruptive lane-changing. ............ 65

5.5  $k$-maps for a moving obstruction traveling at (a) $v = 1$ mph and (b) $v = 20$ mph. .................................................. 66

5.6  Dimensionless bottleneck discharge rate, $\rho$, as a function of the slow vehicle speed $v$ for $n \in \{2, 3, 4\}$ lanes. ......................... 67

5.7  Location of the lane-changing maneuvers relative to the truck (simulation). ............................................................... 67

A-1  Example of §3.5 solved with the multi-lane method. ............... 91

A-2  Crossing between two moving bottlenecks. ........................... 94

B-1  $(t, x)$-diagrams for (a) possible trajectories; (b) approximate trajectory. .................................................. 96

B-2  (a)$(t, x)$ diagrams of a single truck on the uphill; (b) mean of the simplified process. .................................................. 98
List of Tables

2.1 Truck types assumed for analysis . . . . . . . . . . . . . . . . . . . . . 25
3.1 Traffic states for two examples: A and B . . . . . . . . . . . . . . . . 34
Acknowledgments

There is no doubt in my mind that the person who deserves all the credits is my wife Viviana Montenegro. Not only she provided me with the greatest love and support during these demanding years, but she also has been the greatest mom for our daughter Sofía.

I am also very grateful to Carlos Daganzo for everything he taught me during these years. Apart from giving me invaluable academic advice, inspiration and passion for discovery, his unconditional support, care and friendship made me feel at home even though we are thousands of miles away. I will spend my career trying to match the quality of his research, teaching and more importantly, his personal values.

I greatly benefited from the exceptional group of students I had the chance to meet at Berkeley, whose friendship and encouragement will be forever appreciated. I am particularly grateful to Alejandro Lago, Juan Carlos Muñoz, Souyong Ahn, Yu-Wei Li, Anne Goodchild, Yanfeng Ouyang and Monica Menéndez for their inspiration, interesting discussions and feedback they always provided.

Finally, I would like to thank my parents Jorge and Gloria, and my sisters Tota, Anita, Cacao and Mane for their permanent love, care and support, which made us always feel near “home”.

Chapter 1

Introduction

This dissertation studies the dynamics of traffic streams composed by fast and slow vehicles by extending the kinematic wave theory of moving bottlenecks (KW-MB) [1–4]. For decades we have known that heavy vehicles [5–7], roadway geometry [8–12] or lane-changing [13–15] can significantly affect traffic performance, but understanding the mechanisms of cause and effect is still an open question.

On the one hand, existing extensions of the kinematic wave (KW) theory [16, 17] that include multiple user classes [18–28] treat slow vehicles as a fluid rather than a moving obstruction and do not include realistic vehicle acceleration capabilities. The latter is surprising since “free-motion” models for the kinematics of an isolated vehicle are accurate and well understood [29–35]. Additionally, most of these extensions are “single-pipe” models (ie, no lane distinction is made). This severely reduces their comprehensiveness.
On the other hand, microscopic models—used by the vast majority of practitioners worldwide—have not been successfully validated as they are highly sensitive to the input parameters. Moreover, since the number of parameters can be quite large and many of them are not observable, we are left with untractable optimization problems that may admit multiple solutions. In light of the complexity of the task, the Federal Highway Administration recently launched the NGSIM initiative (next generation simulation), to provide a standard framework for developing microscopic models and a test-bed for their validation.

It is the author’s opinion that the main challenge faced by microsimulation is a consistent modeling of discretional lane-changing (ie, lane changes to increase travel speed) and other non-car-following interactions imposed by multi-lane traffic. These interactions tend to be unstable at the microscopic level, in the sense that individual lane-changing decisions may change dramatically with small variation on the initial conditions. At a more aggregate level, however, recent observation shows that lane-changing rates caused by a merge bottleneck are reproducible [36]. This implies that the phenomenon can be explained with simpler models.

In this context, this dissertation develops Hybrid Models that combine the robustness and stability properties of the KW model together with accurate microscopic free-motion models using simple physical principles. To this end, chapter 2 presents a brief background on non-hybrid models: (a) the KW and KW-MB models including numerical solution methods and (b) the free-motion models used in the numerical
demonstrations.

The formulation of the single-pipe hybrid model is developed in chapter 3 where simple numerical examples are included to show the convergence of the model. Although a number of problems can be addressed by this model (e.g., trucks on an upgrade or buses on city streets) the limitations imposed by the single-pipe framework preclude a parsimonious modeling of lane-changing.

In chapter 4 we show that using a multi-lane KW model allows for modeling lane-changing as a continuum with only one additional parameter, as the attributes that trigger lane-changing become endogenous variables (i.e., the speeds on different lanes.) This model is an improved formulation and discretization the model of Munjal and Pipes [37, 38]. A numerical demonstration of convergence is provided.

Chapter 5 includes two example applications, which summarize the main practical findings of this thesis:

1. The upgrade bottleneck application shows the effects of heavy vehicles using the single-pipe method. This example also includes the derivation of a simple expression for the capacity of the road segment. For the case of two vehicle classes, the formula accurately replicates simulation under all conditions. We found that these results differ considerably from the predictions of the Highway Capacity Manual (HCM) [39] based on microsimulation.

2. Lane-changing is modeled as moving bottlenecks using a discrete approximation of the lane-changing flows generated in chapter 4. The example of lane-changing
caused by a slow vehicle reveals an exceptional agreement with empirical data.

This experiment also suggests that the “capacity drop” often observed on merge bottlenecks could be explained by the same physical principle.

Finally, the main conclusions and extensions are included in chapter 6.
Chapter 2

Background

This chapter presents some background on KW and KW-MB models together with the finite-difference schemes used for their numerical solution. Current extensions of the KW model, plus a review on lane-changing and free-motion models are also included.

2.1 The KW model

The KW model is a scalar conservation law for the density, $k(t, x)$, of vehicles on a road at time $t$ and location $x$, supplemented with a fundamental diagram (FD) that gives the flow, $q(k, t, x)$, as a function of the local density $k(t, x)$. The time and space indexes will be omitted as much as possible. For a road without entrances or exists
the conservation law becomes:

$$\frac{\partial k}{\partial t} + \frac{\partial q(k)}{\partial x} = 0.$$  (2.1)

System (2.1) belongs to the family of first order nonlinear hyperbolic partial differential equations (PDEs). When \(q(\cdot)\) is concave, as often assumed in traffic flow, the solution of (2.1) is unique and stable only if an \textit{entropy condition} [40] is imposed; ie, choosing the solution that maximizes the flow [41].

To illustrate the predictions of the model, consider initial values given by

$$k(0, x) = \begin{cases} 
  k_u & \text{if } x < 0, \\
  k_d & \text{otherwise}.
\end{cases} \quad (2.2)$$

System (2.1)-(2.2) is called a “Riemann problem” in the theory of PDEs and constitutes the building-block for developing current numerical solution methods for the KW model. In general, Riemann problems are not trivial, but in the case of the KW model the solution is simple. It can be shown [41] that the flow at \(x = 0\), \(\tilde{q}(k_u, k_d)\), is given by the cell transmission (CT) rule [42]:

$$\tilde{q}(k_u, k_d) = \min\{\lambda(k_u), \mu(k_d), Q\}, \quad (2.3)$$

where \(\lambda(k, t, x)\) and \(\mu(k, t, x)\) are two monotonic functions that uniquely define the FD (see the solid lines in Fig. 2.1a, and \(Q\) is the capacity of the road at \(x = 0\). As a result of the entropy condition the solution depends on \(k_u\) and \(k_d\) as follows:

\[\text{The CT rule applies even for non-concave FD’s.}\]
• (deceleration waves) if \( k_u < k_d \) the discontinuity propagates at a speed \( s \) given by the “shock condition”:

\[
s = \frac{q(k_u) - q(k_d)}{k_u - k_d},
\]

(2.4)

• (acceleration waves or rarefaction fans) if \( k_u > k_d \) and \( q(k) \) is strictly concave, a gradual acceleration of vehicles which induces waves that spread, dissipating the discontinuity as time passes.

Notice that unlike deceleration waves, rarefaction fans have not been observed in the field.

2.1.1 Fundamental Diagrams

It is widely accepted that traffic states can be grouped in two phases or regimes: (i) free-flow regime, given by the increasing portion of the FD, where vehicles do not interact very much; (ii) congested or queued regime, given by the decreasing portion of the FD, where drivers are not able to travel at their desired speed because of vehicles ahead.

The appearance of FD’s obtained in the field depends on the aggregation interval and on the occurrence regime changes, as shown in [43, 44]. The dots in Fig. 2.1a depict a typical scatter plot of loop-detector flow-occupancy data for all lanes on consecutive 30-sec intervals. If 3-min intervals are used, the plot becomes less scattered as shown by the dots on part (b) of the figure. Note how well a triangular FD fits
Figure 2.1: (a) Demand and supply functions; (b) triangular FD. Dots depict empirical data for (a) 30-sec and (b) 3-min intervals for loop-detector # 7 on I-80 @ University Av. (Berkeley, CA) on 09/28/02 from 6 am to 9 pm, westbound.

A recent interpretation for traffic states was provided by Daganzo [42] in the context of numerical methods. He points out that in free-flow, $\lambda(k)$ represents the ability of a segment to send flow to the segment downstream, or the drivers desire to advance; in congestion, $\mu(k)$ gives the ability of a segment to receive flow from the upstream segment. Later, Lebacque [45] argues that the sending ability represents the demand for advancing to the downstream segment while the receiving ability is a simile for the supply for serving upstream demand.

In the continuum, the FD can be expressed as:

$$q(k) = \min\{\lambda(k), \mu(k), Q\},$$

(2.5)

where $Q(t, x)$ is the capacity of the road at $(t, x)$.

On all the numerical examples given in this dissertation we will use triangular
FD’s, which are defined with three observable parameters: free-flow speed $u$, congested wave speed $-w$ and jam density $\kappa$; see Fig. 2.1b. They are the simplest ones consistent with empirical data, which indicates that acceleration and deceleration waves travel at a nearly constant speed and without spreading [44, 46–59]. Notice that the KW model predicts spreading acceleration waves, unless one uses a triangular FD.

It will be convenient to define the equilibrium speed-density relation $V(k) \equiv (k)/k$.

### 2.1.2 Godunov’s method

During the last decade much has been learned about discretizing the KW model with finite difference methods [41, 42, 60–64]. In these methods, the highway is partitioned into small sections (of length $\Delta x$) and time into discrete time steps (of duration $\Delta t$.) The resulting numerical grid ($t_j = j\Delta t, x_i = i\Delta x$) is shown by the dots in Fig. 2.2.

