NON-EQUILIBRIUM DYNAMICS OF SECOND-ORDER TRAFFIC MODELS

A Dissertation
Submitted to
the Temple University Graduate Board

in Partial Fulfillment
of the Requirements for the Degree of
DOCTOR OF PHILOSOPHY

by
Rabie Ramadan
May 2020

Examing Committee Members:
Benjamin Seibold, Advisory Chair, Mathematics
Gillian Queisser, Mathematics
Isaac Klapper, Mathematics
Rodolfo Ruben Rosales, External Member, MIT
ABSTRACT

NON-EQUILIBRIUM DYNAMICS OF SECOND-ORDER TRAFFIC MODELS

Rabie Ramadan

DOCTOR OF PHILOSOPHY

Temple University, May 2020

Professor Benjamin Seibold, Chair

Even though first order LWR models have many limitations, they are still widely used in many engineering applications. Second-order models, on the other hand, address many of those limitations. Among second-order models, the inhomogeneous Aw-Rascle-Zhang (ARZ) model is well-received as its structure generates characteristic waves that make physical sense. The ARZ model — and other $2 \times 2$ hyperbolic systems with a relaxation term — possess a critical phase transition: whenever the sub-characteristic condition (SCC) is violated, uniform traffic flow is unstable, and small perturbations grow into nonlinear traveling waves, called jamitons. The case where the SCC is satisfied has been studied extensively. However, what is essentially unstudied is the question: which jamiton solutions are dynamically stable? To understand which stop-and-go traffic waves can arise through the dynamics of the model,
this question is critical. This dissertation first outlines the mathematical foundations of the ARZ model and its solutions, then presents a computational study illustrating which types of jamitons are dynamically stable, and which are not. After that, a procedure is presented that characterizes the stability of jamitons. The study reveals that a critical component of this analysis is the proper treatment of the perturbations to the shocks, and of the neighborhood of the sonic points. The insight gained from answering the question regarding the dynamical stability of jamitons has many applications. One particular application presented here is deriving an averaged model for the ARZ model. Such a model is as simple to solve (analytically and numerically) as the LWR model, but nevertheless captures the cumulative effects of jamitons regarding fuel consumption, total flow, and braking events.
ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor, Professor Benjamin Seibold, for his expertise, guidance, and patience throughout the research process and the writing of this thesis. Without his assistance, this work would not have been possible. I would also like to express my appreciation for his support outside of my research. His classes were inspiring and vital for my mathematical maturity, and his suggestions regarding my presentation skills, writing style and how to interact with other academics were very valuable.

I would also like to thank my committee members Professors Gillian Queisser and Isaac Klapper for their help and support for my research as well as my studies at Temple University. In addition, I want to give special thanks to the department chair, Professor Brian Rider, the graduate chair, Professor David Futer, the teaching coordinator, Professor Maria Lorenz, and Professor Daniel Szyld, from whom I learned a lot.

Lastly, I want to extend my gratitude to Professor Rodolfo Ruben Rosales, whom I had the pleasure of working with on several occasions. His expertise in non-linear waves was essential for the advancement of my research, and his ideas and insight were very valuable for this thesis.
To my mother and father,
# TABLE OF CONTENTS

ABSTRACT iv

ACKNOWLEDGEMENT vi

DEDICATION vii

LIST OF FIGURES x

LIST OF TABLES xiii

1 INTRODUCTION 1

2 MACROSCOPIC TRAFFIC MODELS 11
   2.1 The Fundamental Diagram 13
   2.2 First Order Macroscopic Models 18
   2.3 Second Order Models 20
      2.3.1 The Payne–Whitham Model 21
      2.3.2 The ARZ Model 23
   2.4 The Sub–Characteristic Condition (SCC) 25

3 JAMITONS 29
   3.1 Specific model functions 31
   3.2 Linear stability of uniform flow 33
   3.3 Traveling wave analysis and jamitons 41

4 COMPUTATIONAL STUDY OF JAMITON STABILITY 58
   4.1 The Godunov Method 58
   4.2 Numerical scheme for the ARZ model with relaxation term 61
   4.3 Results on the stability of jamitons 64
5 STABILITY ANALYSIS OF JAMITON SOLUTIONS 69
5.1 A Perturbation system for single-jamiton waves ............ 70
5.2 Qualitative characterization of the perturbation system .... 74
5.3 Fundamental challenges caused by the sonic point ........... 78
5.4 Quantitative results: Asymptotic stability .................... 80
5.5 Quantitative results: Transient growth ....................... 85

6 AVERAGING OF THE ARZ MODEL 91
6.1 A model for the averaged density ............................ 93
6.2 A fuel consumption model .................................... 100
6.3 Methods for quantifying ARZ cumulative effects using aARZ-1 106
   6.3.1 Fuel consumption ...................................... 106
   6.3.2 Total flow ............................................. 109
   6.3.3 Braking ............................................... 111
6.4 A more realistic example ....................................... 113
6.5 The shortcomings of the aARZ-1 model and possible directions 117

7 CONCLUSION 121
7.1 Summary ................................................... 121
7.2 Discussion ................................................... 125
7.3 Future directions ............................................. 127

REFERENCES 129
LIST OF FIGURES

2.1 Three different fundamental diagram functions: The brown solid line is a plot of the quadratic Greenshields FD, the dashed dark blue line is a plot of the triangular Newell-Daganzo FD, and the solid red line is a plot of a smoothed triangular FD. . . . . . . 17

3.1 Plots of the specific model functions $Q(\rho)$, $h(\rho)$ and $U(\rho)$ (a-c), as well as the derivatives $h'(\rho)$ and $U'(\rho)$ (d-e). A plot of $U''(\rho) + h'(\rho)$ is in panel (f), and it shows where the SCC is satisfied or violated. . . . . . . . . . . . . . . . . . . . . . . . . . . . . 34

3.2 Plots of the growth rate $g_{\tilde{\rho}}(k) = \text{Re}(\sigma)$ as a function of the wave number $k$, for different constant base states $\tilde{\rho}$, as well as the asymptotic growth rate $g_{\tilde{\rho}}^\infty$ as a function of $\tilde{\rho}$. . . . . . . . . . . . . . . . . . . . . . . . . . . . . 38

3.3 A jamiton with sonic density $\rho_S = 70 \text{veh/km}$ and length $L = 274.1 \text{m}$. $\rho(x)$ and $u(x)$ are plotted in the left panels. The jamiton is plotted on top of the fundamental diagram in the right panel. . . . . . . . . . . . . . . . . . . . . . . . . . . . . 48

3.4 A short jamiton with length $L = 15 \text{m}$, which contains 0.89 vehicles. $\rho(x)$ and $u(x)$ are plotted in the left panels. The jamiton is plotted on top of the fundamental diagram in the right panel. . . . . . . . . . . . . . . . . . . . . . . . . . . . . 49

3.5 A long jamiton with length $L > 1 \text{km}$ and with a nearly flat tail. $\rho(x)$ and $u(x)$ are plotted in the left panels. The jamiton is plotted on top of the fundamental diagram in the right panel. . . . . . . . . . . . . . . . . . . . . . . . . . . . . 50

3.6 A plot of some maximal jamitons (blue lines) in the fundamental diagram along with their corresponding sonic points. The magenta curve represents the maximal jamiton envelopes. . . . 51

3.7 Five different jamiton phase planes. The jamiton regions are enclosed by different boundaries and are shaded in blue. . . . . . . . . . . . . . . . . . . . . . . . . . . . . 55
4.1 Classification of 980 jamitons into stable and unstable, displayed in four different phase planes. In each plane, the dashed brown line represents the zero length jamiton, and the dark blue line is the limit of jamitons with infinite length. The two disconnected red regions correspond to the “splitting” and “merging” instabilities, respectively.

4.2 Stability classification in the phase plane ($\bar{\rho}, \bar{Q}$).

5.1 Coefficient functions for (5.10) and a jamiton with $v_S = 12.5\text{m/veh}$ and $v^+ = 8.9\text{m/veh}$. This jamiton has a length of 561m and contains 40 vehicles. Note that here we revert to physical units (vehicles) for the horizontal axis.

5.2 Illustration of the discretization used to approximate (5.10), as described in §5.4. The top (bottom) graphic shows the characteristic curves corresponding to the $u$ ($q$) variable. The $u$-characteristics expand away from the sonic point towards the domain boundaries (where the shock is). The scheme’s time step is selected so that the $q$-characteristics advance by $h$ per time step.

5.3 Classification of 980 jamitons into asymptotically stable vs. unstable, where asymptotic stability is given by $\rho(M) < 1$ (and instability by $\rho(M) > 1$). Here, for each jamiton, $M$ is the one-step update matrix that comes from a discretization with 8000 grid points. Note that the criterion used here can only detect “splitting” instabilities.

5.4 Classification of 980 jamitons according to the transient growth factor (5.23). Three levels of $F$ are displayed, with the thresholds at $F_1 = 10^5$ and $F_2 = 10^{15}$ to yield: stable if $F < F_1$, moderately unstable if $F_1 < F < F_2$, and unstable if $F_2 < F$.

6.1 A plot of an example simulation. The initial conditions are plotted in the top left panel. $\rho(x)$ and $\bar{\rho}(x)$ are plotted in the bottom left panel at the final time $t = 1200$ sec. A scatter plot of the average quantities at $t = 0$ and $t = 1200$ sec is plotted on top of a FD in the right panel.

6.2 A scatter plot of the average quantities at $t = 0$ and $t = 1200$ sec is plotted on top of a FD, summarizing five different simulations.

6.3 A simulation illustrating the comparison between the ARZ model (dark blue line), an LWR model (dotted brown line), and the aARZ-1 model (dotted red line). Plots are made at $t = 0$, $t = 1000$ sec, and $t = 2300$ sec.
6.4 Plots of the FC rate of the four vehicles considered in [8] vs. the vehicle velocity, and the fitted average polynomial $K(s)$ (solid red curve). ................................................................. 101
6.5 A plot of four level curves of the function $F_c(s, a)$. The curves represent the FC rate functions for four different acceleration levels versus speed. ..................................................... 104
6.6 A plot of $\bar{Q}(\bar{\rho}, \eta)$ for $\eta = 0$, $\eta = 0.25$, $\eta = 0.5$, $\eta = 0.75$, and $\eta = 1$. .......................................................... 119
LIST OF TABLES

3.1 A summary of some jamiton related quantities, units associated with them, and related formulas. ........................................ 53

6.1 A summary of total fuel consumption (Liters) resulting from the ARZ, LWR and aARZ-1 models. The three simulations were run until the final time $T_f = 1200$ sec, and are for a ring road of length 8 km. All three simulations share the same initial conditions. ........................................ 109

6.2 A summary of total flow (vehicles) at $x = 0$ resulting from the ARZ, LWR and aARZ-1 models. The three simulations were run until the final time $T_f = 1200$ sec, and are for a ring road of length 8 km. All three simulations share the same initial conditions. ........................................ 111

6.3 A summary of braking time per vehicle per hour (min/veh/hr) resulting from the ARZ, LWR and aARZ-1 models. The three simulations were run until the final time $T_f = 1200$ sec, and are for a ring road of length 8 km. All three simulations share the same initial conditions. ........................................ 113

6.4 A summary of the total fuel consumption, total vehicle flow at $x = 2$ km, and the braking time per hour of the average vehicle, resulting from the ARZ, LWR and aARZ-1 models. The three simulations were run until the final time $T_f = 2400$ sec, and are for a ring road of length 20 km. All three simulations share the same initial conditions. The naive aARZ-1 model is the same as the aARZ-1 model, but where it is treated as a simple LWR model when calculating TFC and BTR. ........................................ 115
6.5 A summary of the relative difference in fuel consumption, total vehicle flow at $x = 2$ km, and the braking time per hour of the average vehicle, between the ARZ model (considered to be the ground truth here) and the LWR model, the $\text{aARZ-1}$ model, and the naive version of the $\text{aARZ-1}$ model. These relative differences are for the results displayed in Table 6.4. . . . . . . 116
CHAPTER 1

INTRODUCTION

The transportation system of any nation is an essential element of its functioning infrastructure. Without it, basic economic axioms would not hold: the optimal distribution of resources and concentrated production of food, goods and energy would not be feasible. In addition, for any country, the transportation system is a vital component in the pursuit of developing a socially vibrant, well-integrated society.

One of the pillars of a complete transportation system is traffic. More than a hundred years ago, cars and other types of vehicles started appearing on our roads alongside horses [53]. Soon after that, they became an essential part of our everyday lives, requiring the development of many new roads and highways. Since then, studying and modeling traffic became a vital and relevant large scale exercise to ensure safety and efficiency of a big sector of the economy.
As a matter of fact, in the last century, exploring and understanding traffic flow dynamics became part of many scientific fields including physics [46], engineering [73] and mathematics [14].

The modeling of vehicular traffic flow via mathematical equations is a key building block in traffic simulation, state estimation, and control. It is also important to note that there are some key differences between the approaches taken by engineers and mathematicians in studying traffic flow. For instance, engineers are more concerned about the predictive accuracy and forecasting of future traffic states, while mathematicians are more interested in understanding the first principles of traffic flow. Moreover, in contrast with traffic engineering, which is an empirical data-driven field, the mathematical modeling approach to traffic flow uses traffic data as a guide to model more complicated traffic phenomena for which fitting statistical models fail to produce reliable results.

Zooming in on the mathematical approach, there is some diversity in the philosophy of modeling traffic flow. Over the last century, there has been several different ideas about how to model traffic. One could, for example, use microscopic/vehicle-based models [79, 69, 6], which are large ordinary differential equations (ODE) systems. In this framework, the behavior of each individual vehicle and its interactions with other vehicles is tracked. Or one could use a cellular description of traffic [68, 17], which divides a road or a
highway into cells and assigns laws on how vehicles advance through these cells. Or mesoscopic models \cite{36, 77, 39} can be used, which employ a statistical mechanics perspective. In this continuum setting, vehicle interactions are modeled in analogy to gas-kinetic theory. Other types of continuum models exist as well, including dispersive \cite{49, 48}, and viscous \cite{44, 45} models. Lastly, macroscopic models \cite{63, 84, 97, 75, 76, 54, 5}, which rely on the methods of continuum mechanics, can be used to model traffic. This can also be referred to as the fluid dynamical approach, because the partial differential equations (PDEs) that come out of this setting are very similar to the PDEs encountered in modeling fluid dynamics.

In microscopic models, each vehicle is its own entity, and the model tries to capture the interaction between all vehicles through the equations of motion. The second order ODE system below is a general microscopic traffic model with \(N\) vehicles:

\[
x''_j(t) = f(t, x_1, x'_1, \ldots, x_N, x'_N) \quad \text{for} \quad j = 1, 2, \ldots, N,
\]

where \(x_j(t)\) is the position of the \(j\)-th vehicle at time \(t\), and \(f\) is some function of time, and the position and velocity of all vehicles in the system. A particular case of the above model is the follow the leader model \cite{28}, where the acceleration of each vehicle depends on the states of the vehicle itself and the vehicle ahead. Such a model can be summarized as follows:

\[
x''_j(t) = f(t, x_j, x'_j, x_{j+1}, x'_{j+1}) \quad \text{for} \quad j = 1, 2, \ldots, N.
\]
In cellular traffic models, a road or a highway is modeled by a one-dimensional array with $M$ cells of length $l$ (the array could have open or periodic boundary conditions). Each cell may either be empty or occupied by a single vehicle at a given time. In addition, each vehicle has a velocity $v = n\Delta v$, where $n \in \{0, 1, 2, 3, \ldots, N_{\text{max}}\}$.

For a given initial configuration of this setup, this model evolves in time through an update process. First the velocity of each vehicle is updated based on how far or close the vehicle ahead of it is. After that, each vehicle advances a number of cells based on its velocity. It is also possible to add a velocity randomization step after every update to try to trigger instabilities in the model. Such a step could be that for each vehicle, we reduce its velocity by $\Delta v$ with probability $p$.

In mesoscopic traffic models, the state of vehicles are represented in phase space, where the coordinates are the position $x$ and the velocity $v$ of the vehicle. This type of modeling is motivated by the kinetic theory of gases [47], and likewise, a one-vehicle distribution function $f(x, v, t)$ is defined so that $\int_c^d \int_a^b f(x, v, t) \, dx \, dv$ gives, at time $t$, the number of vehicles in the road interval $[a, b]$ and in the velocity interval $[c, d]$.

In the case of a single-lane road without on and off-ramps, the one-vehicle distribution function satisfies the kinetic traffic equation:
\[ \frac{\partial f}{\partial t} + v \frac{\partial f}{\partial x} + \frac{\partial}{\partial v} (Af) = \Phi(f, f), \]

where \( \Phi \) is an interaction term that describes the deceleration processes due to the inability to take over slow moving vehicles, and \( A = A(x, t, v) \).

Microscopic models are resolved down to each individual vehicle and its interactions with others, which offers a high level of detail. However, the computational cost associated with them is very high, especially for systems with a large number of vehicles. On top of that, it is often hard for microscopic models to capture any global traffic features or dynamics. Cellular models offer a similar level of detail to microscopic models, but are computationally more efficient. Nevertheless, it is difficult to extract or infer much about driving behaviour from cellular models. Mesoscopic models offer a level of detail that is in between microscopic and macroscopic models. There is not a clear cut answer as to which type of modeling is best when it comes to modeling traffic. Each approach has its advantages and disadvantages, and choosing one approach over another depends on the modeling purpose [38].

The focus of this work will be on continuum models. In particular, macroscopic models which describe the spatio-temporal evolution of the vehicle density (and in some cases other field quantities as well) via hyperbolic conservation laws. Instead of tracking individual vehicles, as in microscopic models, or cells on a road, as in cellular models, macroscopic traffic models keep track of
aggregate quantities. The three main quantities used in these models are the traffic density $\rho$, the traffic velocity $u$, and the traffic flow rate $Q$.

Macroscopic models play a central role in traffic flow theory and practice for several reasons. Some of them are:

- Mathematically, other types of descriptions of traffic reduce or converge to macroscopic models in certain limits, including cellular models [2], gas-kinetic models [39, 2], and microscopic models [4].

