Optimal location of Variable Speed Limit application areas in freeways

Treball realitzat per:
Irene Martínez Josemaría

Dirigit per:
Francesc Soriguera Martí

Màster en:
Enginyeria de Camins Canals i Ports

Barcelona, 13 Setembre 2017

Departament d’Enginyeria Civil i Ambiental
Declaration of Authorship

I, Irene Martínez Josemaría, declare that this thesis titled “Optimal location of Variable Speed Limit application areas in Freeways” is my own work. I confirm that:

1. This work was done wholly while in candidature for the master’s degree in Enginyeria de Camins Canals i Ports at Universitat Politècnica de Catalunya.
2. Where I have consulted the published work of others, this is always clearly attributed.
3. Where I have quoted from the work of others, the source is always given.
4. I have acknowledged all main sources of help.

Signed:

[Signature]

Irene Martínez Josemaría
Acknowledgments

I would first like to thank my advisors Dr. Wen Long Jin at ITS UC Irvine and Dr. Francesc Soriguera at the UPC. They have provided me with full support. Thanks to them I have discovered the research world that I really like. Through the weekly meetings with Dr. Jin, I have learned a lot thanks to his advises and interesting discussions. Thanks for your time and interest.

Second I would like to thank Prof. Rangel and Dr. Balsells, it has been an honor to be awarded with the Mobility Balsells Fellowship. Thanks for the opportunity to come to UC Irvine and funding my stay.

I would also like to acknowledge the people I have met during my stay at UC Irvine, specially all the minis and other Balsells. Borja and Jordi supported me in the challenge of learning LATEX and C programming, which has been essential for this thesis development, reducing the computation time of the simulations considerably (compared to Matlab) and ensuring an easy and understandable presentation of the results of my thesis. Felipe de Souza has also helped me in research and was always willing to discuss the results and find useful papers to contrast them.

On the other hand, I am very grateful to my home university friends for all the support they have provided me during my academic life. This thesis is only the last milestone of a long university period. I would like to express a special thank you to Luis, David, Carlos, Clara and Eli. I am sure they are not only part of my good memories but also part of my future. I would also like to thank all my other friends who have been my side in the goods and in the bads. In a more particular way I would like to thank all the gallinitas. Specially, for all the weekends where we have been traveling and relaxing. These days have proven that even if we are slowly defining our new and own paths in life (and they might not go in the same direction anymore) we can still meet and be united to share our thoughts, desires and peaceful moments and, on top of all, we can always laugh and play together as we were kids.

To sum up, I want to wish all of them the best luck possible on this new adventure (the industry life) that most of them have already started. I wish them luck in any other adventure that life has prepared for them, such as traveling, marriage or maybe having kids some day. I hope that we can learn from each other in all those new stages of life, as we have been learning together about life in general (and Civil Engineering in particular) until now.

I also want to mention Mireia and Sara, who are one of the oldest friends I have and with whom I have celebrated all my successes in life. I also want to share my gratitude to you, because you have also supported me, even when things seemed to go all wrong.

Finally, I want to thank my father and mother for the unconditional support they have showed. They have contributed to my education in an active and passive way, all my life. I am sure that the distance during my stay at UC Irvine was hard to manage, but I am happy to make them proud with my effort and dedication. Somehow, I should also thank my grandparents, they have taught my parents and me a lot of things in live.
Abstract

Variable speed limits (VSL) are an effective method for preventing, eliminating, or delaying the occurrence of capacity drop at active bottlenecks. Many studies are focused on feedback algorithms to adapt the VSL to different traffic conditions. The location of VSL application area is another important design problem that has been largely overlooked in the literature, with only few practical guidelines. It has been suggested that vehicles should achieve the free-flow speed before entering an active bottleneck. On the other hand, the VSL application area cannot be too far away from the bottleneck to avoid queue spillback to upstream off-ramps. This study brings some light into the optimal location of VSL control both theoretically and numerically.

First, I define an open-loop control strategy and develop a simulation tool in a hypothetical lane-drop freeway stretch based on a car-following model with bounded acceleration (BA) that captures the capacity drop phenomenon. The results prove that the traffic breakdown is only prevented when the VSL is applied at certain distances from the lane-drop. Second, it is shown that capacity drop is prevented when the stationary acceleration process is only defined by the BA-model. Moreover, it is proven that to prevent traffic breakdown it is not necessary that vehicles accelerate to free-flow speed before entering the lane-drop. Finally, an analytical formulation is developed to determine the optimal location of the VSL application area. A systematic sensitivity analysis allows to draw some conclusions on the effect of different traffic and geometric factors on this optimal location of the control application area.

**Keywords**: Lane-drop bottleneck, capacity drop, VSL application area, optimal location, bounded acceleration stationary states
Resumen

Los límites variables de velocidad (VSL) son una estrategia de control del tráfico que pretende evitar, eliminar o retrasar los efectos del conocido capacity drop o pérdida de capacidad en los cuellos de botella de las autovías. En la literatura podemos encontrar numerosos estudios sobre algoritmos para adaptar el límite de velocidad a las condiciones del tráfico. Sin embargo, la localización de las zonas de aplicación de estos VSL es otro parámetro de diseño importante, que hasta ahora ha pasado prácticamente desapercibido. En la literatura actual, sólo se mencionan algunas recomendaciones prácticas, por ejemplo, se ha sugerido que los vehículos deben llegar la velocidad libre antes de llegar al cuello de botella. Por otra parte, sabemos que el área de aplicación del control no debería estar muy alejado del cuello de botella para evitar congestionar otros accesos situados aguas arriba. Este trabajo final de máster desarrolla un razonamiento teórico y numérico para determinar la localización óptima del área de aplicación del control.

En primer lugar, se define una estrategia de control abierta (open-loop) y se desarrolla una herramienta de simulación en una autovía hipotética con reducción de carril basada en un modelo car-following (de microsimulación). El modelo considera una aceleración limitada (BA) y reproduce el fenómeno de capacity drop. Con los resultados de la simulación se demuestra que se puede evitar la pérdida de capacidad cuando el proceso de aceleración es estacionario y está definido por el modelo de BA. Además, se demuestra que no es necesario que los vehículos aceleren hasta la velocidad libre antes de entrar en el cuello de botella. Finalmente, se desarrolla una formulación teórica para determinar la localización óptima de las áreas de aplicación del control y se realiza un análisis de sensibilidad, para establecer que efecto tienen los diferentes parámetros (geométricos y del modelo).
Resum

L’ús de límits variables de velocitat (VSL) és una estratègia de control del trànsit que pretén evitar, eliminar o endarrerir els efectes del conegut “capacity drop” o pèrdua de capacitat als colls d’ampolla de les autovies. A la literatura podem trobar nombrosos estudis que se centren en algoritmes per adaptar el límit de velocitat a les condicions del trànsit. No obstant això, la localització de les zones d’aplicació d’aquests VSL és un altre paràmetre de disseny important que, malauradament, mai s’ha estudiat en profunditat en la literatura. Només s’han proposat un parell de recomanacions pràctiques en base a uns resultats empírics. Per una banda, s’ha suggerit que els vehicles han d’arribar a la velocitat lliure abans d’arribar al coll d’ampolla. D’altra banda, sabem que l’àrea d’aplicació del control no pot estar molt lluny del coll d’ampolla per evitar congestionar altres sortides (aigües a dalt) de l’autovia. Aquest treball de final de màster desenvolupa una raonament teòric i numèric per determinar la localització optima de l’àrea de aplicació del control.

En primer lloc, es defineix una estratègia de control oberta (open-loop) i es desenvolupa una eina de simulació en una autovia hipòtica amb reducció de carril basada en un model “car-following” (de microsimulació) que considera acceleració limitada (BA) i que captura el fenomen de “capacity drop”. Amb els resultats obtinguts es demostra que es pot evitar la pèrdua de capacitat quan el procés d’acceleració és estacionari i està definit (controlat) pel model de BA. A més a més, es demostra que no és necessari que els vehicles accelerin fins a la velocitat lliure abans d’entrar al coll d’ampolla. Finalment, es desenvolupa una formulació teòrica per determinar la localització optima de les àrees d’aplicació del control. També es realitza un anàlisis de sensibilitat per concloure quin efecte tenen els diferents paràmetres geomètrics i d’estat del trànsit en la localització optima de l’àrea d’aplicació del control.
## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Declaration of Authorship</td>
<td>iii</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td>v</td>
</tr>
<tr>
<td>Abstract</td>
<td>vii</td>
</tr>
<tr>
<td>Resumen</td>
<td>ix</td>
</tr>
<tr>
<td>Resum</td>
<td>xi</td>
</tr>
<tr>
<td>List of Figures</td>
<td>xv</td>
</tr>
<tr>
<td>List of Tables</td>
<td>xix</td>
</tr>
<tr>
<td>1 Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Main Research Objectives</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Structure of the Document</td>
<td>3</td>
</tr>
<tr>
<td>1.3 Literature Review</td>
<td>3</td>
</tr>
<tr>
<td>1.3.1 Brief overview on Traffic Flow Theory</td>
<td>3</td>
</tr>
<tr>
<td>1.3.2 Brief overview on Control Systems Theory</td>
<td>10</td>
</tr>
<tr>
<td>1.3.3 Current state of Art on Capacity Drop and Bounded Acceleration models</td>
<td>13</td>
</tr>
<tr>
<td>1.3.4 Current state of Art on Variable Speed Limits (VSL)</td>
<td>15</td>
</tr>
<tr>
<td>2 Model and Methodology</td>
<td>19</td>
</tr>
<tr>
<td>2.1 The lane-drop bottleneck</td>
<td>19</td>
</tr>
<tr>
<td>2.2 Hypothesis of the model</td>
<td>20</td>
</tr>
<tr>
<td>2.3 Car-following formulation</td>
<td>22</td>
</tr>
<tr>
<td>2.4 Open-loop control definition</td>
<td>24</td>
</tr>
<tr>
<td>2.5 Programming and simulations</td>
<td>26</td>
</tr>
<tr>
<td>2.5.1 Scenarios to analyze</td>
<td>26</td>
</tr>
<tr>
<td>2.5.2 Languages used and code description</td>
<td>28</td>
</tr>
<tr>
<td>2.6 Spatial and temporal discretization</td>
<td>31</td>
</tr>
</tbody>
</table>

xiii
### Contents

#### 3 Simulation Results
- 3.1 Simulation results without control ................................. 33
  - 3.1.1 Capacity drop presence and magnitude .......................... 33
  - 3.1.2 Traffic dynamics: preliminary results ........................ 38
- 3.2 Simulation results with control .................................... 40
  - 3.2.1 Application area location to avoid capacity drop ............ 40
  - 3.2.2 Key insight to avoid capacity drop .......................... 43

#### 4 Analytical formulation for optimal control location .......... 47
- 4.1 General problem statement ........................................... 47
- 4.2 Analytical solution for CBA .......................................... 48
  - 4.2.1 Sensitivity analysis of $L_{opt}$ for CBA model .................. 49
- 4.3 New formulation for optimal application area .................... 51
  - 4.3.1 Solution of new formulation for CBA case .................... 53
  - 4.3.2 Numerical solutions for TWOPAS and Gipps models .......... 53
  - 4.3.3 Sensitivity analysis of $L_{opt}$ for different BA-models ....... 55

#### 5 Conclusions .............................................................. 59
- 5.1 Summary and relevance of results .................................. 59
- 5.2 Main findings ............................................................. 60
- 5.3 Discussion and suggestions for further research .................. 61

#### Appendix A Complementary Figures ............................... 63
- A.1 Traffic dynamics: cumulative curves ............................... 63
- A.2 Simulation results without control ................................. 65
- A.3 Results of simulation with control and CBA ....................... 68
  - A.3.1 Changing application area of VSL ............................. 68
  - A.3.2 Changing demand .................................................. 72
- A.4 Convergence for TWOPAS and Gipps integration ................. 75
- A.5 Analysis of optimal location for TWOPAS and Gipps models ..... 76

#### Bibliography .............................................................. 79
# List of Figures

1.1 Lane drop bottleneck with VSL application area ending at $x = L_u$. ........................................... 2
1.2 Fundamental Diagram of Traffic. Ref. [1]. ................................................................. 6
1.3 Greenshields speed-density relation. [2]. ........................................................................ 7
1.4 General triangular fundamental diagram (TFD). ............................................................ 8
1.5 Illustration of Edie’s domain along the vehicle’s trajectory. Ref. [3]. ......................... 10
1.6 Basic Control System without feedback (open-loop). ..................................................... 11
1.7 Control System with feedback (closed-loop). ................................................................. 11
1.8 Most used control strategies are either PID controllers (upper figure) or bang-bang controllers (lower figure) ............................................................... 12
2.1 Lane drop bottleneck with leader vehicles in FIFO [3]. ............................................... 20
2.2 Reduction of flow when imposing VSL on the upstream link of a lane-drop stretch. $C_{VSL}$ and $V_{SL_{max}}$ definition. ......................................................................................... 25
2.3 Concepts of demand and supply related to density. .......................................................... 28
3.1 Traffic performance and acceleration pattern for Scenario 2 with CP-model [3]. ........... 34
3.2 Traffic performance and acceleration pattern for Scenario 2. With unbounded acceleration. .............................................................................................................................. 36
3.3 Convergence analysis of CDR and its time appearance. Compare different demand scenarios (1 to 3) for the basic geometry and CBA model. .................................................. 37
3.4 Capacity drop analysis depending on lane-drop length, acceleration ratio, $a_0$, and BA model. ...................................................................................................................... 37
3.5 Cumulative curves at lane-drop start and end. Upstream demand is twice the downstream capacity (Scenario 3 without control) ......................................................... 39
3.6 Spill back of queue at $x = L = 100m$ (i.e. capacity drop appearance) for $a_0 = 2m/s^2$. . 39
3.7 Spill back of queue at $x = 0$ for $a_0 = 2$. ........................................................................ 40
3.8 Spill back of queue at $x = L = 150m$ (i.e. capacity drop appearance) for $a_0 = 2m/s^2$. . 40
3.9 Spill back of queue at $x = 0$ for $a_0 = 2$ and $L = 150m$. .................................................. 41
xvi

LIST OF FIGURES

3.10 Open-loop control with VSL application area at 200m upstream of lane-drop start, i.e. \( L_u = -200m \). Base case with demand Scenario 2. ................................. 42

3.11 Main simulation results observations ................................................................. 45

4.1 (a) Case study in Chapter 3 (b) Envelope of \( f(x) \) for different lane-drop lengths. 49

4.2 For different VSL (a), (b) shows the associated normalized stationary flow \( CV_{SL} \). (c) shows the distance needed to accelerate to \( v_f \), (d) shows the critical length, \( L_c \), to accelerate from \( V_{SL} \) to \( v_2 \). .................................................. 50

4.3 Speed from the BA-model, \( v_c(x) \) and minimum speed, \( \Psi(x) \). Example with \( V_{SL} = 3m/s \) and \( L = 100m \). ................................................................. 52

4.4 Backwards integration from \( v_2 = 25.99m/s \) for TWOPAS model with \( L = 100m \) and \( V_{SL} = 3.611m/s \). ................................................................. 54

4.5 Backwards integration from \( v_2 = 25.99m/s \) for Gipps model with \( L = 100m \) and \( V_{SL} = 3.611m/s \). ................................................................. 55

4.6 Effects of acceleration ratio, \( a_0 \), and speed limit ratio, \( V_{SL}/V_{SL max} \), on critical distance \( (L_c = |L_0| + L) \). Results for Gipps model on freeway with \( l_1 = 2 \). .......................... 56

4.7 Effects of acceleration ratio, \( a_0 \), and speed limit ratio, \( V_{SL}/V_{SL max} \), on critical distance \( (L_c) \). Results for TWOPAS model on freeway with \( l_1 = 2 \) ........................................ 57

4.8 Comparing BA-models. Critical length, \( L_c \), computed for \( V_{SL} = 0.95V_{SL max} \) and \( l_1 = 2 \) upstream lanes. (a) Effects of acceleration ratio, \( a_0 \), (b) Effects of an spatial-averaged acceleration, i.e. the slope of the secant line of \( v_c(x) \) between \( V_{SL} \) and \( v_2 \). (c) Effects of an averaged speed, considering that for all simulations the results are from a CBA-model. ........................................ 58

A.1 Cumulative curves at \( x = 0 \) and \( x = L \) for base case, Scenario 1. ........................ 63

A.2 Cumulative curves at \( x = 0 \) and \( x = L \) for base case, Scenario 2. ........................ 64

A.3 Cumulative curves at \( x = 0 \) and \( x = L \) for base case, Scenario 3. ........................ 64

A.4 Demand Scenario 3, with CBA \( a_0 = 2m/s^2 \) and lane-drop length \( L = 250m \). .......................... 65

A.5 Demand Scenario 3, with Gipps BA-model \( a_0 = 2m/s^2 \) and lane-drop length \( L = 100m \). .......................... 66

A.6 Upstream number of lanes is \( l_1 = 3 \) and \( d_1 \) exceeds downstream capacity in 25%, with CBA \( a_0 = 2m/s^2 \) and lane-drop length \( L = 100m \). .......................... 67

A.7 Control located at \( L_u = 0m \) and demand exceeding 50% of downstream supply. .......................... 68

A.8 Control located at \( L_u = -10m \) and demand exceeding 50% of downstream supply. .......................... 69

A.9 Control located at \( L_u = -20m \) and demand exceeding 50% of downstream supply. .......................... 70

A.10 Control located at \( L_u = -30m \) and demand exceeding 50% of downstream supply. .......................... 71

A.11 Control located at \( L_u = -50m \) and demand exceeding 10% of downstream supply. .......................... 72

A.12 Control located at \( L_u = -50m \) and demand exceeding 50% of downstream supply. .......................... 73

A.13 Control located at \( L_u = -50m \) and demand exceeding 100% of downstream supply. .......................... 74

A.14 Convergence study for numerical integration for Figure 4.4 TWOPAS model. .......................... 75

A.15 Convergence study for numerical integration for Figure 4.5 Gipps model. .......................... 75
A.16 Effects of number of lanes $l_1$ and speed limit ratio $V_{SL}/V_{SL_{max}}$ on critical distance $(L_c = |L_0^{opt}| + L)$. Results for (a) Gipps and (b) TWOPAS model. Acceleration ratio $a_0 = 2$.