Let $k^j_i$ be the numerical approximation of $k(j\Delta t, i\Delta x)$. For a method to converge to the right solution one must be able to express it in conservation form [40]:

$$k_i := k_i - \frac{\Delta t}{\Delta x} [\Phi(k_{i-p}, k_{i-p+1}, \ldots, k_{i+q}) - \Phi(k_{i-p-1}, k_{i-p}, \ldots, k_{i+q-1})]$$  \hspace{1cm} (2.6)

for some flux function $\Phi$ of $p + q + 1$ arguments. We use the update symbol “:=” to indicate that the terms on the RHS are evaluated on the previous time-interval w.r.t terms in the LFS. Otherwise, (2.6) would read: $k_i^{j+1} = k_i^j - \ldots$.
Figure 2.2: (a) Numerical domain of dependence and stencil for Godunov’s method with triangular FD; (b) exact solution of a Riemann problem when \( k_{i+1} \geq k_i \) are congested.

When \( \partial q / \partial k \) changes sign (such as in traffic flow,) Godunov’s method [65] is the best first order numerical scheme. This method provides the entropy solution by solving the Riemann problem between neighboring cells forward in time. Godunov’s method can be written as

\[
k_i := k_i - \frac{\Delta t}{\Delta x} [\tilde{q}(k_i, k_{i+1}) - \tilde{q}(k_{i-1}, k_i)],
\]

(2.7)

ie, it is a conservative method with \( p = 0, q = 1 \) and flux function (2.3).

The stencil of the method is the set of points in the discrete \((t, x)\) plane that are involved in the RHS of (2.6); see arrows in Fig. 2.2a for the Godunov stencil with triangular FD. The numerical domain of dependence of the scheme at \((j, i)\) is the set of grid-points that are connected to \((j, i)\) through a network of stencils; see bold dots in the figure. The domain of dependence of the PDE at \((j, i)\), namely \( \mathcal{D}(j, i) \), is the \((t, x)\) region that could affect the value of \( k^j_i \) according to the PDE; see shaded area
The Courant-Friedrich-Levy (CFL) stability condition states that the numerical domain of dependence must be a subset of $D(j, i)$. This is satisfied by choosing grids such that

$$\frac{\Delta x}{\Delta t} \geq \max_k |\frac{\partial q(k)}{\partial k}|,$$

which means that vehicles must spend at least one time-step inside a cell.

### 2.2 $N$-curves

Newell’s cumulative count curves, $N(t, x)$, give the cumulative number of vehicles that have crossed location $x$ by time $t$. The bivariate function $N(t, x)$ is (i) continuous, (ii) non-decreasing in time, (iii) non-increasing in space, and (iv) constant along a vehicle trajectory. If we let sub-indices represent partial derivatives then $q = N_t$ and $k = -N_x$, and the KW model expresses the existence of a fundamental diagram $F(\cdot)$ such that

$$N_t = F(-N_x, t, x).$$

In this context, the conservation equation becomes the identity $N_{xt} = N_{tx}$. Homogeneous problems with triangular FD can be easily solved exactly using Newell’s minimum principle [66]; inhomogeneous problems with concave FD’s is best solved with Daganzo’s variational method [67] with an error that is uniformly bounded. In
essence, these methods assert that the value of $N$ at any given $(t, x)$ point is given by the minimum of the $N$ values obtained from every point in the domain of dependence of (2.9).

2.3 Theory of Moving Bottlenecks

From an historical perspective, Gazis and Herman [68] were the first to propose a model to predict the impacts of a slow vehicle. They did so for a slow truck on a two-lane road, but their analysis is partial. The KW-MB theory was introduced by Newell [2, 66] for describing the effects of a long slow moving convoy on traffic streams. Lebacque et al [3] analyze the problem in a moving coordinate system and proposed the first numerical method for its solution. Muñoz and Daganzo [4] improve the formulation of the model based on empirical data and show the suitableness of the point-bottleneck approximation.

In the KW-MB theory, any vehicle that moves slower than the traffic stream can become an active moving bottleneck, causing a queue upstream and a free-flow state downstream, given by points $U$ and $D$ in Fig. 2.3a. The slope of the line connecting these points is the speed of the moving bottleneck, denoted $v$, and the speed of cars queueing behind the slow vehicle is $v_U$. The flow of state $D$, $Q_D$, is called the capacity of the moving bottleneck.

Notice that this theory assumes that moving bottleneck trajectories and capacities are defined exogenously; ie, that one knows $v(t)$ and $Q_D(t)$ beforehand. Since moving
bottlenecks in this thesis are part of the traffic stream, one must check that the problem remains well-posed; ie, that the values of $v(t)$ and $Q_D$ are consistent with traffic conditions in contact with the slow vehicle at all times. As noted in [4], only the bold segments in Fig. 2.3a constitute physically possible states. Other states imply some sort of collision: (i) states below $E$ imply $v \leq V(k)$; (ii) states above $DU$ imply that some proportion of the traffic stream will collide with the slow vehicle.

However, issue (ii) above is automatically handled by the theory, as states above $DU$ will immediately transition to either $D$ or $U$. Fig. 2.3b shows the solution on the time-space plane of a problem with a moving bottleneck that enters a freeway flowing at capacity, travels at a constant speed $v$ and leaves $T$ time units later.

Flow $R$ in Fig. 2.3a represents the maximum passing rate at which vehicles pass the moving bottleneck; ie, it gives the flow measured by an observer moving with the moving bottleneck when it is active. Flows $R$ and $Q_D$ are related by a linear
transformation, which for triangular FD’s is

\[ R = Q_D(1 - \frac{v}{u}). \]  

(2.10)

It was shown experimentally [4] that for constant \( v \), there exists a reproducible relation \( R = R(v) \). We can also express this dependency in terms of the capacity \( Q_D \), ie

\[ Q_D = Q_D(v), \quad \text{with} \quad \frac{dQ_D}{dv} > 0. \]  

(2.11)

This experimental evidence has not been explained yet. A psychological interpretation is that drivers are more “afraid” to pass slower moving bottlenecks, which is not very intuitive. Up to now, the only alternative is to determine \( Q_D(v) \) empirically. This implies a tremendous data-collection effort because one needs at least one experiment for each combination of number of lanes, type of geometry and speeds of the moving bottleneck.

In §5.2 we provided a possible explanation based on the bounded accelerations of vehicles passing the moving bottleneck. Because vehicle acceleration is not instantaneous the passing vehicle can cause a temporary obstruction during the maneuver. The slower the moving bottleneck the more severe the obstruction. Since the obstruction occurs on the passing lane only, we require a multi-lane version of the KW model, which is developed in chapter 4.
2.3.1 Discretization of the KW-MB model

The discretization of the KW-MB model has a short history. To the author's knowledge, the first attempt was suggested verbally in Lebacque et al [3]. It corresponds to an extension of Godunov scheme where the flow at the cell containing the (point) bottleneck is redefined. This method is of limited importance because it can only handle homogenous freeways and moving bottleneck trajectories that pass through lattice points.

Giorgi et al [69] uses the ideas in [3] but redefines the cells near the bottleneck so that the CFL stability condition is satisfied. It is not clear to the author that this method would work for arbitrary trajectories and crossings among moving bottlenecks. Additionally, inhomogeneous roads are not handled naturally by the method and requires a special procedure that imposes additional computations.

Later, Daganzo and Laval [70] replace the continuous trajectory of a moving bottleneck, $\phi(t)$, by an approximate step-trajectory, $\phi^a(t)$, restricted to the numerical grid; see Fig. 2.4a. In terms of vehicle number, the error introduced by this approximation is uniformly bounded. This bound tends to zero as the lattice spacing is reduced. This method allows the use of “off-the-shelf” software for solving general inhomogeneous problems involving crossings.

The only disadvantage of this method is shown in part (b) of the figure: the approximate flows, densities and speeds do not converge to the exact ones, even as vehicle counts do; these quantities “flip-flop” between states $C$ and $D'$, the latter
Figure 2.4: (a) Approximate trajectories; (b) appearance of the solution.

defined as the queued state at flow $Q_D$. Thus, to provide estimates one needs to
average their values over multiple cells, introducing a small first-order error.

To overcome this disadvantage, Daganzo and Laval [71] developed another method
that produces smooth flows. This is done by solving “composite Riemann problems”
–ie, two neighboring Riemann problems solved simultaneously– using an extension of
Newell’s minimum principle [66].

Finally, a variational method has recently been developed [67] which outperforms
all others. In this method any traffic problem (which could include moving bottle-
necks) can be solved with unprecedented accuracy and very quickly using dynamic
programming. In the case of triangular FD’s the method is exact. In other cases the
error is uniformly bounded. The method is not used in this dissertation because it
was developed concurrently.
2.4 Extensions to the KW model

Over the past several decades, a steady stream of extensions of the KW model have been proposed. They were motivated by the alleged inability of the model to explain the fluctuations from equilibrium values observed in the field, e.g., capacity drop [48, 49, 72–78], hysteresis [79], capacity of moving bottlenecks [4], propagation of oscillations [51, 80–85], clusters [86–89], persistent traffic jams [48, 90, 91] and fast waves [48].

Recently, it has been shown that KW can indeed explain most of them [43, 44, 53, 58, 82], while capacity drop, fast waves and the capacity of moving bottlenecks remain unexplained. However, Daganzo’s multi-lane models [21, 22] suggest that lane-changing could be at the root of capacity drop and fast waves, revealing that multi-lane models are fundamental for an adequate traffic representation.

Next, a small subset of relevant extensions to the KW model are described.

2.4.1 Second-order models

Lately, under the presumption that first-order models do not provide “interesting non-equilibrium results”, abundant higher-order models [92–97] have been proposed since the seminal works of Payne and Whitham [98, 99]. These models use (2.1) together with another PDE for describing the acceleration of vehicles. Second-order models have been heavily criticized in the literature [100–102]. The formulation of proper numerical solution methods and the modeling of boundary conditions [101,
is still an open question. Furthermore, they ignore the most likely cause of traffic instabilities, ie lane-changing.

2.4.2 Single-pipe multi-class models

Current single-pipe multi-class models [20, 23–27] can be expressed as the conservation equation for each vehicle class, $m$, together with a separate FD for each class, $F_m$:

\[
\frac{\partial k_m}{\partial t} + \frac{\partial q_m}{\partial x} = 0, \quad \text{and} \quad q_m = \alpha_m F_m\left(\frac{k_m}{\alpha_m}\right),
\]

(2.12)

(2.13)

where the $\alpha_m$’s are variables that should capture traffic interactions. The different models published in the literature differ in the assumption to compute the $\alpha_m$’s. Assuming equal spacing among vehicles leads to the models in [23, 27]; equal gap leads to [20]; gap proportional to vehicle length gives the free-flow model in [25]; equal speed gives the models for congested traffic in [25, 104]; finally, user optimum leads to [26].