- Practically, macroscopic models are best-suited for the estimation of traffic states [99, 104] (because of their ability to incorporate sparse GPS data [3, 37]), and for control [74].

- Computationally, a macroscopic description is a natural framework to upscale millions of vehicles to a cell-transmission model [17] with much fewer degrees of freedom (compared with microscopic or cellular models).

- Societally, macroscopic models are more desirable for privacy and data security concerns, because they do not resolve individual vehicles.

The archetype macroscopic model is the *Lighthill-Whitham-Richards (LWR) model* [63, 84]

$$\rho_t + Q(\rho)_x = 0.$$ 

This model describes the evolution of the vehicle density $\rho(x, t)$ where $x$ is the road position and $t$ is time. The *fundamental diagram (FD) function* $Q(\rho) =$
\( \rho U(\rho) \), where the *equilibrium velocity function* \( U(\rho) \) is the bulk flow velocity as a function of density, is motivated by the 1935 measurements by Greenshields [31]. Many other types of FDs have been proposed later [63, 29, 97, 70, 17], and they have been proved to be more accurate. As a matter of fact, real FD data exhibits a substantial spread in the congested regime [43]. More complex traffic models capture this spread [13, 88, 25, 24], but the LWR model does not. Yet, due to its simplicity it nevertheless is widely used. Moreover, as is discussed later in this work, the LWR model emerges as a reduced equation for more complex models.

Furthermore, another critical shortcoming of the LWR model is that it cannot reproduce the phantom traffic jam phenomenon. In this work, we focus on *second-order models* that augment the vehicle density \( \rho(x, t) \) by an independent field variable for the bulk traffic velocity \( u(x, t) \), and describe their evolution via a \( 2 \times 2 \) balance law system: a hyperbolic conservation law system with a relaxation term in the velocity equation. The *Payne-Whitham (PW) model* [75, 102]

\[
\begin{align*}
\rho_t + (\rho u)_x &= 0, \\
u_t + uu_x + p(\rho)_x / \rho &= \frac{1}{\tau} (U(\rho) - u)
\end{align*}
\]

was the first second-order model proposed. Here \( U(\rho) \) is the *desired velocity function*, and \( \tau \) is the relaxation time that determines how fast drivers adjust to their desired velocity \( U(\rho) \). The *traffic pressure* \( p(\rho) \) models preventive driving.
Even though the PW model does capture traffic waves accurately [27, 88], it is generally rejected [18] due to spurious shocks that overtake vehicles from behind. For that reason, other hyperbolic models are preferred. In this work, the focus will be on the inhomogeneous Aw-Rascle-Zhang (ARZ) model, which is another second order model. It reads as follows:

\[
\rho_t + (\rho u)_x = 0,
\]

\[
(u + h(\rho))_t + u(u + h(\rho))_x = \frac{1}{\tau}(U(\rho) - u).
\]

It is important however to note that the fundamental structure of a $2 \times 2$ hyperbolic system, with a relaxation in the second equation, is common to all models of interest in this study. Note that hyperbolic models in general do not resolve zones of strong braking, but rather approximate them by traveling discontinuities (shocks) whose dynamics are described by appropriate jump conditions [23].

Models with the structure described above possess a critical phase transition. If the sub-characteristic condition (SCC) is satisfied, then uniform flow is stable [101]. Conversely, when it is violated, uniform flow is unstable and nonlinear traveling wave solutions exist [59, 40, 71, 27, 88]. The case of the SCC satisfied is well studied [101, 102, 64, 11, 60]. This work will focus on understanding the behavior and stability of solutions when the SCC is violated.

This dissertation is organized as follows. In Chapter 2, the theory behind macroscopic traffic modeling is presented. The fundamental diagram idea is
discussed in §2.1, and derivations of first and second order models appear in §2.2 and §2.3, respectively. §2.4 covers the SCC, and its relationship with linear stability. Chapter 3 covers the traveling wave solutions of the ARZ model. There, §3.1 presents the choices of equations, parameters and model functions used, §3.2 characterizes the nature of instability to uniform flow, and §3.3 outlines a procedure to solve the ARZ model, as well as characterising those traveling wave solutions (jamitons).

In Chapter 4, a systematic computational study of the stability of jamitons is conducted. Developing a numerical scheme to solve the ARZ model is presented in §4.1 and §4.2, and the actual results of the study are given in §4.3. Those results then motivate a stability analysis of those nonlinear traveling waves, presented in Chapter 5. A perturbation system of single-jamiton waves is derived in §5.1, and then analyzed qualitatively in §5.2. In this analysis of the perturbation system, a major challenge, explained in §5.3, is the treatment of the sonic point. After that, quantitative results for the perturbation system are displayed in §5.4 and §5.5.

The results in both chapters 4 and 5 provide valuable insight that allows us to then propose an averaged ARZ model (which solves for an averaged quantity \( \bar{\rho} \)) in Chapter 6. The rationale for the averaged model is outlined in §6.1. Then, it is demonstrated that the averaged model — which is simpler to solve analytically and numerically — is able to capture the cumulative effects of
the ARZ model in terms of fuel consumption, vehicle flow, and braking (§6.3). To carry out the fuel consumption calculations in this chapter, a vehicle based fuel consumption model was needed, and it was established in §6.2. Finally, a conclusion containing a summary of this thesis, a discussion of the implications of the results, and a broader outlook on future work is presented in Chapter 7.
CHAPTER 2

MACROSCOPIC TRAFFIC MODELS

In macroscopic descriptions of traffic flow, the quantities of interest are the traffic density, traffic flow rate or flux, and traffic velocity. If we denote the position along a road or highway by \( x \), and time by \( t \), then we can define the following functions:

- \( \rho(x,t) \) represents the traffic density at position \( x \) and time \( t \). The units associated with this function are vehicles per unit length.

- \( Q(x,t) \) represents the traffic flow rate at position \( x \) and time \( t \). The units associated with this function are vehicles per unit time.

- \( u(x,t) \) represents the traffic bulk velocity at position \( x \) and time \( t \). The
units associated with this function are regular velocity units (unit length per unit time).

In fact, the traffic bulk velocity $u(x,t)$ is defined according to the following relationship:

$$Q(x,t) = \rho(x,t)u(x,t).$$ (2.1)

Getting real-life measurements for these quantities is a crucial task in macroscopic modeling of traffic. Without real data, this class of traffic modeling remains a theoretical endeavour, as the models cannot be reliably constructed or validated. Over the past century, traffic engineers have devised different ways to measure the traffic density and the traffic flow rate. Nowadays, both of these quantities can be accurately measured via single-loop or double-loop sensors [72].

To measure the traffic flow rate $Q$, first pick a sampling time interval $\Delta t$, and a position of the road. Then use the sensor to count the number of vehicles $N$ that pass that location in the chosen time interval. The flow rate can then be calculated as $Q = \frac{N}{\Delta t}$.

On the other hand, traffic density is measured indirectly. In this case, the sensors are used to measure the “Occupancy” $O$, which is the proportion of time (in a given time interval) that a vehicle occupies a certain position on the road. Traffic density is then calculated by the equation $\frac{O}{l}$, where $l$ is a chosen average for vehicle length.
Given relationship (2.1), measuring the traffic density and the flow rate is enough to deduce the bulk traffic velocity $u$. Finally, note that the quantities $\rho$ and $Q$ can be aggregated over multiple lanes.

2.1 The Fundamental Diagram

Using sensor data of traffic density and flow rate allows for the construction of a scatter plot where each point represents a $(\rho, Q)$ pair of measurements at the same location and time interval. The traffic density $\rho$ is plotted on the horizontal axis, and the traffic flow rate $Q$ is plotted on the vertical axis.

On a specific road, the traffic density ranges from 0 (empty road) to $\rho_{\text{max}}$, the maximum density on the road corresponding to bumper-to-bumper traffic. Theoretically, the flow rate will be 0 when the traffic density is 0 (no vehicles are passing through because there are no vehicles). Moreover, the flow rate is expected to be 0 when the traffic density is $\rho_{\text{max}}$ (no vehicles are passing through because the traffic is severely dense to the point where all the vehicles are stuck in place).

Figure 1 in [25] shows an example of such a scatter plot. Scatter plots on other roads might look different but they all show a clear separation between two modes of traffic: low- and high-density traffic. For low traffic densities, which can be described as free-flow traffic, there is a tight relationship between $Q$ and $\rho$. For higher traffic densities, which can be described as congested...
traffic, the relationship between $Q$ and $\rho$ is a lot more loose.

The idea of a *Fundamental Diagram* (FD) in traffic flow comes from trying to find a functional relationship between density and flow rate [31, 63, 97, 70]. In other words, writing the flow rate $Q$ as $Q(\rho)$ gives rise to a FD of traffic. As discussed above, it is a good idea to let $Q(0) = 0$ and $Q(\rho_m) = 0$ when constructing $Q(\rho)$. In addition, the data collected via sensors suggest that a concave down function would be a good fit for a FD. Figure 1 in [25] also shows a possible FD that tries to resemble the data. Note that the FD does a good job approximating sensor data when traffic is in a free-flow state. However, when traffic is in a congested state, the sensor data suggests that the relationship between $Q$ and $\rho$ is not really functional, and therefore the FD has no chance in making a meaningful approximation of $Q$ given $\rho$.

To define the two traffic regimes discussed above more rigorously, consider the following: Given a concave down FD, there is a unique density $\rho_c$ at which a maximum flow rate $Q_m$ is attained (i.e. $Q(\rho_c) = Q_m$). The range of densities in the FD can therefore be split into two intervals: the free-flow regime where $\rho \in [0, \rho_c)$, and the congestion regime where $\rho \in (\rho_c, \rho_m]$. $\rho_c$ is called the critical traffic density. When $\rho < \rho_c$, the flow rate $Q$ is less than $Q_m$ because there are not enough vehicles on the road to achieve the maximum flow rate. When $\rho > \rho_c$, the flow rate $Q$ is also less than $Q_m$ because there are too many vehicles on the road, and the congestion causes the throughput to drop.
Many attempts have been made to draw a fundamental diagram through fitting real-life measurements of flow rate versus traffic density. The first of these measurements were done by Greenshields in the 1930s [31], which led him to propose a quadratic function $Q(\rho)$ to describe the relationship between flow rate and density. Later measurements showed that this was not very accurate, and different functions were proposed [63, 29, 17].

The Greenshields FD comes from the assumption that traffic speed is a linear function of traffic density. It is also assumed that traffic speed $u = 0$ when $\rho = \rho_{\text{max}}$, and $u = u_{\text{max}}$ when $\rho = 0$. This gives rise to the relationship $u(\rho) = u_{\text{max}} \left(1 - \frac{\rho}{\rho_{\text{max}}}\right)$. Using equation (2.1) allows us to write the Greenshields flux function:

$$Q(\rho) = \rho \times u(\rho) = \rho u_{\text{max}} \left(1 - \frac{\rho}{\rho_{\text{max}}}\right). \quad (2.2)$$

The flux function in (2.2) is quadratic and concave down. Also, the critical density for this flux function is always $\rho_{c} = \frac{\rho_{\text{max}}}{2}$. Such a flux function can be uniquely determined with a good approximation of $u_{\text{max}}$ and $\rho_{\text{max}}$ from the data. In addition to the general shortcomings to the idea of FDs, the Greenshields FD has two specific shortcomings:

- Real-life measurements suggest that the critical density is usually lower than $\rho_{\text{max}}/2$.

- Real-life measurements suggest that the relationship between $Q$ and $\rho$ is much closer to linear than quadratic when $\rho < \rho_{c}$.
To overcome these two shortcomings, the Newell-Daganzo fundamental diagram has been proposed [70, 17]. This FD is a piecewise linear continuous function defined as follows:

\[
Q(\rho) = \begin{cases} 
\rho u_{\text{max}} & \text{if } \rho \leq \rho_c \\
\frac{u_{\text{max}} \rho_c}{\rho_c - \rho_m} \rho + \frac{u_{\text{max}} \rho_c \rho_m}{\rho_m - \rho_c} & \text{if } \rho > \rho_c 
\end{cases} \tag{2.3}
\]

The FD of the flux function defined in (2.3) has a triangular shape, and is sometimes referred to as the triangular FD. This flux function can be uniquely determined with a good approximation of \(u_{\text{max}}, \rho_{\text{max}}\) and \(\rho_c\) from the data. The main drawback to the Newell-Daganzo flux function is that it is not differentiable when \(\rho = \rho_c\).

To overcome the lack of smoothness of the Newell-Daganzo FD, there are other proposed flux functions that are concave down, smooth and fitted to sensor data. One idea [25] is to construct a smoothed version of the triangular flux, which gives a FD function,

\[
Q(\rho) = c \left( g(0) + (g(1) - g(0)) \frac{\rho}{\rho_{\text{max}}} \right) - g \left( \frac{\rho}{\rho_{\text{max}}} \right),
\]

where \(g(y) = \sqrt{1 + \left( \frac{y - b}{\lambda} \right)^2}\). \tag{2.4}

The parameters \(c, b\) and \(\lambda\) are defined through the data-fitting process. Figure 2.1 below shows examples of what the FDs discussed above look like. They
Figure 2.1: Three different fundamental diagram functions: The brown solid line is a plot of the quadratic Greenshields FD, the dashed dark blue line is a plot of the triangular Newell-Daganzo FD, and the solid red line is a plot of a smoothed triangular FD.

all share the same $\rho_{\text{max}}$ and $u_{\text{max}}$. For the Daganzo FD, $\rho_c = \frac{\rho_{\text{max}}}{3}$. The other parameters in the smoothed Daganzo FD are chosen to fit the data in [25].
2.2 First Order Macroscopic Models

Macroscopic models provide a description of the time evolution of the vehicle density $\rho(x, t)$ on a road. Here, $x$ is the position along the road (multiple lanes are aggregated into one variable), and $t$ is time. Deriving a first order model comes from the conservation of “mass” principle [100]. In this case, the quantity conserved is the number of vehicles. Starting with a road segment $[x_1, x_2]$, at any time $t$, the number of vehicles in this segment is $\int_{x_1}^{x_2} \rho(x, t) \, dx$.

The rate of change in the number of vehicles in $[x_1, x_2]$ at time $t$ is the inflow flux $Q(x_1, t)$ minus the outflow flux $Q(x_2, t)$. This gives rise to the following equation:

$$\frac{d}{dt} \int_{x_1}^{x_2} \rho(x, t) \, dx = Q(x_1, t) - Q(x_2, t).$$

(2.5)

Using the Fundamental Theorem of Calculus and the fact that $x_1$, $x_2$ and $t$ are arbitrary, equation (2.5) leads to the continuity equation

$$\rho_t + Q_x = 0.$$  

(2.6)

Recall that $Q = \rho u$, so equation (2.6) can be written as $\rho_t + (\rho u)_x = 0$. Even though this continuity equation correctly describes the evolution of $\rho$, it cannot be solved because it contains two unknowns. To overcome that, first order models make the assumption that $Q$ is a function of $\rho$ (or alternatively, $u$ is a function of $\rho$). The resulting model is called a Lighthill-Whitham-Richards...
(LWR) model [63, 84], and it reads as:

\[ \rho_t + (Q(\rho))_x = 0 . \tag{2.7} \]

Note that if \( u = U(\rho) \), then the flux function can be written as \( Q(\rho) = \rho U(\rho) \).

It is important to note that the key ingredient in first order traffic modeling is the design of a good FD which assumes a functional relationship between the flow rate and the traffic density. The LWR model in (2.7) is in fact a scalar hyperbolic conservation law [16], and it admits weak solutions with discontinuities and shocks [51].

First order LWR models are widely used in practice [58, 10, 105] as a macroscopic description of traffic flow for their simplicity and compatibility with some observational facts, like the propagation of traffic disturbances. However, LWR models have many shortcomings. First, they rely on a FD, and as established in §2.1, the relationship between \( Q \) and \( \rho \) is not really a functional one. Moreover, LWR models satisfy a maximum principle [80], meaning that the traffic density at any point on the road, and at any moment in time, cannot be larger than the maximum density on the road at the initial time. We know this is not true, because traffic jams can develop on a road even if the initial traffic density was constant, as has been demonstrated by multiple experiments [92, 90].

These limitations of first order models make it clear that they fail to capture key phenomena in traffic flow, which makes the case for the need of second
order models.

2.3 Second Order Models

In contrast, instead of writing $u$ or $Q$ as function of $\rho$, second order models augment (2.6) by an independent field variable for the bulk velocity $u(x,t)$, and describe their evolution via a $2 \times 2$ balance law system: a hyperbolic conservation law system [52] with a relaxation term in the velocity equation. Due to conservation of vehicles, the density always evolves by the continuity equation, $\rho_t + (\rho u)_x = 0$. In turn, the velocity equation encodes the actual modeling of the vehicle dynamics and interactions, and it is derived from prescribing vehicle acceleration as a function of the main traffic quantities and their derivatives.

A vehicle with trajectory $x(t)$ has acceleration $a$ that can be calculated as follows:

$$a = \frac{d}{dt} \dot{x}(t) = \frac{d}{dt} u(x(t), t) = u_t + uu_x$$  \hspace{1cm} (2.8)

In general, second order models set this vehicle acceleration $a$ to a function $a(\rho, u, \rho_t, \rho_x, u_t, u_x)$. Two different second order models are discussed below.
2.3.1 The Payne–Whitham Model

The *Payne-Whitham (PW) model* [75, 102], which can be written as

\[
\rho_t + (\rho u)_x = 0,
\]
(2.9)

\[
u_t + uu_x + p(\rho)_x/\rho = \frac{1}{\tau}(U(\rho) - u),
\]

was the first second-order model proposed. Here \(U(\rho)\) is the *desired velocity function*, and \(\tau\) is the relaxation time that determines how fast drivers adjust to their desired velocity \(U(\rho)\). The *traffic pressure* \(p(\rho)\) models preventive driving (it is the analog of the "pressure" in fluid dynamics). Moreover, \(p(\rho)\) is assumed to be an increasing function of the traffic density \(\rho\). Note that the PW model is equivalent to setting the acceleration of vehicles to

\[
a(\rho, u, \rho_t, \rho_x, u_t, u_x) = \left(\frac{U(\rho) - u}{\tau}\right) + \left(-\frac{p(\rho)_x}{\rho}\right).
\]
(2.10)

The first term in this acceleration prescription is a relaxation term, which says that a driver’s acceleration is proportional to how far they are from the desired velocity. The second term is there to account for the driver’s reaction to the density and the change in density ahead. In fact, with an increasing pressure function \(p(\rho)\) (i.e. \(p'(\rho) > 0\)), this second term means that drivers slow down when the density ahead is larger.