A.17 Analysis of VSL and geometry influence on $L_0^{opt}$ for TWOPAS model with constant $a_0 = 2m/s^2$.

A.18 Analysis of VSL and geometry influence on $L_0^{opt}$ for Gipps model with constant $a_0 = 2m/s^2$.

A.19 Effects of acceleration ratio $a_0$ on $L_0^{opt}$ for Gipps model. Freeway stretch with $l_1 = 2$ and different lane-drop lengths.
# List of Tables

1.1 Basic concepts and variables of traffic flow theory ........................................ 4

2.1 TFD definition parameters .............................................................................. 22

2.2 Basic scenario to analyze. Geometry and acceleration ratio .............................. 26

2.3 Freeway characteristics derived from basic-scenario and TFD parameters .......... 27

2.4 Scenarios analyzed that activate the lane-drop bottleneck .............................. 27

3.1 Basic scenarios to analyze without control ...................................................... 35

3.2 Scenarios to analyze with control ................................................................... 41

3.3 Scenarios with different geometries ................................................................. 42

3.4 Scenarios with different lane-drop lengths and same demand and number of lanes \( (l_i = 2) \) ........................................ 43

3.5 Results of required \( L_u \) for different simulation scenarios (BA-models and lane-drop lengths) 43
Chapter 1

Introduction

1.1 Main Research Objectives

In the last decades there has been an increase in car use. Since engineers cannot adapt the infrastructures to the constantly increasing demand for both economic and environmental reasons, we need to find another way to improve traffic flow and reduce delay due to congestion. These type of strategies are called traffic management or traffic control. Its main goal is to increase the flow (i.e. number of vehicles per unit of time) on freeways. The aspiration is to achieve more flexibility in the use of the existing highways.

One of the main causes that increase travel time on freeways is a phenomenon called capacity drop. This phenomenon takes place at bottlenecks (freeway stretches that have a reduction of capacity) when they become active. The activation of a bottleneck occurs when an excess of demand of vehicles (greater than the capacity of the bottleneck) starts generating queues upstream of the bottleneck, whereas downstream of the bottleneck free-flow regime is found (i.e. without congestion). When a bottleneck becomes active, the maximum discharge low-rate of the bottleneck can drop by 5% to 20% of its actual capacity [8], [9], [10], [11], [12]. In addition, the occurrence of capacity drop can further starve some off-ramps and prolong vehicles’ travel times due to the upstream propagation of queues [13]. In order to prevent the activation of the critical bottlenecks or mitigate the impacts of the harmful capacity drop, many management strategies have been proposed, most of them are ramp metering control (RMC) [14], Variable Speed Limit control (VSL) [15], or a combination of both [16], [17].

Nowadays, VSL are used in many places, especially in Europe. Their goal is to adapt the road traffic speed to different situations (e.g. weather, accidents, bottlenecks, etc.) in real time. It is claimed that they help to reduce pollution and accident rates by homogenization of the road. The term “homogenization” refers to reduction of variance in speed, occupancy and/or lane changes. Another possible function of VSL is the prevention of traffic breakdown by avoidance of too high densities.
1.1. MAIN RESEARCH OBJECTIVES

The main objective of this MSc thesis is to define an optimal location of the VSL application area for control strategies that aim to prevent, delay or eliminate the capacity drop phenomenon. In the literature, many feedback control strategies have been proposed to do so, but few is said about the location of the control implementation. In the interest of filling this gap in research, the microscopic behavior of cars is simulated in an hypothetical freeway corridor using a model that reproduces capacity drop phenomenon endogenously by considering bounded acceleration (BA).

The capacity drop phenomenon is studied for an active bottleneck in a lane-drop zone, Figure 1.1. The analysis includes a detailed study of the lane-drop effect on traffic performance and the capacity drop ratio is studied for different freeway geometries and BA-models. Furthermore, a management strategy with use of VSL is proposed to improve traffic performance for different scenarios. The micro-simulation results are analyzed by computing the macroscopic characteristics of traffic. Even if the characteristics of the road represent an ideal scenario with no curves, no slopes and homogeneous traffic across lanes, the results bring some insights into the influence of VSL on traffic. To do so different speed limits and different VSL locations (see $L_u$ in Figure 1.1) on the road are considered. The possibility of avoiding or delaying the capacity drop phenomenon in a lane-drop scenario will be analyzed and a trade-off is searched to locate the VSL application area in the most beneficial place.

![Figure 1.1: Lane drop bottleneck with VSL application area ending at $x = L_u$.](image)

The results prove that capacity drop prevention is only achieved under certain conditions, i.e. when the control is located at a certain distance, $L_u$, from the start of the bottleneck. Moreover, the results of the simulations reveal a key insight that allows to develop a theoretical formulation to derive the optimal location of the VSL application area. By constraining the interaction between the acceleration zone downstream of the VSL application area and the acceleration zone inside the lane-drop bottleneck, it allows to define a necessary and sufficient condition to prevent the occurrence of capacity drop. Based on this condition, the minimum distance between the VSL application area and the bottleneck to prevent capacity drop can be calculated. The dependence of such a distance on various traffic and geometry factors is also analyzed. Thereby it is expected that the results of this MSc thesis will help to the efficient use of VSL technology.
1.2 Structure of the Document

The following section of this Introduction presents the literature review. The sections 1.3.1 and 1.3.2 are thought to introduce the reader to the basic concepts of traffic flow theory and control systems. If you are a civil engineer or traffic engineer with knowledge of the fundamental diagram, the conservation of vehicles and the differences between macroscopic and microscopic models you can skip “Brief overview on traffic flow theory” and if you are an engineer or mathematician that is comfortable with control systems and feedback systems you might skip “Brief overview on control system theory”.

Later, the state of the art on bounded acceleration models, capacity drop models and VSL control strategies is presented in sections 1.3.3 and 1.3.4. With this complete background, in Chapter 2 the model used for the micro-simulations is explained in detail and the methodology to implement the VSL control strategy is presented. Then, in Chapter 3, the results of the simulations are presented and discussed. In Chapter 4 an analytic formulation is developed to determine the minimum (i.e. optimal) distance between the VSL application area and the bottleneck. Finally, Chapter 5 contains the conclusions of this MSc thesis and some recommendations for further research.

1.3 Literature Review

1.3.1 Brief overview on Traffic Flow Theory

The basic concepts of traffic flow theory are presented in this section. The objective is to familiarize the reader with traffic flow theory in order to ease the full understanding of the present document. Any Engineer that is familiar with traffic flow theory or traffic engineering can skip this section.

There are basically two approaches to model mathematically traffic flow. One approach, from a microscopic view, studies individual movements of vehicles and interactions between vehicle pairs. This approach considers individual driving behavior but the computation of many problems becomes easily mathematically intractable, due to the complexity of the system. The dynamics between pairs of vehicles is difficult to track when a considerable volume of traffic flow is considered. The first microscopic models where developed in the 60’s by General Motors [18], [19]. The second approach studies the macroscopic features of traffic flows. Macroscopic models are more suitable for modeling traffic flow in complex networks since less supporting data and computation are needed. These models consider only the fundamental variables (flow, density and speed) as averaged values in the freeway that depend only on location, $x$, and time, $t$. Table 1 presents some basic definitions of the variables and parameters commonly used in traffic flow theory.

A homogeneous traffic flow means that the whole roadway has the same traffic parameters at any location. On the other hand, traffic flows are considered inhomogeneous when the roadway
has different parameters at different locations. Traffic flow can be studied in whole networks, where traffic dynamics may appear due to the different origin and destiny demands. Nevertheless, the main objective of traffic flow theory is to study single roads under specific conditions (e.g. merges, diverges, different number of lanes, different demand patterns and different types of bottlenecks).

<table>
<thead>
<tr>
<th>Concept</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flow, $q$</td>
<td>The number of vehicles crossing a section during a certain period of time</td>
</tr>
<tr>
<td>Time mean speed, $v_t$</td>
<td>The mean speed of a certain number of vehicles that crosses a fixed section during a period of time</td>
</tr>
<tr>
<td>Space mean speed, $v_s$</td>
<td>The mean speed of vehicles in a certain stretch of the road (for a fixed time)</td>
</tr>
<tr>
<td>Density, $k$</td>
<td>The number of vehicles in a certain stretch of the road over the length of this stretch</td>
</tr>
<tr>
<td>Capacity, $C$</td>
<td>Maximum flow in a road. It is usually defined per lane. It has been proven that there is a range of capacity values for all roads in the world (2000 - 2300 veh/h/lane). The actual value depends on the road characteristics: slope, curves, country, etc.</td>
</tr>
<tr>
<td>Critical density, $k_c$</td>
<td>Density at which we observe capacity (maximum flow).</td>
</tr>
<tr>
<td>Jam density, $k_j$</td>
<td>The maximum density in a road. Represents the number of vehicles when they are stopped (usually considering a safety distance between vehicles)</td>
</tr>
<tr>
<td>Jam spacing, $s_j$</td>
<td>The distance between the front bumper of vehicle $n$ and the front bumper of vehicle $n - 1$. It is the inverse value of $k_j$.</td>
</tr>
<tr>
<td>Free-flow speed, $v_f$</td>
<td>The speed that vehicles have when they travel as fast as they want (no speed limit nor other enforcement). It is equivalent to say the maximum speed of the road.</td>
</tr>
<tr>
<td>Shock wave speed, $w$</td>
<td>Speed at which information is propagated in congested traffic. This speed is negative.</td>
</tr>
<tr>
<td>Occupancy, $occ$</td>
<td>Proportion of time that a detector is occupied. Can be treated as a measure of density.</td>
</tr>
</tbody>
</table>

Table 1.1: Basic concepts and variables of traffic flow theory

Macroscopic traffic flow models

The study of macroscopic models starts based on the assumption that high-density streams of vehicles behave like an incompressible fluid. This is equivalent to assume that individual drivers react to the presence of other vehicles in such a way that only the stream or total flow characteristics need to be considered. In other words, we can describe the traffic performance with a small number of variables, such as flow rate, $q$, the traffic density, $k$, and the travel speed, $v$. Those variables can be obtained from overall characteristics of the traffic. These variables depend on the location and time; the macroscopic analysis does not consider the individual location and
speed of vehicles, but an aggregated (i.e. averaged) speed and density. The relationship between these three variables is the fundamental relation of traffic flow, also known as the \textit{constitutive law}, Eq. 1.1. Traffic flow models are based on this fundamental relation and the conservation of vehicles, Eq. 1.2. This two features are true overall, under any circumstances and conditions.

In steady flow conditions (when the the fundamental characteristics of traffic are time-independent), a speed-density relation can be observed in traffic, see Eq. 1.3 The fundamental diagram of traffic flow (FD) can be defined with Eq. 1.1 and Eq. 1.3. The FD is a graphical representation of the flow-density relation (i.e. fundamental relation of traffic), see Figure 1.2. This relation represented is only valid for steady states and does not capture the traffic dynamics at a microscopic level. Such steady states are defined by a number of vehicles traveling at same speed during a long period of time. This is an approximation, because no purely steady traffic can be observed in real roads. Nevertheless this approximation is good enough if traffic at a specific location is constant during about $\sim 10\text{min}$ (i.e. has the same mean speed and same mean density). The shape of the FD depends on the definition of the speed function, $\eta(k)$.

\begin{equation}
q = kv \tag{1.1}
\end{equation}

\begin{equation}
k_t + q_x = 0 \tag{1.2}
\end{equation}

\begin{equation}
v = \eta(k) \tag{1.3}
\end{equation}

The mean speed used in the fundamental relation is the space mean speed, $\bar{v}_s$, and not the time mean speed, $\bar{v}_t$. The time mean speed is always equal or higher than the space mean speed (see Eq. 1.4) because in the first one all vehicles have the same weight, independently from their speed (sectional speeds are averaged), but in $\bar{v}_d$ the vehicles that are slower have a higher weight. The reason for this is that the slower vehicles remain a longer time period on the road. This is intuitively understood in an extreme example: Consider a stretch where there is a vehicle with speed $v = 0$. This vehicle's speed will be counted and averaged for the space mean speed, while it will never cross a section (because it is stopped) and thus will not be taken into account in the time mean speed average.

\begin{equation}
\bar{v}_t = \bar{v}_s + \frac{\sigma^2}{\bar{v}_s} \tag{1.4}
\end{equation}

\textbf{LWR model}

Many continuum traffic flow models can be described by a system of hyperbolic PDEs. One of the most extended and comprehensive macroscopic models is the so called Lighthill, Whitham and Richards (LWR) model, developed in \cite{20}, \cite{21}. The LWR model is a first-order model and it has been solved for the homogeneous roadway rigorously. Additionally, some empirical solutions
have been found to the inhomogeneous LWR model. This model relies on the assumption that there exists an equilibrium speed-density relationship (Eq. 1.3) and thus assumes steady state traffic. The second assumption is the conservation of number of vehicles, which is true under any circumstances (can be proved mathematically and graphically). Basically it is the superposition of Eq. 1.1 to Eq. 1.3. Thereafter, traffic flow can be described by Eq. 1.5 and the conservation of number of vehicles as a first order quasi-linear partial differential equation (Eq. 1.6). This equation was solved for the first time around 1950’s. The bottom line is that this LWR-model can be used for different density-speed relations and thus is valid for any FD shape. Nevertheless, the LWR-model cannot explain some empirical features of traffic breakdown and may present unrealistically high acceleration rates.

\[ q(t,x) = k(t,x)v(t,x) = k(t,x) \cdot \eta(k(t,x)) = \phi(k(t,x)) \] (1.5)

\[ \frac{\partial k(t,x)}{\partial t} + \frac{\partial \phi(k(t,x))}{\partial x} = 0 \] (1.6)

Greenshields traffic flow relation

The first and most simple relation between speed and density is proposed by Greenshield [22, 2]. He postulated a linear relationship, as shown in Figure 1.3. The equation is simply a linear function with maximum value (i.e. free-flow speed) at \( k = 0 \), and minimum value (i.e. 0km/h) for jam density, \( k_j \):
1.3. LITERATURE REVIEW

The concept of flow was not known back when Greenshields collected the traffic flow data and defined this fundamental relation. The throughput of vehicles was described as density-vehicles per hour. Since the speed-density relation is a linear function, the fundamental relation of traffic flow is a parabolic function. In this case, the maximum flow (i.e. Capacity of the freeway) is the maximum of the parabola (see Eq. 1.8) and it can be proven that the related density (i.e. critical density) is $k_j/2$. Therefore, the associated critical speed is the one in Eq. 1.9.

$$C = \frac{v_f k_j}{4} \quad (1.8)$$

$$v_c = v_f - \frac{v_f}{2} = \frac{v_f}{2} \quad (1.9)$$

**Triangular Fundamental Diagram**

One of the most used fundamental diagrams is the triangular fundamental diagram (TFD), which is defined by means of only three parameters. It is obtained by considering the following speed-density relationship [23], [24]:

$$V(k) = \min \left\{ v_f, w \left( \frac{k_j l}{k} - 1 \right) \right\} \quad (1.10)$$

The three parameters in the TFD are the free-flow speed ($v_f$), the per-lane jam density ($k_j$) and the shock wave speed in congested states ($w$). These concepts are explained in Table 1.1. Correspondingly, the flow-density relation is presented in Eq. 1.11 which depends on the number of lanes $l$. 

$$V(k) = v_f - \frac{v_f k}{k_j} \quad (1.7)$$
1.3. LITERATURE REVIEW

Figure 1.4: General triangular fundamental diagram (TFD).

\[ q(x, t) = \phi(k(x, t)) = \min\{v_f k(x, t), w(k_j l - k(x, t))\} \] (1.11)

It is important to note that considering a triangular FD implies that the critical speed, \( v_c \), is exactly the free-flow speed, i.e. the capacity of each lane is \( C = v_f k_c \). Where the critical density for each lane is computed from Eq. (1.12). Clearly the capacity of the road, is proportional to the number of lanes (Eq. 1.13).

\[ k_c = \frac{w}{v_f} + w k_j \] (1.12)

\[ C_{tot} = lC = \frac{v_f w}{v_f + w} k_j l \] (1.13)

Similarly to any other FD, traffic states with \( k \leq l k_c \) are uncongested states (also called free-flow states), whereas higher densities than the critical freeway density are related to congested states, see Figure 1.4. There is another concept worth to be explained: the time-gap, \( \tau_0 \). The time-gap is \( k_j w \) and is useful to describe congested states in the FD, see Eq. (1.14)

\[ k = \frac{\tau_0 l - q}{w} \]
\[ v = \frac{q}{\tau_0 l - q} w \] (1.14)

Car following models

The driver’s behavior under car-following (CF) is a crucial topic in the study of vehicular microscopic traffic theory. In particular, this subject is becoming essential in developing the adaptive
cruise control strategies with autonomous vehicles. Car-following theories have been developed by numerous researchers, but they often are associated with the work of Herman [25], [26] at the research laboratories of General Motors Corporation. The authors pointed out that the “follow the leader” types of problems are the product of the psychological behavior of the drivers as they respond to certain stimuli. The basic CF-model considered is one in which it is assumed that a driver will attempt to keep relative speed between his vehicle and the one immediately ahead as small as possible. The stimulus will be a change in the speed of the vehicle ahead, and the lagged response will be a change in the speed of the following vehicle. The model is constrained to minimize differences in the relative speed between the two vehicles. Other ways to define a CF-model are based on imposing a minimal spacing between vehicles, for example the DMV (or 2 seconds) rule, which is generally used in California: A good rule for following another vehicle at a safe distance is to allow yourself at least the length of a car between your vehicle and the vehicle ahead for every ten miles per hour of speed at which you are traveling. This basically means that the safe distance that the following vehicle needs to maintain changes depending of his own speed. This rule is equivalent to Eq. 1.15, where the distance is the spacing minus the length of the vehicle, \( d(t) = s(t) - L_0 \).

\[
d(t) = 1.36v(t) \sim 2v(t) \tag{1.15}
\]

\[
k(x,t) = -\frac{1}{\partial X(t,N)} \tag{1.16}
\]

A CF-model can also be derived from a macroscopic model. Because the definition of spacing \( s(x,t) \) for each vehicle can be obtained from the road geometry and traffic conditions: spacing is inversely proportional to density, \( k(x,t) \). Moreover, density is the spatial derivative of the trajectory (Eq. 1.16). However, the conversion of a microscopic model to a macroscopic model is easier. When the trajectory of all vehicles has been determined with a CF-model, the fundamental variables can be computed with Edie’s formula, in Eq. 1.17. Where \( \Omega(t,N) \) is the area of Edie’s domain, see Figure 1.5. Edie’s formula [27] discretizes and numerically calculates the density, speed and flow throughput along a vehicle’s trajectory.

\[
\Omega(t,N) = \frac{\Delta t}{8} [X(t + \Delta t, N - \Delta N) - X(t + \Delta t, N + \Delta N) + 2 (X(t, N - \Delta N) - X(t, N + \Delta N)) + X(t - \Delta t, N - \Delta N) - X(t - \Delta t, N + \Delta N)]
\]

\[
k(t,N) = \frac{\Delta t \Delta N}{\Omega(t,N)}
\]

\[
q(t,N) = \frac{X(t + \Delta t, N - X(t - \Delta t, N)}{2\Omega(t,N) \Delta N}
\]

\[
v(t,N) = \frac{q(t,N)}{k(t,N)} \tag{1.17}
\]
1.3.2 Brief overview on Control Systems Theory

Control System Modeling

A control system is composed by a controller and a plant. The plant is the control object (i.e. the system that has its intrinsic rules to operate). The controller measures an input, processes the information and generates an action on the plant. Usually the input is related to the objective of the control. It can be an observation or an objective itself.