The problem with these models is that they cannot address the interactions among different user classes with enough physical coherence, since (i) slow vehicles are treated as a fluid rather than moving obstructions, and thus lane-changing cannot be included explicitly; and (ii) the arrangement of vehicles across lanes is ignored (it is very different to have cars and trucks mixed or segregated in different lanes). In many
cases FD’s are chosen for numerical rather than physical reasons, which has led to numerous claims based on spurious foundations. 2

2.4.3 Multi-lane single-class models

The continuum multi-lane extension of the KW model was already devised in 1971 by Munjal and Pipes [37]. For an $n$-identical-lane road the model reads:

$$\frac{\partial k_\ell}{\partial t} + \frac{\partial q_\ell}{\partial x} = \Phi_\ell, \quad \ell = 1 \ldots n$$  (2.14)

where $k_\ell$ and $q_\ell$ give the density and flow on lane $\ell$. Lane-changing rates onto lane $\ell$, $\Phi_\ell$, are assumed given by a simplified form of Gazis et al’s model [80]:

$$\Phi_\ell = \alpha(k_\ell' - k_\ell), \quad \forall \ell, \ell' \text{ contiguous},$$  (2.15)

where $\alpha$ is a constant in units of time$^{-1}$. It is clear from (2.15) that lane-changing traffic has priority over the through traffic on the target lane because they are computed ignoring all possible competing flows.

Although this “priority problem” is somewhat unrealistic, it allows to simplify the formulation to a single equation that can be solved with stream-lined procedures when no regime changes are allowed [105]. To include regime changes, an unrealistic parabolic FD on each lane is assumed in that reference, leading to “numerical insights” on the steady-state conditions of the system that may be unrealistic. Munjal and Pipes’ model was first extended in [106], which added a source terms to include ramp

---

2For example, non-concave and different “congested branch” for each user class is used in [23] whereas empirical evidence supports one concave user-independent branch.
flows explicitly, defined $\alpha$ as a linear function of $(k_\ell - k_{\ell'})$ and incorporated the complete Gazis et al.'s model including a time-lag. Notably, the first discretization of the multi-lane model is presented, but the numerical scheme used at that time (Lax-Wendroff) is unable to treat properly boundary conditions in the KW model. This reference also develops a second-order and a two-dimensional formulation, but the results favor the simpler first-order model.

No other first-order macroscopic multi-lane models have been developed maybe because a multi-lane model alone cannot unveil new insights about traffic flow. Other approaches have been used too. The stochastic models in [28, 107] aim at describing the steady-state conditions rather than traffic dynamics. Gas-kinetic and second-order models are also available [108, 109] but suffer from the drawbacks mentioned in §2.4.1 and require several additional parameters.

The following section is devoted to Daganzo’s models [18, 19, 21, 22] that suggests that lane-specific formulations are valuable when different user classes are considered and/or driver psychology is taken into account.

### 2.4.4 Multi-lane multi-class models

The only publications under this category are Daganzo’s models [18, 19, 21, 22], devoted to solve two specific problems. For tractability, only two user types and two lane types are included, and traffic rules are very simple but physically grounded. In this simplified setting, the Riemann problems can be solved graphically.
• In [18, 19] only one user class is allowed to use all lanes while the other uses only a subset. Using the physical analogy of flows between reservoirs (the Incremental Transfer (IT) principle) they are able to streamline the solution of the Riemann problems.

• In [21, 22] fast and slow vehicles (rabbits and slugs) are considered. Slugs travel slower than rabbits and are not allowed to use the median lanes. A “reversed-\(\lambda\)-shaped” FD is used together with psychological considerations, in order to reproduce several puzzling phenomena, such as capacity drop, fast waves and hysteresis.

Note that in the “rabbits and slugs” model the effects of lane-changing are underestimated since it is assumed that rabbits can accelerate infinitely fast to the speed on the target lane. Chapter 4 explores a more parsimonious multi-lane model where lane-changing vehicles can be treated as temporary moving bottlenecks.

\subsection*{2.4.5 Bounded accelerations}

Lebacque [110] incorporates gravity into the KW model by including a single upper bound for vehicle acceleration. This model assumes only one vehicle type that share the same acceleration bound independent of the geometry. An undesirable feature of this formulation is that the scale-invariance of the numerical scheme is lost; ie, the results depend on the mesh size. Note that [69] successfully combined this approach with moving bottlenecks for the case of simple bottleneck trajectories.
Although this work does not provide a framework for modeling moving bottle-necks, it is the only publication aiming at including realistic accelerations into the KW model.

\subsection*{2.5 Lane changing}

When Lighthill and Whitham formulated the KW model, they did so for a single lane where lane-changing is not an issue. To extend the model to two-lane roads they proposed to double the flow for a given density per lane. This assumption is only reasonable if lane-changing does not cause an interruption on the target lane. It is also an optimistic assumption when comparing its predictions with the weaving-section formulas in the HCM.

The following are some recent examples that manifest the importance that lane-changing can have on traffic:

- Daganzo [21] uses lane-changing to explain the hysteresis phenomenon observed in [79], a long-lasting puzzle in the traffic community. He argues that (i) the lack of vehicle conservation (on the target lane) is the cause of hysteresis; and (ii) very high densities are caused by drivers on the target lane, who are motivated to travel closer to each other in order to impede other vehicles from “cutting in” in front.

- It was observed in [52] that prior to the activation of a merge bottleneck there
is a drop in the speed of passing lanes. It is hard to imagine a reason other than the first lane-changes from the shoulder lane to explain such a speed drop.

- The above is consistent with the merge bottleneck data in [36] where an increase in lane-changing counts occurs simultaneously with a lower bottleneck discharge rate. Interestingly, they observed a reproducible relation between the ramp-metering rate and lane-changing flows, and that high capacities can be restored by controlling the ramp-metering rate.

- Ways to control the spatial distribution of lane-changing near ramps were suggested in [111] in order to confine the friction caused by lane-changing to less congested areas.

- In [84, 112] lane-changing is mentioned as a candidate to explain the growth in the amplitude of stop-and-go waves propagating upstream.

Unfortunately, there are no continuum models that capture the reductions in capacity caused by the bounded accelerations of lane-changing vehicles. Rather, the literature is generous on microscopic models [13, 113–117], models based on neuronal networks [118, 119], steady-state stochastic models [107, 120], gas-kinetic models [108] and discrete choice models [121].
2.6 Free motion models

Free-motion models for the kinematics of an isolated vehicle are well understood [29–35]. They are based on Newtonian mechanics and capture the physical limitations imposed by roadway geometry on the engine for typical driver behavior.

The free motion model incorporated in TWOPAS [33] is used in this dissertation but any other can be used. It gives the maximum possible (or “desired”) acceleration of a vehicle, \( a(\cdot) \), as a function of its current speed \( v(t) \), and the following parameters:

1. vehicle characteristics: weight-to-power ratio, \( W \), and frontal area-to-power ratio, \( A \),
2. grade at every location, \( G(x) \), and altitude, \( h(x) \).

The acceleration for trucks, \( a \), is given by:

\[
a = \frac{\beta a_p}{\beta + 1.5 S(a_p)(a_p - a_c)} \quad \text{(in ft/s}^2,\text{)}
\]  

(2.16)

where \( S(\cdot) \) returns the sign of its argument; the rest of the parameters are defined as
follows:

\[
\beta = \begin{cases} 
0.4v & \text{if } v < 10, \\
10 & \text{otherwise,}
\end{cases} 
\]  

(2.17a)

\[
a_p = \frac{a_c + 15.368C_p/Wv}{1 + 14.080/Wv^2},
\]  

(2.17b)

\[
a_c = -.2445 - .0004v - .021C_d v^2/A - 222.6C_p/Wv - gG,
\]  

(2.17c)

\[
C_p = 1 - .00004h,
\]  

(2.17d)

\[
C_d = (1 - .00000688h)^{4.255},
\]  

(2.17e)

where \( h \) is expressed in ft, \( v \) in ft/s, \( W \) in lb/HP, \( A \) in ft\(^2\)/HP and \( G \) as a decimal; the acceleration of gravity is \( g = 32.17 \text{ ft/s}^2 \).

For the numerical analysis in this dissertation we chose the two truck types on the following Table:

<table>
<thead>
<tr>
<th>Veh Type</th>
<th>W (lb/HP)</th>
<th>A (lb/ft(^2))</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Heavy trucks</td>
<td>228</td>
<td>682</td>
</tr>
<tr>
<td>2. Light trucks</td>
<td>140</td>
<td>312</td>
</tr>
</tbody>
</table>

Table 2.1: Truck types assumed for analysis

Fig. 2.5a depicts the accelerations predicted by (2.16-2.17) for light and heavy trucks and \( G = 2 \% \).

For cars, the free motion model takes a linear form that only depends on the car’s maximum speed, \( v_{max} \), and zero-speed acceleration, \( a_0 \):

\[
a = a_0(1 - v/v_{max}) - gG.
\]  

(2.18)
The car type used in the numerical demonstrations of §5.2 is a high performance car, defined with $v_{max} = 142.7 \text{ ft/s}$ and $a_0 = 14.1 \text{ ft/s}^2$.

### 2.6.1 Crawl speed

The crawl speed $v_c$ is the steady state speed of a vehicle on an infinite upgrade, where the speed drops to a point where the engine lacks power to accelerate. For a particular vehicle type, it is solely a function of $G$.

To compute the crawl speed one can combine (2.17d) with (2.17c), set the latter equal to zero and solve for $v$. We found that the following polynomials give an accurate estimate of the crawl speed (in mph) for the relevant truck types for $0.01 \leq G \leq 0.09$:

$$v_c(G) = \begin{cases} -0.09G^3 + 2G^2 - 18G + 71, & \text{for heavy trucks}, \\ 0.44G^2 - 9.4G + 71, & \text{for light trucks}. \end{cases} \quad (2.19)$$

Fig. 2.5b depicts (2.19) for these truck types.
For the case of cars $v_c$ can be obtained analytically setting (2.18) equal to zero end solving for $v$. One gets

$$v_c = v_{\text{max}}(1 - \frac{g}{a_0}G), \text{ for cars.} \quad (2.20)$$

To get an idea, take as a rough approximation $g/a_0 \approx 2$ and $G = 4\%$ to see that on a typical freeway uphill, cars lose only 10\% of their speed.
Chapter 3

Single-Pipe Formulation

3.1 Introduction

This chapter presents the formulation of the single-pipe hybrid model, which extends the KW-MB by computing moving bottleneck trajectories endogenously and realistically. The causes that force a subset of vehicles to drive slower, e.g., gravitational forces imposed by roadway or safety factors may be explicitly incorporated.