The PW model (2.9) is a non-linear system of conservation laws with two characteristic speeds. They can be calculated by first writing (2.9) in conser-
Conservation form:

\[ \rho_t + (\rho u)_x = 0 , \]

\[ q_t + \left( \frac{q^2}{\rho} + p(\rho) \right)_x = \frac{1}{\tau} (\rho U(\rho) - q) \]

where \( q = \rho u \) is the conserved field. The flux function in the conservation law system (2.11), and its Jacobian are:

\[
F \left( \begin{array}{c} \rho \\ q \end{array} \right) = \left( \begin{array}{c} q \\ \frac{q^2}{\rho} + p(\rho) \end{array} \right), \quad \nabla F = \left( \begin{array}{cc} 0 & 1 \\ -\frac{q^2}{\rho^2} + p'(\rho) & \frac{2q}{\rho} \end{array} \right)
\]

The characteristic speeds of (2.9) are the eigenvalues of \( \nabla F \):

\[ \lambda^{(1)} = u - \sqrt{p'(\rho)} , \quad \lambda^{(2)} = u + \sqrt{p'(\rho)} . \]

With an increasing pressure function \( p(\rho) \), we have \( p'(\rho) > 0 \), which means that \( \lambda^{(1)} \) and \( \lambda^{(2)} \) are distinct and real, which implies that system (2.9) is strictly hyperbolic [52].

The PW model addresses both shortcomings of first order LWR models discussed in §2.2:

- It does not assume a functional relationship between \( Q \) and \( \rho \).
- It captures the traffic instability phenomenon, where a constant base state could generate traffic waves (traffic jams) [27, 88].

Nevertheless, the PW is generally rejected for two main drawbacks [18]:


• It is possible for the model to generate negative speeds \((u < 0)\), which is unrealistic in traffic flow applications.

• It is possible for the information introduced by the vehicle dynamics to move faster than vehicle speeds \((\lambda^{(2)} > u)\). This is also unrealistic as drivers should only be affected by the state of traffic ahead.

2.3.2 The ARZ Model

To address the drawbacks of the PW model, Aw and Rascle [5], and independently, Zhang [106] proposed what is now regarded as the ARZ model

\[
\rho_t + (\rho u)_x = 0 ,
\]

\[
(u + h(\rho))_t + u(u + h(\rho))_x = \frac{1}{\tau}(U(\rho) - u) .
\]

(2.14)

This is another \(2 \times 2\) hyperbolic balance law system where the first equation is still the continuity equation. Here, \(h(\rho)\) is called the hesitation function, which is assumed to be strictly increasing. \(U(\rho)\) is again the desired velocity function, which is strictly decreasing.

Note that the ARZ model is equivalent to setting the acceleration of vehicles to the function

\[
a(\rho, u, \rho_t, \rho_x, u_t, u_x) = \left( \frac{U(\rho) - u}{\tau} \right) + \left( - h(\rho)_t - uh(\rho)_x \right) .
\]

(2.15)

As with the acceleration function prescribed in the PW model, the first term here is again a relaxation term, and the second term accounts for drivers’
reactions to their current states as well as the states ahead. In fact, with an increasing hesitation function \( h(\rho) \) (i.e. \( h'(\rho) > 0 \)), the second term means that drivers slow down when they observe, in their own frame of reference, that the traffic density is increasing.

The assumption that \( h \) is strictly increasing yields a hyperbolic system that overcomes the shortcomings to the PW model [5]. To verify that, we first write (2.14) in conservation form. Here, the conserved fields are \( \rho \) and \( q = \rho(u + h(\rho)) \). The ARZ model then reads as:

\[
\begin{align*}
\rho_t + (q - \rho h(\rho))_x &= 0, \\
q_t + \left( \frac{q^2}{\rho} - q h(\rho) \right)_x &= \frac{1}{\tau} \left( \rho(U(\rho) + h(\rho)) - q \right).
\end{align*}
\]

(2.16)

The flux function in the conservation law system (2.16), and its Jacobian are:

\[
F \begin{pmatrix} \rho \\ q \end{pmatrix} = \begin{pmatrix} q - \rho h(\rho) \\ \frac{q^2}{\rho} + q h(\rho) \end{pmatrix}, \quad \nabla F = \begin{pmatrix} -\rho h'(\rho) - h(\rho) & 1 \\ \frac{-q^2}{\rho^2} - q h'(\rho) & \frac{2q}{\rho} - h(\rho) \end{pmatrix}.
\]

(2.17)

The characteristic speeds of (2.14) are given by solving for the eigenvalues of \( \nabla F \). This corresponds to solving the following characteristic equation for \( \lambda \):

\[
\left( \lambda + \rho h'(\rho) + h(\rho) \right) \left( \lambda - \frac{2q}{\rho} + h(\rho) \right) + \frac{q^2}{\rho^2} + q h'(\rho) = 0,
\]

which is equivalent to solving

\[
\lambda^2 + \left( 2h(\rho) + \rho h'(\rho) - \frac{2q}{\rho} \right) \lambda + \frac{q^2}{\rho^2} + h(\rho)^2 - q h'(\rho) + \rho h(\rho) h'(\rho) - \frac{2q}{\rho} h(\rho) = 0,
\]
which implies that
\[
\lambda = \frac{2q - 2h(\rho) - \rho h'(\rho) \pm \sqrt{\rho^2 h'(\rho)^2}}{2} = \frac{q}{\rho} - h(\rho) - \frac{\rho h'(\rho)}{2} \pm \frac{\rho h'(\rho)}{2}.
\]

Therefore, the eigenvalues of \( \nabla F \) are

\[
\lambda_1 = \frac{q}{\rho} - h(\rho) - \rho h'(\rho) = u - \rho h'(\rho), \quad \lambda_2 = \frac{q}{\rho} - h(\rho) = u.
\]

With the assumption that \( h(\rho) \) is strictly increasing, it follows that \( h'(\rho) > 0 \) and therefore we have \( \lambda_1 \leq \lambda_2 = u \). This proves that information cannot propagate faster than vehicle velocities. In addition, when \( \rho \neq 0 \) (non-empty traffic), \( \lambda_1 \) and \( \lambda_2 \) are distinct, making the ARZ model strictly hyperbolic.

### 2.4 The Sub–Characteristic Condition (SCC)

The ARZ model (2.14) (and the PW model (2.9)) share the relaxation term in the second equation. This term is expected to push the traffic velocity \( u \) towards the desired velocity \( U(\rho) \) when \( \tau \to 0^+ \). Notice that when this happens, the traffic dynamics evolve according to the first equation in the model (the continuity equation), which is an LWR model with \( Q(\rho) = \rho U(\rho) \), and with one characteristic speed \( \mu = Q'(\rho) = U(\rho) + \rho U''(\rho) \).

The sub-characteristic condition (SCC) [101, 102, 64, 11] for a second order model is given as follows:

**Definition 2.1.** For a given \( \rho \) and \( u = U(\rho) \), the SCC is satisfied if the characteristic speed \( \mu = \mu(\rho) \) of the reduced LWR model is in between the two
characteristic speeds $\lambda_1 = \lambda_1(\rho, u)$ and $\lambda_2 = \lambda_2(\rho, u)$ of the second order model, i.e. $\lambda_1 < \mu < \lambda_2$.

For the ARZ model (2.14), we know that $\lambda_1 = u - \rho h'(\rho)$ and $\lambda_2 = u$. So, the SCC is satisfied when

$$U(\rho) - \rho h'(\rho) < U(\rho) + \rho U'(\rho) < U(\rho)$$

(2.19)

or

$$-\rho h'(\rho) < \rho U'(\rho) < 0.$$  

(2.20)

The second inequality in (2.20) is always satisfied if $U(\rho)$ is chosen to be a decreasing function. In that case, the SCC reduces to the inequality:

$$U'(\rho) + h'(\rho) > 0.$$  

(2.21)

The significance of the SCC in this context comes from its relationship to the concept of Linear Stability. To discuss that, first notice that the ARZ model and its corresponding reduced LWR model (the one where $Q(\rho) = \rho U(\rho)$) share a class of solutions: solutions where $\rho(x,t) = \rho_0 = \text{constant}$ and $u(x,t) = u_0 = U(\rho_0) = \text{constant}$. These solutions are called base state solutions, and they represent uniform traffic flow, where vehicles are equi-spaced. In addition, the velocity $u$ is also uniform and is at the desired velocity level. Linear stability is then defined as follows:

**Definition 2.2.** The base state $(\rho_0, u_0)$ is linearly stable if any infinitesimal perturbations to it decay in time.
There is a fundamental relationship between linear stability of uniform flow and the SCC. In fact, if the SCC is satisfied, then uniform flow is stable [101, 102, 64, 11]. Conversely, when it is violated, uniform flow is unstable and nonlinear traveling wave solutions exist [59, 40, 71, 27, 88]. The relationship between the SCC and linear stability of uniform flow can be summarized in Whitham’s Theorem below:

**Theorem 2.1 (Whitham).** The base state \((\rho_0, u_0)\) is linearly stable \(\iff\) \((\rho_0, u_0)\) satisfies the SCC.

As discussed in Section 2.2, one of the advantages of second order models (e.g. ARZ model) is their ability to capture the phenomenon of traffic instability. Therefore, the ability to distinguish linearly stable base states from linearly unstable base states is of significant importance. To do that directly requires the calculation of the growth factor for basic wave perturbations of the form \(e^{ikx}\), which is done in more detail in a later chapter (§3.2) for the ARZ model with specific model functions. Even though such a calculation provides insight into the nature of stability of these models, it is a cumbersome calculation. On the other hand, Theorem 2.1 above allows one to make the same distinction by only checking the SCC, which is a much simpler calculation.

The case of the SCC being satisfied has been studied extensively for 2 \(\times\) 2 systems with a relaxation term [101, 102, 64, 60]. The more general case of \(N\times N\) systems has also been studied in detail [11]. In particular, the SCC is related
to positive diffusion when conducting a Chapman-Enskog expansion of the model [49, 34]. Many of the established results about the SCC extend beyond traffic flow models and apply to general hyperbolic conservation law systems with stiff relaxation terms. For example, it is proven that the equivalence in Theorem (2.1) can be extended from relating the SCC and linear stability to relating the SCC and stability. In other words, the solution of a uniform base state, with small perturbations, converges to the starting base state as $t \to \infty$.

In addition, other results of non-uniform base states are also established when the SCC is satisfied. For instance, in various special circumstances [64, 60], the solution to some models that satisfy the SCC everywhere converges to the solution of the reduced model in the limit as $\tau \to 0^+$. And more generally, if $\tau$ is not close to zero, the solution would then converge to the solution of the reduced model with an additional nonlinear “viscosity” term, where the order of that term is $O(\tau)$.

In this work, however, the focus will be on understanding the behavior and stability of solutions when the SCC is violated.
CHAPTER 3

JAMITONS

While the general results and methodologies apply to a wide class of second-order models with relaxation (including the PW model (2.9) and generic second-order models [55, 24]), the focus of this work is on the inhomogeneous Aw-Rascle-Zhang (ARZ) model [5, 106]. Again, in non-conservative form it reads as

\[
\rho_t + (\rho u)_x = 0,
\]

\[
(u + h(\rho))_t + u(u + h(\rho))_x = \frac{1}{\tau}(U(\rho) - u).
\] (3.1)

In the homogeneous ARZ model, the field \( w = u + h(\rho) \) can be interpreted as a convected quantity moving with the flow (the hesitation function reduces the empty road velocity \( w \) by \( h(\rho) \)). Given this interpretation we postulate that \( u + h(\rho) \) is a “property” of each vehicle, so that \( q = \rho(u + h(\rho)) \) is the density for a conserved quantity. Since \( \rho \) is conserved as well, this leads to the
following conservation forms for the equations

\[\rho_t + (q - \rho h(\rho))_x = 0,\]
\[q_t + \left(\frac{q^2}{\rho} - qh(\rho)\right)_x = \frac{1}{\tau} \left(\rho(U(\rho) + h(\rho)) - q\right).\]  
\(3.2\)

With these conservation forms we can now allow shocks (or contact discontinuities) in the solutions, which must satisfy the Rankine-Hugoniot jump conditions [23]

\[s[\rho] - [\rho u] = 0,\]
\[s[\rho(u + h(\rho))] - [\rho u^2 + \rho u h(\rho)] = 0.\]  
\(3.3\)

Here \([\zeta]\) denotes the jump of the variable \(\zeta\) across the discontinuity, and \(s\) is the speed. In addition, the Lax entropy conditions [23] must be satisfied. Specifically, one family of characteristics goes through the discontinuity, while the other converges into it (for a shock), or is parallel to it (for a contact).

Furthermore, in real traffic flow the shocks are compressive (i.e., as vehicles go through a shock, the density increases) and move slower than the vehicles. Thus, to guarantee that the entropy conditions produce this behavior, the following assumptions on \(h(\rho)\) are made [88]:

- \(h(\rho)\) is a strictly increasing function, or \(h'(\rho) > 0\)
- \(\rho h(\rho)\) is a strictly convex function, or \((\rho h(\rho))'' = 2h'(\rho) + \rho h''(\rho) > 0\)
The characteristic speeds of (3.2) are:

\[ \lambda_1 = \frac{q}{\rho} - h(\rho) - \rho h'(\rho) = u - \rho h'(\rho) , \quad \text{and} \quad \lambda_2 = \frac{q}{\rho} - h(\rho) = u , \]  

(3.4)

where the characteristic speed \( \lambda_1 \) is genuinely nonlinear (associated with shocks and rarefactions), while the characteristic speed \( \lambda_2 \) is linearly degenerate (associated with contacts).

### 3.1 Specific model functions

While the analysis and general results derived below hold for generic models (3.2), the computational study and the illustrative graphs are presented for a specific choice of model functions. As in [88], the following choices are made:

- The maximum vehicle density for a single lane corresponding to bumper-to-bumper traffic is \( \rho_{\text{max}} = 1/7.5 \text{m} \).
- The maximum vehicle speed, theoretically attained on an empty road, is \( u_{\text{max}} = 20 \text{m/s} \).
- A fundamental diagram function \( Q(\rho) \), which is a smoothed version of the Newell-Daganzo triangular flux [70, 17], is constructed as follows:

\[
Q(\rho) = c \left( g(0) + (g(1) - g(0)) \frac{\rho}{\rho_{\text{max}}} - g \left( \frac{\rho}{\rho_{\text{max}}} \right) \right) ,
\]

where \( g(y) = \sqrt{1 + \left( \frac{y-b}{\lambda} \right)^2} \),
\[ c = 0.078 \rho_{\text{max}} u_{\text{max}}, \quad b = \frac{1}{3}, \quad \lambda = \frac{1}{110}. \]

The parameters are chosen to have the function fit real sensor data [88].

- The \textit{desired velocity function} is then given as \( U(\rho) = Q(\rho)/\rho. \)
- The \textit{hesitation function} is \( h(\rho) = \frac{8}{s} \sqrt{\frac{\rho}{\rho_{\text{max}} - \rho}} \)
- The relaxation time is \( \tau = 3\text{s}. \)

Note that all the values above are for a single lane. When considering multi-lane traffic, realistic values can be obtained by scaling \( \rho \) and \( Q \) by the number of lanes.

The functions above and their derivatives are plotted in Figure 3.1 below. Notice that \( Q(\rho) \) is strictly concave. This is a usual assumption in traffic flow. In particular, it guarantees a nice shock theory for the reduced equation, which is an LWR model where the shocks satisfy the Lax entropy condition strictly [56]. In addition, it is clear from Figure 3.1c that \( U(\rho) \) is a decreasing function of \( \rho \), which is another frequent assumption in traffic modeling. This means that given equilibrium driving conditions, the higher the traffic density, the lower the traffic velocity.

The hesitation function \( h(\rho) \) defined above is strictly increasing. It possesses an inflection point when \( \rho \approx 0.2 \rho_{\text{max}} \), and it becomes singular as \( \rho \to \rho_{\text{max}} \). This ensures that \( h'(\rho) \) is large for low and high densities. The rationale behind this design choice is the SCC. Recall that the satisfaction of
the SCC \( (U'(\rho) + h'(\rho) \geq 0) \) is equivalent to having stable uniform flow. The choices above make \( h'(\rho) \) large for low and high densities. This allows for the generation of the phase transition presented by Kerner [42]: stable uniform flow for low densities and synchronized flow with a complex sequence of narrow jams for higher densities. This choice is also in line with observational evidence of stop-and-go waves. When \( \rho \) is too low, the free flow dynamics of traffic prevent the rise of stop-and-go waves. Note that also when \( \rho \) is close to \( \rho_{\text{max}} \), traffic is too constricted for any development of such waves.

In Figure 3.1f, the quantity \( U'(\rho) + h'(\rho) \) is plotted as a function of \( \rho \). This quantity crosses the horizontal axis twice, dividing the domain \([0, \rho_{\text{max}}]\) into three regions: two exterior regions where the SCC is satisfied, and one interior region where the SCC is violated.