To model a control system, we will need two different algorithms: (i) controller and (ii) plant dynamics. In a dynamical system (i.e. plant) we have one or more state-variables and a rate of change of those variables. Therefore it is common to threat system dynamics with differential equations, i.e. trying to define the rate of change of the state variable as a function of the variable itself. Usually, the plant dynamics can be described by a linear system (Eq. 1.18) and the observation can also be described by a linear function of the state variable, $x(t)$ and the controller $u(t)$, see Eq. 1.19. Moreover, if these parameters $A, B, C, D$ are time independent, the system is called linear-time-independent system and can be described by Eq. 1.20. However, as we have shown in the previous sections, traffic flow theory is usually based on non-linear equations.

$$\frac{dx(t)}{dt} = A(t)x(t) + B(t)u(t)$$  \hspace{1cm} (1.18)

$$y(t) = C(t)x(t) + D(t)u(t)$$  \hspace{1cm} (1.19)
\[ \begin{aligned}
\frac{dx(t)}{dt} &= Ax(t) + B u(t) \\
y(t) &= C x(t) + D u(t)
\end{aligned} \] (1.20)

**Open and Closed loop systems**

An open-loop control system is similar to a look-up-table (see Figure 1.6). These tables are build based on the experience of engineers. This strategy is not responsive to changing demands or special events. In the case of VSL control it is equivalent to display a specific speed limit independent of the traffic conditions (weather, density or demand).

![Control System without feedback (open-loop).](image)

On the other hand, a closed-loop control has a feedback. There is a state variable that is changed depending on other parameters. The goal of the feedback controller is to regulate an output by tracking the differences of the input to an input reference, \( r \), see Figure 1.7. The feedback is used to compensate disturbances and uncertainty of dynamic processes, it compares the input and the output in order to adjust the controller in an appropriated way. In other words, it studies how disturbances on the inputs (measurement errors or actual disturbances) affect the output. This is basically the study of the equilibrium and the disturbance attenuation properties of the system. In the transportation world, an extended example of feedback system is the cruise controller in vehicles, that adapts the speed of the vehicle in order to keep it constant at the value of the driver’s choice (reference speed). The controller compensates changes in the slope of the road by measuring the speed and adjusting the throttle accordingly.

![Control System with feedback (closed-loop).](image)

There are many types of feedback control systems. The most commons are proportional (P) controller, integral (I) controller, derivative (D) controller, and a combination of them (e.g. PI, PID controllers), bang-bang controller (also called on-off controller), optimal controller with
objective function (i.e. minimization problem that depends on the signal) and other intelligent and adaptative controllers.

For example, considering a PI (proportional-integral) controller is a simplification of the widely used PID controllers (proportional-integral-derivative). These type of control systems have been successfully applied to design the ALINEA ramp metering algorithm [28] and are also commonly used in the mechanical engineering industry (Chapter 10 in [29]). The derivative term, that is being ignored in the PI controller, is in charge of improving the transient performance of the controller.

![Figure 1.8: Most used control strategies are either PID controllers (upper figure) or bang-bang controllers (lower figure).](image)

This type of closed-loop control is composed by two differential equations, where the first one is the inputs differential equation and the second one is the outputs differential equation. The output $\hat{u}(t)$ is a PI-controller, defined in Eq. 1.21. Its differential form (Eq. 1.22) is the first differential equation needed for the control system. Where $e(t)$ is the error of the state variable (i.e. density).

$$\hat{u}(t) = K_p e(t) + K_i \int e(\tau) d\tau$$  \hspace{1cm} (1.21)

$$\frac{d}{dt} \hat{u}(t) = -K_p \frac{d}{dt} k(t) + K_i (\bar{k} - k(t))$$  \hspace{1cm} (1.22)

While closed-loop controls are more flexible and robust (if designed well), they are also more expensive to implement and may give more problems (if the design is not well enough). A big problem of feedback control is the appearance of positive feedback, since the controller becomes unstable. Otherwise, the close-loop are more efficient than open-loop, because they can adapt the controller to the output. A better description of general closed-loop control properties can be found in [29].
1.3.3 Current state of Art on Capacity Drop and Bounded Acceleration models

There are a lot of phenomena that characterize traffic flow behavior. One of the most known and still not completely understood is the so called capacity drop [9], [10], [30]. This phenomenon generates a reduction of the mainstream flow rates once congestion occurs on a freeway. The flow-rate at the downstream location of the traffic jam is lower than the pre-queue capacity. This drop in flow rate is the capacity drop. This phenomenon is specially harmful because it happens when the capacity of the roads is most needed [14]. The relatively constant drop in discharge flow has been observed to be around 5 – 20% [31]. Capacity drop is observed when bottlenecks become active, i.e. with the formation of a queue upstream to the location where the drop is observed [9], [30]. The capacity drop has been studied since 1990 and it can occur at many bottleneck locations, such as curves, upgrades, tunnels, lane-drops, work or accidents, merges etc. In [30] it was suggested that capacity drop occurs some minutes (∼ 10 minutes) after the queue has started to form. Nevertheless, a complete understanding of this mechanism behind the queue formation and subsequent traffic breakdown is still elusive.

In fact, there is no overall agreement on which factor ensues capacity drop when a bottleneck becomes active. Some authors consider that capacity drop may be triggered by microscopic phenomena, such as heterogeneous lane behavior, lane-changes or slow vehicles merging into the freeway [32], [33]. On the other hand, BA has also been considered to be ensuing capacity drop, this was first suggested in [10] and more recently discussed in [34].

In [31] the authors observed that the capacity drop is associated with extensive queues in the shoulder lane, important reductions on speeds and an increase in number of lane changes in some merge bottlenecks. They reasoned that drivers in the median lane may have decelerated as a cautionary behavior. This type of driver behavior was previously observed at freeway diverge bottlenecks [35]. The mechanism of these drops was initiated by a queue in the freeway shoulder lane. It was completed by disruptive lane changing that occurred as drivers maneuvered around this shoulder lane queue. However, it was argued that lane changing alone might not explain the capacity drop. In contrast, in [10], it was conjectured that capacity drop or the reduced flow is a consequence of the way drivers accelerate away from the queue. Moreover, in [9], [30], a long gradual acceleration zone was observed downstream to an active bottleneck.

Many efforts have been devoted to model the characteristics of capacity drop and behavioral mechanisms for its occurrence at active bottlenecks. Some researchers have tried to capture the phenomenon by using reverse lambda shaped fundamental diagrams [36], because the capacity drop is also featured by the observation of discontinuous flow-density relations that are usually described as reverse lambda [36], [37], [38], [39]. Nevertheless, it has been argued that such FD is incomplete. [10] explained that the reverse lambda shape is due to unobserved steady states. In other words, traffic states at certain ranges of density were not captured with the measurements. This is reasonable, because a discontinuous fundamental diagram definition would lead to infinite characteristic wave speeds [41], which contradicts the fact that information travels at a finite speed along traffic. Thereafter, other researchers are trying to reproduce capacity drop with a continuous FD. For example, in [42] the methodology to introduce capacity drop in a macroscopic
first-order model is to reduce the supply function of cells located at a downstream location of a congested one. In [43] a similar theory is developed, where a phenomenological kinematic wave theory is proposed. Nevertheless, these models lack to reproduce microscopic mechanisms and the drop ratio is an exogenous parameter (i.e. needs to be calibrated for each specific site). In [44] some other first-order traffic flow models that try to incorporate capacity drop phenomenon in their formulation are revised.

In order to take into account the vehicles behavior, some microscopic models try to reproduce heterogeneous lane behavior and lane changes. For example, [45] and [46] are able to replicate the occurrence and magnitude of capacity drop by considering moving bottlenecks. However, these models are usually complex. On the other hand, other models try to replicate capacity drop based on restricting the acceleration of vehicles. Many continuum models lead to infinite acceleration and deceleration rates, such as the LWR-model, which motivates the implementation of BA. However, since [47] very few continuum traffic flow models explicitly incorporate BA. The BA-LWR model developed by Lebacque [47], [48] is a two-phase continuum model. This model is difficult to apply and there is no numerical solution method for it. In [49], the BA-LWR model was extended for fixed and moving bottlenecks but is not able to reproduce capacity drop. Several BA criteria have been proposed in the literature. The constant bounded acceleration (CBA) with acceleration ratio of \( a_0 \) is the simplest one; other models include a speed-acceleration relationship e.g. TWOPAS model (Eq. 1.23) and Gipps model (Eq. 1.24), [50], [51]. In practice, any speed-acceleration function can be defined, as long as (i) the acceleration rate is non-negative, (ii) the acceleration rate is bounded and (iii) the acceleration-speed relation is non-increasing.

\[
A_{max}(v(t)) = a_0 \left( 1 - \frac{v(t)}{v_f} \right) \tag{1.23}
\]

\[
A_{max}(v(t)) = 2.5a_0 \left( 1 - \frac{v(t)}{v_f} \right) \sqrt{\frac{1}{40} + \frac{v(t)}{v_f}} \tag{1.24}
\]

The most recent study based on the LWR-model that includes BA [52], is able to reproduce capacity drop at a continuous lane-drop bottleneck. Jin demonstrates that the Riemann problem can be uniquely solved. This first-order model can reproduce capacity drop in an endogenous way. Moreover, a instantaneous continuous standing wave is observed between the LWR stationary states inside the lane-drop and the bounded acceleration stationary states in the downstream acceleration zones. After this first-order model is developed, in [3] a CF-model is derived with the methodology presented in [53]. It considers the BA-LWR model [48] and the lane-changing model [32]. Moreover, it is extended to other BA functions. It is behavioral, since the dropped capacity can be endogenously calculated from the FD, freeway geometry, acceleration process and lane-changes. This model considers important simplifications with the intention to identify in a clear manner the causes of capacity drop. It assumes (i) first-in-first-out through all lanes (i.e. no overtaking), such that the effective leader of each vehicle might be on a different lane, (ii) that all lanes are equally used (i.e. lane density for all lanes), and therefore all lanes have the same flow, (iii) non-constant BA (i.e. \( A_{max} = A(v) \)) but does not consider bounded deceleration. Hypothesis (ii) might not be true as a general rule, but it is needed to ensure the simplicity of
1.3. LITERATURE REVIEW

Precisely this CF-model [3] is the one that will be used in this MSc thesis to develop both the simulation tool and the analytical definition of optimal location of VSL. Thereafter, a deeper explanation of this model is presented in Chapter 2.

1.3.4 Current state of Art on Variable Speed Limits (VSL)

Several control strategies have been proposed in the literature to successfully increase the system’s discharge-flow rate at active bottlenecks where capacity drop occurs. The most popular traffic management strategy is the ramp metering control (RMC), which aims to improve the freeway performance by regulating the inflow from the on-ramps. RMC has been widely studied [14], but it fails to avoid the capacity drop when the mainstream demand is too high. Thus, a mainline traffic flow control strategy (MTFC) needs to be defined. The most spread of these strategies is VSL. It was first introduced in Germany in the early 1970s [5] and a decade later in the Netherlands [54]. Nowadays, VSL is a popular freeway traffic control in Europe and some of the American States [15]. Its goal is to limit road traffic speed in real time, by displaying a specific speed at the variable message signs (VMS). This speed is meant to adapt to different situations: weather, accidents, congestion, etc.

From the traffic operation perspective [17], two types of strategies with VSL can be found in the literature: (i) homogenization effect at a vehicular level (i.e. between lanes and within each lane) and (ii) restriction of mainline flow to prevent traffic breakdown. The first type of strategies impose speed limits around the critical speed (i.e. the speed at which capacity is observed) and are grounded on the belief that lower speed limits promote the reduction in speed, flow and occupancy differences. The reduction of pollution and accident rates by homogenization of the road is widely reported. Nevertheless, its effects on freeway efficiency raise controversy [55]. The main idea behind the second type of MTFC is to reduce the traffic flow entering a critical bottleneck by lowering the speed of the vehicles at an upstreamer location (i.e. substantially lower than critical speed), so as to prevent, eliminate, or delay the occurrence of capacity drop. This type of MTFC is considered a storing strategy and has similar characteristics as a RMC [7]. There are two types of implementation approaches that pursue this mainline metering. The literature review on this different implementations is extracted from [56], where the author of this work is second author:

- The first type of implementation ([57], [58], [7]) assumes that VSL can be used as a mainline metering mechanism by imposing very low speed limits upstream of bottleneck locations (i.e. down to 10km/h in [58] or around 20km/h in [7], [16]). This type of implementation that aims a permanent flow restriction to prevent capacity drop is usually designed in coordination with RMC. The simulations of this strategy in test corridors result in reductions of 20% of the total travel time. In [7] and [16] the authors use the macroscopic second-order traffic flow model included in the METANET simulator [59], [60] and [61]. Their model considers that a VMS will influence the fundamental diagram by linearly scaling down the flow – density diagram by the ratio $b$ (see Figure 1.9d). Other authors use
the same second-order traffic flow model but VSL is included by assuming that the same flow – density relationship prevails and that the desired speed is the minimum between the one corresponding to the experienced density and the other caused by the displayed speed limit (see Figure 1.9b), [57]. On the other hand, [58] use the AIMSUN microscopic traffic simulator [62], which implements the Gipps’ models for car following [51] and lane changing [63]. All these models consider that a significant flow restriction can be achieved by imposing low speed limits. Such expectation comes from: (i) the reasoning based on traffic flow principles, assuming that drivers’ behavior (i.e. the spacing-speed relationship) will not be modified by the existence of the speed limit, (ii) some empirical data in [6] where there seems to be a reduction in flow for the 40mph (i.e. 64km/h) speed limit, although details regarding such reduction are limited, as the authors were focusing on the capacity increase due to VSL, not on its reduction, and (iii) engineering experience showing that bottlenecks are created in some roads that operate at high speed limits (≥ 80km/h) followed by sections with lower enforced speed limits.

- The second type of implementation is based on an instantaneous imposition of very low speed limits over a extended freeway stretch. This type includes the SPECIALIST algorithm, developed and successfully tested in the Netherlands [64], [65]. By instantaneously lowering the speed limit, the flow is reduced proportionally while the density is kept constant. This type of reduction of flow is not permanent as the first type, but even though temporary it is enough in the conducted experiment to resolve shock-wave jams and recover capacity. A similar approach could also be applied to an infrastructural bottleneck with a fixed location [69].

The main drawback of VSL is the queue formation and consequent shock wave that propagates upstream (Figure 1.1). This queue might block off-ramps and worsen the traffic efficiency due to queue spillback mechanism [13], [16]. Therefore, it can be inferred that the location of the VSL application area is a crucial design issue: the closer the VSL application area to the bottleneck, the later will the queue affect the upstream off-ramps. Nevertheless, there is a common belief that vehicles need to accelerate to the free-flow speed before entering the bottleneck; thus the VSL application area cannot be too close to the bottleneck. This would mean that an optimal location of the VSL application area exists. Furthermore, in [67] the existence of an optimal location for VSL-control was observed, although their simulations were centered on use of VSL as a traffic harmonization management strategy (i.e. type (i) of VSL). This study was done with a microscopic model, based on cognitive risk-based simulation. On the other hand, flow-restriction MTFC studies in the literature are focused on feedback control algorithms [1], [7], [16], [65], but few is said about the existence of a required (nor optimal) distance from the VSL-control to the bottleneck. Moreover, the acceleration process exiting the VSL application area has not been studied either.

Some studies assume that ∼ 700m are needed to accelerate from the exit of the VSL application area to the bottleneck [16], [68]. This assumption (when cited) is based on Figure 2.7 of [69], which shows speed clusters measured at different locations downstream of a queue, indicating that the acceleration process exiting a queue might be stationary. The results suggest that ∼ 700m are needed to recover speed of ∼ 70km/h (in the analyzed site v_f ∼ 90km/h
1.3. LITERATURE REVIEW


Nevertheless, to the authors’ knowledge, there is no systematic study that confirms or discusses this data. In [7] it is assumed that vehicles must accelerate to the critical speed in order to pass the bottleneck without its activation. The necessary length to do so is also based on [69]. In a more recent study [66], the acceleration length from the end of the VSL application area to the bottleneck is computed to prevent “semi-congested” traffic from entering the bottleneck. Therefore the targeted speed before entering the bottleneck is the free-flow speed. They considered a constant acceleration ($\sim 0.5m/s^2$) and a $v_f \sim 110km/h$; varying VSL from a conservative 0km/h to 60km/h, the acceleration stretches varied considerably (i.e. 350m to 1km). The feedback-control VSL strategies in the literature usually impose minimum speed limit (e.g. 50km/h in [71], [72] or 40km/h [56]), thus following [66] discussion a convenient design would be of $L_u \sim 810m$, higher than the $\sim 700m$ proposed in [16], [68].