In this chapter we assume that roadway geometry affects only a subset of vehicles (trucks,) whose “desired speed”, \( \hat{v} \), can be obtained from vehicle and road characteristics by means of a free-motion model. Fig. 3.1 shows a flow chart of the hybrid model. The model is Markovian since the state of each user class in the future can be obtained with conditions in the present. The state variables are the densities of the KW stream and the speeds and positions of slow vehicles. The constrained free-
motion (CFM) rule takes $\hat{v}$ and the speed of cars immediately downstream, $v^*$, and gives the minimum of the two, which will be the actual speed of trucks, $v$, during $[t, t + \Delta t)$. Knowing $v$ allows for (i) updating the state of moving bottlenecks and (ii) applying the KW-MB theory to update the state of cars, provided we know the capacity of the slow vehicle at that time, $Q_D(t)$.

We are confident that a model based on the figure will make robust predictions because its core component, the KW-MB theory, is macroscopic and stable in nature; and the CFM rule involves a minimum operation. Consequently, the hybrid system benefits from the practical advantages of stable macroscopic models: (i) input parameters are readily measurable in the field; (ii) rough input estimates yield reasonable predictions; and (iii) the model output will be valid if individual components have been validated.

Next, we formulate the method for computing trajectories endogenously. The convergence of the method is demonstrated by means of two numerical examples.

![Figure 3.1: Schematic representation of the hybrid model.](image-url)
3.2 The CFM rule

As mentioned in §2.3, one must check that traffic states in contact with the moving bottleneck remain on the bold segments in Fig. 2.3a. Traffic states on the remainder of the FD imply collisions: (i) states above line $DU$ imply car-truck collisions; and (ii) states below $E$ imply truck-car collisions. Item (i) is automatically handled by the KW-MB theory, as states above $DU$ will immediately transition to either $D$ or $U$. The purpose of the constrained free-motion (CFM) rule is to handle (ii) by accounting for the influence that traffic stream may have on slow vehicles.

As shown in §2.6 the free motion model gives the desired acceleration of a vehicle, $a(\cdot)$, as a function of its current speed $v(t)$, its weight-to-power ratio, $W$, its frontal area-to-power ratio, $A$, and the grade at the current location, $G(\phi(t))$. The desired speed of the vehicle, $\hat{v}(t)$, can therefore be defined in discrete time as

$$\hat{v} := v + a \Delta t,$$

(3.1)

The CFM rule simply takes the minimum between the desired speed predicted by the free motion model and the speed of cars immediately downstream of $\phi$, denoted $v^*(t)$; ie,

$$v = \min\{v^*, \hat{v}\}, \quad \text{and}$$

(3.2)

$$\phi := \phi + v \Delta t.$$

(3.3)

$^1$The altitude, $h(x)$, is held constant at 100 ft for simplicity.
Note that (3.2) implies instant deceleration capabilities. While this is unrealistic, it induces small errors since deceleration forces are 4 or 5 times greater than acceleration forces for a typical vehicle [110]. Additionally, deceleration maneuvers cannot affect the discharge rate of a bottleneck, as this maneuver takes place inside its queue.

### 3.3 Implementation

The only additional parameter needed for numerical implementation is the moving bottleneck capacity, $Q_D(t)$. In keeping with simplicity, we take $Q_D(t)$ as the capacity of the unblocked lanes, ie

$$Q_D = Q\frac{n-1}{n},$$

(3.4)

where $n$ is the number of lanes in the freeway. Although [4] observed that $Q_D$ increases with $v$, (3.4) can be used as a rough approximation.

From the available alternatives to carry out the KW-MB procedure [67, 69–71] we chose the method in [70] because of its simplicity and software availability. In this method we assume that $\phi(t)$ and $Q_D(t)$ are constant during the interval $[t, t + \Delta t)$. Thus, one simply defines a static bottleneck of capacity $Q_D(t)$ in a cell, $i$, near to the moving bottleneck, and we use

$$\hat{q}(k_i, k_{i+1}) = \min\{\lambda(k_i), \mu(k_{i+1}), Q_D\},$$

(3.5)

instead of (2.3)– but only in cell $i$. The density update (2.7) remains unchanged.
Recall that the KW-MB method in [70] does not produce smooth flows (see Fig. 2.4b) and therefore one has to average several cells in order to have a reliable estimate for $v^*(t)$. Formally, we use

$$v^* = V\left(\frac{1}{4} \sum_{m=i}^{i+4} k_m\right),$$

where $V(k)$ is the speed-density relation consistent with the FD in use.

The next section shows that the method converges to the KW-MB solution.

### 3.4 Numerical demonstration

Consider a level $n$-lane freeway segment of length 0.4 mi flowing at capacity state $C$ in Fig. 3.2 when, at $(t, x) = (0,0)$ a slow vehicle enters the segment. At the same time but at $x = 0.4$ an incident occurs for 30 sec imposing a capacity $Q_J$. Each lane obeys a triangular FD as in Fig. 3.2 with $u = 60$ mph, $w = -u$ mph and $\kappa = 150$ vpmpl. We use $w = -u$ because our KW numerical solver is then exact, and therefore the results isolate the errors introduced by the approximation in [70]. It is assumed that $\hat{v}$ is constant (=30 mph) in order to obtain the exact solution by hand.

To see how (3.2) provides the desired KW solution, we examine the $N$-curves along a truck trajectory, denoted $N_\phi(t) = N(t, \phi(t))$. As opposed to car trajectories –where the vehicle number $N$ is constant– $N_\phi$ increases whenever cars pass the truck; ie, when the moving bottleneck travels slower than traffic. The numerical estimation
of $N_\phi$ was carried out with the moving-observer formula [122]:

$$N_\phi := N_\phi + k_i (v^* - v) \Delta t,$$

(3.7)

where it is clear that the passing rate, $\dot{N}_\phi = \partial N_\phi / \partial t$, increases with ($v^* - v$).

In order to test the method on all possible conditions, we consider the following two cases; see Fig. 3.2:

- **Case A**: when the jam imposes a speed $v_J \leq \hat{v}$. The truck is forced to slowdown and travel at the same speed of cars inside the queue. Notice that in this case $\dot{N}_\phi = 0$.

- **Case B**: when $\hat{v} \leq v_J \leq v_U$. The truck is not forced to slow down but is embedded in a queue where cars can pass it. Note that now $\dot{N}_\phi > 0$. 

**Figure 3.2**: ($k, q$)-diagram showing Cases A and B.
3.4.1 Case A: truck is slowed by a queue

To illustrate this case we use an example with \( n = 2 \) lanes and \( v_J = 12 \) mph. The exact solution is shown in Fig. 3.3a. The moving bottleneck trajectory is the bold line on the figure while the resulting state variables are shown in the following table.

<table>
<thead>
<tr>
<th>State</th>
<th>Example A</th>
<th>Example B</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( q )</td>
<td>( k )</td>
</tr>
<tr>
<td>( C )</td>
<td>9,000</td>
<td>150</td>
</tr>
<tr>
<td>( J )</td>
<td>3,000</td>
<td>250</td>
</tr>
<tr>
<td>( D )</td>
<td>4,500</td>
<td>75</td>
</tr>
<tr>
<td>( D' )</td>
<td>4,500</td>
<td>225</td>
</tr>
<tr>
<td>( U )</td>
<td>7,500</td>
<td>175</td>
</tr>
<tr>
<td>( R )</td>
<td>2,225</td>
<td>0</td>
</tr>
</tbody>
</table>

Part (a) of the figure shows the exact \((t, x)\) solution obtained by hand. The bold line is the truck trajectory \( \phi \). Part (b) presents the \((N, t, x)\) surface using \( \Delta t = 1 \) sec. The bold line is the \( N_\phi \) curve.

Parts (c-d) present the numerical density maps for two values of \( \Delta t \). It is apparent how the solution converges to the exact solution of part (a) as the mesh is refined. Notice how the truck queue “flip-flops” between \( C \) and \( D' \), as expected from Fig. 2.4b. The reader can verify from part (c) that inside this queue there are two cells in state \( C \) for every cell in state \( D' \); thus the average flow is \((2Q + Q_{D'})/3 = 7,500 \) vph as in the exact solution.

Part (e) of the figure shows the numerical estimate of \( N_\phi \) for different mesh-sizes, where convergence is clear. Note that the three solutions only differ while the truck
is inside the queue, but are practically identical elsewhere. As expected, the passing rate is zero inside the queue and maximum, \( Q_R = 2,250 \) vph, when the moving bottleneck is active.

To complete the description, part (f) of the figure shows the \( N \)-curves \( N_s(t) \) measured at the five evenly-spaced locations \( x_s, s = 1 \ldots 5 \), shown as dotted lines on part (d). The vertical lines connecting the two figures conveniently show the changes in state for \( N_2 \).

### 3.4.2 Case B: moving bottleneck joins a fast queue without slowing

An illustration in this case uses \( n = 6 \) lanes and \( v_J = 35 \) mph. The resulting state variables are also shown on Table 3.1. The \( k \)-map of the numerical solution with \( \Delta t = .5 \) sec are shown in Fig. 3.4a. Given that \( \hat{v} \leq v_J \), in the exact solution \( \phi \) is a straight line, as illustrated by the solid line in the figure. It is apparent how the truck does not affect the jam, as predicted by the KW-MB theory. As expected, now the passing rate \( \dot{N}_\phi \) is positive inside the jam. According to the moving observer formula, \( \dot{N}_\phi \) inside the jam should be

\[
\dot{N}_\phi = q_J - k_J v = 19,895 - 568 \cdot 30 = 2,842 \text{ vph}
\]

as in part (b) of the figure.
3.4.3 Realistic accelerations

In real applications (such as the upgrade bottleneck in §5.1,) one would like to have $v^*$ given by a free-motion model. Fig. 3.5 shows such an example for a heavy truck with $v(0) = 5$ mph under the traffic conditions of example A. Notice that the area of the initial downstream state $D$ is larger than the one after exiting the queue. This happens because the truck exits the queue at $v_J = 12$ mph, which is greater than $v(0)$. This effect –only captured by hybrid models– will turn out to be important.