### 3.2 Linear stability of uniform flow

Before analyzing the stability of nonlinear waves in Chapter 5, we now discuss important aspects related to the stability of uniform flow: base state solutions of (3.1) in which \( \rho = \tilde{\rho} \) and \( u = U(\tilde{\rho}) \) are constant in space and time. The linear stability analysis itself is a well-established normal models analysis [44, 27, 35] and we briefly outline the key steps. Consider infinitesimal wave perturbations (where \( k \) is the wave number and \( \sigma \) the complex growth rate)
Figure 3.1: Plots of the specific model functions $Q(\rho)$, $h(\rho)$ and $U(\rho)$ (a-c), as well as the derivatives $h'(\rho)$ and $U'(\rho)$ (d-e). A plot of $U'(\rho) + h'(\rho)$ is in panel (f), and it shows where the SCC is satisfied or violated.
of the base state, 

\[ \hat{\rho} = \hat{R} e^{ikx+\sigma t} \quad \text{and} \quad \hat{u} = \hat{U} e^{ikx+\sigma t}, \]

substitute the perturbed solution \( \rho = \tilde{\rho} + \hat{\rho} \) and \( u = U(\tilde{\rho}) + \hat{u} \) into (3.1), and consider only constant and linear terms. This leads to the system

\[
\begin{bmatrix}
\sigma + ik\psi & ik\hat{\rho} \\
\sigma \phi + ik\psi \phi - \frac{\xi}{\tau} & \sigma + ik\psi + \frac{1}{\tau}
\end{bmatrix}
\begin{bmatrix}
\hat{R} \\
\hat{U}
\end{bmatrix}
= \begin{bmatrix}
0 \\
0
\end{bmatrix},
\]

for the perturbation amplitudes, where \( \psi = U(\tilde{\rho}) > 0, \phi = h'(\tilde{\rho}) > 0, \) and \( \xi = U'(\tilde{\rho}) < 0 \). Nontrivial solutions can only exist if the matrix in (3.5) has a vanishing determinant, which requires

\[ \sigma = -ik\psi + ik\frac{1}{2}\rho\phi - \frac{1}{2\tau}(1 + \Gamma), \]

where \( \Gamma \) satisfies

\[ \Gamma^2 = 1 - k^2\tau^2\rho^2\phi^2 - 2ik\tau\rho(\phi + 2\xi). \]

Writing \( \Gamma = \Lambda_1 + i\Lambda_2 \) in terms of its real and imaginary part yields the two equations

\[ \Lambda_1^2 - \Lambda_2^2 = 1 - k^2\tau^2\rho^2\phi^2 \quad \text{and} \quad \Lambda_1\Lambda_2 = k\tau\rho(\phi + 2\xi), \]

which then leads to the following quadratic equation for \( z = (\Lambda_1)^2:\)

\[ z^2 - (1 - \beta^2k^2)z - \gamma^2k^2 = 0. \]
Here \( \beta = \tau \tilde{\rho} \phi \) and \( \gamma = \tau \tilde{\rho} (\phi + 2 \xi) \). The positive solution of (3.6), as a function of \( k \), is

\[
z^+(k) = \frac{1}{2} \left( 1 - \beta^2 k^2 \right) + \sqrt{\left( 1 - \beta^2 k^2 \right)^2 + 4 \gamma^2 k^2} = \frac{1}{2} \left( 1 - \beta^2 k^2 \right) + \sqrt{\left( 1 + \beta^2 k^2 \right)^2 + 4 (\gamma^2 - \beta^2) k^2}.
\]

This function has the following properties:

(i) \( z^+(0) = 1 \).

(ii) \( \lim_{k \to \infty} z^+(k) = (\gamma/\beta)^2 \), which follows from (3.7) and the asymptotic \( (k \gg 1) \) formula:

\[
\sqrt{(1 - \beta^2 k^2)^2 + 4 \gamma^2 k^2} \sim \beta^2 k^2 \sqrt{1 + 2(\gamma^2 - \beta^2) \beta^{-4} k^{-2}} \\
\sim \beta^2 k^2 + (2(\gamma/\beta)^2 - 1).
\]

(iii) It is strictly monotonic if \( |\gamma| \neq |\beta| \), i.e., it is strictly increasing if \( |\gamma| > |\beta| \) and strictly decreasing if \( |\gamma| < |\beta| \). This fact follows from (3.7), because the sign of the term \( 4(\gamma^2 - \beta^2) k^2 \) determines the slope of \( z^+(k) \): if \( |\gamma| = |\beta| \), it is constant, and if the term is positive (negative), the function goes up (down) with \( k \).

The growth rate of normal modes is

\[
g_{\tilde{\rho}}(k) = \text{Re}(\sigma) = -\frac{1}{2\tau} (1 + \text{Re}(\Gamma)) = -\frac{1}{2\tau} (1 + \Lambda_1) = -\frac{1}{2\tau} \left( 1 \pm \sqrt{z^+(k)} \right).
\]

Linear stability, i.e., \( \text{Re}(\sigma) \leq 0 \), is equivalent to \( z^+ \leq 1 \) (only the negative root of \( \sqrt{z^+} \) could cause positive growth). Hence, stability holds exactly if
\[ |\gamma| < |\beta|, \text{ or equivalently } \phi + \xi > 0, \text{ or equivalently } \]

\[ h'(\tilde{\rho}) + U'(\tilde{\rho}) \geq 0 . \tag{3.9} \]

This last condition is exactly what the SCC yields as well.

To recapitulate, for the inhomogeneous ARZ model (3.1), there are exactly two possibilities: the stability condition holds, or it is violated. If the stability condition (the SCC) (3.9) holds, then all basic wave perturbations \( e^{ikx} \) have non-positive growth rates, and solutions are linearly stable. If (3.9) is violated, then all waves grow. Moreover, the rate of growth \( g_{\tilde{\rho}}(k) \) is an increasing function of the wave number \( k \), that has \( g_{\tilde{\rho}}(0) = 0 \), and approaches (as \( k \to \infty \)) the asymptotic growth rate

\[ g_{\tilde{\rho}}^\infty = \lim_{k \to \infty} g_{\tilde{\rho}}(k) = \frac{1}{\tau} (|\gamma/\beta| - 1) = \frac{1}{2\tau} (|1 + 2\xi/\phi| - 1) = \frac{1}{\tau} \left( \frac{-U'(\tilde{\rho})}{h'(\tilde{\rho})} - 1 \right). \]

Figure 3.2 shows the growth rate functions \( g_{\tilde{\rho}}(k) \) for the specific model given in §3.1, with stable base states in panel 3.2a and unstable base states in panel 3.2b. In the latter, one can clearly see the strict increase of \( g_{\tilde{\rho}} \) with \( k \), and the asymptotic limit \( g_{\tilde{\rho}}^\infty \). Panel 3.2c shows a plot of the asymptotic growth rate \( g_{\tilde{\rho}}^\infty \) as a function of \( \tilde{\rho} \). This function has kinks because \( \sigma \) has two solutions, and the growth rate is defined as the larger real part of these \( \sigma \)’s. In particular, the kinks occur when the \( \sigma \) with the larger real part switches from one solution to the other.

Clearly, base states that satisfy (3.9) are well-behaved. However, with
(a) Growth rates $g_{\tilde{\rho}}(k)$ for different $\tilde{\rho}$ that satisfy (3.9), i.e., are linearly stable.

(b) Growth rates $g_{\tilde{\rho}}(k)$ for different $\tilde{\rho}$ that violate (3.9), i.e., are linearly unstable.

(c) Asymptotic growth rate (worst case)

$$g_{\tilde{\rho}}^\infty = \lim_{k \to \infty} g_{\tilde{\rho}}(k)$$ as a function of $\tilde{\rho}$.

Figure 3.2: Plots of the growth rate $g_{\tilde{\rho}}(k) = \text{Re}(\sigma)$ as a function of the wave number $k$, for different constant base states $\tilde{\rho}$, as well as the asymptotic growth rate $g_{\tilde{\rho}}^\infty$ as a function of $\tilde{\rho}$. 
regards to modeling phantom traffic jams and jamitons, we are particularly interested in base states that violate (3.9). These require some more careful discussion. While instabilities to uniform states are ubiquitous in science and engineering, having a growth rate that is increasing for all wave numbers is unusual. The much more common scenario (for example, fluid instabilities moderated by viscosity or surface tension [21]) is that medium waves are unstable and short waves (i.e., $k$ large) are stable again, yielding a critical wave number $k^*$ of maximal growth. In that case, one can argue that out of infinitesimal perturbations, in which all wave lengths are present, the linearized dynamics will single out the ones with dominant growth. Hence, the wave number $k^*$ will be selected to first enter the nonlinear regime.

However, arguments of this type do not work for (3.1) because, as we have shown, its growth function $g_\rho(k)$ has no maximum. Rather, the shorter the waves in the perturbation, the faster their growth. It should be stressed that despite this behavior, the linearized model for (3.1) is mathematically well-posed: for any final time $t$, the amplification of normal modes is bounded by $\exp(t g_\rho^\infty)$. Still, from an application perspective, properly answering the question of which wave lengths dominate once an augmented perturbation leaves the linear regime is important, but it is more challenging than in the usual situation.

While the PDE model (3.1) has no maximum wave number, reality does,
namely the vehicle scale. Specifically, wave numbers beyond a $k_{\text{max}}$, given by the minimum spacing between vehicles, have no practical meaning. One possible way to exclude features on such unphysically short length scales is to add a small amount of viscosity to the ARZ model (3.1), as in Kerner-Konhäuser [44, 45] for the PW model (2.9). In Figure 3.2b, this would change the functions $g_\beta(k)$ to drop off once $k$ gets close to the vehicle scale. Similarly, the numerical discretization of the PDE (3.1) on grids that are never finer than the vehicle scale will produce a wave number cut-off via the numerical viscosity of the method [56].

Another possibility (employed here in Chapter 4) is to consider small perturbations, rather than infinitesimal perturbations, and provide a model for the noise. Specifically, we argue that on real roads, perturbations of all wave lengths $k \in [0, k_{\text{max}}]$ will act: $k < k_{\text{max}}$ due to small variations in road features, wind, etc.; and $k \approx k_{\text{max}}$ due to variabilities across vehicles. The simplest such noise model is one where all wave numbers $k \in [0, k_{\text{max}}]$ appear with equal amplitudes, and perturbations with $k > k_{\text{max}}$ do not occur.

Because the growth function tends to have a plateau near $k_{\text{max}}$ (see Figure 3.2), this linear growth/noise model will yield that all wave numbers $k$ near but below $k_{\text{max}}$ will be amplified to reach the nonlinear regime at the same time. This is not unrealistic, as it means that noise close to the vehicle scale will dominate before systematic nonlinear wave effects kick in.
As a final remark we wish to point out that once solutions of the ARZ model (3.1) leave the linear regime (around a uniform base state), the nonlinear dynamics tend to turn those vehicle-scale waves into oscillations with shocks that then collide and merge to form nonlinear wave structures of much smaller amplitude to wave-length ratios. However, those nonlinear transient dynamics are extremely complicated, and this insight is merely based on our observations from numerous highly resolved computations (like those done in Chapter 4). What we will study, though, is the stability of true traveling wave solutions of (3.1) (jamitons) in the situation when the SCC (3.9) is violated (see Chapter 5).

3.3 Traveling wave analysis and jamitons

It is important to stress that macroscopic models (without explicit lane changing) can equivalently be written in Lagrangian variables. In (3.2) the equations are cast in Eulerian variables $\rho(x,t)$ and $q(x,t)$. The Lagrangian formulation, as used in [30, 88], employs the variables $\nu(\sigma,t)$ and $u(\sigma,t)$, where $\sigma$ is the (continuous) vehicle number, defined so that

$$d\sigma = \rho \, dx - \rho u \, dt,$$

(3.10)
and $v = 1/\rho$ is the specific traffic volume, i.e., the road length per vehicle (a proxy for vehicle spacing). In these variables the ARZ model reads as

$$v_t - u_\sigma = 0,$$

$$\left(u + \hat{h}(v)\right)_t = \frac{1}{\sigma}(\hat{U}(v) - u),$$  \hspace{1cm} (3.11)

where $\hat{h}(v) = h(1/v)$ and $\hat{U}(v) = U(1/v)$. The four assumptions discussed earlier on the model functions in Eulerian variables

$$\frac{dU}{d\rho} < 0, \quad \frac{d^2Q}{d\rho^2} < 0, \quad \frac{dh}{d\rho} > 0, \quad \frac{d^2}{d\rho^2} \rho h(\rho) > 0 \quad (3.12)$$

translate to the following assumptions in Lagrangian variables:

$$\frac{d\hat{U}}{dv} > 0, \quad \frac{d^2\hat{U}}{dv^2} < 0, \quad \frac{d\hat{h}}{dv} < 0, \quad \frac{d^2\hat{h}}{dv^2} > 0.$$  \hspace{1cm} (3.13)

For simplicity, the hats in those functions are omitted unless explicitly required for clarity. The characteristic speeds of (3.11) are

$$\lambda_1 = h'(v) , \quad \text{and} \quad \lambda_2 = 0,$$  \hspace{1cm} (3.14)

and the associated Rankine-Hugoniot shock jump conditions (which are equivalent to the Eulerian shock jump conditions in (3.3)) are

$$m [v] - [u] = 0,$$

$$[u] + [h(v)] = 0,$$  \hspace{1cm} (3.15)

where $-m$ is the propagation speed of the shock in the Lagrangian variables (in the Eulerian frame $m$ is the flux of vehicles through the shock). Note that, for contact discontinuities, the conditions are $m = 0$ and $[u] = 0$. 
In this work, both types of equivalent descriptions are employed. The Lagrangian model (3.11) is simpler, and better suited for any mathematical analysis. However, the quantities in the Eulerian model (3.1) are easier to visualize physically, so the computational analyses and results are presented in those variables.

Jamiton solutions can now be constructed via the Zel’dovich-von Neumann-Döring (ZND) theory [26]. One starts out with a traveling wave ansatz. In Eulerian variables, one seeks for solutions \( \rho(x,t) = \rho(\eta), \ u(x,t) = u(\eta) \) of (2.16) that depend on the single variable \( \eta = \frac{x-st}{\tau} \). In Lagrangian variables, one considers solutions \( v(\sigma,t) = v(\chi), \ u(\sigma,t) = u(\chi) \) of (3.11), where \( \chi = \frac{\sigma+mt}{\tau} \). Here \( s \) is the traveling wave speed in the road frame, while the Lagrangian wave speed \( -m \) relates to the mass flux \( m \) of vehicles through the wave.

We start with the Lagrangian formulation [88]. The traveling wave ansatz leads to

\[
\frac{m}{\tau} v'(\chi) - \frac{1}{\tau} u'(\chi) = 0 , \tag{3.16}
\]

\[
\frac{m}{\tau} u'(\chi) + h'(v(\chi)) \frac{m}{\tau} v'(\chi) = \frac{1}{\tau} (U(v(\chi)) - u(\chi)) . \tag{3.17}
\]

Equation (3.16) yields that

\[
m v - u = -s , \tag{3.18}
\]

where \( s \) is a constant of integration. Using (3.18) to substitute \( u \) by \( v \) in (3.17),
we obtain the scalar first-order *jamiton ODE*

\[ v'(\chi) = \frac{w(v(\chi))}{r'(v(\chi))}, \]  

(3.19)

where the two functions \( w \) and \( r \) are defined as

\[ w(v) = U(v) - (mv + s) \quad \text{and} \quad r(v) = mh(v) + m^2 v. \]

Because \( h'(v) < 0 \) and \( h''(v) > 0 \), the denominator in (3.19) has exactly one root, the *sonic value* \( v_S \) (occurring at the *sonic point*), such that \( h'(v_S) = -m \). The ODE (3.19) can be integrated through \( v_S \) if the numerator in (3.19) has a simple root at \( v_S \) as well. This leads to the *Chapman-Jouguet condition* [26]

\[ mv_S + s = U(v_S), \]  

(3.20)

which yields a relationship between the constants \( m \) and \( s \) as follows:

\[ m = -h'(v_S) \quad \text{and} \quad s = U(v_S) - mv_S. \]  

(3.21)

One therefore has a one-parameter family of smooth traveling wave solutions, parameterized by \( v_S \), each being solutions of (3.19).

Shocks that move with the same speed \(-m\) can be inserted into these smooth profiles. The first condition in (3.15) implies that the quantity \( mv - u \) is conserved across the shock (in addition to being conserved along the smooth parts by (3.18)). Both conditions in (3.15) together imply that \( r(v) \) is conserved across shocks. Hence, when integrating (3.19), one can at any
value $v^-$ insert a shock that jumps to a value $v^+$ with $r(v^+) = r(v^-)$ and continue integrating (3.19) from there. Moreover, for those shocks to satisfy the Lax entropy conditions [102], one can only jump downwards, i.e., $v^+ < v_S < v^-$. This in turn requires that the smooth jamiton profile $v(\chi)$ must be an increasing function. Using L’Hôpital’s rule in (3.19) at the sonic point yields that

$$0 < \frac{w'(v_S)}{r''(v_S)} = \frac{U'(v_S) - m}{mh''(v_S)} = \frac{U'(v_S) + h'(v_S)}{mh''(v_S)},$$

which means that the SCC is violated. In other words, as shown in [88], jamiton profiles with shocks can exist if and only if the SCC is violated.

The construction in Eulerian variables is analogous, albeit a bit more technical (cf. [27]). The traveling wave ansatz leads to

$$-s\rho' + (\rho u)' = 0,$$

$$(u - s - \rho h'(\rho))u' = U(\rho) - u.$$

Integrating the first equation yields

$$\rho(u - s) = m,$$

(3.22)

which allows one to substitute $\rho$ via $u$ and vice versa. The second equation becomes the jamiton ODE

$$u'(\eta) = \frac{(u - s)(U(\rho) - u)}{(u - s)^2 - mh'(\rho)},$$

(3.23)
where \( \rho = \frac{m}{u - s} \). The Chapman-Jouguet condition (matching roots of numerator and denominator) leads to the relations:

\[
m = \rho_S^2 h'(\rho_S) \quad \text{and} \quad s = U(\rho_S) - \rho_S h'(\rho_S).
\]

(3.24)

Shock and entropy conditions are then implemented analogous to the Lagrangian situation.

With these rules, jamiton solutions can be constructed (in either choice of variables). For a given choice of \( v_S \) (and thus uniform propagation speed), any pattern of solutions to (3.19) connected by shocks (satisfying the above conditions) results in a traveling wave solution. The jamitons between any two shocks can be arbitrarily short (with a small variation around \( v_S \)), or may be arbitrarily long. In fact, it is not even required for the jamitons between shocks to have the same length [27, 88].