In summary, several assumptions have been done in the available literature on the required acceleration span between the control application area and the bottleneck. In particular, (i) it is assumed that a distance is needed between the VSL application area and the bottleneck, (ii) vehicles should be able to accelerate to a certain speed (usually the free-flow speed) before entering the bottleneck; (iii) as a consequence, the smaller the VSL, the larger distance is needed.
between the VSL application area and the bottleneck. Nevertheless, all these conjectures were
done without an analysis on whether this stretch is either necessary or sufficient. In this study,
however, it will be demonstrated that the first assumption is correct under some situations, but
the second and third assumptions are generally incorrect.


Chapter 2

Model and Methodology

2.1 The lane-drop bottleneck

The car-following model (CF-model) used to perform the simulations in this MSc thesis has been developed recently by Jin [3] and will be used in a lane-drop bottleneck. When the upstream part of the freeway has a higher number of lanes than in the downstream part of the road we talk about a lane-drop stretch. This lane-drop is a bottleneck itself due to the reduction of capacity. Many studies have already considered the lane-drop bottleneck problem and have tried to (a) reproduce capacity drop and (b) define a VSL control to reduce congestion, e.g. [73].

The flow in a lane-dropped road can be described in a generalized way as Eq. 2.1 [74], [43]. This can be particularized with TFD, Eq. 2.2. Note that $l(x)$ is not a constant function. The geometry of the lane-drop considered in this model is the following: The upstream link (with $l_1$ number of lanes) is considered from $x = -\infty$ to $x = 0$ and the lane drop zone, of length $L$, starts at $x = 0$ (Figure 1.1). The downstream link, consequently is defined from $x = L$ to $x = \infty$ and has $l_2 = l_1 - 1$ lanes. In other words, the number of lanes at each location can be described by Eq. 2.3.

\[
\begin{align*}
\phi(x,t) &= \begin{cases} 
D(x^-,t) & \text{if } D(x^-,t) \leq S(x^+,t) \\
\min\{S(x^+,t); C^-\} & \text{if } D(x^-,t) > S(x^+,t)
\end{cases} \quad (2.1) \\
\phi(x,t) &= \min\{v_f(x,t)k(x,t); (k_j \cdot l(x) - k(x,t))w\} \quad (2.2) \\
l(x) &= \max\left\{l_2; \min\left\{l_1; l_1 - \frac{x}{L} \Delta l\right\}\right\} \quad (2.3)
\end{align*}
\]
2.2 Hypothesis of the model

There are several things that must be addressed before presenting in detail the model formulation. This section of the thesis presents all the assumptions and implications that these hypothesis might have on the results of the simulation.

The first question that must be answered is why a microscopic simulation is needed. The macroscopic analysis is the most common approach to define different VSL strategies to increase flow discharge rate or recover the capacity drop. However, information about what is happening in real time to each vehicle is lost. With macroscopic analysis many insights of the VSL strategy might remain unnoticed. In order to be able to capture the temporal transition a microscopic model is needed. Considering stationary states from macroscopic models, as in [73], is a first step of the analysis, but a deeper understanding is aimed when using a CF-model. Additionally, the introduction of autonomous vehicles in the roadways is foreseen for the next decades. For this reason, a model that tracks individual vehicle’s trajectories is of special interest in order to develop an individual-vehicle control strategy in the future. In the same way that the microscopic model captures the temporal transition, developing this analysis in a lane drop bottleneck allow us to capture a spatial transition as well.

The first simplification in the model used [3] is to consider FIFO (first in first out) through all lanes. In this way, the effective leader of a vehicle might be on a different lane, see Figure 2.1. As a consequence, no overtaking maneuvers are considered in the model. Additionally, this means that all lanes travel at the same average speed. Moreover, it is assumed that all lanes are equally used, i.e. consider the same lane density for all lanes. Therefore, trough the fundamental relation of traffic flow theory, all lanes have the same flow. This hypothesis is not fulfilled in Europe, because it is compulsory to overtake vehicles on the left and thus the median lane is always faster than the shoulder lane. Nevertheless, this assumption might be more realistic in other places (i.e. USA), were no overtaking rules exist.

Another hypothesis is that lane changes do not influence traffic conditions. In other words, there is always enough space between vehicles to merge between lanes, when a lane drops. Nevertheless, in real life, if vehicles are to close to change lanes, a lane change might cause the following vehicles in other lanes to slow down. In this case, the model is not trustworthy for very high densities. A maximum acceptable density could be computed in order to determine when this model might not be representative any more. A different way to introduce the effect of lane changes is to consider a reduction of lanes due to lane-changing activity. This idea is presented.
2.2. HYPOTHESIS OF THE MODEL

in [33] and is based on the fact that a car is using two lanes during the lane changing maneuver. Therefore, the effective number of lanes at each location can be redefined considering the lane-changing intensity at that specific location, with equation (2.4). This effect is not considered in this MSc thesis and is left for further research, as an extension of the model.

\[ \hat{l}(x) = \frac{l(x)}{1 + \iota(x)} \]  

(2.4)

Traditionally, the speed is determined by the density, \( v = \eta(k) \). Nevertheless, in this model the speed is computed from the spacing (Eq. 2.5), The reason for this procedure is clear: the density is unknown without the vehicle’s trajectories (density is calculated from Edie’s formulas, as explained in Chapter 1.3.2) and those trajectories cannot be obtained unless the speed is known. In fact, the so called multilane spacing is considered instead of the classical spacing in this model. The classical spacing is the distance between the two front bumpers of consecutive vehicles. In this model, due to the FIFO assumption, the leader of a vehicle might not be in the same lane, thus the definition of the effective spacing is essential. The multilane spacing considers the number of lanes on the road stretch analyzed and the geometry (i.e. number of lanes, \( l_1 \) and \( l_2 \)). The basic idea to calculate this multilane spacing is to discretize the concept of average spacing of a stretch of highway to a specific vehicle spacing.

\[ v = \eta(k) = \eta(1/s) = \theta(s) \]  

(2.5)

The model is based on micro-simulation, therefore the spacing must be computed for each pair of vehicles. Assuming that the position of two consecutive vehicles is known: vehicle \( n \) is at \( x = a \) and the following vehicle \( (n+1) \) is at position \( x = b \). Then the multilane spacing can be easily computed by multiplying the distance between them \( (a - b) \) by the number of lanes, \( l \). This concept is slightly more complicated in the lane drop zone, as the number of lanes is not an integer. For this reason the concept of cumulative lane-miles is introduced (or cumulative lane-kilometers), which is defined by \( Z(x) \) in Eq. 2.6, where \( l(y) \) is the number of lanes in the highway at location \( y \). Thus the multilane spacing can be calculated from Eq. 2.7, where \( X(n) \) is the location of vehicle \( n \), for a specific pair of vehicles or from Eq. 2.8 if a finer discretization of vehicles \( (\Delta N) \) is considered, where \( X_N \) is the spacing and \( X(N) \) is the position of vehicle \( N \).

\[ Z(x) = \int_{y=0}^{y=x} l(y) dy \]  

(2.6)

\[ \xi(n + 1) = Z(X(n)) - Z(X(n + 1)) \]  

(2.7)

\[ \xi(N) = -l(X)X_N = \frac{Z(X(N - \Delta N)) - Z(X(N))}{\Delta N} \]  

(2.8)

The model considers bounded acceleration, \( A_{max} = A(v) \), [73]. This hypothesis consists on a modification of the two-phase model [18]. Similarly to [18], in [52] the speed of each vehicle
can be determined either by the LWR-model, considering the surrounding traffic conditions (i.e. multilane spacing,) or by its own conditions (BA) using Eq. 2.9. The main difference to \cite{48} is that the maximum acceleration is not necessarily a constant ($A_{max} \neq a_0$). When the speed is determined by the surrounding traffic, it can be considered the first-order phase of the model. Thereafter, the speed is determined by an algebraic equation. On the other hand, when the vehicle’s speed is bounded by $A(v)$, the speed is obtained from a differential equation. In this thesis, apart from the CBA model, TWOPAS and Gipps bounded acceleration models are considered, Eq. 1.23 and Eq. 1.24, respectively. On the other hand the model does not consider bounded deceleration, thus deceleration to the imposed speed limit is instantaneous.

\[
\begin{align*}
    v_t + v v_x & \leq A(v); \quad v = \theta(\xi) \\
    v_t + v v_x & = A(v); \quad v \leq \theta(\xi)
\end{align*}
\]  \hspace{1cm} (2.9)

Finally, a constraint on the VSL display needs to be considered. The speed limit, $u$, should be a discrete value, i.e. 45 $km/h$ or 60 $km/h$. This is the way to ensure that the drivers are able to read and understand easily the traffic signs. In this study the discrete nature of the speed limit is partially ignored. The VSL is always rounded to the lowest integer value (obtained from a continuous approach). Nevertheless, this is not necessarily a comprehensive value (i.e. a multiple of 5). This is not really relevant when I2V (infrastructure to vehicle) communications will be extended and when autonomous car penetration rate is high enough.

\section{2.3 Car-following formulation}

For simplicity, a triangular fundamental diagram (TFD) is assumed, where speed density relation is defined by Eq. 1.10. In this type of model, traffic states depend on three parameters: free-flow speed, $v_f$, per lane jam density, $k_j$, and the shock wave speed in congested traffic, $w$ (to refresh concepts see Figure 1.4 in the Introduction Chapter). In this thesis all three parameters are considered fixed for all simulations (see Table 2.1).

\begin{center}
\begin{tabular}{lcc}
Variable & Value & Units \\
$v_f$ & 30 & m/s \\
$w$ & 5 & m/s \\
$k_j$ & 1/7 & veh/m \\
\end{tabular}
\end{center}

\textbf{Table 2.1: TFD definition parameters}

The general formulation for homogeneous roads of the TFD (Eq. 1.11) can be particularized for a road with lane-drop by considering one upstream and one downstream diagrams. Inside the lane-drop intermediate TFD can be considered, since each location has a fixed number of lanes, $l(x)$, even if this number is not an integer, Eq. 2.10

\[Q = \phi(x,t) = \min\{v_f k(x,t), (k_j l(x) - k(x,t)) w\} \]  \hspace{1cm} (2.10)
2.3. CAR-FOLLOWING FORMULATION

The car-following model used in this thesis \[3\] is obtained from the conversion of the purely macroscopic LWR model, Eq. 1.6, to the microscopic model. This conversion method is developed in \[53\]. It is based on the nonstandard \[76\] second-order formulation of the LWR. After introducing the multilane spacing in LWR-model, the model presents macroscopic and microscopic features, Eq. 2.11, because it involves a macroscopic variable, \(k(t,x)\), and a microscopic variable, \(\xi(N)\). Moreover, the bounded acceleration (Eq. 2.9) is introduced to this nonstandard second-order formulation (Eq. 2.12).

\[
\frac{\partial k(t,x)}{\partial t} + \frac{\partial k(t,x)}{\partial t} \theta(\xi(N)) = 0 \tag{2.11}
\]

\[
v_t + vv_x = \min \left\{ \frac{\eta(k, \xi(N)) - v}{\epsilon} \right\} \tag{2.12}
\]

where \(\epsilon = \lim_{\Delta t \to 0^+} \Delta t\) is a hyper-real infinitesimal number used in nonstandard analysis.

According to \[53\], Eq. 2.12 is equivalent to Eq. 2.13 where \(\theta(\xi)\) is the speed-spacing relation, see Eq. 2.3 and Eq. 2.8. Using the symplectic Euler method to discretize the position, \(X(t,N)\), speed and acceleration can be written as Eq. 2.14. Where \(X(t,N)\) and \(X_{tt}(t,N)\) are the speed and acceleration rate of vehicle \(N\) at time \(t\), respectively. Then, setting \(\epsilon = \Delta t\) a time-discrete and vehicle-continuous car-following model is obtained (Eq. 2.15).

\[
X_{tt} = \min \left\{ A(v) ; \frac{\theta(\xi(N)) - X_t}{\epsilon} \right\} \tag{2.13}
\]

\[
\begin{align*}
X_{tt}(t,N) &= \frac{X_t(t+N, N) - X_t(t,N)}{\Delta t} \\
X_t(t,N) &= \frac{X(t+N, N) - X(t,N)}{\Delta t}
\end{align*}
\tag{2.14}
\]

\[
\begin{align*}
X_t(t+N, N) &= \min \left\{ X_t(t,N) + \Delta t \cdot A(X_t(t,N)); \theta(\xi(X(t,N))) \right\} \\
X(t+N, N) &= X(t+N, N) + \Delta t \cdot X_t(t+N, N)
\end{align*}
\tag{2.15}
\]

Furthermore, the multilane spacing, \(\xi(N)\), is discretized similar to Eq. 2.8. Thereafter, the car-following model becomes time- and vehicle-discrete, and can be written by Eq. 2.16. Where \(\theta(\xi)\) is defined by Eq. 2.17. Note that when considering equally used lanes, density is equal for all lanes and therefore the multilane spacing is Eq. 2.18.

\[
\begin{align*}
X_t(t+N, N) &= \min \left\{ X_t(t,N) + \Delta t \cdot A(X_t(t,N)); \theta \left( \frac{Z(X(t,N) - \Delta N) - Z(X(t,N))}{\Delta N} \right) \right\} \\
X(t+N, N) &= X(t+N, N) + \Delta t \cdot X_t(t+N, N)
\end{align*}
\tag{2.16}
\]
\[ \theta(\xi(x)) = \min \{ v_f; w(k_j \xi(x) - 1) \} \]  
(2.17)

\[ \xi(x) = s(x)l(x) = \frac{l(x)}{k(x)} \]  
(2.18)

### 2.4 Open-loop control definition

In an open-loop control system, the control signal \( u \) is constant (i.e. time-independent). In this study, the control signal is the speed limit imposed, i.e. \( VSL \) from now on. This means that the speed limit is not dependent on traffic conditions. Consequently, the control strategy will not be optimal, but depending on its definition, it is able to achieve the objective (to prevent, eliminate or delay the capacity drop). Since the interest of the present work is to study the optimal location of the control, we will center the research in studying whether the control is able to avoid capacity-drop or delay it, depending on its location and value.

The optimal location to place the variable message signs (VMS) is as close as possible to the lane-drop bottleneck and, on the other hand, at least at a distance \( L_u \) that allows to prevent traffic breakdown. As aforementioned, it has been thought that \( L_u \) is the required length to accelerate to a certain speed before entering the bottleneck. This speed has been assumed around the critical speed or the free-flow speed, depending on the source (see Section 1.5.2 for more details). Since we are studying traffic flow based on a triangular fundamental diagram (TFD), the critical speed is the free-flow speed.

The length needed to accelerate to a certain speed can be obtained from basic kinematic equations, if CBA is considered, Eq. 2.19 through 2.21. Moreover, considering the TFD parameters in Table 2.1, an acceleration ratio of 2\( m/s^2 \) and an initial speed \( (v_0) \) of 4\( m/s \), Eq. 2.21 leads to a required length to accelerate to free-flow speed of 220\( m \). With a lower acceleration rate (i.e. 1\( m/s^2 \)), \(|L_u|\) increases to 440\( m \); while targeting a speed of 20\( m/s \) (i.e. 70\( km/h \)), based on [66], reduces the required length to 95\( m \). Note that all these values are significantly shorter than the naive guess of \( \sim 700 \)\( m \) used in several sources [7], [16], [68]. With different simulations (i.e. different BA-models, acceleration ratios, \( a_0 \), and VSL application locations) we aim to determine which is required length to avoid capacity drop. Moreover, we will consider whether the geometry has also an influence on the necessary distance to avoid traffic breakdown.

\[ \begin{align*}  
x &= x_0 + v_0 t + \frac{1}{2} a(v) t^2 \\
v &= v_0 + a(v) t^2 \\
t &= \frac{v_f - VSL}{a_0} \\
L_{ff} &= -x_0 = \frac{v_f^2 - VSL^2}{2a_0} 
\end{align*} \]  
(2.19-2.21)
2.4. OPEN-LOOP CONTROL DEFINITION

The regimen inside the VSL application area is considered a congested state and the speed limit is used to control throughput (reduce it) for a period of time. The concept of reducing flow through reducing the speed limit, has been largely discussed (e.g. [56]), because few empirical studies support this idea. To achieve this goal very low speed limits should be imposed. The applicability of this strategy might be questionable when speeding is common in the selected roadway. Nevertheless, with the introduction of autonomous vehicles this strategies might be of easier application and will prove better results.

For the present research, it is essential to define a VSL that is able to accommodate downstream supply without ensuing capacity drop. Therefore, the main goal of the control is to reduce the supply in a section upstream from the bottleneck (i.e. control the demand at the bottleneck entrance). The control will generate a congested, but controlled (because queues and delays are known) traffic state. The election of such speed limit can be deduced from the TFD (Figure 2.2). It is straightforward to see that there is a maximum VSL for a specific downstream supply, this VSL\(_{\text{max}}\) is the speed that allows a throughput as high as the downstream capacity (i.e. \(C_2\)), see Eq. 2.22. For lower speed limits, the throughput that exits the control can be computed with Eq. 2.23.

\[
VSL_{\text{max}} = \frac{C_2 w}{k_j l_1 w - C_2} = \frac{l_2 v_f w}{l_1 (w + v_f) - l_2 v_f} \quad (2.22)
\]

\[
C_{VSL} = \frac{VSL \cdot k_j l_1}{VSL + w} \quad (2.23)
\]

Figure 2.2: Reduction of flow when imposing VSL on the upstream link of a lane-drop stretch. \(C_{VSL}\) and \(VSL_{\text{max}}\) definition.
2.5 Programming and simulations

2.5.1 Scenarios to analyze

We should analyze the situation without control and the system performance under different control strategies (i.e. different VSL values and $L_u$ locations). The simulations that need to be performed should be analyzed also for different initial conditions, i.e. demand levels, different freeway geometries and BA-models. The problem is defined by the boundary conditions (demand and supply in the upstream and downstream links). These conditions can be either steady boundary conditions (BC) or change over time (time-dependent demand). The strategy (i.e. VSL) that should be implemented depends mostly on these BC. A naturally solution for a link that is always oversaturated is to implement an open-loop control, as they are much more efficient. Nevertheless, during periods of changing demand it might useful to implement a closed-loop system, because it is difficult to estimate when the peak demand appears and when it is reduced.