3.5 Discussion

Single-pipe hybrid models can be applied to any problem where the single-pipe assumption is reasonable; eg, trucks on an uphill or buses on a rural road. However, empirical evidence on congested off-ramps or diverge bottlenecks with $n > 1$ indicate that a multi-pipe approach is more realistic [59]. The multi-pipe approach for example can capture the following:

1. in real life not all lanes may be affected by the moving bottleneck. It is the author’s experience that in freeways of 4 or 5 lanes a moving bottleneck rarely affects more than 3 lanes. This may be because trucks usually use the right lane and vehicles exiting in the next off-ramp would rather queue behind the moving bottleneck than passing it,

2. discretionional lane-changing, where the driver tries to travel on the fastest neigh-
boring lane. In this case, one needs to know traffic conditions on different lanes, which are not endogenous variables in the KW model.

The above motivates the multi-pipe formulation in the next chapter.
Figure 3.3: (a) \((t, x)\) diagram of the exact solution for example A; (b) numerical surface \((N, t, x)\) using \(\Delta t = 1\) sec. (c) \(k\)-map of numerical solution with \(\Delta t = 2\) sec; (d) with \(\Delta t = 0.5\) sec; (e) numerical \(N\)-curves along \(\phi\) for different mesh-sizes; (f) numerical \(N\)-curves at 5 locations.
Figure 3.4: (a) \((t, x)\) density map of numerical solution for example B with \(\Delta t = .5\) sec; (b) numerical \(N_\phi\).

Figure 3.5: A realistic example of a heavy truck with \(v(0) = 5\) mph under the conditions of example A.
Chapter 4

Multi-Pipe Formulation

4.1 Introduction

This chapter presents a multi-lane KW model that enables representing multi-pipe traffic streams with hybrid models. Existing formulations of the multi-lane model [37, 38, 106] do not take advantage of the latest numerical methods [41, 42] for solving the continuum model, inducing unrealistic priorities among competing flows and inadequate treatment of the propagation of discontinuities.

To overcome these difficulties §4.2 and 4.3 present an extension of the CT rule [60] where competing flows are a result of supply-demand interactions. The proposed formulation requires only one additional parameter –the capacity for changing lanes– which is readily observable in the field. This formulation exhibits desirable convergence properties for the case of a lane-drop, as shown in §4.4. The final section shows
how moving bottlenecks can be incorporated, and simple examples that include interactions are presented in appendix A. Notice that the full potential of the resulting model is presented in chapter 5, where each individual lane-changing maneuver is treated as a moving obstruction.

4.2 The continuum formulation

Consider a freeway with $n$ identical lanes where identical vehicles flow. \(^1\) Let $k = [k_1, \ldots, k_n]$ be a vector composed of the density on each individual lane $k_\ell(t, x)$, $\ell = 1, 2 \ldots n$, and let $q_\ell(t, x)$ be the flow on lane $\ell$. We saw in §2.4.3 that the conservation equation for a single lane can be expressed as \([37]\):

$$\frac{\partial k_\ell}{\partial t} + \frac{\partial q_\ell}{\partial x} = \Phi_\ell, \quad \ell = 1 \ldots n,$$

where $\Phi_\ell(k, t, x)$ is the net lane-changing rate onto lane $\ell$, in units of vehicles per unit time per unit distance (the word “rate” will be used to indicate such units). Flow transfers occur between adjacent lanes only; ie, the set of movements that a driver faces at any time has, at most, three components: all movements $\ell \ell'$ such that $\ell' \in \{\ell - 1, \ell, \ell + 1\}$, if all adjacent lanes exist. \(^2\)

\(^1\)In this thesis driving on the right is assumed.
\(^2\)It is understood that any variable with sub-index $\ell \ell'$ gives a quantity measured in one direction; ie, from lane $\ell$ to lane $\ell'$. 
4.2.1 Solution method

Let $\Phi_{\ell \ell'}(k, t, x), \ell \neq \ell'$ be the lane-changing rate from lane $\ell$ to lane $\ell'$, so that

$$\Phi_{\ell} = \sum_{\ell' \neq \ell} \Phi_{\ell \ell'} - \Phi_{\ell \ell'}.$$  \hfill (4.2)

To understand these quantities physically consider the time-space region $A$ in Fig. 4.1a. A lane-changing is a trajectory that crosses only one boundary of $A$ (5 in the figure). The number of lane-changes onto lane $\ell$ (black dots in Fig. 4.1a) is given by

$$\sum_{\ell' \neq \ell} \int_{A} \Phi_{\ell \ell'} dt dx,$$

and corresponds to the number of trajectories that only cross the upper or right boundary of $A$ (2 in the figure); similarly, the lane-changes from lane $\ell$ only cross the lower or left boundary of $A$ (3 in the figure).

To solve (4.1)-(4.2) we assume that a “cell-transmission” rule balances people’s desires for changing lanes with the actual possibilities for doing so. Towards this end, let $L_{\ell \ell'}(k, t, x), \ell \neq \ell'$ be the desired lane-changing rate on $\ell \ell'$ (ie, demand for lane-changing), $T_{\ell}(k_{\ell}, t, x)$ the desired through flows on $\ell$, and $\mu_{\ell}(k_{\ell}, t, x)$ the receiving function on $\ell$. The competition for the available supply among the conflicting movement demands determines the actual lane-changing rates $\Phi_{\ell \ell'}$ and through flows $q_{\ell}$. This competition can be carried out for each lane independently since conflicting movements belong to adjacent lanes only. This procedure can be characterized by a vector-valued function of four arguments and three outputs, ie:

$$\mathcal{F}(L_{\ell-1, \ell, \ell}, T_{\ell}, L_{\ell+1, \ell, \ell}, \mu_{\ell}) \rightarrow (\Phi_{\ell-1, \ell, \ell}, q_{\ell}, \Phi_{\ell+1, \ell})$$  \hfill (4.3)
For stability, the function $F$ should satisfy an entropy-like condition that maximizes the flow. It should also reflect sensible priority rules, which depend upon the nature of the lane-changing maneuver (discretionary or mandatory). For the case of discretionary lane-changing we chose to extend the incremental transfer (IT) principle [19] as explained next.  

4.3 The discrete formulation

As in most discrete traffic flow models, we use the finite-difference method, where the highway is partitioned into small sections of length $\Delta x$ and time into discrete time steps of duration $\Delta t$, related by

$$\Delta x = \bar{u}\Delta t,$$

---

3Mandatory lane-changing is not analyzed in this dissertation as it requires the introduction of additional parameters; eg, how far away from an off-ramp will drivers be lined-up on the exit lane.
for stability, where $\tilde{u}$ is the fastest characteristic in (4.1). Additionally, we partition the freeway in cells $(i, \ell)$ as shown in Fig. 4.1b, where $i$ is the section along the roadway and $\ell$ is the lane index. The resulting numerical grid is denoted $(t_j = j\Delta t, x_i = i\Delta x, \ell)$.

For discrete variables we will use subscript $i$ and superscript $j$ to denote the value at the discrete point $(t_j, x_i)$; eg, $k_{i\ell}^j$ is the discrete approximation of $k_{\ell}(t_j, x_i)$. These indexes will be omitted as much as possible. In the discrete, the conservation equation becomes

$$\frac{k_{i\ell}^{j+1} - k_{i\ell}^j}{\Delta t} + \frac{q_{i\ell}^j - q_{i-1,\ell}^j}{\Delta x} = \sum_{\ell' \neq \ell} \Phi_{i-1,\ell'\ell}^j - \Phi_{i\ell'\ell}^j, \quad \forall \ell. \tag{4.4}$$

This equation it is ready for stepping through time, as there is only one term with time index $j + 1$. At each iteration one defines $L$, $T$ and $\mu$ for every cell using $\{k^j\}$, then computes the lane-changing rates and through flows with the IT principle [19]. In the IT principle the total supply $\mu_\ell$ is viewed as a reservoir being filled gradually with pipes coming from upstream. This is shown in Fig. 4.2, where each cell has one reservoir per demand type. Small reservoirs give the lane-changing demand while big ones, the through demand. The IT principle states that each reservoir will send fluid downstream until upstream reservoirs are empty or the receiving reservoir is full. Appendix A1 describes the algorithm.

Up to this point the formulation has been as general as possible, but now we further specify functional forms. Lane-changing demand rates are assumed to be
Markovian as they can be obtained from the current state of the freeway; i.e.,

$$L_{\ell \ell'} = \pi_{\ell \ell'} k_{\ell}, \quad \forall \ell, \forall \ell' \neq \ell,$$

(4.5)

where $\pi_{\ell \ell'} = \pi(k_{\ell}, k_{\ell'})$ is the fraction of drivers that would like to change from lane $\ell$ to lane $\ell'$ in the next unit of time. The Markovian assumption ensures that the probability for changing lanes in the next $\Delta t$ time units is $\pi_{\ell \ell'} \Delta t$, as $\Delta t \to 0$. It follows that the probability of staying in the same lane in the next $\Delta t$ time units is $1 - \sum_{\ell' \neq \ell} \pi_{\ell \ell'} \Delta t$. Therefore, the demand for through flow is given by

$$T_{\ell} = \lambda(k_{\ell})(1 - \sum_{\ell' \neq \ell} \pi_{\ell \ell'} \Delta t) \quad \forall \ell,$$

(4.6)

where $\lambda(\cdot)$ is the sending function for lane $\ell$.\footnote{The sending function for a lane should be roughly equal to the sending function of the whole segment divided by the number of lanes.} The term $\pi_{\ell \ell'} \Delta t$ can be interpreted as a choice function in discrete-choice modeling [123], provided that $\Delta t \to 0$. For simplicity, it is assumed proportional to the (positive) speed difference between lanes,
\[ \Delta v_{\ell\ell'}, \text{ ie} \]
\[ \pi_{\ell\ell'} \Delta t = \frac{\Delta v_{\ell\ell'} \Delta t}{\tau}, \quad \forall \ell, \forall \ell' \neq \ell. \quad (4.7) \]

The parameter \( \tau \) has units of time and should be comparable to the time to complete a lane-changing maneuver. Notice that \( \tau^{-1} \) is an upper bound to \( \pi_{\ell\ell'} \); as such, it can be interpreted as the capacity at \((x, \ell)\) for changing lanes, and should be readily observable in the field.

### 4.4 Convergence

In this section we examine the example of a lane-drop to show the convergence of the model. As it is customary in this dissertation, the fundamental diagram on each lane is assumed triangular with sending function \( \lambda(\cdot) \) and receiving functions \( \mu(\cdot) \), ie:

\[ \lambda(k_{\ell}) = uk_{\ell}, \quad (4.8a) \]
\[ \mu(k_{\ell}) = w(\kappa - k_{\ell}). \quad (4.8b) \]

As we did in chapter 3, the parameter values used in this section are: \( u = -w = 60 \) mph, \( \kappa = 150 \text{ vpmpl} \) and \( \tau = 3 \text{ sec} \).