One construction procedure of a jamiton can be as follows:

1. Choose a sonic density \( \rho_S \) (or a sonic specific volume \( v_S \)) that violates the SCC (2.21).

2. Calculate \( m \) and \( s \) using their corresponding formulas in (3.24) or (3.21).

3. Pick a valid downstream state \((\rho^+, u^+)\) where \( \rho^+ > \rho_S \) (or \((v^+, u^+)\) where \( v^+ < v_S \)).

4. Find the corresponding upstream state \((\rho^-, u^-)\) (or \((v^-, u^-)\)) using the Rankine-Hugoniot jump conditions (3.3) or (3.15).
5. Solve the jamiton ODE (3.23) or (3.19) from $u^+$ to $u^-$ (or $v^+$ to $v^-$) to get the jamiton profile in the $\eta$ variable (or the $\chi$ variable).

6. The $\rho$ profile (or the $u$ profile) follows from (3.22) or (3.18).

7. Use the chain rule (and/or (3.10)) to write the solution in terms of the $x$ variable.

Given that (3.22) is satisfied for any jamiton, we can write an equation for the flow rate at any point of a jamiton as

$$Q = \rho u = s\rho + m. \quad (3.25)$$

Using the formulas in (3.24), the above flow rate can be rewritten as

$$Q = s(\rho - \rho_S) + \rho_S U(\rho_S). \quad (3.26)$$

This means that if a jamiton were to be plotted in the $(\rho, Q)$ plane, it would be a straight line segment with slope $s$, the traveling speed of the jamiton, and would pass through the point $(\rho_S, \rho_S U(\rho_S))$. This point is on the the curve $\rho U(\rho)$, which represents the equilibrium flow rate. The end points of the jamiton segment are $(\rho^+, \rho^+ u^+)$ and $(\rho^-, \rho^- u^-)$, representing the two jamiton states across the shock.

Figure 3.3 below illustrates a jamiton profile along with its corresponding jamiton line plotted on top of the equilibrium fundamental diagram. This particular jamiton has a sonic density $\rho_S = 70$ veh/km, and a traveling wave
Figure 3.3: A jamiton with sonic density $\rho_S = 70$ veh/km and length $L = 274.1$ m. $\rho(x)$ and $u(x)$ are plotted in the left panels. The jamiton is plotted on top of the fundamental diagram in the right panel.

speed $s = 0.54$ km/hr. The jamiton profile in the left panels is plotted on a periodic domain, with the shock visible inside of it. The length of this jamiton is $L = 274.1$ m. The number of vehicles $N$ in this jamiton is calculated as:

$$ N = \int_0^L \rho(x) dx = 13.7109 \text{ veh} . $$

Notice that the number of vehicles in a jamiton is not restricted to integer values because we have a fully continuous model.

In fact, for any sonic density that violates the SCC, a jamiton of any length can be constructed. Very short jamitons do not make physical sense, but they
are nevertheless valid solutions of (3.1). Figure 3.4 below shows an illustration of a short jamiton with length 15 m, containing 0.89 vehicles.

**Figure 3.4:** A short jamiton with length $L = 15$ m, which contains 0.89 vehicles. $\rho(x)$ and $u(x)$ are plotted in the left panels. The jamiton is plotted on top of the fundamental diagram in the right panel.

In addition, arbitrarily long jamitons can also be constructed. These jamitons tend to have a nearly flat tail, as illustrated in Figure 3.5 below. The jamiton in this figure has a length that is larger than 1 km.

Moreover, the sonic density alone determines the traveling speed of a jamiton. This means that jamitons with the same sonic density $\rho_S$ have the same traveling speed $s$. In fact, the traveling speed of jamitons decreases as the sonic density increases. To prove that, first write $s$ as a function of $\rho_S$ using
Figure 3.5: A long jamiton with length $L > 1$ km and with a nearly flat tail. $\rho(x)$ and $u(x)$ are plotted in the left panels. The jamiton is plotted on top of the fundamental diagram in the right panel.

\[(3.24): s(\rho_S) = U(\rho_S) - \rho_S h'(\rho_S).\] Then calculate the derivative of $s(\rho_S)$:

\[
s'(\rho_S) = U'(\rho_S) - h'(\rho_S) - \rho_S h''(\rho_S)
= (U'(\rho_S) + h'(\rho_S)) - (2h'(\rho_S) + \rho_S h''(\rho_S))
= (U'(\rho_S) + h'(\rho_S)) - (\rho_S h'(\rho_S))''.
\]

The first term in (3.28) is negative because any sonic density $\rho_S$ violates the SCC. The second term in (3.28) is positive because the function $\rho h(\rho)$ is strictly convex, which is a restriction on $h$ discussed in §3.1. This implies that $s'(\rho_S)$
is negative, and therefore $s$ is a decreasing function of $\rho_s$.

For a given sonic density, the “longest” jamiton is a theoretical jamiton with infinite length. These jamitons are valid solutions to the ARZ model, and are referred to as maximal jamitons. They can indeed be represented in the fundamental diagram as finite straight line segments. Figure 3.6 shows plots of some of these maximal jamitons, along with the maximal jamiton envelopes. These envelopes represent a closed boundary, where any point of

![Figure 3.6: A plot of some maximal jamitons (blue lines) in the fundamental diagram along with their corresponding sonic points. The magenta curve represents the maximal jamiton envelopes.](image-url)
any jamiton has to fall inside of it. This figure also shows the fact that $s$ decreases as $\rho_S$ increases. This is evident in the plot where the slopes of the jamiton lines decrease as the corresponding sonic density increases.

As discussed above, every jamiton has many characteristic quantities associated with it. Table 3.1 below summarizes some of these quantities and presents some of their related formulas.

One of these quantities alone cannot uniquely define a jamiton. For example, many different jamitons share the same length, and many different jamitons share the same traveling speed. However, two non-redundant jamiton quantities can uniquely pick a jamiton. An example of two redundant quantities would be $\rho_S$ and $s$: Since $s = U(\rho_S) - \rho_S h'(\rho_S)$, knowing $s$ doesn’t add any extra information if $\rho_S$ is already known. On the other hand, an example of two non-redundant quantities would be $\rho_S$ and $L$. For instance, there is only one jamiton with $\rho_S = 50$ veh/km and $L = 250$ m.

Therefore, any valid “phase space” that spans the jamiton solutions has to be two-dimensional. In addition, for any of these phase planes, there are boundaries, outside of which no jamitons exist. Depending on the phase plane, these boundaries may represent the infinite jamitons, the infinitesimal jamitons, or a transition from linear stability to instability (SCC satisfied vs. violated).

Figure 3.7 below shows five different jamiton phase planes, where some
<table>
<thead>
<tr>
<th>Quantity</th>
<th>Description</th>
<th>Units</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho_S$</td>
<td>sonic density</td>
<td>veh/km</td>
<td>-</td>
</tr>
<tr>
<td>$v_S$</td>
<td>sonic specific volume</td>
<td>km/veh</td>
<td>$1/\rho_S$</td>
</tr>
<tr>
<td>$\rho^+$</td>
<td>downstream density</td>
<td>veh/km</td>
<td>-</td>
</tr>
<tr>
<td>$\rho^-$</td>
<td>upstream density</td>
<td>veh/km</td>
<td>-</td>
</tr>
<tr>
<td>$v^+$</td>
<td>downstream specific volume</td>
<td>km/veh</td>
<td>$1/\rho^+$</td>
</tr>
<tr>
<td>$v^-$</td>
<td>upstream specific volume</td>
<td>km/veh</td>
<td>$1/\rho^-$</td>
</tr>
<tr>
<td>$A_s$</td>
<td>shock amplitude</td>
<td>veh/km</td>
<td>$\rho^- - \rho^+$</td>
</tr>
<tr>
<td>$s$</td>
<td>traveling wave speed</td>
<td>km/hr</td>
<td>$U(\rho_S) - \rho_S h'(\rho_S)$</td>
</tr>
<tr>
<td>$m$</td>
<td>mass flux of vehicles</td>
<td>veh/hr</td>
<td>$\rho_S^2 h'(\rho_S)$</td>
</tr>
<tr>
<td>$L$</td>
<td>jamiton length</td>
<td>km</td>
<td>$\int_{\rho^-}^{\rho^+} \frac{1}{\rho^2} \left( \frac{m h'(\rho) + m^2}{\rho U'(\rho) - (m + s \rho)} \right) d\rho$</td>
</tr>
<tr>
<td>$N$</td>
<td>number of vehicles</td>
<td>veh</td>
<td>$\int_0^L \rho(x) dx$</td>
</tr>
<tr>
<td>$\bar{v}$</td>
<td>average specific volume</td>
<td>km/veh</td>
<td>$L/N$</td>
</tr>
<tr>
<td>$\bar{\rho}$</td>
<td>average density</td>
<td>veh/km</td>
<td>$N/L$</td>
</tr>
<tr>
<td>$\bar{Q}$</td>
<td>average flow rate</td>
<td>veh/hr</td>
<td>$s \bar{\rho} + m$</td>
</tr>
</tbody>
</table>

Table 3.1: A summary of some jamiton related quantities, units associated with them, and related formulas.
boundaries are plotted, and the jamiton region is shaded in blue. The first phase plane shown in 3.7a is \((\rho_S, L)\). Since for any \(\rho_S\) that violates the SCC a jamiton of any length exists, the jamiton region is a vertical strip that stretches from 0 to \(\infty\) in the \(L\) axis, and is bounded by the SCC in the \(\rho_S\) axis. The next three phase planes \((\bar{\rho}, s)\), \((\rho_S, \rho^+)\) and \((\rho_S, \bar{\rho})\), shown in 3.7b, 3.7c and 3.7d respectively, have bounded jamiton regions. In each of them, the jamiton region is enclosed by two boundaries: the long jamiton limit where the jamiton length tends to \(\infty\) as it is approached, and the short jamiton limit where the jamiton length tends to 0 as it is approached. Moreover, any point in the four phase planes discussed so far uniquely determines a jamiton solution. Notably, the phase plane in 3.7d highlights an important fact: the average density \(\bar{\rho}\) of any jamiton is always smaller than its sonic density \(\rho_S\). A proof of this fact is presented in [89].

The phase plane \((\bar{\rho}, \bar{Q})\) in 3.7e is different from the others. For instance, the boundaries representing long and short jamiton limits both coincide with the equilibrium curve \(\rho U(\rho)\). On this common boundary, each point corresponds to two different jamitons: one with zero length, and another with infinite length. The jamiton region in this phase plane is enclosed between the equilibrium curve and the lower envelope of the maximal jamitons shown in Figure 3.6. In addition, each point inside the jamiton region corresponds to two different jamitons, but each point on the lower envelope boundary corresponds to a
Figure 3.7: Five different jamiton phase planes. The jamiton regions are enclosed by different boundaries and are shaded in blue.
single unique jamiton. This means that for a given $\bar{\rho}$ there is only one jamiton
that minimizes the average flow rate $\bar{Q}$. In fact, this unique jamiton has a
sonic density $\rho_s$ that satisfies the following equation [89]:

$$\bar{\rho} = -\frac{m'(\rho_s)}{s'(\rho_s)} = \frac{\rho_s \left( 2h'(\rho_s) + \rho_s h''(\rho_s) \right)}{\left( h'(\rho_s) + \rho_s h''(\rho_s) \right) - U'(\rho_s)}. \quad (3.29)$$

And it follows that the average flow of that jamiton is

$$\bar{Q} = m(\rho_s) + \bar{\rho} s(\rho_s) = \rho_s^2 h'(\rho_s) + \bar{\rho} \left( U(\rho_s) - \rho_s h'(\rho_s) \right). \quad (3.30)$$

While all of the jamitons discussed above constitute feasible traveling wave
solutions of the ARZ model (3.2), it does not mean that all such profiles would
be dynamically stable under perturbations. In fact, both numerical evidence
(see Chapter 4) as well as intuition dictate that neither very short, nor very
long jamitons should be stable. The former is unlikely to be stable because
they can be thought of as a small (sawtooth) perturbation of the constant $\rho_s$
(or $v_s$) state which is unstable because the SCC is violated (see §3.2); and the
latter because their long tail will itself be close to a constant which, if that
state violates the SCC, will be dynamically unstable. In other words, jamitons
that are too short merge and have longer waves form between them, and long
jamitons have new instabilities grow in their tails. It is only the middle range
of jamitons (not too short and not too long) that is expected to be dynamically
stable. Only those should arise in practice.

This dynamic stability (of the jamitons themselves) has not been studied
before. This will be explored next, by first conducting a computational study in Chapter 4 that confirms the intuition above and quantifies it; and then deriving and analyzing linear perturbation equations for the jamiton solutions in Chapter 5.
CHAPTER 4

COMPUTATIONAL STUDY

OF JAMITON STABILITY

To understand the dynamic stability of jamitons, we conduct a systematic study of the ARZ model (3.2) via direct numerical computation. After constructing a periodic jamiton as outlined in §3.3, we insert that profile as an initial condition into a numerical scheme (§4.2) and investigate whether the profile is maintained under small perturbations (§4.3).

4.1 The Godunov Method

The goal in approximating a hyperbolic conservation law numerically is finding an approximation of a weak solution (with potential discontinuities or shocks) that satisfies an entropy condition. Finite volume methods [57] do a
good job in doing that for non-linear hyperbolic systems of conservation laws of the form $U_t + F(U)_x = 0$.

$U(x, t)$ defined on $\Omega \subseteq \mathbb{R}^2$ is a weak solution of the hyperbolic system if it satisfies

$$\int_0^\infty \int_{-\infty}^\infty \phi [U_t + F(U)_x] \ dx dt = 0$$

(4.1)

for every $C^1$ function $\phi : \Omega \rightarrow \mathbb{R}$ with compact support. Applying integration by parts to (4.1), we get:

$$\int_0^\infty \int_{-\infty}^\infty \phi_t U + \phi_x F(U)_x \ dx dt = -\int_{-\infty}^\infty \phi(x, 0) U(x, 0) \ dx$$

(4.2)

Defining $\phi$ to be 1 on the rectangle $[x_1, x_2] \times [t_1, t_2]$ and zero outside the rectangle $[x_1 - \epsilon, x_2 + \epsilon] \times [t_1 - \epsilon, t_2 + \epsilon]$, the expression in (4.2) becomes:

$$\int_{x_1}^{x_2} U(x, t_1) - U(x, t_2) \ dx + \int_{t_1}^{t_2} F(U(x_1, t)) - F(U(x_2, t)) \ dt = 0$$

(4.3)

The Godunov method [93], which is a finite volume method, is based on the integral form of the hyperbolic conservation law in (4.3). Given the grid cell $C_i = [x_i - \Delta x/2, x_i + \Delta x/2]$, where $\Delta x$ is the cell size, define the approximation of $U$ at the cell centers by the cell average as

$$U_i^n \approx \frac{1}{\Delta x} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} U(x, t_n) \ dx$$

(4.4)

to get a piecewise constant function $\tilde{U}(x, t_n)$ defined by the above $U_i^n$'s. Then, to find $U_i^{n+1}$, first solve for $\tilde{U}^n(x, t_{n+1})$, and average again over each cell to get

$$U_i^{n+1} \approx \frac{1}{\Delta x} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \tilde{U}(x, t_{n+1}) \ dx$$

(4.5)
To recapitulate, The Godunov method is an iterative method of three consecutive steps. Given $U(x, t_n)$, Find $U^n_i$ by averaging over each cell. Then, evolve resulting piecewise constant function in time to $t_{n+1}$, and average again to get $U^{n+1}_i$. And then repeat this process until reaching the final time. This method can also be summarized with one update rule, which is usually how it is implemented. To get from $U^n_i$ to $U^{n+1}_i$ in the Godunov scheme, one only needs to use the equation

$$U^{n+1}_i = U^n_i - \frac{\Delta t}{\Delta x} \left( F^n_{i+\frac{1}{2}} - F^n_{i-\frac{1}{2}} \right) \quad (4.6)$$

where $F^n_{i+\frac{1}{2}}$ and $F^n_{i-\frac{1}{2}}$ are called numerical fluxes, defined as follows

$$F^n_{i+\frac{1}{2}} = F\left( \hat{U}(U^n_i, U^n_{i+1}) \right) \quad \text{and} \quad F^n_{i-\frac{1}{2}} = F\left( \hat{U}(U^n_{i-1}, U^n_i) \right) \quad (4.7)$$

where $\hat{U}(U_L, U_R)$ is the solution of the Riemann problem between the initial states $U_L$ and $U_R$, evaluated at the cell boundary. So, to implement Godunov’s method, we need to determine the numerical flux at the interface between any two cells, which requires solving the Riemann problem between any two left and right states. In fact, it is not required to solve the Riemann problem exactly, and an approximation suffices [33].

An important detail in the Godunov method is that the time step $\Delta t$ should be chosen in a way to avoid the interaction between waves from neighboring cells. To achieve that, choose a $\Delta t$ that satisfies the CFL condition

$$\Delta t \leq \frac{\Delta x}{s_{\text{max}}}, \quad (4.8)$$
where $s_{\text{max}}$ is the maximum possible wave speed.

The Godunov method is a consistent numerical scheme, which means that it solves the integral form of the conservation form, if the numerical flux function $\mathcal{F}(U_L, U_R)$ is consistent [57], or $\mathcal{F}(U, U) = F(U)$. Using the consistency of the scheme and the Lax-Wendroff theorem [50], it is shown that if the Godunov method converges as the grid is refined, then it converges to a weak solution of the conservation law. This is not enough since there are possibly many weak solutions, but we want to converge to the physically correct unique entropy solution. The Lax-Wendroff theorem does not guarantee that the solution converges to the solution that satisfies the entropy condition. However, the Godunov method always converges to the entropy solution if the Riemann solution used to define the numerical flux function satisfies the entropy condition [57].

4.2 Numerical scheme for the ARZ model with relaxation term

The ARZ model (3.2) is a system of hyperbolic conservation laws with a relaxation term. The hyperbolic part of the system can be solved using a finite volume scheme based on an approximate Riemann solver [57]. In particular, we use first order Godunov scheme. To find the numerical flux at
the cell boundaries, we use the HLL approximate Riemann solver [32], which guarantees that the numerical fluxes satisfy the entropy condition [52].