Additionally, depending on the downstream congestion, the insights gained from the VSL analysis will be easier or more difficult to understand. If downstream is congested, the acceleration zone needed to avoid capacity drop might be different. On the other hand, if the discharge rate of downstream is lower than the dropped capacity, the phenomenon of capacity drop cannot be observed. Many difficult questions arise when considering downstream congestion. Consequently, the scenarios analyzed in this MSc thesis are different upstream demand levels and a constant downstream supply (i.e. $C_2$). If upstream demand is lower or equal to the downstream supply, no queues will be formed nor congestion will appear. These scenarios do not need to be analyzed, as there will bring no additional information.

### Table 2.2: Basic scenario to analyze. Geometry and acceleration ratio.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Value</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lanes upstream ($l_1$)</td>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td>Lanes downstream ($l_2$)</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>Lane drop length ($L$)</td>
<td>100</td>
<td>m</td>
</tr>
<tr>
<td>Acceleration ratio ($a_0$)</td>
<td>2</td>
<td>$m/s^2$</td>
</tr>
</tbody>
</table>

In Table 2.2 the base freeway geometry is defined and base acceleration ratio is chosen. These values represent the basic problem that will be analyzed. Additionally, other scenarios (i.e. changing geometry and BA-model) will be considered. To ensure the capture of the effect on $L_u^{opt}$ of different geometries, acceleration ratios or the BA-models, only one parameter will be changed at a time, keeping the other values as the basic-scenario described in Table 2.2. In is important to remember through all the document that the lane-drop is always considered of one lane reduction, i.e. $l_2 = l_1 - 1$. Therefore, the geometry of the stretch analyzed can be defined with only two parameters, $l_1$ and $L$.

The maximum acceleration rate can be defined as a function of the speed as in TWOPAS or Gipps models, see Eq. 1.23 and Eq. 1.24 or as a constant, i.e. CBA. Obviously, other FD...
definition will have different results, but the insights obtained from this specific analysis will can
easily extended to other configurations. Moreover, the main characteristics of the FD and traffic
flow can be computed for the basic-scenario, see Table 2.3.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Equation</th>
<th>Value</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Downstream critical density ($k_{c2}$)</td>
<td>$\frac{wk_jl_2}{v_f + w}$</td>
<td>0.0204</td>
<td>veh/m</td>
</tr>
<tr>
<td>Downstream capacity ($C_2$)</td>
<td>$v_f \cdot \frac{wk_jl_2}{v_f + w}$</td>
<td>0.6122</td>
<td>veh/s</td>
</tr>
<tr>
<td>Upstream capacity ($C_1$)</td>
<td>$v_f \cdot \frac{wk_jl_1}{v_f + w}$</td>
<td>1.224</td>
<td>veh/s</td>
</tr>
<tr>
<td>Per lane jam spacing ($s_{j}$)</td>
<td>$\frac{1}{k_j}$</td>
<td>7</td>
<td>m</td>
</tr>
</tbody>
</table>

Table 2.3: Freeway characteristics derived from basic-scenario and TFD parameters.

Steady boundary conditions

The scenarios that are specially interesting to be analyzed are the ones that activate the lane-drop bottleneck. This are presented in Table 2.4. With the values of Table 2.1 and Table 2.2
we can compute a first estimate of the required length, $L_u^0$, i.e. distance to accelerate from the
maximum speed limit ($VSL_{max}$) to free-flow speed, considering CBA. This expressio is derived
from Eq.2.21. It can be observed that this fist estimate of $L_u$, Eq. 2.24, does not depend on the
demand level, but on the $VSL$ and the downstream number of lanes, $l_2$.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Demand ($d$)</th>
<th>Supply ($s$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scenario 1</td>
<td>$1.1C_2$</td>
<td>$C_2$</td>
</tr>
<tr>
<td>Scenario 2</td>
<td>$1.5C_2$</td>
<td>$C_2$</td>
</tr>
<tr>
<td>Scenario 3</td>
<td>$2C_2$</td>
<td>$C_2$</td>
</tr>
</tbody>
</table>

Table 2.4: Scenarios analyzed that activate the lane-drop bottleneck

$$L_u^0 = \frac{1}{2a_0} \left( v_f^2 - \left( \frac{l_2v_fw}{l_1(w + v_f) - l_2v_f} \right)^2 \right)$$  \hspace{1cm} (2.24)

Considering steady boundary conditions and no initial congestion (neither upstream nor
downstream) is equivalent to consider the initial density conditions in Eq.2.25. These are the BC
that will be considered for the simulation program. However, since upstream demand is higher
than downstream supply (as in all scenarios proposed in Table 2.4), the density will increase
inside the lane-drop and queues will be formed. Some preliminary studies have been done for
initially congested downstream link (see Eq. 2.27), but the results exceed the scope of this thesis.
To consider congested states, the initial condition of density can be derived mathematically for
each location from the pair of $d(x)$ and $s(x)$, in Figure 2.3. To understand better the concepts
of demand and supply revise [74].
2.5. PROGRAMMING AND SIMULATIONS

$$k(0, x) = \begin{cases} 
  k_1 = \frac{d_1}{v_f} & \text{if } x < L_u \\
  0 & \text{if } L_u \leq x \leq L \\
  k_2 = \frac{d_2}{v_f} & \text{if } L < x
\end{cases}$$

(2.25)

Unsteady boundary conditions

Changing demand over time. This is a more realistic approach. Where demand will increase and then decrease again. The idea is to simulate peak of demand. This is specially interesting to analyze the closed-loop performance and important to be able to compute the total delay, when free-flow regime in the whole highway is recovered. This approach is left for further research, although some ideas have already been developed by the author during these months of research.

2.5.2 Languages used and code description

In this section of the MSc thesis I will explain the process of the program that I have developed to perform the simulations and produce results that will be analyzed in Chapter 3.

The simulations are done with $C^{++}$ programming and the post-process is done with Matlab. All the Figures in this thesis are produced from the own developed programming code, unless stated otherwise. The reason for choosing $C^{++}$ is to speed up the simulations. In order to gain accuracy, the time-step ($\Delta t$) is reduced a lot and therefore Matlab simulation times were too long.
2.5. PROGRAMMING AND SIMULATIONS

Code for no-control system

1. Definition of number of vehicles, duration of the simulation and $\Delta N$. Time step is defined by $\Delta N$ and the condition of no collision. See Eq. 2.30.

2. Definition of the VSL application area (it starts at $L_s = L_u - 50m$ and ends at $L_u$). But not relevant since the free-flow speed is displayed in VMS (i.e. no influence on traffic). Therefore, $VSL = v_f$.

3. Definition of initial density in upper and lower links (Eq. 2.25). Definition of maximum acceleration rate.

4. Computation of the location of the first vehicle. Its initial position is $x = L$. Computation of the speed of the first vehicle. This speed, $v_1$, will be constant through all the simulation period. $v_1$ is determined by the downstream situation: if there is downstream congestion $v_1 < v_f$ (Eq. 2.27). Otherwise $v_1 = v_f$.

5. Second vehicle's position at $t = 0$ is slightly upstream (i.e. 50m) of the start of the lane-drop zone ($x_2(0) = L_s - 50m$).

6. Definition of initial speed of second and other vehicles, $v_i(0)$, following Eq. 2.28. Upstream density, $k_1$, can be computed because we consider uncongested upstream link. Therefore, $v_i(0) = v_f \forall i$ vehicles.

7. The multilane spacing of the all other vehicles can be calculated from Eq. 2.18. Thereafter, the location of all vehicles at time step $t = 0$ is clearly defined. Note that there is a maximal density on the link, and consequently a minimal spacing $s_{j_u} = 7/2 = 3.5m$. Also note that minimal multilane spacing is equal to lane jam spacing $s_j = 1/k_j = 7m$.

8. Definition of 3 matrix: Position $X$, speed $X_t$ and acceleration rate $X_{tt}$. Each of them has as many rows as time steps and as many columns as vehicle discretization ($N/\Delta N$). The columns are not the actual vehicles analyzed, as we compute $X, X_t$ and $X_{tt}$ for each $\Delta N$ vehicle. Note that $\Delta N$ does not need to be necessarily an integer number. Indeed, $\Delta N < 1$. Different discretizations are analyzed. From $\Delta N = 1/2$ to $\Delta N = 1/64$.

9. First row of $X$ and $X_t$ can be filled with the speeds and locations computed in previous steps. Also the first column (since we know the speed and location for all time-steps for the first vehicle (the leader).

10. For next row (next time-step) the speed is computed considering BA and LWR stationary states (Eq. 2.16) for every vehicle.

11. Filling all rows for increasing time step $\Delta t$.

12. Compute density and flow, from Edie’s formulas, Eq. 1.17 [27].

$$VSL = \min \left\{ \frac{w s_2}{k_j l_1 w - s_2}; v_f \right\}$$  \hspace{1cm} (2.26)
\[ v_1(t) = v_1(s_2) = \min \left\{ \frac{ws_2}{k_1 s_2}; v_f \right\} \] (2.27)

\[ v_2(0) = \eta(k_1) = \min \left\{ w \left( \frac{l_1 k_j}{k_1} - 1 \right); v_f \right\} \] (2.28)

**Code for open-loop control system**

1. Definition of number of vehicles, duration of the simulation and \( \Delta N \). Time step is defined by \( \Delta N \) and the condition of no collision. See Eq. 2.30.

2. Definition of the VSL application area (it starts at \( L_s = L_u - 50 \text{m} \) and ends at \( L_u \)). We do not consider vehicles in the zone where we have no control (Eq. 2.25), because those vehicles would affect the efficiency of control.

3. Computation of VSL (Eq. 2.26). This mainly depends on the supply downstream \( (s_2) \) and the TFD characteristics (Table 2.1), because is chosen from macroscopic analysis. From Table 2.4 we know that \( VSL_{max} \) is Eq. 2.22. Nevertheless, due to numerical errors, we always impose a slightly lower VSL, by rounding down the speed limit to the nearest integer. For example, in the basic-scenario \( VSL_{max} = 13.5 \text{km/h} \) and the maximum VSL that will be considered in the simulations is \( VSL = 13 \text{km/h} \), i.e. a 3.7\% smaller.

4. Definition of initial density in upper and lower links (Eq. 2.25). Definition of maximum acceleration rate.

5. The location of the first vehicle at \( t = 0 \) is slightly upstream (i.e. 50\text{m}) of the start of the VSL-controlled zone \( (x_1(0) = L_s - 50 \text{m}) \). In order that all the vehicles are affected by the control. When analyzing an open-loop scenario we consider always downstream uncongested regime. Thereafter, those vehicles in downstream link will not affect the upstreamer vehicles and there is no need to model them.

6. Computation of the speed of vehicles. For initial condition, we consider free-flow states for the whole link. Therefore \( v_1 = v_2 = v_i = v_f \).

7. Second vehicle’s position at \( t = 0 \) depends on \( k_1(0) \). We compute the multilane spacing from the initial density \( (\xi_1) \) and define the location of the each vehicle at: \( x_{i+1}(0) = x_i(0) - \xi_1 \). Being \( \xi_1 \) computed by Eq. 2.18. Thereafter, the location of all vehicles at time step \( t = 0 \) is clearly defined. Note that the minimal multilane spacing is equal to lane jam spacing \( s_j = 7 \text{m} \). Nevertheless, the maximal density on the link may be higher (several lanes) and consequently a minimal critical spacing \( s_{min} = 7/l(x) \).

8. Definition of 3 matrix: Position \( X \), speed \( X_t \) and acceleration rate \( X_{tt} \). See no-control equivalent step. First row of \( X \) and \( X_t \) can be filled with the speeds and locations computed in previous steps.
9. First column of $X_t$ is computed considering $\xi \gg \xi_1$. Therefore, the leader travels at $v_1 = v_f$ except at the control application area, were he has $v_1 = VSL$ speed. Then, the position can be integrated for each time-step (filling first column of $X$).

10. For next rows for increasing time step $\Delta t$, the speed and location are computed considering BA and LWR stationary states (Eq. 2.16).

11. Compute density and flow, from Edie’s formulas, Eq. 1.17 [27].

### 2.6 Spatial and temporal discretization

The definition of $\Delta t$ and $\Delta N$ may influence the results of the computation. Thus, we need to define those values such that the model is reliable. A convergence study is done. Moreover, the relation between the time and space discretization depends on the condition of collision-free car-following model. Imposing a collision-free model is equivalent to impose that the multilane spacing is always equal or bigger than the jam spacing $\frac{1}{k_j}$. It can be proved that a car-following model is collision-free if and only if Eq. 2.29 is fulfilled, see [3] for more details. In other words, in the case of a TFD, the relation of $\Delta t$ and $\Delta N$ must fulfill Eq. 2.30.

\[
\frac{\Delta N}{\Delta t} \geq \max l(x) \cdot \max Q(k) \frac{1}{1 - \frac{k}{k_j}} \quad (2.29)
\]
\[
\Delta t \leq \frac{\Delta N}{\max l(x) \cdot w \cdot k_j} \quad (2.30)
\]

However, in this MSc thesis the relation chosen to ensure convergence is presented in Eq. 2.31. This is one to avoid high computational cost. To ensure convergence with the collision free condition (Eq. 2.30) $\Delta N$ needs to be reduced a lot. However, the key discretization for the convergence has been proven to be $\Delta t$. Therefore, changing the ratio between $\Delta t$ and $\Delta N$ is a easy way to ensure convergence without compromising the computational cost.

\[
\Delta t = \frac{\Delta N}{15l_1 w k_j} \quad (2.31)
\]
Chapter 3

Simulation Results

3.1 Simulation results without control

3.1.1 Capacity drop presence and magnitude

Aiming a better understanding of capacity drop and the traffic mechanisms that the model is able to reproduce, the simulation tool in a lane-drop freeway stretch is used for different demand scenarios, BA definitions and geometries of the freeway. The first question to be addressed is to determine the presence of the capacity drop phenomenon. It is particularly interesting to establish the drop magnitude and also the time when the capacity drop appears. The first question can be answered with an asymptotic analysis: compare the actual discharge rate to theoretical the capacity of the bottleneck, once the queue has been formed. The second question, is a time question, thus the time-discretization is crucial to ensure that the results have converged.

First, all the Scenarios (presented in Tables 3.1 to 3.4) are analyzed without control strategy, to understand the mechanisms behind a lane-drop bottleneck and later, an analysis of the control should be carried out for the same scenarios. Since we are interested in analyzing the capacity drop phenomenon, the scenarios analyzed will always have an uncongested downstream link, i.e. supply is downstream capacity \( C_d \). The basic scenarios are presented in Table 3.1 and are equivalent to the ones presented in Table 2.4

Depending on the initial density of the upstream link (i.e. \( k_1 \)), the shock wave that will be formed between the upstream and downstream stationary states will be different. Moreover, any scenario with \( k_1 \leq k_{c2} \) will not generate a queue nor the subsequent capacity drop, from a theoretical point of view. Thus, no backward travelling shock wave is observed. On the other hand, scenarios 1 through 3 (Table 3.1) will generate different shock waves. When the demand is higher the propagation speed of congestion will be faster. Since the inter of this part of the research is to analyze the start of congestion, the upstream link should have a initial free-flow stationary state. If it would be already congested, the mechanisms behind capacity drop would not be easy to read.
3.1. SIMULATION RESULTS WITHOUT CONTROL

Figure 3.1: Traffic performance and acceleration pattern for Scenario 2 with CF-model [3].

It is important to highlight that the scenarios presented in Table 3.1 are related to a specific geometry of the upstream and downstream links. Since $l_1 = 2$, all Scenarios 1 through 3 consider free-flow states in the upstream link. For other geometries, these high demands might generate congested states in upstream link, which lie behind the scope of this work. To analyze other geometries, other demand levels (i.e. upstream densities) should be considered, see Table 3.3.

For scenarios 1 through 3 the simulation results reveal that once capacity drop takes place, the throughput in the downstream link is reduced. Since acceleration is bounded, vehicles need to travel a certain length in order to accelerate to the free flow speed when they exit the lane-drop bottleneck, and a horizontal shock wave (i.e. standing wave) can be observed at the lane-drop zone.

In Figure 3.1, the Scenario 2 is analyzed. The BA model considered is CBA with acceleration ratio $a_0 = 2m/s^2$ (the basic scenario from Table 2.2). In Figure 3.1a, the shock-wave that separates the free-flow state in upstream link from the congestion in the same link is observed. It starts at $x = L$ and propagates the queue in the upstream direction. This shock wave is clearer...
3.1. SIMULATION RESULTS WITHOUT CONTROL

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Upstream density ( (k_1) )</th>
<th>Downstream density ( (k_2) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scenario 1</td>
<td>( 1.1 k_{cd} )</td>
<td>0</td>
</tr>
<tr>
<td>Scenario 2</td>
<td>( 1.5 k_{cd} )</td>
<td>0</td>
</tr>
<tr>
<td>Scenario 3</td>
<td>( 2 k_{cd} )</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 3.1: Basic scenarios to analyze without control.

than the horizontal stationary-wave, because deceleration is not bounded and vehicles reduce the speed instantaneously. Moreover, it can be observed that when the queue spills back \( x = 0 \), the states in the different parts of the freeway (upstream link, lane-drop zone and downstream link) become stationary. In Figures 3.1b and 3.1c, present the normalized flow rate and normalized speed, respectively. Flow is normalized over downstream capacity \( (C_2) \) and speed is normalized over free-flow speed \( (v_f) \). In the Figure 3.1b, we observe how capacity drop is present, since the flow at the downstream link is smaller than 1. The dropped ratio is not the same for the first vehicles than for the laters ones. We observe that the flow is stable after the shock wave (that propagates the congested states to upstream locations) crosses \( x = 0 \) m.

Comparing the traffic evolution from the simulation program with BA to another simulation case where acceleration is not bounded (see Figure 3.2), we observe that capacity drop is not reproduced in the second case. While unrealistic high acceleration rates are observed at the exit of the lane-drop in Figure 3.2d. This would mean that capacity drop is triggered by the way that vehicles accelerate after exiting the lane-drop, as [10] predicted. Thereafter, we confirm that the model used [3] and presented in detail in Chapter 2 reproduces capacity drop in an endogenous way.