#### 4.4.1 The numerical experiment

Consider a 0.4-mile, 2-lane freeway segment that has a lane-drop on lane 2 at \( x = 0.4 \text{ mi} \); see Fig. 4.3a. At \( t = 0 \) all lanes have optimal density \( k_0 = 75 \text{ vpmpl} \) and
the input demand is also $k_0$ for $t > 0$.

The numerical $N$-curves predicted by our method ($\Delta t = 0.2$ sec) for all lanes combined are shown in part (c), which are measured at the 5 evenly-spaced locations shown in part (a). This figure shows that this solution consists of an instantaneous transition between states $C$ and $B$; see part (b) of the figure. This solution coincides with the (single-pipe) KW solution and indicates that the multi-lane model does not affect the discharge rate of the bottleneck. ⁵

When we look at individual lanes separately, we observe reasonable patterns. Consider part (d) of the figure, which shows the $k$-maps of the numerical solution for each lane and for both lanes combined. Note how the density on lane 2 near the bottleneck is approx 140 vpm and decreases gradually upstream. The opposite happens on lane 1 with a density of $k_0$ at the bottleneck and increases gradually upstream. This behavior is reasonable, as common experience indicates that vehicles on the ending lane tend to drive slower the closer to the bottleneck they are.

To show the convergence of the model one must show that it is stable and consistent [40]. To this end, let $M$ be the discrete model operator, and $\mathcal{K}$ the continuum model operator presented in this chapter.

⁵This coincidence is a property of triangular FD’s and should not be true in general.
4.4.2 Stability

Let $N_0(x)$ be the initial $N$-profile used in the previous example; ie, the $N$-values along the road at $t = 0$, which in this case are given by $N_0(x) = 150x$. Also let a perturbed profile $\tilde{N}_0(x) = 152x$. If $M^j$ is the operator obtained after $j$ iterations of the numerical scheme (4.4-4.7), stability means that for all $j$:

$$\| N_0 - \tilde{N}_0 \| \leq \epsilon \Rightarrow \| M^j N_0 - M^j \tilde{N}_0 \| \leq \epsilon.$$  \hspace{1cm} (4.9)

for some norm $\| \cdot \|$. Condition (4.9) states that errors in the input data do not grow with time.

Fig. 4.4a shows the values of the errors $\| M^j N_0 - M^j \tilde{N}_0 \|$ for the $L_\infty$-norm, $\Delta t = 0.5$ sec and $t_j \leq 30$ sec. It is clear that (4.9) is satisfied since the errors decay as time passes.

4.4.3 Consistency

Let $M_{\Delta t}$ be the operator obtained after one iteration of $M$ with time-step $\Delta t$, and let $K_{\Delta t}$ be the exact continuum operator up to time $\Delta t$. The model is consistent if

$$\| M_{\Delta t} N_0 - K_{\Delta t} N_0 \| \rightarrow 0$$  \hspace{1cm} as $\Delta t \rightarrow 0.$ \hspace{1cm} (4.10)

Condition (4.10) indicates that the solution of $M$ tends to the continuum solution as the grid is refined.
Fig. 4.5 shows the $k$-maps of the lane-drop example for different mesh sizes. It becomes clear that the solution tends to Fig. 4.3d, obtained with $\Delta t = 0.2$ sec. Fig. 4.4b shows the total number of lane-changes (on the entire freeway during the whole simulation period) as a function of $\Delta t$, which also converges to approx. 17.5 as $\Delta t \to 0$.

4.5 Moving bottlenecks in multi-lane models

This section combines the single-pipe ideas for treating moving bottlenecks of chapter 3 with the multi-lane model presented in this chapter. It turns out that including moving bottlenecks in the multi-lane framework is more parsimonious because the physical problem is better represented. For example, we know that the flow-gap in front of the moving bottleneck originates only on the lane $\ell$ occupied by the moving bottleneck. Therefore, the problem reduces to a moving bottleneck on lane $\ell$ only. The capacity of the moving bottleneck is thus implicitly defined as that of the remaining lanes.

As we did on chapter 3, we approximate $\phi$ by a sequence of static bottlenecks; ie, if cell $(i + 1, \ell)$ contains the moving bottleneck we use

$$\Phi_{i\ell'\ell} = 0, q_{i\ell} = 0, \quad \forall \ell'$$

in lieu of (4.3). The CFM rule of §3.2 applies unchanged, with the caveat that downstream traffic speed is now lane-specific.
For illustration purposes, Appendix A2 shows two applications of this method. The first shows a moving bottleneck joining a queue, and the second a moving bottleneck passing another one. From these examples it can be seen that “pencil and paper” solutions to apparently simple problems can be extremely difficult to obtain, as lane-changing flows create different shock waves in different lanes and shocks tend to be curves rather than straight lines.
Figure 4.3: (a) A 2x1 lane-drop; (b) exact solution in the FD; (c) numerical $N$-curves; (d) numerical $k$-map with $\Delta t = 0.2$ sec.
Figure 4.4: (a) Error propagation showing stability of $\mathcal{M}$; (b) total number of lane-changes as a function of $\Delta t$. 
Figure 4.5: $k$-maps of the lane-drop example for $\Delta t = 2$ sec (top), $\Delta t = 1$ sec (middle) and $\Delta t= 0.5$ sec (bottom.)
Chapter 5

Applications

This chapter includes two example applications that show the possibilities of the models developed in chapters 3 and 4. The upgrade bottleneck application shows the effects of heavy vehicles using the single-pipe method. We found that our results differ considerably from the predictions of the HCM based on microsimulation.

In the second application lane-changing is modeled as moving bottlenecks using a discrete approximation of the lane-changing flows generated in chapter 4. The example of lane-changing caused by a slow vehicle reveals an exceptional agreement with empirical data. This experiment also suggests that the “capacity drop” often observed on merge bottlenecks could be explained by the same physical principle.
5.1 Upgrade bottleneck

In this section we use the single-pipe method to estimate the steady state flow on a of freeway with \( n \) identical lanes with an uphill in the middle section. Each lane obeys a triangular fundamental diagram with \( u = 60 \) mph, \( w = -15 \) mph and \( \kappa = 150 \) vpmpl so that \( Q = n \frac{wu}{w+u} \kappa \) is the capacity of the flat segments. The uphill starts at \( x_o = .5 \) mi for \( L \) mi and has \( G \) percent grade. There are two user types, cars and trucks, in proportions \( 1 - r \) and \( r \ll 1 \), respectively. We assume that cars and trucks share the same free-flow speed \( u \) on the flat segments. If not impeded by a queue, cars are able to maintain free-flow speed \( u \) on the uphill while trucks are forced to slow down due to gravity.

As we did on chapter 3, the capacity of the moving bottlenecks is taken as the capacity of the remaining unblocked lanes, ie

\[
Q_D = Q \frac{n-1}{n},
\]

which implies that trucks only use one lane. This is a reasonable assumption for the case of a single truck type and allows the derivation of closed-form steady-state expressions.

5.1.1 Simulation

We applied the hybrid model of chapter 3 for all the combinations of \( n = \{1, 2, 3\} \) lanes, \( L = \{.1, .4, .6, .9, 1.3\} \) mi, \( G = \{2, 4, 6\} \) % grade, \( r = 1, \ldots 25 \) % and the two
Figure 5.1: (a) $(t, x)$ density map of a sample simulation; (b) approximate trajectory for analytical solution.

truck types defined in §2.6. The capacity was obtained by dividing the number of vehicles crossing some location by the simulation period of two hours.

Fig. 5.1a shows the $k$-map of an extract of a sample simulation. Truck arrivals to the uphill are depicted as white circles in the figure. As expected, trucks change speed as they climb the uphill and reach the flat road again; notice how the second truck climbs slower than the first one because it reached the bottom of the hill at a slower speed; the third truck remains inactive inside the queue but activates when it becomes the head of the queue. Capacity losses are due to the traffic state downstream of the moving bottlenecks propagating forward and causing the voids in flow indicated in the figure.
5.1.2 Analytical solution

The approximate solution presented here assumes that all truck trajectories are identical and equal to the “approximate” trajectory in Fig. 5.1b. This trajectory has only two slopes, the free-flow speed $u$ and the truck’s crawl speed, $v_c(G)$. Appendix B shows that given $L$ and $G$, the dimensionless capacity of the upgrade $\rho(n, r, L, G)$ can be approximated by

\begin{align*}
\rho &= \frac{\rho_{\text{min}}}{1 - e^{-rQ_U T (1 - \rho_{\text{min}})}} , \quad (5.2a) \\
T &= L \frac{w + v_c}{w v_c} ; \quad (5.2b) \\
Q_U &= Q_D + \frac{w v_c}{w + v_c} \kappa ; \quad (5.2c)
\end{align*}

where $T$ is the duration of the queueing episode caused by a truck, $Q_U$ is the flow queueing behind a truck moving at $v_c$ and $\rho_{\text{min}} \approx Q_U / Q$.

Consideration of (5.2) shows that it has all the desired properties in the limit cases. In particular, as $r$ increases $\rho$ stabilizes at $\rho_{\text{min}}$, which can also be expressed as

\[ \rho_{\text{min}} = 1 - \frac{w(u - v_c)}{nu(w + v_c)}. \quad (5.3) \]

Note that $\rho_{\text{min}} > 0.7$ for the worst case scenario of heavy trucks $G = 6.5\%$ and $n = 1$ lanes.

It turns out that (5.2) matches simulation results very accurately under all conditions studied here; see for example bottom row in Fig. 5.2. In the entire data set the average absolute difference between the $\rho$ values obtained with simulation and
analytically is .01; see the top row in the figure. As expected, the bigger differences arose for short upgrade, but remain within a 5 %; see the middle row in the figure.

5.1.3 Comparison with HCM

The HCM 2000 [39] capacity formulas for upgrade sections are based on the microsimulation model FRESIM, developed by the Federal Highway Administration. Using the notation of this paper and neglecting recreational vehicles, the HCM capacity expression reads

\[ \rho = \frac{1}{1 + r(E_T - 1)}, \]

(5.4)

where \( E_T = E_T(L, G, r) \) is the equivalents of trucks, ie, the number of cars that produce the same capacity reduction than a truck. The values for \( E_T \) are tabulated for different values of \( r, L \) and \( G \) in exhibit 23-9 of the HCM for the average truck, which has a weight-to-power ratio “between 125 and 150 lb/hp”.

It was not possible to reconcile HCM values with those presented here. HCM recommends \( \rho \) curves that tend to zero as \( r \to 1 \) instead of stabilizing as in (5.3). Moreover, there are cases were \( \rho \) is increasing in some range of \( r \). This is shown in Fig. 5.3 for two typical cases, where the HCM curve is compared to (5.2) for heavy and average trucks. There may be several reason for the disagreement. First, the validation of microsimulation models is a tremendous task and nothing guaranties that, if conditions change, the output will be reasonable; second, the number of
lanes is not considered in (5.4); finally, the free motion models in FRESIM may need validation (as suggested in [124].)