The HLL Riemann solver leads to the following definition of the numerical flux function

$$
\mathcal{F}(U_L, U_R) = \begin{cases} 
F(U_L) & \text{if } s_L \geq 0 \\
\hat{F} & \text{if } s_L < 0 < s_R \\
F(U_R) & \text{if } s_R \leq 0
\end{cases} 
$$

where

$$
\hat{F} = \frac{s_R \mathbf{F}(U_L) - s_L \mathbf{F}(U_R) + s_L s_R (U_R - U_L)}{s_R - s_L} 
$$

and $s_L$ and $s_R$ are estimates of the wave speeds of the left and right states respectively. There are many ways to approximate these wave speeds, but one simple choice [19] for $2 \times 2$ systems can be as follows

$$
s_L = \min\{\lambda_1(U_L), \lambda_1(U_R)\} ,
$$

$$
s_R = \max\{\lambda_2(U_L), \lambda_2(U_R)\} .
$$

The above definition of the wave speeds in (4.11) guarantees that $s_L < s_R$ for the ARZ model.

Given the grid cell $C_i = [x_i - \Delta x/2, x_i + \Delta x/2]$, where $\Delta x$ is the cell size, let

$$
U^n_i = \begin{bmatrix} 
\rho^n_i \\
\rho u^n_i 
\end{bmatrix}, \quad \text{and} \quad F^n_{i+\frac{1}{2}} = \begin{bmatrix} 
(F\rho)_i^n \\
(Fq)_i^n 
\end{bmatrix} ,
$$
denote the approximate solution (cell average) in cell $C_i$ and the numerical flux at the boundary between cells $C_i$ and $C_{i+1}$, respectively, at time $n\Delta t$ ($n$-th time step).

A numerically robust treatment of the relaxation term is achieved by treating it implicitly, resulting in the semi-implicit update rule

$$
\begin{bmatrix}
\rho_{i}^{n+1} \\
q_{i}^{n+1}
\end{bmatrix}
= 
\begin{bmatrix}
\rho_{i}^{n} \\
q_{i}^{n}
\end{bmatrix}
- \frac{\Delta t}{\Delta x} \begin{bmatrix}
(F_{\rho})_{i+\frac{1}{2}}^{n} \\
(F_{q})_{i+\frac{1}{2}}^{n}
\end{bmatrix}
- \begin{bmatrix}
(F_{\rho})_{i-\frac{1}{2}}^{n} \\
(F_{q})_{i-\frac{1}{2}}^{n}
\end{bmatrix}
+ \frac{\Delta t}{\tau} \begin{bmatrix}
0 \\
\rho_{i}^{n+1} (U(\rho_{i}^{n+1}) + h(\rho_{i}^{n+1})) - q_{i}^{n+1}
\end{bmatrix}.
$$

(4.12)

This ensures stability even when $\tau$ is small. Note that, because the implicit term appears only in the $q$-equation and because it is linear in $q_{i}^{n+1}$, the formally semi-implicit numerical scheme is actually fully explicit and the update step can be conducted in two sub-steps:

1) Update the $\rho$ component explicitly:

$$
\rho_{i}^{n+1} = \rho_{i}^{n} - \frac{\Delta t}{\Delta x} \left( (F_{\rho})_{i+\frac{1}{2}}^{n} - (F_{\rho})_{i-\frac{1}{2}}^{n} \right).
$$

2) Now, with $\rho_{i}^{n+1}$ known from the first step, update

$$
(1 - \frac{\Delta t}{\tau}) \ q_{i}^{n+1} = q_{i}^{n} - \frac{\Delta t}{\Delta x} \left( (F_{q})_{i+\frac{1}{2}}^{n} - (F_{q})_{i-\frac{1}{2}}^{n} \right) + \frac{\Delta t}{\tau} \rho_{i}^{n+1} \left( U(\rho_{i}^{n+1}) + h(\rho_{i}^{n+1}) \right).
$$
4.3 Results on the stability of jamitons

Using the numerical scheme described above, we conduct a computational investigation of the stability of jamitons (of the ARZ model (3.2) with the specific model functions and parameters described in §3.1). Specifically, we classify the jamitons as follows: Evolve the solution up to some large final time, while regularly adding small perturbations. Then a jamiton is classified as stable if the jamiton profile is (within a tolerance) maintained at the final time, and unstable otherwise.

To classify a given jamiton $J_0 = [\rho_0(x), u_0(x)]^T$ (of length $L_0$, with sonic density $\rho_S$, upstream density $\rho_0^+$, and speed $s_0$), we set up a periodic domain of length $4L_0$ with initial conditions $[\rho_{ic}(x), u_{ic}(x)]^T = [\rho_0(x \mod L_0), u_0(x \mod L_0)]^T$, i.e., the initial profile is four consecutive jamitons $J_0$ with shocks in between. We discretize using 10,000 grid cells, and run the numerical scheme (from §4.2) up to $t_{\text{final}} = 3,000$ (seconds; we omit units below).

During the numerical solution process, a small smooth perturbation is added to the vehicle velocity field $u = q/\rho - h(\rho)$ in each step. The perturbation in the $n$-th step is

$$p^n(x) = \sqrt{\Delta t} c(t) \frac{1}{\sqrt{\ell}} \sum_{\nu=1}^{\ell} \xi^n_\nu \sin \left( \frac{2\pi \nu x}{L_0} \right),$$

where the $\xi^n_\nu \in \mathcal{N}(0, 1)$ are normally distributed random numbers with mean zero and standard deviation 1. As in the Euler-Maruyama method [67], the
additive noise is scaled with $\sqrt{\Delta t}$. The value $\ell$ is chosen so that the highest frequency mode has wavelength $\frac{L_0}{\ell}$ that is not below the vehicle length $1/\rho_{\text{max}}$, i.e., $\ell = \lfloor L_0 \rho_{\text{max}} \rfloor$. In other words, we have white noise exactly until the vehicle scale, which is well-resolved by the numerical scheme. Finally, the noise scale is $c(t) = \frac{1}{100} u_{\text{max}}$ for $t \leq 100$, and $c(t) = \frac{1}{1000} u_{\text{max}}$ for $t > 100$.

The rationale for this larger initial “thermal noise” is, like in probabilistic optimization techniques, to make it easier for the solutions to escape their initial configuration in case it is only mildly unstable.

Once the solution at $t_{\text{final}}$ is found, we first determine the number of shocks. If that number is not equal to 4, we immediately classify the jamiton $J_0$ as unstable. Otherwise, we check the jamiton speed $s$ by plotting the points $(\rho(x_i, t_{\text{final}}), \rho(x_i, t_{\text{final}})u(x_i, t_{\text{final}}))$ for $i = 1, \ldots, 10000$ in the fundamental diagram (FD), and calculate $s$ as the least squares best fit slope of these data points (see §3.3 for the reason why $s$ is the slope in the FD). If $|s - s_0| > 0.5 \text{m/s}$, we classify $J_0$ as unstable. Otherwise, we classify $J_0$ as stable.

This process is now conducted (and run in parallel on a HPC cluster) for 980 different jamitons that are sampled as follows. First we sample 35 values of $\rho_S$ equidistant in the $\rho$-interval where the SCC is violated. Then, for each $\rho_S$, we pick 28 values of $\rho^+$ in $[\rho_S, \rho_M]$, where $\rho_M$ is the upstream density corresponding to the infinite jamiton [88].

The results of this classification are displayed in Figure 4.1. Each of the
Figure 4.1: Classification of 980 jamitons into stable and unstable, displayed in four different phase planes. In each plane, the dashed brown line represents the zero length jamiton, and the dark blue line is the limit of jamitons with infinite length. The two disconnected red regions correspond to the “splitting” and “merging” instabilities, respectively.
four panels shows the same results, but in four different phase planes. Each jamiton is uniquely determined by two parameters: (i) the sonic density $\rho_S$ or equivalently the wave speed $s$; and (ii) the downstream shock density $\rho^+$, or equivalently, the average density $\bar{\rho}$ across the jamiton, or equivalently, the jamiton length $L$. Panels 4.1a, 4.1c, and 4.1d have the $\rho_S$ on the horizontal axis, and $L$, $\rho^+$, and $\bar{\rho}$ respectively, on the vertical axis. Panel 4.1b displays $s$ vs. $\bar{\rho}$. In each quantity except $L$, the jamiton region (where the SCC (3.9) is violated) spans an interval. The dashed brown curve corresponds the zero-length jamiton limit (in which $\rho_S = \rho^+ = \bar{\rho}$), while the solid dark blue curve represents the limit of infinitely long jamitons. Inside that jamiton domain, the 980 investigated jamitons are displayed as colored dots: stable jamitons are light blue; unstable jamitons are red. Note that the void regions visible in Panel 4.1a (top left), Panel 4.1b (bottom left), and Panel 4.1d (bottom right), also possess jamitons that were not simulated due to the sampling strategy of the 980 examples.

The results display intriguingly clear patterns: there appear to be two smooth curves inside the jamiton region that separate the stable from the unstable jamitons. Specifically, there are two unstable regions separated by a stable region: short jamitons which perturbations cause to coalesce into bigger ones (a “merging” instability); and long jamitons in which the long tail is linearly unstable and sheds growing waves (a “splitting” instability).
This last characterization of these two mechanisms is based on observing the time-evolution of the computations, as well as the stability analysis below.

Figure 4.2 below is another classification of jamiton stability, where the results are displayed in the (\(\bar{\rho}, \bar{Q}\)) phase plane. Note that every point in the jamiton region of this phase plane corresponds to two different jamitons. Here, the results display a different pattern from the one discussed in the four phase planes in figure 4.1. In this phase plane, two regions are observed: a narrow strip just above the lower envelope in which jamitons are stable, and everything above that where jamitons are unstable.

![Figure 4.2: Stability classification in the phase plane \((\bar{\rho}, \bar{Q})\).](image)

Figure 4.2: Stability classification in the phase plane \((\bar{\rho}, \bar{Q})\).
CHAPTER 5

STABILITY ANALYSIS OF JAMITON SOLUTIONS

We now move towards a mathematical analysis of the dynamic stability of jamitons. For this we switch to the Langrangian variables introduced in §3.3. Consider a given jamiton \([v_0(\sigma, t), u_0(\sigma, t)]^T\) with sonic specific volume \(v_{s_0}\), and Lagrangian length (which is actually the number of vehicles in the jamiton) \(N_0\). We start by writing the (Lagrangian) ARZ model (3.11) in the frame of reference of this jamiton, which has a propagation speed \(-m_0 = h'(v_{s_0})\). Thus we introduce the variables \(\chi = \frac{\sigma + m_0 t}{\tau}\) (the same variable used in §3.3 to construct the jamitons) and the non-dimensional time \(t_* = \frac{t}{\tau}\) (for consistency with the scaling used for \(\chi\)). Because of that last choice, any instability growth rate computed with these variables needs to be scaled by \(\tau\) to recover physical
units.

In the coordinates defined above, equations (3.11) become

\[ v_t + (m_0v - u)_\chi = 0 , \]

\[ (u + h(v))_t + m_0 (u + h(v))_\chi = U(v) - u . \]

This system is in conservative form, with conserved quantities \( v \) and \( q = u + h(v) \). The characteristic speeds of (5.1) are

\[ \lambda_1 = m_0 + h'(v) \quad \text{and} \quad \lambda_2 = m_0 . \]

The Rankine-Hugoniot shock jump conditions associated with (5.1) are

\[ (-m_0 + \tilde{m}) [v] + [u] = 0 , \]

\[ [u] + [h(v)] = 0 . \]

where \( \tilde{m} \) is the shock speed in the \( \chi-t_* \) frame. Contacts require \( \tilde{m} = m_0 \) and \( [u] = 0 \).

5.1 A Perturbation system for single-jamiton waves

We now formulate a linear perturbation system of (5.1). There are two fundamental differences to the linear perturbation analysis for uniform flow presented in §3.2. First, because the jamiton profile is non-constant, we obtain a variable-coefficient linear system. Second, because the jamiton contains
a shock, we must introduce a perturbation to the shock’s position as an additional variable (a variable not needed for perturbations of smooth solutions).

As we will see below in more detail, both aspects render this analysis significantly more complicated than the one in §3.2.

Here we consider the stability of periodic jamiton profiles with one shock per period, under periodic perturbations of fixed period. Note that this setup excludes the possibility of jamitons merging by means of adjacent shocks approaching each other. Hence, we only study the “splitting instability” for long jamitons, not the “merging instability” for short jamitons (see §4.3).

Consider a periodic jamiton profile \([v_0(\sigma, t), u_0(\sigma, t)]^T\) of length \(N_0\) between shocks, and write it as \([v_0(\chi), u_0(\chi)]^T\) — a solution of (5.1) on \([0, N_0]\) with the shock placed at 0. Now write \(v(\chi, t) = v_0(\chi) + \delta v(\chi, t_\star)\) and \(u(\chi, t_\star) = u_0(\chi) + \delta u(\chi, t_\star)\), where \(\delta v\) and \(\delta u\) are infinitesimal perturbations. Substituting into (5.1) yields the linear system for \(\delta v\) and \(\delta u\):

\[
\delta v_{t_\star} + (m_0 \delta v - \delta u)_\chi = 0 ,
\]

\[
(\delta u + h'(v_0) \delta v)_{t_\star} + m_0 (\delta u + h'(v_0) \delta v)_\chi = U'(v_0) \delta v - \delta u .
\]

We also need to track the infinitesimal perturbation of the shock position \(\chi = \mu(t_\star)\). We do so by implementing the Rankine-Hugoniot conditions (5.3) in a way consistent with solving (5.1) on \([0, N_0]\) with periodic boundary conditions. This then generates boundary conditions for (5.4). The first equation in (5.3)
yields

\[(\dot{\mu} - m_0)\left([v_0] + [\delta v] + \mu[v_0\chi]\right) + [u_0] + [\delta u] + \mu[u_0\chi] = 0.\]

Expanding this equation, ignoring terms beyond \(O(\mu)\), and using that \([u_0] - m_0[v_0] = 0\) and \(u_0\chi - m_0v_0\chi = 0\), we obtain

\[\dot{\mu}[v_0] - m_0[\delta v] + [\delta u] = 0.\]

The second equation in (5.3) becomes

\[[u_0] + [\delta u] + \mu[u_0\chi] + [v_0] + [h'(v_0)\delta v] + \mu[h(v_0)\chi] = 0.\]

Again, ignoring terms beyond \(O(\mu)\) and using that \([u_0] + [h(v_0)] = 0\), we get

\[[\delta u] + [h'(v_0)\delta v] + \mu[u_0\chi + h'(v_0)\chi] = 0.\]

In this setup the bracket notation denotes \([\zeta] = \zeta(0^+) - \zeta(N_0^-)\). Therefore, we have derived the following variable-coefficient linear model for \(\delta v\) and \(\delta u\) on \([0, N_0]\), with boundary conditions that involve the shock position perturbation \(\mu\):

\[\delta v_t + (m_0\delta v - \delta u)\chi = 0,\]  
\[\left(\delta u + h'(v_0)\delta v\right)_t + m_0\left(\delta u + h'(v_0)\delta v\right)\chi = U'(v_0)\delta v - \delta u,\]  

with boundary condition:  
\[\left[\delta u + h'(v_0)\delta v\right] = -\mu \left[u_0\chi + h'(v_0)\chi\right],\]

where \(\mu\) satisfies the ODE:

\[\dot{\mu} = \frac{m_0[\delta v] - [\delta u]}{[v_0]} .\]  
(5.6)
We conduct two further simplifications to the model. First, we transform it to characteristic form by writing it in terms of the Riemann variables $\delta u$ and $\delta q = \delta u + h'(v_0)\delta v$. Second, we replace the shock perturbation variable $\mu$ by a Robin b.c. for the PDE, as follows. Differentiating the boundary conditions $[\delta q] = -\mu [u_0 x + h'(v_0)\chi]$ with respect to time yields

$$\frac{d}{dt_x} [\delta q] = -[u_0 x + h'(v_0)\chi] \dot{\mu} = -\frac{[u_0 x + h'(v_0)\chi]}{[v_0]} \left( \left[ -\frac{m_0}{h'(v_0)} \delta q \right] - \left[ \left( 1 + \frac{m_0}{h'(v_0)} \right) \delta u \right] \right).$$

Using the definition of $\delta q$ above, and the second equation in (5.5), we can write $\delta q_x$ as:

$$\delta q_x = -m_0 \delta q_x + \left( -\frac{h'(v_0) - U'(v_0)}{h'(v_0)} \right) \delta u + \left( \frac{U'(v_0)}{h'(v_0)} \right) \delta q . \quad (5.7)$$

This then allows us to obtain the Robin boundary conditions for the PDE. Altogether, we get the following system

$$\delta u_t + (m_0 + h'(v_0)) \delta u_x = \left( \frac{m_0 h'(v_0) - h'(v_0) - U'(v_0)}{h'(v_0)} \right) \delta u + \left( \frac{U'(v_0) - m_0 h'(v_0)}{h'(v_0)} \right) \delta q ,$$

$$\delta q_t + m_0 \delta q_x = \left( -\frac{h'(v_0) - U'(v_0)}{h'(v_0)} \right) \delta u + \left( \frac{U'(v_0)}{h'(v_0)} \right) \delta q , \quad (5.8)$$

with boundary condition

$$\delta q_x(0) + k_L \delta q(0) = \delta q_x(N_0) + k_R \delta q(N_0) + c_L \delta u(0) + c_R \delta u(N_0) . \quad (5.9)$$

The coefficients are computable from the jamiton functions as

$$k_L = K(0) , \quad k_R = K(N_0) , \quad c_L = -C(0) , \quad \text{and} \quad c_R = C(N_0) ,$$
where

\[ K(\chi) = \frac{1}{m_0} - \frac{[u_0\chi + h'(v_0)\chi]}{[v_0]} \frac{1}{h'(v_0(\chi))} \frac{h'(v_0(\chi)) + U'(v_0(\chi))}{m_0h'(v_0(\chi))}, \]

\[ C(\chi) = \frac{[u_0\chi + h'(v_0)\chi]}{[v_0]} \left( \frac{1}{m_0} + \frac{1}{h'(v_0(\chi))} \right) + \frac{h'(v_0(\chi)) + U'(v_0(\chi))}{m_0h'(v_0(\chi))}. \]

5.2 Qualitative characterization of the perturbation system

We now adopt a short notation for the jamiton perturbation system (5.8), with b.c. (5.9), by writing \((u, q)\) and \((x, t)\) in place of \((\delta u, \delta q)\) and \((\chi, t_*)\), and introducing coefficient functions to obtain:

\[ u_t + b_1(x)u_x = a_{11}(x)u + a_{12}(x)q, \]

\[ q_t + b_2q_x = a_{21}(x)u + a_{22}(x)q, \]

with b.c.

\[ (q_x + k_L q)(0) = (q_x + k_R q)(N_0) + c_L u(0) + c_R u(N_0). \]

The characteristic speed \(b_2 = m_0 > 0\) is constant and positive. In turn, \(b_1(x)\) vanishes at the sonic point \(x_S\), it is negative for \(x < x_S\) and positive for \(x > x_S\). Hence, the only in-going characteristic is at \(x = 0\), for \(q\) (consistent with a single b.c.). The function \(a_{11}(x)\) crosses from negative to positive at \(x_S\) as well, and it is always negative for \(x < x_S\); it may or may not cross back to
negative for some $x > x_S$. Finally, $a_{22}(x) < 0$ everywhere. Figures 5.1 and 5.2 display the functions and characteristic curves, respectively, for an example jamiton.