It is of special interest to calculate the dropped ratio, this is done in a working paper [3] and therefore is not explained in detail this MSc thesis. [3] presents a formulation to compute the dropped capacity for the CBA model, Eq. 3.1. It proves that the capacity drop ratio only depends on the road geometry, including \( l(x) \) and \( L \), the fundamental diagram (free-flow speed, shock wave speed and jam density) and the bounded acceleration, \( a_0 \), and is independent from the initial conditions, or boundary conditions.

\[
a_0 = \frac{\Delta l w^3 k_j (C^-)^2}{L(l_2 w k_j - C^-)^3} \tag{3.1}
\]

As aforementioned, the analysis of convergence of capacity drop magnitude and time appearance is an important issue. Figure 3.3 shows how the capacity drop magnitude does not depend on the initial conditions of traffic, as already discussed in [3]. The convergence to a certain value is achieved when reducing the space discretization. Additionally, in the right hand side of Figure 3.3, the convergence to the time appearance of the capacity drop is presented. The appearance of the phenomenon depends on the BC and is faster for higher demand leves, as predicted.

A sensitivity analysis of the dropped ratio is done for a freeway stretch with \( l_1 = 2 \) upstream lanes. In Figure 3.4, the capacity dropped ratio (CDR) is shown for different lane-drop lengths and BA-models. The CDR is higher for TWOPAS and Gipps model than for CBA. Additionally, lower \( a_0 \) ensue higher drops on flow ratio, whereas longer lane-drop length, \( L \), ensue lower CDR.
3.1. SIMULATION RESULTS WITHOUT CONTROL

Figure 3.2: Traffic performance and acceleration pattern for Scenario 2. With unbounded acceleration.

From the author’s perspective, it is of special interest to analyze the speed function over space and the acceleration and deceleration patterns, for different vehicles. The following observations are present in any geometry and for all BA-models:

- The second vehicle, starts decelerating inside the lane-drop zone and starts accelerating exactly when it leaves the lane-drop zone. The locations where vehicles start to decelerate due to geometry and density conditions vary for different lane-drop lengths, but the longer the lane-drop is, the more downstream is this location.

- Since the deceleration process generates a backward travelling shock wave (Figure 3.1a), later vehicles decelerate at an upstreamer location from the point where the first vehicle did. The backward travelling shock-wave is better observed (a single and instantaneous deceleration) once the queue spills back from the lane-drop start. The time when this occurs depends on the geometry (i.e. requires more time for a longer lane-drop) and BA-model. Some preliminary results on these traffic dynamics are presented in Section 3.1.2.
3.1. SIMULATION RESULTS WITHOUT CONTROL

Figure 3.3: Convergence analysis of CDR and its time appearance. Compare different demand scenarios (1 to 3) for the basic geometry and CBA model.

Figure 3.4: Capacity drop analysis depending on lane-drop length, acceleration ratio, $a_0$, and BA model.

- The acceleration of each vehicle starts also at an upstreamer position and soon the acceleration rate becomes only space-dependent (stationary). The time when this occurs is when the queue spills back at $x = 0m$.

- When the stationary state is reached, the speed profile of the vehicles is concave inside the lane-drop and convex outside of it (Figure 3.1c). For non-constant BA the speed-profile is
flatter and the speed of the vehicles exiting the lane-drop is lower, because acceleration rates are lower for increasing speed. Since free-flow speed in never achieved in TWOPAS and Gipps models, the acceleration distance to free-flow speed is longer (theoretically infinite) than for CBA.

3.1.2 Traffic dynamics: preliminary results

Since we are using a car-following model, it is worth to analyze the transition states, even though its not the main focus of this MSc thesis. A way to do so is using cumulative curves of vehicles. This methodology was first introduced by [77]. In Figure 3.5 the cumulative count of vehicles is presented for Scenario 3. In this case, the cumulative vehicles curve is computed for section $x = 0$ and the relative shifted cumulative count of vehicles at $x = L$. By relative shifted it is meant the actually measured cumulative count of vehicles shifted to the left by the free flow travel time between the section analyzed (i.e. $x = L$) and the reference section (i.e. $x = 0m$). With this kind of figures different transition periods can be defined. A characteristic of this cumulative curves is that its slope represents the flow of vehicles through the section. It can be observed that from vehicles 2 to twenty-something the flow rate is slowly being reduced, until all vehicles experience the same throughput: the dropped capacity. Moreover, the horizontal distance between both curves represents the delay of each vehicle, which also increases with time until it is constant for all vehicles. Similar cumulative counts of vehicles can be found in Annex A1, Figures A.1 to A.3 for different demand levels.

The moment when in Figure 3.5 the slope of cumulative count at $x = 0m$ (blue) becomes parallel to the cumulative count at $x = 100m$ (red) is the time of spillback from the lane-drop zone to the upstream link. Note that this is the moment when all vehicles will have the same delay, i.e. a stationary state is achieved. On the other hand, we see that the downstream flow is low from the first moment. Therefore, it can be inferred that capacity drop appears almost instantaneously. For future research it will be interesting to study whether this is a characteristic of the model or from the geometry analyzed. It has been reported [30] that capacity drop appears approximately 10 minutes after the queues start to form (i.e. when vehicles start to experience some delay). Whether the measurement of this data was in the appropriate location or the model is too sensitive/reactive is still under research.

Another way to analyze the evolution of flow at different locations through time is to plot a figure for that location. This is done in the following Figures were the right vertical axes represents the time that each vehicle (i.e. $x$-axis) crosses the section and the left vertical axis represents the flow for each vehicle at that specific location. For a lane-drop length of 100m, Figures 3.6 and 3.7 represent respectively the analysis at location $x = L$ and $x = 0$. Thereafter, Figure 3.6 is a measure of the time when capacity drop appears (i.e. the second vehicle already experience a reduction in flow). Moreover, the black arrows in Figure 3.6 show the time where the CDR value becomes stable and does not change for later vehicles ($\sim 25s$). On the other hand, Figure 3.7 shows the flow rate for vehicles at the start of the lane-drop and measures the time when the spill-back arrives at this location. Here, we observe a change of slope in the blue circle-line when this occurs and some more time is needed in order that the dropped ratio is
3.1. SIMULATION RESULTS WITHOUT CONTROL

Figure 3.5: Cumulative curves at lane-drop start and end. Upstream demand is twice the downstream capacity (Scenario 3 without control).

observable. The dropped ratio is the same for both locations but is observed at different timings (we accept the $\sim 0.03\%$ difference as a numerical error).

Figure 3.6: Spill back of queue at $x = L = 100m$ (i.e. capacity drop appearance) for $a_0 = 2m/s^2$.

The same analysis can be done for different scenarios (shown in Figures 3.8 and 3.9) with different a lane-drop lengths. For a $L = 150m$ the dropped ratio is lower (as discussed with Figure 3.4) and it is observable later in time than the scenario with $L = 100m$. On the other hand, the time when the first reduction of flow is observed (i.e. Time Capacity Drop) is almost the same for both cases.
3.2 Simulation results with control

3.2.1 Application area location to avoid capacity drop

The main idea to avoid capacity drop using VSL is to reduce the throughput of vehicles before the lane-drop. The objective is to reduce the flow without penalizing much the users affected by the VSL-control. From the simulations without control, we know that when demand at the upstream part of the bottleneck is lower or equal than the downstream supply (i.e. capacity) the phenomenon of capacity drop is not triggered. To ensure that users are better-off with the delay caused by the control than the delay derived from the capacity drop, queue diagrams can be used. Queuing diagrams help determining the total delay for different control strategies and for no-control scenario.
3.2. SIMULATION RESULTS WITH CONTROL

The VSL imposed for this simulations is based on Eq. 2.22. To avoid numerical errors, the speed limit is imposed slightly lower than $V_{SL_{\text{max}}}$. The control is chosen slightly lower (13 km/h), since numerical errors in the model can lead to false capacity drop if the VSL is too close to the maximum value. Such numerical errors can also be contained at a VSL close to the maximum value if the time step is small enough, but this entails a really high computational cost. After performing some simulations with the car-following model and VSL control, we observe that when vehicles exit the control application area, they need to travel a certain length in order to accelerate to the free-flow speed. An horizontal shock wave can be observed. As a first tentative for the simulation scenarios, we consider $L_u^0$, see Eq. 2.24 and Table 3.2. When imposing $V_{SL} = 13 \text{ km/h} = 3.611 \text{ m/s}$ in the simulations at $L_u = 200 \text{ m} \approx L_u^0$, capacity drop is avoided for all Scenarios 1 through 3 (see Figure 3.10). Here a iteration process starts trying to define a closer location of the control to the lane-drop that still avoids capacity drop. This is done for all different scenarios, see an example in Annex 3.2.1 with Figures A.7 to A.10.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Upstream density ($k_u$)</th>
<th>$V_{SL_{\text{max}}}$ [km/h]</th>
<th>$L_u^0$ [m]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scenario 1</td>
<td>$1.1 k_{cd}$</td>
<td>13.5</td>
<td>221.5</td>
</tr>
<tr>
<td>Scenario 2</td>
<td>$1.5 k_{cd}$</td>
<td>13.5</td>
<td>221.5</td>
</tr>
<tr>
<td>Scenario 3</td>
<td>$2 k_{cd}$</td>
<td>13.5</td>
<td>221.5</td>
</tr>
</tbody>
</table>

Table 3.2: Scenarios to analyze with control

The main difference on traffic flow when using a VSL strategy that avoids capacity drop to the scenario without control is that in the first case, traffic states are stationary. Furthermore, we observe a standing wave at the end of the control $x = L_u$ and a rarefaction wave that propagates the queue upstream (depending on the demand level of the upstream link, see Figures A.11 to A.13 in Annex A.3.2)

Since the results show that the location does not depend on the demand level, but rather on the speed limit imposed, different scenarios than the proposed in Table 3.2 need to be defined and analyzed. An interesting analysis would be to change the geometry (number of lanes) and
3.2. SIMULATION RESULTS WITH CONTROL

Figure 3.10: Open-loop control with VSL application area at 200m upstream of lane-drop start, i.e. $L_u = -200m$. Base case with demand Scenario 2.

thereafter use different $VSL$. The required length to accelerate to free-flow speed, $L_{ff}$, will vary consequently (Table 3.3). Moreover, since CDR is influenced by the lane-drop length it is plausible that its avoidance might also be related to $L$. Nevertheless, this is a hypothesis that has never been explored in the past (mainly because the causes of the capacity drop ratio are not completely understood yet). Changing $L$ will not influence $VSL_{max}$, nor the necessary length to accelerate to free-flow speed, see Table 3.3. Nevertheless, the simulation results show that increasing $L$ reduces the required distance $|L_u|$ to prevent capacity drop. Other Scenarios worth analyzing are changing the BA model and acceleration ratio, since we have proven that the CDR also depends on them.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Demand $(d)$</th>
<th>$l_i$</th>
<th>$L$ [m]</th>
<th>$VSL_{max}$ [km/h]</th>
<th>$L_u^0$ [m]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scenario 4</td>
<td>$1.25C_2$</td>
<td>3</td>
<td>100</td>
<td>24</td>
<td>214</td>
</tr>
<tr>
<td>Scenario 5</td>
<td>$1.12C_2$</td>
<td>4</td>
<td>100</td>
<td>32.4</td>
<td>205</td>
</tr>
<tr>
<td>Scenario 6</td>
<td>$1.11C_2$</td>
<td>5</td>
<td>100</td>
<td>39.3</td>
<td>196</td>
</tr>
</tbody>
</table>

Table 3.3: Scenarios with different geometries
3.2. SIMULATION RESULTS WITH CONTROL

Scenario Demand \( (d) \) \( l_1 \) \( VSL_{\text{max}} \) \( \text{km/h} \) \( L_u^0 \) \( \text{m} \)
---
Scenario 2 1.5\( C_2 \) 2 100 13.5 221.5
Scenario 2a 1.5\( C_2 \) 2 150 13.5 221.5
Scenario 2b 1.5\( C_2 \) 2 250 13.5 221.5
Scenario 2c 1.5\( C_2 \) 2 500 13.5 221.5

Table 3.4: Scenarios with different lane-drop lengths and same demand and number of lanes \( (l_1 = 2) \)

For all simulation cases there is a critical distance \( L_u \) for which the capacity drop is avoided, but closer controls to the bottleneck did not succeed on preventing it. A summary of some of the simulations is presented in Table 3.5. The results show that the location of the VMS determines the relation and interaction between two different acceleration processes: (i) lane-drop acceleration due to queue formation upstream of the end of the lane-drop and (ii) exiting-control acceleration from the queue formed by the VSL control. For the constant geometry, increasing the acceleration ratio reduces the optimal distance, whereas considering TWOPAS or Gipps model increased the critical length with respect to the CBA-model (e.g. more than 350m for a 100m long lane-drop with \( l_1 = 2 \) stretch and \( a_0 = 2\text{m/s}^2 \)). Moreover, the most sensitive variable is the lane-drop length, \( L \). By increasing it, the necessary acceleration length before entering the lane-drop could be reduced, even to 0m in some cases.

<table>
<thead>
<tr>
<th>BA-model</th>
<th>Acceleration ratio</th>
<th>Lane-drop length</th>
<th>Required location of application area</th>
</tr>
</thead>
<tbody>
<tr>
<td>CBA</td>
<td>1 ( [\text{m/s}^2] )</td>
<td>100 ( [\text{m}] )</td>
<td>( \sim -240 )</td>
</tr>
<tr>
<td>CBA</td>
<td>2 ( [\text{m/s}^2] )</td>
<td>100 ( [\text{m}] )</td>
<td>( \sim -70 )</td>
</tr>
<tr>
<td>CBA</td>
<td>2 ( [\text{m/s}^2] )</td>
<td>150 ( [\text{m}] )</td>
<td>( \sim -20 )</td>
</tr>
<tr>
<td>CBA</td>
<td>2 ( [\text{m/s}^2] )</td>
<td>250 ( [\text{m}] )</td>
<td>0</td>
</tr>
<tr>
<td>CBA</td>
<td>2 ( [\text{m/s}^2] )</td>
<td>500 ( [\text{m}] )</td>
<td>0</td>
</tr>
<tr>
<td>CBA</td>
<td>3 ( [\text{m/s}^2] )</td>
<td>100 ( [\text{m}] )</td>
<td>( \sim -10 )</td>
</tr>
<tr>
<td>TWOPAS</td>
<td>2 ( [\text{m/s}^2] )</td>
<td>100 ( [\text{m}] )</td>
<td>( \sim -450 )</td>
</tr>
<tr>
<td>Gipps</td>
<td>2 ( [\text{m/s}^2] )</td>
<td>100 ( [\text{m}] )</td>
<td>( \sim -160 )</td>
</tr>
</tbody>
</table>

Table 3.5: Results of required \( L_u \) for different simulation scenarios (BA-models and lane-drop lengths).

3.2.2 Key insight to avoid capacity drop

The results for all geometries and BA-models are similar, and only one case will be analyzed in detail for the sake of brevity: a freeway stretch with \( l_1 = 2 \) and a continuous lane-drop of 100m with a CBA of \( 2\text{m/s}^2 \). From Eq. 2.22, the maximum speed limit that should be imposed is \( VSL_{\text{max}} = 3.75\text{m/s} \) (i.e. 13.5km/h). The results (Figure 3.11a, 3.11b) show how capacity drop is avoided when placing the control at \( L_u = -70\text{m} \), but not when the location is \( x = -60\text{m} \). This is a lower acceleration length than the theoretical 221.5m (to free-flow speed) but also than the theoretical 92m (to \( v = 70\text{km/h} \)).

Regarding Figure 3.11c,1, an interaction between two different acceleration processes is observed, when breakdown is not prevented. Furthermore, comparing it to the speed profile from
3.2. SIMULATION RESULTS WITH CONTROL

the avoided capacity drop case (i.e. Figure 3.11.2) reveals some interesting insights. The curvatures are characteristically different. In the first case, the speed profile is different for each vehicle until traffic becomes stationary. The traffic states become stationary when the queue spills back to the upstream link. At this point, the speed profile becomes concave inside the lane-drop and convex once the vehicle exits the lane-drop (similar to the case without control, Figure 3.11). This acceleration pattern can be divided in two zones: zone 1 form \( x = 0 \text{m} \) to \( x = L \) and zone 2 form \( x = L \) to \( x = L + L' \). This zones are equivalent to the lane-drop zone and acceleration zone described in Section 4 of [52]. In the second case, where capacity drop is prevented, the speed profile is convex at any location and is always stationary, the acceleration length is only one zone that starts at \( x = L_u \) and ends at \( x = L + L'' \).

Even more interesting insights are obtained when comparing several locations of the control (Figures 3.11.1-f). In these figures, \( t = 0 \text{s} \) belongs to the moment when the first vehicle (leader) crosses section \( x = 0 \text{m} \). Figures 3.11 and 3.11 show 20th-vehicle’s characteristics for different simulations: the speed profile (3.11d) and the flow and time when the 20th-vehicle of the simulation exists the lane-drop (3.11e).

In the control locations where capacity drop is avoided, the speed profile is convex during the whole acceleration process (before, inside and after the lane-drop). On the other hand, in the cases where traffic breakdown occurs, the speed profile of the vehicles is initially convex, but at a certain location it suddenly falls to a lower value and starts increasing in a concave shape (Figure 3.11d). It is not exactly the same for all cases, but lies always inside the lane-drop zone. This location turns out to be the point where the speed in the CF-model is determined by the LWR-model (because it is smaller than the speed that the BA-model would describe). The sudden fall in speed can be explained as a flow reduction (i.e. capacity drop appearance). This change is sudden, because the speed in the model is defined by a minimum of two functions (Eq. 2.16), which does not consider transition-states. In other words, the speed is determined either by the steady states defined by the LWR-model (i.e. TFD) or by the stationary states in the BA-model. When the vehicle exits the lane-drop zone, the acceleration process changes and the speed profile is convex again.

In Figure 3.11 it is shown that when the control does not prevent the breakdown, the CDR is not affected by the VMS-location. On the other hand, the location of the control does affect the time appearance of capacity drop, but not significantly, i.e. \( \approx 2 \text{s} \) delay, when shifting the control 50m upstream. Therefore, this delay does not affect significantly the time when a vehicle exits the lane-drop (Figure 3.11b). Thereafter, control strategies aiming to delay capacity drop might be called in question.