5.2 Disruptive lane-changing (DLC)

In chapter 4 we presented a theory for predicting lane-changing flows on multi-lane traffic. Since lane-changing flows are continuous variables and accelerations are not bounded in this model, it fails to reflect what we believe is the major impact of lane-changing: the temporary blocking on a target lane by a slow-moving lane-changer.

When a lane is interrupted by a lane-changing vehicle coming from a slower lane, it creates a gap in flow in front of it that propagates forward in space on that lane only. This gap has the potential for propagating indefinitely, inducing a permanent loss in capacity. To see this, Fig. 5.4 shows a time-space diagram of a single lane flowing at a speed $v_d$ and interrupted by lane-changing vehicle with initial speed $v_a$. Two cases need to be considered: if $v_d$ is less than the free-flow speed $u$ then the moving bottleneck will be able to catch up with traffic and the flow gap will eventually be recovered, see part (a) of the figure; however, if $v_d = u$ it is clear from part (b) that the loss in flow is permanent. We call this a disruptive lane-change (DLC).

One could argue that these flow gaps may be filled with vehicles from neighboring lanes. However, when all the lanes are flowing at capacity, the gap cannot be filled without creating a similar gap on other lanes; ie, the loss in flow cannot be recovered. Therefore, the location of the maneuver should play an important role in determining
the amount of disruption. We expect maneuvers “near” the bottleneck to be more
likely to result in a DLC.

Here we combine the hybrid theory of chapter 3 with the multi-lane model of
chapter 4 into a model that can predict the effects of DLC without introducing any
new parameters. This theory can be used with bottlenecks of any nature, eg a lane-
drop, a slow truck or an on-ramp. This implies that our proposed physical principle
has the potential for explaining the discharge rate of several bottleneck types.

After the formulation of the method in the following section, we present an exam-
ple that accurately predicts the capacity (and other effects) of a moving bottleneck
caused by a slow truck.

\subsection{5.2.1 Formulation}

In this section we “quantize” the continuum lane-changing process of chapter 4.
To model each jump discretely let $\eta_{x\ell\ell'}(t)$ be the number of lane-changes occurred on
the freeway up to time $t$ from lane $\ell$ to lane $\ell'$ at location $x$, ie,

$$\eta_{x\ell\ell'}(t) = \int_0^t \Phi_{\ell\ell'}(s, x)ds. \tag{5.5}$$

where $\Phi_{\ell\ell'}(t, x)$ is the lane-changing rate of movement $\ell\ell'$ and $\eta_{x\ell\ell'}(t)$ is real. There are
several ways to quantize (5.5) with integer jumps but the most attractive one is using
an inhomogeneous Poisson process with rate $\Phi_{\ell\ell'}$ given by the multi-lane model of
chapter 4. \footnote{Notice that lane-changing flow transfers occur continuously, just as in chapter 4, not in discrete
jumps of one vehicle as it would appear more natural. However, implementing the latter procedure}

\footnote{Notice that lane-changing flow transfers occur continuously, just as in chapter 4, not in discrete
jumps of one vehicle as it would appear more natural. However, implementing the latter procedure}
number simulated is always discrete. The advantages of the Poisson approach are: (i) the lane-changing process is Markovian since drivers react to traffic conditions in the present, while the past seems irrelevant; (ii) there is no need to define additional parameters; and (iii) it is simple to implement. ²

The implementation of this idea is as follows. Each time that a lane-changing rate is positive, we perform a Bernoulli trial with probability $\Phi_{\ell_0} \Delta t \Delta x$. If the outcome is a success, we introduced a moving bottleneck on the target lane with initial speed $v_\ell$, and a vehicle type possibly drawn from typical vehicle mixes. The following numerical experiments assume that the lane-changing vehicles are equally fast cars.

5.2.2 Flow downstream of a moving obstruction

Consider again a long and flat homogeneous $n$-lane facility flowing at capacity. The four parameters that define the model are $u = 60$ mph, $w = -15$ mph, $\kappa = 150$ vpmpl and $\tau = 6$ sec. At $(t, x) = (0, 0)$ a slow truck enters the segment and travels at a constant speed $v \leq u$. Fig. 5.5 shows the $k$-maps for one simulation, two bottleneck speeds, $v = 1$ and 20 mph. Note from the white regions on the figure how passing maneuvers reduce the flow downstream of the moving obstruction.

It can be seen from Fig. 5.5a that the locations of the lane-changing maneuver (indicated as bold dots in the figure for all maneuvers) plays a key role. The closer to

²Notice that a deterministic method is less realistic because it is not memoryless.
the bottleneck the bigger the flow gap. This is in agreement with the intuition in the opening section, as lane-changes near the bottleneck tend to become DLCs. There is a critical distance from the bottleneck where lane-changing no longer has an effect. This is not observed in part (b) of the figure because most of the maneuvers take place very close to the bottleneck.

We repeated this numerical experiment for all \( v \in \{0, 5, 10, 20, \ldots, 50\} \) and \( n \in \{2, 3, 4\} \), and determined a normalized bottleneck discharge rate, \( \rho \), for each one of these cases and constructed the chart in Fig. 5.6. The normalization constant is the capacity of the unblocked lanes. Previous theories would predict \( \rho = 1 \), which neglected the dependency between capacity and truck speed. The empirical data collected in [4] and [39] are indicated in Fig. 5.6 as circles and a square, respectively, while simulation results are shown as curves. It can be seen that our model agrees with available observations quite accurately in the range \( v > 30 \) mph.

For \( 0 < v < 30 \) mph our predictions are somewhat intriguing as the system capacity slightly increases when the speed of the truck decreases. One would expect DLCs to be more severe as \( v \to 0 \) since \( v \) is also the initial speed of the lane-changing maneuvers close to the truck; let’s call this the “initial speed effect”. However, careful consideration shows that the maneuver’s spatial distribution is more extended as \( v \to 0 \) as noted in Fig. 5.5. In order to compare both parts of this figure with the same scale, Fig. 5.7 shows the location of the lane-changing maneuvers relative to the truck. It is clear how lane-changing maneuvers tend to occur farther from the truck.
as it travels slower. This implies that a smaller proportion of the maneuvers will be DLCs. We can conclude that this effect overcompensates the “initial speed effect”, allowing for higher discharge rates as \( v \to 0 \). Of course, this curious result is a direct consequence of the lane-choice model considered here (4.7) and field validation needs to be performed, but it clearly highlights the importance of lane-changing location relative to the bottleneck.

The empirical data for \( v = 0 \) (square in the figure) corresponds to an incident that blocks one lane on a two-lane facility as reported in [39]. As commented in [4] this low value could be caused by the “rubbernecking” of drivers passing through the incident. Since this effect is not included in our model it is natural to observe discrepancies. Interestingly, this discrepancy shows that the nature of the bottleneck (incident or lane-drop in this case) plays an important role because they induce different driver behavior.
Figure 5.2: Capacity plots showing agreement between simulation and approximation.
Figure 5.3: Comparison of HCM values and eq. (5.2) for two typical cases.

Figure 5.4: (a) Recoverable flow gap; (b) disruptive lane-changing.
Figure 5.5: $k$-maps for a moving obstruction traveling at (a) $v = 1$ mph and (b) $v = 20$ mph.
Figure 5.6: Dimensionless bottleneck discharge rate, $\rho$, as a function of the slow vehicle speed $v$ for $n \in \{2, 3, 4\}$ lanes.

Figure 5.7: Location of the lane-changing maneuvers relative to the truck (simulation.)
Chapter 6

Conclusions and extensions

This dissertation developed the formulation of hybrid models, a class of traffic flow models that combine macroscopic and microscopic models to represent heterogeneous traffic streams. These models merge together a continuum KW stream with models describing the acceleration capabilities of each individual slow vehicle. These models are stable and can be validated by parts, which enables fast and effective calibration/validation processes compared to the requirements of microscopic models.

Arguably, the major contribution is a theory based on disruptive lane-changing (DLC) that has the potential for explaining the discharge rate of the most common bottleneck types. The validity of this theory was demonstrated with the multi-lane hybrid model of §4 that requires only one additional parameter. We found that the loss in capacity imposed by the bounded accelerations of lane-changing vehicles is able to explain the relationship between the discharge rate of a moving bottleneck
and its speed, as observed empirically in [4]. Furthermore, we found that the spatial distribution of lane-changes plays a key role, particularly for low moving bottleneck speeds.

Another contribution is the derivation of a closed-form expression (5.2) for the steady-state capacity of upgrade bottlenecks, as a function of the length and grade of the uphill and the percentage of trucks. These formulas match simulation results with surprising accuracy. We found an important disagreement with the HCM formulas, especially under large truck proportion. A possible explanation may be that the HCM formulas are based on microsimulation, suggesting a revision of the manual’s methodology. Truck management schemes near bottlenecks caused by geometry can now be systematically devised; eg, truck-only lanes, truck metering or truck bunching. More generally, formula (5.2) –with suitable parameter values– applies to any moving bottleneck problem provided that they constitute a time-independent proportion of the traffic stream. Examples include freeway segments with general geometric profiles and the effects of buses, turning vehicles or bicycles on city streets.

It is natural to conjecture that DLC theory may explain the “capacity drop” phenomenon observed on merge bottlenecks; ie, as soon as a queue forms on the freeway a lower bottleneck discharge rate is observed. This would imply that effective ITS solutions can be designed with the ability of maximizing system capacity by controlling the variables that induce the drop in capacity. Research is needed to design strategies able to (i) control the spatial and temporal occurrence of lane-
changing maneuvers (eg, variable message signs upstream of the merge,) and (ii) control the arrival pattern of heavy vehicles to the bottlenecks.

The framework proposed in this dissertation can be extended in a number of directions. Of particular interest is the modeling of congested urban networks since the boundary conditions imposed by roundabouts and signalized/unsignalized intersections can be handled very naturally with multi-lane modeling. In such context, moving bottlenecks can be used to represent buses, trucks, bicycles, lane-changing vehicles, turning vehicles or snow plows.

A major practical feature that remains to be explored is using the “variational principle” [67] as the exact numerical solution method for the models presented here. The main challenge is extending the variational principle to the multi-lane case, in a way consistent with the IT principle. This would make possible the efficient and realistic simulation of large transportation networks for the first time, opening the door for real-time simulation for assisting system-wide control strategies.


[66] G F Newell. A simplified theory of kinematic waves in highway traffic, i general


[82] MJ Cassidy and M Mauch. An observed traffic pattern in long freeway queues. 