Qualitatively, the solutions of (5.10) behave as follows:

- The solutions are generally wave-like in nature, because (5.10) is an advection-reaction system.

- Waves enter the $q$-field at $x = 0$ and are transported with the $q$-field to the right with constant speed $b_2$, while being dampened by the $a_{22}$-term and modified (via the $u$-field) through the $a_{21}$-term.

- The $q$-field constantly feeds into the $u$-field via the $a_{12}$-term.

- For $x < x_S$, the $u$-field is transported towards $x = 0$ and dampened by the $a_{11}$-term.

- For $x > x_S$, the $u$-field is transported towards $x = N_0$ and amplified/dampened by $a_{11}$-term.

- The outgoing characteristics at $x = 0$ ($u$) and $x = N_0$ ($u$ and $q$) combine via (5.9) and feed back into $q$ at $x = 0$.

Our goal is now to (a) characterize the dynamic stability of the given jamiton by means of the behavior of the solutions of its associated perturbation system (5.10) (including boundary conditions), and (b) use this insight to
Figure 5.1: Coefficient functions for (5.10) and a jamiton with \(v_S = 12.5\text{m/veh}\) and \(v^+ = 8.9\text{m/veh}\). This jamiton has a length of 561m and contains 40 vehicles. Note that here we revert to physical units (vehicles) for the horizontal axis.
explain and understand the computational results of the fully nonlinear ARZ model (3.2) presented in Chapter 4. To that end, we start by establishing that there are (at least) two distinct notions of (in)stability that must be considered here.

First, asymptotic stability under infinitesimal perturbations (studied in §5.4). This is captured by the $t \to \infty$ behavior of linear model (5.10): if for any i.c. $[u(x,0), q(x,0)]^T$ the solution decays exponentially as $t \to \infty$, then this notion of stability is met. Strong linear instability occurs when there is a positive feedback mechanism that produces an exponential growth of an initial perturbation in time, eventually driving the full model (3.2) out of the linear regime, no matter how small the initial (non-zero) perturbation is. At the borderline between these two behaviors, the solutions to the linear system may remain bounded for all time, or grow/decay at a sub-exponential rate.

The second notion of stability is given by the maximum transient growth criteria (studied in §5.5). Because (5.10) is non-normal, even if asymptotic stability applies, an initially small perturbation may be amplified significantly at transient times, before eventually dying off as $t \to \infty$ [96]. However, if that amplified perturbation becomes sufficiently large, nonlinear effects will take over in the full ARZ model (3.2). In this scenario, how far the system ends up from equilibrium depends both on the transient growth factor (see below) and the magnitude of the perturbations.
5.3 Fundamental challenges caused by the sonic point

In the same way as the original inhomogenous ARZ model may look misleadingly innocuous (“just a hyperbolic system with a relaxation term”), yet develops extremely complex dynamics if the SCC is violated, the jamiton perturbation system (5.8) may look innocent as well — and also that impression would be false. The fact that the characteristic speed \( b_1 \) transitions from negative to positive at \( x_S \) (a direct consequence of this being a sonic point), causes fundamental structural challenges.

It may seem rather natural to attempt to study (5.8) by expanding its solutions using eigenmodes, and seek solutions to the eigenvalue problem

\[
\begin{align*}
\lambda u &= -b_1(x)u_x + a_{11}(x)u + a_{12}(x)q , \\
\lambda q &= -b_2 q_x + a_{21}(x)u + a_{22}(x)q .
\end{align*}
\]

(5.12)

However, the right hand side operator here is non-normal; and it is well known that for non-normal operators, spectral calculations can be extremely unreliable [94, 95, 22].

Furthermore, the presence of the sonic point makes the situation substantially worse, even if one were to have access to “exact” computations. To illustrate the issue consider the simple model problem

\[ u_t + (xu)_x = \frac{3}{4} u , \quad -1 < x < 1 . \]

(5.13)
The exact solution of (5.13) is easily obtained using characteristics:

\[ u = u_0(x \, e^{-t}) e^{-\frac{t}{4}}, \quad (5.14) \]

where \( u_0 \) is the initial data. This clearly is a stable situation by any “physically reasonable” definition. On the other hand, if we look for eigenfunctions by separating \( u = \phi(x) e^{\lambda t} \), we find that: \( \phi = |x|^\alpha \), with \( \alpha = - (\lambda + \frac{1}{3}) \) and any \( \lambda \) with \( \text{Re}(\lambda) < \frac{1}{4} \), is an acceptable square-integrable eigenfunction. Even worse: every eigenvalue has infinite multiplicity (apply \( \frac{d^n}{dx^n} \) to the eigenvalue equation with the solutions above).

Thus from a naive eigenvalue calculation one would conclude that an exponential instability occurs! But here, with an exact solution, the situation is clear: the presence of a sonic point allows the existence of solutions that are not smooth. Then stability and growth/decay rates depend on the smoothness restrictions imposed. While \( L^2 \) yields instability, \( L^\infty \) or \( H^1 \) yield stability, but with different bounds on the decay rates. Thus, in a numerical computations one would have to worry about what restriction (if any) the computation enforces as the resolution increases.

Because of these issues we refrain from using the approach in (5.12), and instead characterize (in)stability via alternative ways that do not use eigenmode expansions.
5.4 Quantitative results: Asymptotic stability

The $t \to \infty$ behavior of the solutions of the jamiton perturbation system (5.10) (incl. b.c.) depends on a delicate balance of growth vs. decay effects. And because those are governed by the functions $a_{ij}(x)$, $b_i(x)$, and the b.c. constants, we do not attempt a fully analytical characterization here. Instead, we formulate a sequence of approximations to the solutions of (5.10) and analyze their behavior. Specifically, we formulate the following approximation scheme.

We discretize the spatial domain into a regular grid:

$$\{x_0, \ldots, x_m\} = \{0, h, 2h, \ldots, N_0 - h, N_0\}$$

and conduct time steps of size $\Delta t = h/b_2$, see Figure 5.2. We denote the grid approximations

$$U^n_j \approx u(jh, n\Delta t) \quad \text{and} \quad Q^n_j \approx q(jh, n\Delta t),$$

and denote the full state vector at time $n\Delta t$ by $Y^n = [U^n, Q^n]^T$, where

$$U^n = [U^n_1, \ldots, U^n_m]^T \quad \text{and} \quad Q^n = [Q^n_1, \ldots, Q^n_m]^T.$$

An update matrix for the transport part of (5.10) (incl. b.c., but neglecting the $a_{ij}$-terms) is obtained via tracking characteristics: for each grid point $x_j = jh$, determine the associated foot point $\hat{x}_j$ as the solution of the ODE

$$\dot{x}(s) = -b_1(x(s)) \quad \text{with} \quad x(0) = x_j, \quad \text{evaluated at} \quad s = \Delta t.$$

Then,

$$U^{n+1}_j = \Psi^n(\hat{x}_j),$$
Figure 5.2: Illustration of the discretization used to approximate (5.10), as described in §5.4. The top (bottom) graphic shows the characteristic curves corresponding to the $u$ ($q$) variable. The $u$-characteristics expand away from the sonic point towards the domain boundaries (where the shock is). The scheme’s time step is selected so that the $q$-characteristics advance by $h$ per time step.
where $\Psi_{U^n}(x)$ is the piecewise-linear interpolant based on the grid data $U^n$. Due to the clever choice of time step, the $q$-update can be solved exactly via

$$Q_j^{n+1} = Q_{j-1}^n \quad \text{for all } j \geq 1,$$

The b.c. are used to update $Q_0^{n+1}$ according to the following equation

$$Q_0^{n+1} = \frac{1}{k_L - h^{-1}} \left( (k_R - h^{-1}) Q_{m-1}^n + c_L \Psi_{U^n}(\tilde{x}_0) + c_R \Psi_{U^n}(\tilde{x}_m) \right).$$

We denote the update matrix summarized above by $M_1$.

A second matrix for the growth/decay part (i.e., neglecting the advection terms) is formulated as follows:

$$[U_j^{n+1}, Q_j^{n+1}]^T = \exp(\Delta t A(x_j)) \cdot [U_j^n, Q_j^n]^T,$$

where $A(x)$ is the $2 \times 2$ matrix formed by the $a_{ij}(x)$ values. We denote this resulting update matrix $M_2$.

One step of the numerical scheme, written as

$$Y^{n+1} = M \cdot Y^n,$$

is given by the update matrix $M = M_2 \cdot M_1$. This first order method is carefully designed to not incur any slow drifts. Because the scheme is linear with time-independent coefficients, the $t \to \infty$ behavior of the solutions is fully characterized by its one-step update matrix $M$, specifically by its spectral radius $\rho(M)$ (largest modulus of the eigenvalues of $M$). In particular, asymptotic
stability (of the approximation) is given exactly if

\[ \rho(M) < 1. \]  \hspace{1cm} (5.15)

Once \( M \) is set up, the stability condition (5.15) can be checked via Matlab’s numerical linear algebra routines, resulting in a systematic classification of jamitons into asymptotically stable vs. unstable.

A caveat in this approach is that for any choice of grid size \( h \), we check the asymptotic stability of an approximation to (5.10). However, because we have a convergent sequence of approximations, we approach the true answer for (5.10) as \( h \to 0 \). Moreover, for any \( h > 0 \), the approximation slightly overestimates stability due to the scheme’s numerical diffusion (which vanishes as \( h \to 0 \)), resulting in a too small but growing (as \( h \to 0 \)) unstable jamiton region.

Figure 5.3 displays the results. It shows the classification of the same jamitons as in Figure 4.1 into asymptotically stable and unstable using the asymptotic stability criterion: \( \rho(M) < 1 \) (unstable: \( \rho(M) > 1 \)), where for each jamiton, \( M \) is the one step update matrix that comes from a discretization with 8000 grid points. Comparing those results to the nonlinear system results in Figure 4.1, we indeed see that (i) only the splitting instability (long jamitons) can be captured; and (ii) the unstable region is underestimated. This last aspect is likely also affected by the fact that asymptotic stability does not account for transient growth effects; which we consider next.
(a) Classification of asymptotic stability in the phase plane \((\rho_S, L)\).

(b) Classification of asymptotic stability in the phase plane \((\bar{\rho}, s)\).

(c) Classification of asymptotic stability in the phase plane \((\rho_S, \rho^+)\).

(d) Classification of asymptotic stability in the phase plane \((\rho_S, \bar{\rho})\).

**Figure 5.3:** Classification of 980 jamitons into asymptotically stable vs. unstable, where asymptotic stability is given by \(\rho(M) < 1\) (and instability by \(\rho(M) > 1\)). Here, for each jamiton, \(M\) is the one-step update matrix that comes from a discretization with 8000 grid points. Note that the criterion used here can only detect “splitting” instabilities.
5.5 Quantitative results: Transient growth

Even if the system (5.10) is asymptotically stable, small perturbations may be amplified significantly at transient times. Via asymptotic arguments we can argue that the dominant wave amplitude growth mechanism is the growth of the $u$-field as it travels between the sonic point $x_S$ and the right domain boundary $N_0$. The argument is motivated by the Wentzel-Kramers-Brillouin (WKB) [7] expansion of the eigenvalue problem of the system (5.10). This WKB analysis is a high-frequency asymptotic expansion, where $\epsilon$ is the small asymptotic parameter. The WKB ansatz looks for asymptotically large eigenvalues of the form $\frac{\lambda}{\epsilon}$, and high-frequency solutions of the form

$$u(x) = A(x) \exp \left( \frac{\Theta(x)}{\epsilon} \right),$$

$$q(x) = B(x) \exp \left( \frac{\Theta(x)}{\epsilon} \right).$$  (5.16)

The eigenvalue problem being solved here can be written as

$$b_1(x)u_x = \left( a_{11}(x) - \frac{\lambda}{\epsilon} \right) u + a_{12}(x)q,$$

$$b_2 q_x = a_{21}(x)u + \left( a_{22}(x) - \frac{\lambda}{\epsilon} \right) q.$$  (5.17)

Substituting the functions $u(x)$ and $q(x)$ in (5.16) into the eigenvalue problem (5.17), we get the following equations

$$b_1(x)A'(x) + b_1(x) \frac{\Theta'(x)}{\epsilon} A(x) = a_{11}(x)A(x) - \frac{\lambda}{\epsilon} A(x) + a_{12}(x)B(x),$$

$$b_2 B'(x) + b_2 \frac{\Theta'(x)}{\epsilon} B(x) = a_{21}(x)A(x) + a_{22}(x)B(x) - \frac{\lambda}{\epsilon} B(x).$$  (5.18)
Moreover, we make a regular expansion ansatz for $A$ and $B$ as

$$A(x) = A_0(x) + \epsilon A_1(x) + \epsilon^2 A_2(x) + \ldots,$$

$$B(x) = B_0(x) + \epsilon B_1(x) + \epsilon^2 B_2(x) + \ldots.$$ 

If both $A_0(x) \neq 0$ and $B_0(x) \neq 0$, then the system in (5.18) is not solvable. In fact, the two solutions of (5.18) are the following:

1. $B_0(x) = 0$, and
   $$A_0(x) = C_1 \exp \left( \int \frac{a_{11}(x)}{b_1(x)} \, dx \right),$$
   $$B_1(x) = \frac{b_1(x) a_{21}(x) A_0(x)}{\lambda(b_1(x) - b_2)}.$$  
   (5.19)

2. $A_0(x) = 0$, and
   $$B_0(x) = C_2 \exp \left( \int \frac{a_{22}(x)}{b_2} \, dx \right),$$
   $$A_1(x) = \frac{b_2 a_{42}(x) B_0(x)}{\lambda(b_2 - b_1(x))}.$$  
   (5.20)

where the $C_i$’s are constants of integration. Therefore, the general solution to (5.18) is a superposition of the two solutions (5.19) and (5.20) presented above. Note that the first solution (5.19) has $u \gg q$ and $q$ is “attached” to $u$, while the second solution (5.20) has $q \gg u$ and $u$ is “attached” to $q$. Moreover, since $a_{22}(x) < 0$ everywhere, and $a_{11}(x) > 0$ when $x > x_S$, the first solution (5.19) will dominate over the second solution (5.20) to the right of the sonic point.

Similarly, we can consider the high frequency solutions of (5.10), which are observed to have the strongest growth. Here we are looking at solutions
that are rapidly varying in space and time. In this situation the behavior is
dominated by the left hand side of (5.10), and we can see that such solutions
generally consist of a superposition of two waves: the “u-wave”, dominated
by the excitation in u, and the “q-wave”, dominated by the excitation in q.
Consider first the u-wave. Then, because \( u \gg q \) (for a similar reason as the
one shown in the WKB expansion of (5.17)), we can simplify the equations to
obtain

\[
\begin{align*}
    u_t + b_1(x)u_x & \approx a_{11}(x)u, \\
    q_t + b_2 q_x & \approx a_{21}(x)u.
\end{align*}
\]

From this we can see that q is “slaved” to u (since the homogeneous part of
the solution to the second equation should be considered as belonging to the
q-wave). A similar argument applies to the q-wave; however, the u-wave will
dominate because \( a_{11} > 0 \) to the right of \( x_S \), while \( a_{22} < 0 \).

Hence, neglecting the q-wave (and its influence on u) we obtain that u
evolves (approximately) according to the characteristic equations

\[
\frac{dx}{dt} = b_1(x), \quad \text{and} \quad \frac{du}{dt} = a_{11}(x)u. \tag{5.21}
\]

The speed \( b_1(x) \) vanishes at \( x_S \), but so does the growth rate \( a_{11} \), resulting in
an overall finite net growth. By the chain rule, the characteristic equations
lead to the ODE

\[
\frac{du}{dx} = \frac{a_{11}(x)}{b_1(x)} u, \tag{5.22}
\]
with normalized i.c. \( u(x_S) = 1 \), to estimate the transient amplification factor \( F \). Solving the ODE yields

\[
F = \exp \left( \int_{x_S}^{N_0} \frac{a_{11}(x)}{b_1(x)} \, dx \right). \tag{5.23}
\]

This quantity can be computed via quadrature, using L'Hôpital's rule at/near \( x_S \). However, note that the arguments above do not apply across the sonic point, even though the integrand is not singular, because the parameterization of the characteristics by \( x \) (i.e. \( \frac{dt}{dx} = \frac{1}{b_1(x)} \)) implicit in the calculation above breaks down there.