When analyzing the effect of control at locations with \( L_u < -70 \text{m} \) in 3.11, the flow ratio in the downstream link is slightly lower than its capacity (i.e. \( \approx 2\% \) smaller, see 3.11), but this is expected since the VSL chosen is slightly under the theoretical \( VSL_{\text{max}} \). As expected, the avoidance of capacity drop allows a reduction of the time needed to cross section \( x = L \) (Figure 3.11) and, consequently, any other section downstream from the control. This reduction is \( \approx 10 \text{s} \) for vehicle 20th, but over half a minute for vehicles 80th and further. Consequently, the overall time savings are noteworthy.
3.2. SIMULATION RESULTS WITH CONTROL

Figure 3.11: Main simulation results observations.
Chapter 4

Analytical formulation for optimal control location

4.1 General problem statement

From the simulations, it has been depicted that as long as the speed of the vehicles inside the lane-drop zone is determined by the BA-model and not by the LWR-stationary states, the capacity drop can be avoided. If this is achieved, the traffic states inside and downstream from the control are stationary, i.e. they are time-independent. Moreover, it can be shown that flow is also space-independent from the conservation of vehicles (Eq. 4.1). The fundamental variables of these stationary states can be calculated, considering that the flow is determined by the imposed speed limit (Eq. 2.23) and the acceleration in the stationary state is defined by the BA-model (Eq. 4.2). Therefore, the speed can be obtained by solving the ordinary differential equation (ODE), whose initial condition is defined by the speed limit at the unknown location $L_u$.

$$\frac{dk(x)}{dt} + \frac{dq(x)}{dx} = 0 \Rightarrow \frac{dq(x)}{dx} = 0 \quad (4.1)$$

$$v_s(x)\frac{dv_s(x)}{dx} = a_s(x) = A(v_s(x)) \quad (4.2)$$

For the speed to be determined by the BA-model, the congested speed of the TFD (i.e. associated to LWR-model) must be higher than the speed defined by the BA-model, for each vehicle and at each location (Eq. 2.16). In other words, Eq. 4.3 must be fulfilled for every density, $k_s(x)$. Since there will be a stationary state, the stationary density can be calculated from the fundamental relation of traffic (Eq. 1.1). Developing these equations allows to state the general optimization problem for $L_u$ (Eq. 4.1). This formulation is a complicated minimization problem with an ODE whose initial condition (IC) is precisely the function to minimize.
v_\ast(x) \leq V(k_\ast(x)) = w\left(\frac{l(x)k_j}{k_\ast(x)} - 1\right) = w\left(\frac{l(x)k_jv_\ast(x)}{CVSL} - 1\right) \quad (4.3)

L_{\text{opt}}^u = \min_{L_u} |L_u|

s.t.
\frac{dv_\ast(x)}{dx} = \frac{A(v_\ast(x))}{v_\ast(x)}

v_\ast(L_u) = VSL

v_\ast(x) \geq \Psi(x) = \frac{wCVSL}{wl(x)l_j - CVSL}, \forall x \in [L_u, L] \quad (4.4)

4.2 Analytical solution for CBA

In the case of CBA, the non-linear ODE can be integrated directly considering the initial condition \(v_\ast(L_u) = VSL\), resulting in Eq. 4.5. Introducing Eq. 2.23 to the third condition in Eq. 4.4, we obtain that Eq. 4.6 must be fulfilled for all \(x \in [L_u, L]\). The minimization problem will lead to \(L_{\text{opt}}^u = \min \{f(x)\}\), see Eq. 4.7. For the scenario with speed limit \(VSL = 13\text{km/h} = 3.61\text{m/s}\) and geometry with \(l_1 = 2\) and \(L = 100\text{m}\) we can compare the analytical results, see Figure 4.1, with the simulation results studied in detail in Section 3.2.2. In the corresponding \(f(x)\) we observe that the location that avoids capacity drop is exactly \(L_u = -65.75\text{m}\).

\[v_\ast(x) = \sqrt{\frac{VSL^2 + 2a_0(x - L_u)}{2a_0}}\]  

\[\sqrt{VSL^2 + 2a_0(x - L_u)} \geq \frac{wCVSL}{wl(x)l_j - CVSL} = \frac{wVSLl_1}{l(x)(VSL + w) - l_1VSL} \quad (4.6)\]

\[L_{\text{opt}}^u \leq f(x) = x + \frac{VSL^2}{2a_0}\left(1 - \left(\frac{wl_1}{(VSL + w)(l(x) - l_1VSL)}\right)^2\right), \forall x \in [L_u, L] \quad (4.7)\]

Conceptually, the end of the control application area should be at \(x \leq 0\), thereafter \(L_{\text{opt}}^u \leq 0\). Since \(f(x)\) is convex in \(x\) and it can be proven that \(f(x = 0) = 0\), the interest on fulfilling Eq. 4.7 is actually restricted to \(x = L\). Therefore, a function of \(L_u(L)\) can be defined, when considering the envelope of different lane-drop lengths functions \(f(x = L)\). It can be concluded that \(L_{\text{opt}}^u\) is defined by \(f(x = L)\), see Eq. 4.8 and Figure 4.1.

It is clear that \(L_{\text{opt}}^u\) linearly increases with \(L\) until \(L_{\text{opt}}^u = 0\). Moreover, this envelope has slope 1:1, which means that a constant “critical” acceleration length, \(L_c\), exist. This critical distance is is the distance between the end of the control application area and the end of the lane-drop stretch. It is a constant once the acceleration rate (\(a_0\)), number of lanes and \(VSL\) are defined. For the studied configuration, this required length, \(L_c = 165.7\text{m}\), is shorter than the
4.2. ANALYTICAL SOLUTION FOR CBA

4.2.1 Sensitivity analysis of $L_u^{opt}$ for CBA model

Considering different lane-drop geometries with the CBA $a_0 = 2m/s^2$, and varying the control speed ($VSL$), several observations can be highlighted. Figure 4.2a shows the $VSL$ used for the length to accelerate from $VSL$ to free-flow speed (i.e. 221.7m). Moreover, this $L_c$ is proven to be the required acceleration length from $VSL$ to the congested speed of the downstream fundamental diagram associated to the throughput that the control defines, i.e. $C_{VSL}$. This targeted speed, $v_2$, can be computed by Eq. 4.9 and it is graphically presented in Figure 2.2. Thereafter, vehicles do not necessarily need to reach free-flow speed when they enter the lane-drop, nor when they exit it.

$$ L_u^{opt} = \min \left\{ 0, L + \frac{VSL^2}{2a_0} \left( 1 - \left( \frac{wl_1}{(VSL + w)(l_2 - l_1VSL)} \right)^2 \right) \right\} \tag{4.8} $$

$$ v_2 = \frac{wl_1VSL}{l_2(w + VSL) - l_1VSL} \tag{4.9} $$

Moreover, considering $v_2$, the third condition in 4.4 can be rewritten, since $C_{VSL}$ can be described with $v_2$ and $l_2$. See condition in Eq. 4.10

$$ \Psi(x) = \frac{wl_1VSL}{l(x)(w + VSL) - l_1VSL} = \frac{wl_2v_2}{l(x)(w + v_2) - l_2v_2} \tag{4.10} $$
4.2. ANALYTICAL SOLUTION FOR CBA

Figure 4.2: For different VSL (a), (b) shows the associated normalized stationary flow $C_{VSL}$, (c) shows the distance needed to accelerate to $v_f$, (d) shows the critical length, $L_c$, to accelerate from $VSL$ to $v_2$.

computation of the corresponding critical lengths (4.2d). For each lane-drop configuration (i.e. number of upstream lanes) the maximum speed limit is computed from Eq. 2.22 and some other speed limits are computed as a percentage of this $VSL_{max}$. Note that the case analyzed in Section 3.2.2 would be a point in the blue diamond-line close to $0.96VSL_{max}$. Reducing the speed limit will reduce the flow-rate at downstream locations (see Figure 4.2b). This fact is more notorious for freeways with high number of lanes. Additionally, lowering $VSL$ will increase the necessary length to accelerate to free-flow speed, see 4.2c. Fortunately, lowering the imposed speed limit reduces the actual required length to accelerate to a speed where downstream traffic is stable. Moreover, this critical distance is surprisingly independent on the freeway geometry. Consequently, the critical distance to avoid capacity drop is smaller with lower VSL and does not depend much on the number of lanes.

Regarding Figure 4.2, some observations should be highlighted: (i) when the speed imposed is $VSL_{max}$ the more lanes, the shorter $L_c$ is, (ii) for these speed limits the length to accelerate to $v_f$ and the critical length to avoid capacity drop are the same (iii) around $VSL = 0.95VSL_{max}$
4.3. NEW FORMULATION FOR OPTIMAL APPLICATION AREA

all road geometries require the same acceleration stretch $L_c \sim 160m$ to avoid capacity drop, (iv) for lower VSL the fewer number of lanes, the shorter the required acceleration stretch is and (v) for freeways with $l_1 \geq 4$ the optimal application area does not vary substantially with the geometry.

Considering a different acceleration ratio for the CBA model modifies the required distance between the lane-drop and the VSL application area. This is an intuitive result, but the point to which the definition of $a_0$ is critical is surprising. The required length may vary over 230m, when changing $a_0$ from 1$m/s^2$ to 3$m/s^2$. Indeed, in Eq. [4.8] it can be observed that the critical length is inversely proportional to the acceleration ratio.

Since the BA in real life decreases with increasing speed, results with CBA $2m/s^2$ are rather optimistic, while results with lower CBA might be more realistic. In fact, a recent study [52] was able to calibrate the CBA considering real data from [70]. The acceleration ratio obtained was $0.4m/s^2$, which is surprisingly low, as the authors had predicted [70]. Since $a_0$ is inversely proportional to the critical length $|L_u^{opt}| + L$, the results of Figure 4.2 should be multiplied by $\sim 5$ to obtain more realistic values.

4.3 New formulation for optimal application area

In this section we will prove that the minimization problem statement in Eq. [4.4] is equivalent to a simpler mathematical problem. The main motivation to develop this new formulation is the observation that capacity drop is avoided when vehicles reach at least a speed of $v_2$ when they exit the lane-drop stretch. The definition of this speed and the critical length, $L_c$, might simplify the calculation of the optimal location. Whether this targeted speed is independent from the BA-model and that this criteria is necessary and sufficient to avoid capacity drop, needs to be proven mathematically. To do so, a good understanding of all the equations involved in the original problem statement is needed. Considering the analytical results of the CBA case and the shape of the analytical functions of speed (considering BA-model or LWR-model, see Figure 4.3), a list of several observations can be highlighted:

1. If $A(v_*(x))$ is a decreasing or constant function in speed and the speed $v_*(x)$ is increasing in $x$, the slope of $v_*(x)$ is decreasing on $x$, i.e. the speed profile is convex.

2. The speed, $v_*(x)$, increases to the free-flow speed, $v_f$, in the way that is defined in the BA-model. For TWOPAS and Gipps model, $v_f$ is never achieved and thus $v_*(x)$ will have an horizontal asymptote.

3. Since $A(v_*(x))$ does not depend on $x$, the speed profile $v_*(x)$ has a fixed shape and can be moved along $x$ depending on the boundary conditions (BC).

4. The initial condition $v_*(L_u) = VSL$ determines the start of the increasing speed.

5. It can be proven that $\Psi(x)$ is a continuous concave function between $0 < x < L$, since Eq. [2.3] is a linear decreasing function on $x$ inside the lane-drop ($x \in [0, L]$). See definition of $\Psi(x)$ in Eq. [4.10].
6. From Eq. 2.23 $\Psi(0) = VSL$ because $l(x = 0) = l_1$.

7. Considering the second part of Eq. 4.10 $\Psi(L) = v_2$. Where $v_2$ is the speed on the congested branch of the downstream TFD that allows a throughput of $C_{VSL}$. This speed ensures that the traffic flow is stable (but not necessarily uncongested) at the downstream link, see Figure 2.2.

8. From the third condition, we know that the maximum value of $L_u$ is 0, i.e. $L_u \leq 0$.

![Figure 4.3: Speed from the BA-model, $v_*(x)$ and minimum speed, $\Psi(x)$](image)

From observations 1 and 5 to 7, is easy to see that if $v_*(x = 0) \geq VSL$ and $v_*(x = L) \geq v_2$, the condition $v_*(x) \geq \Psi(x)$ is always true. Moreover, if $v_*(x = 0) \geq VSL$, the second restriction of Eq. 4.4 is automatically true. Thus, the minimization problem can be written as in Eq. 4.11:

$$L_{opt}^u = \min_{L_u} |L_u|$$

s.t.

$$\frac{dv_*(x)}{dx} = \frac{A(v_*(x))}{v_*(x)}$$

$$v_*(0) \geq VSL$$

$$v_*(L) \geq v_2$$

(4.11)

Moreover for smaller $L_u$ the stationary speed $v_*(x = L)$ will be greater, from observations 1 and 3. Therefore to minimize $|L_u|$ the condition should be $v_*(x = L) = v_2$, as long as $L_u$ is non-positive (condition 8). With all this discussion, the problem statement can be rewritten: $L_{opt}^u$ is the minimum value between 0 and the solution of $g(L_u) = 0$ (see Eq. 4.12).
4.3. NEW FORMULATION FOR OPTIMAL APPLICATION AREA

\[ g(L_{u}^{opt}) = v_{s}(L_{u}^{opt}) - VSL = 0 \]

where
\[ \frac{dv_{s}(x)}{dx} = \frac{A(v_{s}(x))}{v_{s}(x)} \text{ with } v_{s}(L) = v_{2} \] (4.12)

s.t. \( L_{u}^{opt} \leq 0 \)

In this section we have proven that, vehicles do not necessarily need to reach free-flow speed when they enter the lane-drop, nor when they exit it. The only case when the vehicles need to reach free-flow speed at the end of the lane-drop is when the speed of the control is the maximum speed that can be imposed \((VSL_{\text{max}})\). Moreover, it is clear that the longer the lane-drop is, the more flexible is the location of \( L_{u} \) (i.e. the optimal value is closer to 0m). Finally, the optimization problem for other BA-models, such as Gipps and TWOPAS, needs to be solved numerically.

### 4.3.1 Solution of new formulation for CBA case

In this section we prove that the results for the minimisation problem with ODE and initial condition on \( L_{u} \) is equivalent to the non-linear zero of function problem, for CBA. The solution of the ODE in (4.12) is solved in Eq. (4.13). Therefore, the nonlinear function to obtain \( L_{u}^{opt} \) is described in Eq. (4.14), which can be solved analytically (see Eq. (4.15)).

\[ v_{s}(x) = \sqrt{v_{2}^{2} + 2a_{0}(x - L)} \] (4.13)

\[ g(L_{u}^{opt}) = \sqrt{v_{2}^{2} + 2a_{0}(L_{u}^{opt} - L)} - VSL = 0 \] (4.14)

\[ L_{u}^{opt} = \frac{VSL^{2} - v_{2}^{2}}{2a_{0}} + L \] (4.15)

In the case study analyzed in detail, we considered the case with \( VSL = 13km/h = 3.611m/s \) and thus \( L_{u}^{opt} = -65.74m \), from Eq. (4.15). This is how the simulations had predicted and also the result of the minimization problem (see Figure 4.1a). On summary, for a CBA, \( L_{u}^{opt} \) depends on \( a_{0} \), the geometry (i.e. number of lanes and length), the shock wave speed and the control speed, since \( v_{2} = v_{2}(w, VSL, l_{1}, l_{2}) \).

### 4.3.2 Numerical solutions for TWOPAS and Gipps models

The simulation results presented in Chapter 3 showed that the required distance increases from \( L_{u} \sim 70m \) to \( L_{u} \sim 450m \) for TWOPAS model and to \( L_{u} \sim 160m \) for Gipps model when the acceleration ratio and the geometry are fixed (see Table 3.5). Changing \( a_{0} \) will affect the \( A(v) \) function and consequently further modify \( L_{u}^{opt} \). Using a simple numerical integration method to
4.3. NEW FORMULATION FOR OPTIMAL APPLICATION AREA

solve Eq. 4.12, the values from the simulation can be confirmed, see Figure 4.4 for TWOPAS model and Figure 4.5 for Gipps model.

To solve this numerical problem, Matlab is used since the computational cost is considerably low. From the ODE in Eq. 4.12, the following numerical explicit integration can be defined:

\[ v_*(x_{i+1}) = v_*(x_i) - A(v_*(x_i)) \frac{dx}{v_*(x_i)} \]

where \( v_*(x_0) = v_*(L) = v_2 \) and \( x_{i+1} = x_i - dx \) (4.16)

Since it is an explicit integration method a convergence analysis is required to ensure that the definition of \( dx \) is small enough to integrate with enough precision. The convergence study is presented in the Annex A with Figure A.14 and A.15 for TWOPAS and Gipps model, respectively. It is shown that with a 1m spatial discretization the relative error is less than 2%. However, since the computational cost is relatively low and to avoid doing a convergence analysis for all computations of \( L_u \) in different scenarios, the Figures of this MSc thesis are the result of considering a spatial discretization of 1mm.

It is important to highlight that \( VSL_{\text{max}} \) cannot be implemented in the theoretical TWOPAS and Gipps models, since in those models the free-flow speed is never achieved. This means that from a theoretical perspective, traffic breakdown could not be avoided. Moreover, in real life, vehicles can indeed accelerate to \( v_f \) and, consequently, the validity of those models might be put into doubt. From all this, it can be concluded that the definition of a proper \( A(v) \) is crucial to obtain the optimal control location. Thereafter, further research in this area is still elusive;
4.3. NEW FORMULATION FOR OPTIMAL APPLICATION AREA

Figure 4.5: Backwards integration from $v_2 = 25.99 \text{m/s}$ for Gipps model with $L = 100 \text{m}$ and $VSL = 3.611 \text{m/s}$.

from the authors perspective, the analysis of real trajectories and speeds in acceleration zones is of special interest.