Appendix A

Complements for the multi-lane model

A.1 The IT principle

This appendix formalizes the extension of the IT principle described in §4.2, which was characterized by a vector-valued function of four arguments and three outputs, i.e:

\[ F(L_{\ell-1,\ell}, T_{\ell}, L_{\ell+1,\ell}, \mu_{\ell}) \rightarrow (\Phi_{\ell-1,\ell}, q_{\ell}, \Phi_{\ell+1,\ell}). \]  

(A-1)

In order to simplify the exposition we redefine the inputs as \( \lambda_{\ell\ell'} \) given by

\[
\lambda_{\ell\ell'} = \begin{cases} 
L_{\ell\ell'} \Delta x & \text{if } \ell \neq \ell', \\
T_{\ell} & \text{if } \ell = \ell'. 
\end{cases}, \forall \ell, \ell'. 
\]  

(A-2)
In this way, all the inputs have the same units (flow). It is also convenient to define

$$\alpha_{\ell \ell'} = \frac{\lambda_{\ell \ell'}}{\sum_n \lambda_{n\ell'}}, \quad (A-3)$$

as the proportion of the total demand coming from movement $\ell \ell'$. In this way, the problem is to split the total supply $\mu_\ell = \mu(k_\ell)$ to satisfy upstream demands. As mentioned in §4.2, the procedure must satisfy two conditions: (i) be a reasonable approximation of traffic behavior, and (ii) maximize total flow at every location. In the case of the IT principle, this problem can be carried out with the following computational procedure.

**Computational procedure**

Given are $\lambda_{\ell-1, \ell}$, $\lambda_{\ell \ell}$, $\lambda_{\ell+1, \ell}$ and $\mu_\ell$. For each $\ell$ do:

1. **Step 1.** Compute auxiliary flows $f_{\ell' \ell} = \min\{\lambda_{\ell' \ell}, \alpha_{\ell' \ell} \mu_\ell, Q\}, \forall \ell'$

2. **Step 2.** Set $\mu_\ell := \mu_\ell - \sum_{\ell'} f_{\ell' \ell}$. If $\mu_\ell = 0$ stop, else

3. **Step 3.** Set $\lambda_{\ell' \ell} := \lambda_{\ell' \ell} - f_{\ell' \ell}, \forall \ell'$. If $\forall \ell': \lambda_{\ell' \ell} = 0$, stop, else reset $\alpha_{\ell' \ell}$ with (A-3) and go to Step 1.$\diamond$

At each iteration, $m$, this algorithm allocates as much supply as possible to each movement subject to the maximum allocation given by the demand proportion. However, if $\mu_\ell > \sum_n \lambda_{n\ell}$ then there will always be a movement such that

$$\lambda_{\ell' \ell} \leq \alpha_{\ell' \ell} \mu_\ell,$$

for some movement $\ell \ell'$, \quad \((A-4)\)

and some supply will be lost if another movement could have used it. This would violate the flow-maximizing character of the entropy condition, possibly yielding an
unstable model. To overcome this limitation, the algorithm will iterate until the remaining supply is split between the other movements, as sought. Notice that this algorithm converges in at most \( m = 3 \) iterations. Finally, the outputs are obtained using

\[
q_\ell = \sum_m f_{m\ell},
\]

and,

\[
\Phi_{\ell'\ell} = \sum_m f_{m\ell'}/\Delta x, \quad \ell \neq \ell.
\]

(A-5) \hspace{1cm} (A-6)

A.2 Examples of moving bottlenecks in the multi-lane model

This appendix shows two applications of moving bottlenecks in the multi-lane model. The first shows a moving bottleneck joining a queue, and the second shows one moving bottleneck passing another one.

A.2.1 A single moving bottleneck join a queue

Fig. A-1 shows the example on §3.5—a heavy truck with \( v(0) = 5 \) mph under the conditions of Case A on that section. Although the figure is similar to the one-pipe solution of Fig. 3.5, there are some discrepancies on the back-of-the-queue (BOQ) of the jam \( J \). This is expected since the jam propagates differently on each lane. In
fact, the average across lanes could never display a sharp shock (as in Fig. 3.5) unless both lanes behave equally; ie, when the flow downstream of the moving bottleneck is the same on both lanes. Consequently, the truck trajectory is slightly different from the single-pipe one as the truck joins the BOQ at a different time.

A.2.2 Passing

It is expected that when slow vehicles are close enough interactions may occur. Although complex interactions could be incorporated within the proposed framework, this is not our goal because the stability of the overall model could be jeopardized. Instead, we present the simplest case of a slow vehicle that changes lanes to overtake an even slower one.

Crossings among point moving bottlenecks are naturally handled by the multi-lane model. All one needs to do is to make sure that the two moving bottlenecks are on different lanes while the crossing takes place. For simplicity, we assume that

**Figure A-1:** Example of §3.5 solved with the multi-lane method.
when a moving bottleneck wants to change lanes, it can do so immediately, so that the problem reduces to using the appropriate “$\ell$” index in (4.11).

As an example, consider the same 2-lane freeway with initial and boundary data $k_1= 60 \text{ vpmpl}$, which is lower than the optimal density of 75 vpmpl. We include two active moving bottlenecks on lane 1, $M_{B_r}, r = 1, 2$ with constant desired speed $v^*_r$ and infinite acceleration capabilities. Initially, $M_{B_2}$ is downstream of $M_{B_1}$. The latter travels at speed $v_1 < v^*_2$, which could be lower than $v^*_1$ if it is queueing behind $M_{B_2}$; see top-left diagram in Fig. A-2. We assume that $M_{B_1}$ will start the lane-changing as soon as the gap between the two bottlenecks reaches some critical value, $d$, ie, when

$$|\phi_2 - \phi_1| = d.$$  \hspace{1cm} (A-7)

At this point $M_{B_1}$ will move to lane 2 and begin passing, which is denoted by the thin dotted line in the figure. After the passing is completed, $M_{B_1}$ may (or may not, as in the figure) return to the original lane. The remaining parts of Fig. A-2 present the numerical $k$-maps obtained with such a simple crossing model when $v^*_1 = 40 \text{ mph}$, $v^*_2 = 5 \text{ mph}$, $d = 0.05 \text{ mi}$ and $\Delta t = 0.1 \text{ sec}$. Several interesting properties of the solution can observed:

- since the discharge rate of $M_{B_1}$ is greater than the initial density of 60 vpmpl on lane 2, the BOQ of $M_{B_2}$ begins to grow faster on this lane and recedes on lane 1; see points “1” and “2” in the figure;

- this causes lane-changing from lane 2 to lane 1, inducing the curved shocks
shown in the figure;

- as soon as MB₁ reaches the BOQ, it decelerates to the speed inside the queue and becomes inactive until it changes lanes; see segment “2-3” in the figure;

- after changing lanes, MB₁ remains inactive until passing MB₂. After this point some lane-changing occurs to lane 1.

Finally, this example show that “pencil and paper” solutions to apparently simple problems can be extremely difficult to obtain, mainly because of the lane-changing flow, which create different shock waves in different lanes and shocks tend to be curves rather than straight lines.
Figure A-2: Crossing between two moving bottlenecks.
Appendix B

Analytical Solution of the Upgrade Bottleneck

This appendix presents the analytical solution for the capacity of upgrade bottlenecks under the presence of only two vehicle classes, cars and trucks as described in §5.1. The complexity of the problem resides in the infinite number of possible truck trajectories that can arise on a given segment, even though a single truck type is assumed. This is because the speed of trucks at the bottom of the upgrade determines the actual trajectory, and can vary from free flow-speed $u$ to the crawl speed of trucks, $v_c(G)$. It follows that the set of truck trajectories is bounded by the free-flow and crawling trajectories as shown in Fig. B-1a.

Truck trajectories are supposed identical and equal to an “approximate” trajectory, ie, one in between the free-flow and crawling trajectories. Fig. B-1b superim-
Figure B-1: $(t, x)$-diagrams for (a) possible trajectories; (b) approximate trajectory.

poses these two curves and depicts the approximate trajectory, which is defined with the parameters $L, u$ and $v_c(G)$. In this setting, one can apply renewal theory because truck trajectories are independent of one another and so are truck inter-arrival times, $h$.

As follows from Fig. B-2, the mean of the process at $x_o$, $m(h)$, is given by $rhQ_U, h \leq T$ and $r[TQ_U + (h - T)Q], h \geq T$, where

$$T \doteq L \frac{w + v_c}{wv_c} , \hspace{1cm} \text{(B-1)}$$

is the duration of the disturbance caused by a single truck. It is easy to verify that for triangular fundamental diagrams we have

$$Q_U = Q_D + \frac{w v_c}{w + v_c} \kappa . \hspace{1cm} \text{(B-2)}$$

Let the random variable $h$ have mean $\bar{h}$ and probability density function $f(h)$. It
follows that the time until the next truck arrival is \( f(h) = m(h)e^{-m(h)} \), ie

\[
f(h) = \begin{cases} 
  rQ_U e^{-rQ_U h} & \text{when } h \leq T, \\
  e^{-Tr(Q-Q_U)}rQ e^{-rQ_U h} & \text{when } h > T.
\end{cases}
\]

(B-3)

Noting that \( \int x^\beta e^{-\beta x} dx, \beta > 0 \), equals \( \frac{1}{\beta}(1 - e^{-\beta T}) - Te^{-\beta T} \) on \([0, T]\) and \( \frac{1}{\beta}e^{-\beta T} + Te^{-\beta T} \) on \([T, \infty)\), and taking the expectation of (B-3) gives

\[
\tilde{h} = \left[ \frac{e^{-rQ_U T}}{Q} + \frac{(1 - e^{-rQ_U T})}{Q_U} \right]/r.
\]

(B-4)

Since the mean number of vehicles in a truck-led platoon is \( 1/r \), the capacity of the upgrade can be expressed as \( 1/(r\tilde{h}) \) and its dimensionless capacity as \( \rho = 1/(r\tilde{h}Q) \). Therefore,

\[
\rho = \frac{\rho_{\text{min}}}{1 - e^{-rQ_U T}(1 - \rho_{\text{min}})},
\]

(B-5)

where \( \rho_{\text{min}} = Q_U/Q \) is the minimum capacity; ie, when the queue is permanent. Notice how (B-5) exhibits of the correct properties in the limit cases: (i) \( \rho \to 1 \) as \( r \to 0 \) or \( T \to 0 \), and (ii) \( \rho \to \rho_{\text{min}} \) as \( r \to 1 \) or \( T \to \infty \).
Figure B-2: (a) \((t, x)\) diagrams of a single truck on the uphill; (b) mean of the simplified process.