4.3.3 Sensitivity analysis of $L_{opt}$ for different BA-models

In this section we analyze the $L_{opt}$ obtained from the numerical integration of the speed considering TWOPAS BA-model and Gipps model. Similarly to the CBA-model, when reducing the speed limit imposed, the required distance from the end of the VSL application area to the start of the lane-drop, decreases. Nevertheless, the influence of the parameters is not quantitatively the same.

For both models, the more lanes the freeway has, the smaller is $|L_{opt}|$, for a given speed limit. However, we also observe that the number of lanes has not a big influence on the critical length. See Figure A.16 in Annex A5. A specific example with the basic scenario (with lane-drop length of $L = 100 \text{m}$ and CBA with $2 \text{m/s}^2$) can also be found in the same Annex A5, see Figure A.17 for TWOPAS model, see Figure A.18 for Gipps model.

Consequently, the most influential parameters on the required distance to avoid capacity drop are the acceleration ratio and the speed limit imposed. In Figures 4.6 and 4.7 the effect of these two parameters on the critical length for Gipps and TWOPAS models, respectively, are presented. From Figures 4.6 and 4.7 it is important to highlight that:

i The tendencies of the critical distance, $L_c$, are similar for both BA-models.

ii For TWOPAS the value of $VSL$ is more influential than for Gipps model on the required
4.3. NEW FORMULATION FOR OPTIMAL APPLICATION AREA

$L_u$, since it has a more pronounced curvature (the contour line are closer to each other).

iii The distances required to avoid capacity drop are significantly shorter in Gipps model than for TWOPAS model (in some geometries less than half the distance).

iv The lower the acceleration ratio, the longer the required distance.

v The larger the acceleration ratio, the less influence it has on the required distance.

Figure 4.6: Effects of acceleration ratio, $a_0$, and speed limit ratio, $VSL/VSL_{max}$, on critical distance ($L_c = |L_u^opt| + L$). Results for Gipps model on freeway with $l_1 = 2$.

As an example of the second observation, we can observe again the Figures A.17 and A.18. In TWOPAS model, if $VSL \leq 0.7VSL_{max}$ the optimal location of the application area is 0m. On the other hand, for Gipps model, the optimal location of the control application area is 0m for $VSL \leq 0.8VSL_{max}$.

The left hand side Figure A.19 shows that in Gipps model, the lane-drop length has also a linear relation to $L_u^opt$, this is expected from the discussion in previous sections and is similar to the CBA and TWOPAS models. The acceleration ratio, however, is not linear to $L_u$, see Figure A.18 right. With this example we confirm observation (v) of Figure 4.8.
4.3. NEW FORMULATION FOR OPTIMAL APPLICATION AREA

Figure 4.7: Effects of acceleration ratio, $a_0$, and speed limit ratio, $V_{SL}/V_{SL_{max}}$, on critical distance ($L_c$). Results for TWOPAS model on freeway with $l_1 = 2$.

To compare TWOPAS and Gipps model results to CBA model we cannot compare directly with the acceleration ratio, $a_0$ (as done in Figure 4.8a). For TWOPAS or Gipps model the acceleration process is slower than for CBA model. Consequently, the acceleration length much larger. In order to compare critical lengths between models, we should define a new averaged acceleration ratio for TWOPAS and Gipps model. Two ideas are proposed to do so:

- Approximating the acceleration process by a spatial averaged acceleration, i.e. computing the secant of the speed-profile between the two known points: $v(L_u) = V_{SL}$ and $v(L) = v_2$. This spatial averaged acceleration is presented in Eq. 4.17 and Figure 4.8b. It has $s^{-1}$ units.

- Approximating the acceleration process by a CBA. To compute the acceleration ratio associated, $a_0^*$, Eq.4.15 can be used. Logically, this is equivalent to use CBA with a lower $a_0$. The results are presented in Figure 4.8c.

$$a_s = \frac{v_2^2 - V_{SL}^2}{L_c}$$

$$a_s = \frac{v_2^2 - V_{SL}^2}{L - L_{opt}^a}$$

(4.17)
4.3. NEW FORMULATION FOR OPTIMAL APPLICATION AREA

Figure 4.8: Comparing BA-models. Critical length, $L_u$, computed for $VSL = 0.95VSL_{max}$ and $l_1 = 2$ upstream lanes. (a) Effects of acceleration ratio, $a_0$, (b) Effects of an spatial-averaged acceleration, i.e. the slope of the secant line of $v_*(x)$ between $VSL$ and $v_2$. (c) Effects of an averaged speed, considering that for all simulations the results are from a CBA-model.

The most relevant result is that both approximations of a new $a_0$ (the first applied to the three models and the second applied only to TWOPAS and Gipps models) reveals that all BA-models are equivalent. In other words, for any Gipps model or TWOPAS model, we can find an $a_0$ for CBA that will result in the same optimal location. Therefore, it can be concluded that the BA model is not as relevant as expected. This makes sense, because the criteria to define $L_u^{opt}$ is to fulfill two conditions (at $x = L_u$ and at $x = L$) and the actual acceleration pattern (i.e. speed profile shape) is irrelevant while it is convex. This is a promising result, since calibrating a CBA model with data is much easier than calibrating other models (we do not need the speed along a trajectory but at two specific locations).

In summary, it has been shown that the two main factors that influence the critical length are the BA and the speed limit. In reality a proper definition of $a_0$ for a CBA is enough to ensure a good understanding of the acceleration process. Moreover, since there is a maximum speed limit related to a maximum distance between control and lane-drop, a conservative approach would be to design the VSL control application area to this $VSL_{max}$. Since the BA-models are equivalent when choosing the appropriate acceleration ratio, the definition of optimal location of application area is Eq. (4.18) when $a_0$ is appropitally calibrated.

$$L_u^{opt} = \frac{VSL^2 - v_2^2}{2a_0} + L$$  \hspace{1cm} (4.18)
Chapter 5

Conclusions

5.1 Summary and relevance of results

The objective of this thesis is to determine whether there is an optimal location of the VSL application area. Moreover, we aim to explain numerically and analytically how this optimal location ($L_{opt}$) can be defined. On one hand, the location of the control should avoid capacity drop in a lane-drop corridor. On the other hand it should be as close as possible to the bottleneck, in order to reduce the effects of queue spillbacks to close off-ramps.

To do so, a car-following model developed by Dr. Jin [3] is used to develop a simulation tool in C++. The model introduces BA as an additional restriction to the LWR-model. The results of the simulation prove that the model is able to reproduce capacity drop in an endogenous way. This drop ratio ($\approx 5 - 34\%$) depends on the freeway geometry and the BA-model. Then, an open-loop VSL-control to restrict the mainline flow in the interest of preventing, delaying or eliminating capacity drop is defined. This VSL is implemented in the simulation tool and shows how the location of the VSL application zone is crucial to achieve the avoidance of the capacity drop phenomenon. On the other hand the results reveal that the use of VSL to delay the appearance of the harmful phenomenon might be called into question. Since this work is centered on the design of the optimal application area of the control, a closed loop VSL-control to eliminate capacity drop is left for future research.

In Chapter 3, analyzing the speed profile, a better understanding of the mechanisms underlying traffic breakdown is achieved. When stationary traffic is achieved after the capacity drop appearance, the vehicles accelerate from the upstream congestion to the free-flow speed with a concave speed profile inside the lane-drop and a convex one outside of it. When the breakdown is avoided with the use of VSL, the vehicles accelerate outside of the control application area in a convex speed profile, which is stationary and can be obtained by solving the ODE in Eq. 4.2. It has been derived that ensuring that all vehicles have the convex speed profile (i.e. determined by the BA-model) is a way of preventing capacity-drop.
With the insights gained, in Chapter 4, a mathematical minimization problem of the distance between the control application area and the lane-drop is presented, where the variable to minimize is the initial condition of the ODE. Moreover, based on the analytical solution of the minimization problem with CBA, it has been proven that ensuring a convex stationary speed profile is equivalent to impose that the speed of the vehicles exiting the lane-drop is at least $v_2$, defined by the downstream geometry and the control speed in Eq. [4.9] This allows an easier formulation of the problem that can be solved analytically for CBA and numerically with an explicit integration and a non-linear zero of functions for TWOPAS and Gipps model. A sensitivity analysis of the optimal application area is done for the different BA-models. It is shown that for each freeway geometry, $a_0$ and $VSL$, a minimum constant span of distance $L+ |L_0^\text{opt}|$ can be defined to avoid capacity drop, this is called the critical length $L_c$. In the case of CBA the critical length is not very sensitive to the number of lanes of the freeway.

Furthermore, it is proven that the BA-model has a significant influence on the optimal distance between the control and the start of the lane-drop. The lower the BA ratio is, the longer must be this stretch. Moreover, BA-models that consider an acceleration-speed relation require even longer length than a CBA model (e.g. for a 100m lane-drop and a $a_0 = 2m/s^2$, TWOPAS requires $L_u = -414m$ and Gipps requires $L_u = -154m$, whereas for a CBA $L_u = -61m$ is sufficient). On the other hand, increasing the lane-drop length reduces $L_0^\text{opt}$ (even to 0m), which is one of the most promising results of this thesis. Because this means that VSL-controls may be located immediately upstream of the bottleneck if the lane-drop length is long enough. This “critical length” depends on the BA-model definition, which, for a specific CBA, can be approximated by only one parameter (the displayed VSL). Moreover, it has been shown that for lower speed limits, this $L_c$ is reduced and, consequently, the most restrictive lane-drop length can be calibrated from the freeway geometry (i.e. $VSL_{\text{max}}$ only depends on the freeway geometry and the TFD parameters). Moreover, since lane-drops are usually long transitions ($\sim 1km$), the VMS location of actual VSL implementations could be improved.

### 5.2 Main findings

From the simulation results in Chapter 3 and analytical formulation derived in Chapter 4 to compute the optimal location of the control application area it is proven that:

i There exists a minimum distance between the VSL application area and the bottleneck in order to prevent the occurrence of capacity drop, even though this minimum distance could be zero under certain conditions.

ii Vehicles do not need to accelerate to the free-flow speed before entering the bottleneck, but can continue their acceleration process inside and downstream of the bottleneck.

iii The larger the speed limit, the longer is the critical distance to avoid capacity drop, and designing the VSL application area should be based on the maximum VSL, not a really small VSL.
iv The condition to avoid capacity drop is that: (a) the speed is determined by the bounded acceleration, and (b) that the speed achieved at the end of the bottleneck is at least the speed, $v_2$, on the congested branch of the fundamental diagram associated to the flow restricted by the speed limit, i.e. $C_{VSL}$.

v Different BA-models are equivalent when the appropriate acceleration ratio, $a_0$, is chosen. Therefore, the critical distance can be calculated analytically for the equivalent CBA with Eq. 4.18.

These main conclusions fill a gap on the literature and will help to define a better control, by improving a fundamental part of the VSL design: its location.

5.3 Discussion and suggestions for further research

The CF-model used in this thesis does not reproduce traffic instabilities (such as stop & go phenomena). This means that capacity drop must not necessarily be triggered by those instabilities but by other phenomena (e.g. such as the bounded acceleration). In the theoretical TWOPAS and Gipps models for BA the free-flow speed is never achieved. Thereafter the maximum speed limit defined by the bottleneck geometry cannot not be implemented, because from a theoretical perspective of the insights gained in this thesis, traffic breakdown could not be avoided. In real life, vehicles can indeed accelerate to $v_f$ and, consequently, the validity of those models might be put on doubt. Thereafter, the definition of a proper $A(v)$ is crucial to obtain the optimal control location. Moreover, it has been proven that to determine the critical length an equivalent CBA model can be defined. This is specially interesting because it is easier to calibrate and $L_{opt}^u$ can be obtained through an algebraic equation.

On the other hand, the hypotheses of the car-following model used in this paper (no overtaking, same flow distribution across lanes, TFD, instantaneous deceleration) are a simplification of a real world freeway performance. In the future, it is of interest to contrast the numerical results obtained in this work with other simulation programs to verify weather the results hold for other models.

Moreover, in the results presented in this thesis the effect of lane changes has not been taken into account. As discussed in [3], lane changes will reduce the effective number of lanes because vehicles occupy more than one lane for a specific period of time. In [3] it is proven that the location of this lane changes may affect the capacity drop ratio. It could be argued that the lane changing intensity, $\iota(x)$, may modify the effective lane-drop effect and therefore the optimal location of the VSL application area may be influenced by the shape of this function $\iota(x)$. Weather this is true and to which extent the location is affected by the lane changes is left for future research.

Finally, the results of this thesis motivate the author to develop a closed-loop control strategy based on the knowledge gained. The closed-loop strategy should consider VMS in several sections of the upstream link to implement the required speed limit (obtained from downstream and queue
conditions) at the associated optimal location. This will be a breakthrough on the implementation of VSL-control and is of special interest since the delay (due to information transmission) between the imposition of the VMS and its effect on traffic flow could be reduced. The design of a closed-loop strategy will also be useful for a future analysis considering connected vehicles. A high market penetration rate of vehicle-to-infrastructure (V2I) communication could be useful to pass the VSL control on the roadside to the connected vehicles through V2I technology. This leads to many open questions that will be addressed in future research.
Appendix A

Complementary Figures

A.1 Traffic dynamics: cumulative curves

In this section we present 3 graphs for the basic scenario of $l_1 = 2$ and $L = 100m$ with three different deland levels. The Figures present the cumulative curves at $x = 0m$ and $x = L = 100m$. The scenarios refered are related to the Table 3.1. The simulations done here are without control.

Figure A.1: Cumulative curves at $x = 0$ and $x = L$ for base case, Scenario 1.
Figure A.2: Cumulative curves at $x = 0$ and $x = L$ for base case, Scenario 2.

Figure A.3: Cumulative curves at $x = 0$ and $x = L$ for base case, Scenario 3.
A.2 Simulation results without control

In the following graphs we present 3 completeley different scenarios. Non of them presents any control strategy.

Figure A.4 has a lane-drop length of 250m and $l_1 = 2$ upstream lanes. The BA-model is of constant acceleration ratio of $a_0 = 2\text{m/s}^2$. Demand level is 50% higher than downstream capacity. This is Scenario 2b in Table 3.4 but without control.

Figure A.5 has a lane-drop length of 100m and $l_1 = 2$ upstream lanes. The BA-model is Gipps model with acceleration ratio of $a_0 = 2\text{m/s}^2$. Demand level is 50% higher than downstream capacity. This is Scenario 2 in Table 3.1 but with Gipps model.

Figure A.6 has a lane-drop length of 100m and $l_1 = 3$ upstream lanes. The CBA-model is considered with acceleration ratio of $a_0 = 2\text{m/s}^2$. Demand level is 25% higher than downstream capacity. This is Scenario 4 in Table 3.3 but with without control.
Figure A.5: Demand Scenario 3, with Gipps BA-model $a_0 = 2\text{m/s}^2$ and lane-drop length $L = 100\text{m}$. 
Figure A.6: Upstream number of lanes is $l_1 = 3$ and $d_1$ exceeds downstream capacity in 25%, with CBA $a_0 = 2\text{m/s}^2$ and lane-drop length $L = 100\text{m}$.
A.3 Results of simulation with control and CBA

A.3.1 Changing application area of VSL

Graphs that exemplify the influence on the location of the control to avoid capacity drop. The geometry used is $l_1 = 2$ and $L = 150m$, whereas the acceleration model used is CBA with acceleration ratio $a_0 = 2m/s^2$. Demand upstream is $1.5C_2$.

Figure A.7: Control located at $L_u = 0m$ and demand exceeding 50% of downstream supply.
Figure A.8: Control located at \( L_u = -10 \text{m} \) and demand exceeding 50\% of downstream supply.
Figure A.9: Control located at $L_u = -20m$ and demand exceeding 50% of downstream supply.
A.3. RESULTS OF SIMULATION WITH CONTROL AND CBA

Figure A.10: Control located at \( L_u = -30 \text{m} \) and demand exceeding 50% of downstream supply.
A.3. RESULTS OF SIMULATION WITH CONTROL AND CBA

A.3.2 Changing demand

Graphs that exemplify the effect of demand on the shock wave speed between upstream link and control location. The geometry and acceleration ratio is from the "base" scenario, i.e. $l_1 = 2$, $L = 100m$ and CBA with $a_0 = 2m/s^2$. The location of the control is $L_u = -50m$.

In none of the cases capacity drop can be avoided, since $|L_u| + L = 150m$. But we can observe the upstream propagation of the queue due to VSL control is different depending on the demand level.

![Graphs showing trajectories of vehicles, normalized speed, flow rate over space, and acceleration rate.](image)

Figure A.11: Control located at $L_u = -50m$ and demand exceeding 10% of downstream supply.
Figure A.12: Control located at $L_u = -50m$ and demand exceeding 50% of downstream supply.
Figure A.13: Control located at $L_u = -50\,m$ and demand exceeding 100% of downstream supply.
A.4 Convergence for TWOPAS and Gipps integration

In the left part of both figures, it is observed that $dx = 10m$ is clearly not a fine enough discretization, but $1m$ discretization starts to be reasonable (relative error less than 2%).

Figure A.14: Convergence study for numerical integration for Figure 4.4 TWOPAS model.

Figure A.15: Convergence study for numerical integration for Figure 4.5 Gipps model.
A.5 Analysis of optimal location for TWOPAS and Gipps models

Figure A.16: Effects of number of lanes $l_1$ and speed limit ratio $V_{SL}/V_{SL_{max}}$ on critical distance ($L_c = |L_{opt}^{\text{up}}| + L$). Results for (a) Gipps and (b) TWOPAS model. Acceleration ratio $a_0 = 2$

Figure A.17: Analysis of VSL and geometry influence on $L_{opt}^{up}$ for TWOPAS model with constant $a_0 = 2\text{m/s}^2$.

In Figure A.19, the results are computed for a freeway stretch with two lanes upstream and only one lane downstream (i.e. $l_1 = 2$) and different lane-drop lengths.
Figure A.18: Analysis of VSL and geometry influence on $L_{\text{opt}}^u$ for Gipps model with constant $a_0 = 2\text{m/s}^2$.

Figure A.19: Effects of acceleration ratio $a_0$ on $L_{\text{opt}}^u$ for Gipps model. Freeway stretch with $l_1 = 2$ and different lane-drop lengths.
Bibliography