An important fact is that the quantity \( F \) can be computed without solving the jamiton ODE. This is achieved by parameterizing the jamiton in terms of \( v_S \) and the left shock state \( v_{N_0} = v(N_0) \). Then, because \( a_{11} \) and \( b_1 \) are functions of \( x \) only via the jamiton \( v(x) \), one can apply a change of variables to replace \( x \)-integration by \( v \)-integration. The Jacobian for the transformation follows from the jamiton ODE \( (3.19) \). This yields the formula

\[
F = \exp \left( \int_{v_S}^{v_{N_0}} \frac{m_0 h''(v)}{h'(v)(h'(v) + m_0)} - \frac{m_0(h'(v) + U'(v))}{h'(v)(U(v) - m_0v - s_0)} \, dv \right).
\]

Figure 5.4 shows the stability classification via this criterion for the same jamitons studied in Figure 4.1. As in Figure 5.3, we do not capture merging instabilities. For the splitting instability, we consider two thresholds for the amplification factor: \( F_1 = 10^5 \) and \( F_2 = 10^{15} \). Classifying jamitons below the \( 10^5 \) amplification factor as stable is consistent with the magnitude of noise
(a) Classification of jamitons according to $F$ in the phase plane $(\rho_S, L)$.

(b) Classification of jamitons according to $F$ in the phase plane $(\bar{\rho}, s)$.

(c) Classification of jamitons according to $F$ in the phase plane $(\rho_S, \rho^+)$.

(d) Classification of jamitons according to $F$ in the phase plane $(\rho_S, \bar{\rho})$.

**Figure 5.4:** Classification of 980 jamitons according to the transient growth factor (5.23). Three levels of $F$ are displayed, with the thresholds at $F_1 = 10^5$ and $F_2 = 10^{15}$ to yield: stable if $F < F_1$, moderately unstable if $F_1 < F < F_2$, and unstable if $F_2 < F$. 
in the nonlinear computation (§4.3), which was roughly $10^{-5}$. The results show that the stability boundary in Figure 4.1 is not reproduced perfectly, but reasonably well. An interesting advantage of this measure of “instability” is that it is not just a yes/no criterion, but rather provides a measure of the “badness” of the instability. One key missing piece in this criterion is that it does not characterize the “pumping” mechanism of perturbations from $q$ into $u$ at/near the sonic point. Hence, we do not know how large the perturbation magnitude really is near $x_S$. 
CHAPTER 6

AVERAGING OF THE ARZ MODEL

As discussed earlier in this work, the study of traffic modeling is not restricted to macroscopic models. Different types of models have different advantages, and picking one type over the other is often motivated by the scale of the project. For example, microscopic models can resolve interactions on the vehicular scale, while the main advantage to macroscopic models is their ability to study traffic on a large scale.

Second order macroscopic models (e.g. ARZ model (3.1)) have the advantage of capturing jamitons, which have been observed in real life. The scale of stable jamitons (see Chapter 4) is not quite as small as the vehicle scale, but still smaller than urban scales. Going forward, we will make a distinction
between three different scales of traffic:

- **Vehicle scale**: 10–75 m: Heterogeneities and specific braking patterns matter, and collisions might happen. Microscopic models are best suited to describe this scale.

- **Jamiton scale**: 150–800 m: Systematic wave component of the flow are of interest. Second-order models (ARZ) can capture this scale.

- **Urban scale**: 10–100 km: Large scale flow patterns are of concern.

For urban scale projects, first order LWR models are usually used for their simplicity and computational efficiency: these models are “stable”, which means that instabilities never grow, and it is enough to resolve simulations with large grid cells ($\Delta x = 25–100\text{m}$). Such coarse resolutions cannot be used, for example, in simulating the second-order ARZ model. For such a model to be meaningful, it has to be able to resolve the jamiton scale, which requires grid cells with $\Delta x < 4\text{m}$. With such fine resolutions, large scale projects could easily end up with more than 20,000 grid cells, and 50,000 time steps, which consumes vast computational resources.

Therefore, one goal is to find models for the urban scale that are simpler and more computationally efficient than the ARZ model, but nevertheless capture the cumulative effect of instabilities and waves (jamitons).
6.1 A model for the averaged density

A topic of interest when discussing second order models is the average of the traffic density over a stretch of the road. Define the quantity

$$\bar{\rho}(x_0, t) = \int_{x_0-\delta}^{x_0+\delta} w(x) \rho(x, t) \, dx$$

(6.1)
to be the average traffic density at position $x_0$ and time $t$. The function $w(x)$ is a weight function, and $2\delta$ is the window where the averaging happens. The scale of $\delta$ here is chosen to be in between the jamiton scale and an urban scale. The above quantity is of particular interest because it can describe traffic on a scale that is even larger than the traffic instabilities in macroscopic models, but still be useful when quantifying effective traffic features on an urban scale.

Given that $\rho(x, t)$ is the solution of the ARZ model (3.1), what is $\bar{\rho}(x, t)$ a solution of? More precisely, is there a set of equations that $\bar{\rho}(x, t)$ satisfy?

The idea of deriving a model for $\bar{\rho}$, or “averaging the ARZ model” is inspired by Reynolds’s averaging [83, 85, 86] of the Navier-Stokes equations in fluid dynamics. To carry this out rigorously, one could start by defining the quantities of interest as

$$\rho = \bar{\rho} + \rho' \quad \text{and} \quad u = \bar{u} + u' ,$$

where the bar-operator is defined as

$$\bar{Z}(x, t) = \int_{x-\delta}^{x+\delta} w(x) Z(x, t) \, dx ,$$
and having

\[
\overline{\rho} = 0 \quad \text{and} \quad \overline{u} = 0 ,
\]

and then substituting these quantities into the ARZ model and averaging the equations themselves. The averaging discussed above happens over space, but an average over time, or both time and space could also be done. For an illustration of this process, the averaging of a first order LWR model has been done in [87]. This process has proved to be especially challenging for the second-order ARZ model. Other ideas to derive a model for \( \overline{\rho} \) can also be found in [9, 103, 65, 66].

In this work, the focus will be on deriving a model for \( \rho \) in an ad-hoc fashion, following the insights gained in Chapters 4 and 5, as well as the results from many simulations of the ARZ model.

The first main insight comes from answering the research problem of which jamitons are dynamically stable. Any solution to the ARZ model that contains unstable jamitons will not persist for a long time: short jamitons will “merge” and long jamitons will “split” (see §4.3). This means that jamiton solutions that sustain themselves for a while have to be stable jamitons. Figure 4.2 gives valuable information for the relationship between stable jamitons and their average quantities \( \overline{\rho} \) and \( \overline{Q} \): for a given \( \overline{\rho} \), the stable jamitons are the ones near the lowest possible \( \overline{Q} \). In other words, stable jamitons cluster near the lower jamiton envelope when plotted in the \((\overline{\rho}, \overline{Q})\) phase plane.
The second main insight comes from observing which jamitons arise from uniform flow in simulations. To do that, we run simulations of the ARZ model (3.2) using the HLL scheme described in §4.2 starting with the initial conditions
\[ \rho(x,0) = \rho_0 \quad \text{and} \quad u(x,0) = u_0 = U(\rho_0) . \]

Similar to what is described in §4.3, noise is added to the model with every time step to induce jamitons and “kick” any jamitons that arise. The output of this simulation would be matrices approximating \( \rho(x,t) \) and \( u(x,t) \).

To get the average quantities, \( \bar{\rho}(x,t) \) is first calculated using equation (6.1), with \( \delta = 2 \text{ km} \). In addition \( w(x) \) is chosen to be a truncated Gaussian with mean \( x \), standard deviation \( \delta/3 \), and defined over \( [x - \delta, x + \delta] \). \( w(x) \) is also normalized to have \( \int_{x-\delta}^{x+\delta} w(\tilde{x})d\tilde{x} = 1 \). Second, using the fact that \( Q(x,t) = \rho(x,t)u(x,t) \), and the same averaging operator, \( \bar{Q}(x,t) \) is calculated.

The simulations are run on a periodic domain of size 8 km until reaching the final time \( T_f = 1200 \text{ seconds} \). A scatter plot of \( (\bar{\rho}(x,0),\bar{Q}(x,0)) \) and \( (\bar{\rho}(x,T_f),\bar{Q}(x,T_f)) \) are then plotted. If \( \rho_0 \) violates the SCC, jamitons will arise and the average flow rate drops.

Figure 6.1 shows an example of one of these simulations, where the initial conditions with \( \rho_0 = 59 \text{ veh/km} \) are shown in the top left panel. This initial density violates the SCC, so the perturbations grow into jamitons, and the ones that survive at \( t = T_f \) are plotted in the bottom left panel. The dashed
red line in the same panel plots \( \bar{\rho}(x, T_f) \). The right panel plots the averaged quantities in the fundamental diagram. The red dots correspond to the points \((\bar{\rho}(x, 0), \bar{Q}(x, 0))\). They all coincide and fall exactly on the equilibrium curve because of how the initial conditions are defined. The blue dots correspond to the points \((\bar{\rho}(x, T_f), \bar{Q}(x, T_f))\). These points mostly coincide, and have a similar average density to \( \rho_0 \), but have an average flow rate that is much lower, placing them near the lower jamiton envelope.

**Figure 6.1:** A plot of an example simulation. The initial conditions are plotted in the top left panel. \( \rho(x) \) and \( \bar{\rho}(x) \) are plotted in the bottom left panel at the final time \( t = 1200 \) sec. A scatter plot of the average quantities at \( t = 0 \) and \( t = 1200 \) sec is plotted on top of a FD in the right panel. This observation holds for all initial uniform flows that violate the SCC.
Starting with uniform flow, where the average quantities lie on the equilibrium curve, evolving the initial conditions according to the ARZ model results in jamitons where the average quantities cluster around the lower envelope. Figure 6.2 summarizes five different simulations. The red dots refer to the averaged quantities at $t = 0$ and the blue dots refer to the same quantities at $t = T_f$. The arrows connect the red and blue points corresponding to the same simulation.

![Figure 6.2](image)

**Figure 6.2:** A scatter plot of the average quantities at $t = 0$ and $t = 1200$ sec is plotted on top of a FD, summarizing five different simulations.

Based on the observations and insights discussed above, the simplest model
that can be proposed for \( \bar{\rho} \) is a first order LWR model, with a special flux function (fundamental diagram) \( \bar{Q} \), which for a given \( \bar{\rho} \) is defined to be the smaller of \( \bar{\rho}U(\bar{\rho}) \) and the corresponding flow rate on the lower jamiton envelope. This first order LWR model can be written as

\[
\bar{\rho}_t + \left( \bar{Q}(\bar{\rho}) \right)_x = 0.
\]

(6.2)

We will refer to the model in (6.2) as aARZ-1 from now on.

To test the accuracy of aARZ-1, we compare it to the exact solution given by directly averaging the solution of the ARZ model, as well as the solution of the LWR model with the function \( \rho U(\rho) \) as the flux function. We choose, as an example, a single lane periodic road of length 20 km. The initial conditions are chosen such that \( \rho_0(x) \) contains a uniform flow segment, a discontinuity, and a segment with linearly varying density. In addition, for the ARZ model, we choose \( u_0(x) = U(\rho_0(x)) \). These initial conditions are shown in the top panel of Figure 6.3.

To get the exact solution for \( \bar{\rho} \), we apply the same averaging operator discussed above, and shown in equation (6.1), to the ARZ solution \( \rho(x,t) \). We then compare that to the solution of the aARZ-1 model, \( \bar{\rho}_{aARZ1}(x,t) \), and the solution of the LWR model \( \rho_{LWR}(x,t) \).

Figure 6.3 below is an illustration of this comparison. It shows the ARZ solution of this simulation at \( t = 0 \) (initial condition), \( t = 1000 \) sec, and \( t = 2300 \) sec. The last two snapshots show a clear jamiton solution pattern.
Figure 6.3: A simulation illustrating the comparison between the ARZ model (dark blue line), an LWR model (dotted brown line), and the aARZ-1 model (dotted red line). Plots are made at $t = 0$, $t = 1000$ sec, and $t = 2300$ sec.
Below each of these snapshots of the ARZ solution, the “average” quantities are plotted. First, $\bar{\rho}$, the averaged density of the ARZ model using the averaging operator (6.1) is displayed, then $\bar{\rho}_{aARZ}(x,t)$, the solution of the aARZ model with the same initial conditions, and lastly, $\rho_{LWR}(x,t)$, the solution of the LWR model (same initial conditions as well).

The results in Figure 6.3 make it clear that an LWR model with the flux function corresponding to the equilibrium curve $\rho U(\rho)$ does a very poor job in approximating $\bar{\rho}$. However, the aARZ model does a significantly better job in approximating $\bar{\rho}$, especially in the regions where a sequence of varying jamitons exist. Note that the solution to aARZ may contain shocks, whereas $\bar{\rho}$ never does because of the nature of the smooth averaging that is applied to $\rho$.

The goal in finding a model that approximates $\bar{\rho}$ is to avoid having to solve the ARZ model, and instead rely on $\bar{\rho}$ to extract the cumulative effects of jamitons. Below, we will discuss how to quantify, given $\bar{\rho}$, the total fuel consumption, the total flow of vehicles, and braking time of the average vehicle. But first, we will derive a fuel consumption model from the literature.

### 6.2 A fuel consumption model

Our aim is to quantify the overall fuel consumption (FC) of all vehicles on the highway in various situations. The relationship between the speed of a vehicle and its FC efficiency (Liters/km) is discussed in [1, 8]. Vehicles consume
Figure 6.4: Plots of the FC rate of the four vehicles considered in [8] vs. the vehicle velocity, and the fitted average polynomial $K(s)$ (solid red curve).

more fuel per distance traveled when they are driving at very low speeds (a certain amount of fuel is used to just keep the engine, and accessories, running), or at very high speeds (more energy is needed to overcome air drag).

Figure 41 in [8] provides fuel consumption efficiency data (Liters/km) for four types of vehicles (Ford Explorer, Ford Focus, Honda Civic, and Honda Accord), as functions of the vehicle speed. Multiplying the FC efficiency by the vehicle speed yields the FC rate (Liters/hr). The FC rates for the four vehicles as functions of the vehicle speed are shown in Figure 6.4. We average these
four curves with equal weights, and approximate the resulting data points via a sixth order polynomial using least squares (which is accurate up to a 5% error). This average FC rate function reads as:

\[
K(s) = (5.7 \times 10^{-12})s^6 - (3.6 \times 10^{-9})s^5 + (7.6 \times 10^{-7})s^4 \\
- (6.1 \times 10^{-5})s^3 + (1.9 \times 10^{-3})s^2 + (1.6 \times 10^{-2})s + 0.99 .
\]  

Units in the above equation have been omitted for notational efficiency (s is in km/hr and K is in Liters/hr). The function is given by the red curve in Figure 6.4.

To carry out calculations that take acceleration into account, we wish to find FC rates at different acceleration levels, similar to the zero acceleration FC rate function \(K(s)\) in Figure 6.4. Figure 1 in [81] shows the FC rate of vehicles at three different acceleration levels — \(a = -0.9 \text{ m/s}^2\), \(a = 0 \text{ m/s}^2\), and \(a = 0.9 \text{ m/s}^2\) — and for vehicle speeds ranging from 0 to 120 km/hr.

Looking at the data, there are three obvious limitations. The first limitation is the discrepancy between the zero acceleration FC rate in [81] and the function \(K(s)\). This is mostly because different vehicles were used to calculate each function. The data in [81] are the average of eight early and mid 1990s light duty vehicles and trucks, and the data in [8] used to calculate \(K(s)\) are for three mid-2000s light duty vehicles and a 2004 SUV. The second limitation of the data in [81] is that they do not go as far as \(K(s)\) in terms of vehicle speeds. This is not a big limitation for the models we use, but we would like
to present a setup that works for more generic models. The third limitation is that the data only contain three different acceleration levels.

To remedy the first limitation, we neglect the actual FC rates reported at [81], and we instead calculate the ratios between the instantaneous FC rate at $a = 0.9 \text{ m/s}^2$ and the same rate at $a = 0 \text{ m/s}^2$ at different speeds. Similarly, we calculate ratios between the FC rate at $a = 0 \text{ m/s}^2$ and $a = -0.9 \text{ m/s}^2$. Then, we use $K(s)$ to construct FC rate functions at $a = 0.9 \text{ m/s}^2$ and $a = -0.9 \text{ m/s}^2$, while preserving the already calculated ratios.

To remedy the second limitation, we extrapolate the newly constructed FC rate functions using a third degree polynomial. This preserves the convexity of these functions at high speeds. And finally, to remedy the third limitation, we use interpolation to calculate the FC rate at acceleration levels different from $a = 0.9 \text{ m/s}^2$ and $a = -0.9 \text{ m/s}^2$.

Figure 6.5 shows the constructed FC rate functions discussed above, along with $K(s)$ and an interpolated FC rate function at an acceleration level $a = 0.5 \text{ m/s}^2$. This figure is essentially a plot of some level curves of a function $F_C(s, a)$, which for a given vehicle speed $s$, and vehicle acceleration $a$, approximates the FC rate of that vehicle.

Note that the fuel consumption model derived here is coming from vehicles that are now considered quite old. This means that current values of fuel consumption by vehicles might differ. However, the calculations done here
can be carried out using any fuel consumption model of choice. Note that some of the features in the fuel consumption model are physical in nature. For example, we expect all fuel consumption rates (Liters/hr) to be convex when plotted versus vehicle speed, because of air drag for high speeds and basic vehicle needs for low speeds. Another note regarding the fuel consumption model here is that it does not capture any energy regenerative braking that some electric vehicles are nowadays able to do.

**Figure 6.5:** A plot of four level curves of the function $F_c(s,a)$. The curves represent the FC rate functions for four different acceleration levels versus speed.
The goal here is to quantify the total FC \( TFC \) (in Liters) by all vehicles on a stretch of the highway/road \([0, Z]\), during the time interval \([0, T]\). This can be calculated using the formula

\[
TFC = \int_0^T \int_0^Z \rho(x,t) F_c(u(x,t), a(x,t)) \, dx \, dt .
\] (6.4)

Here \( a(x,t) \) is the acceleration of vehicles at position \( x \) and time \( t \). A formula for \( a \) was derived in §2.3.1:

\[
a = u_t + uu_x .
\] (6.5)

For an LWR model where \( u = U(\rho) \), and \( \rho_t = -Q'(\rho)\rho_x \), the acceleration \( a \) can be reformulated as

\[
a = -\rho \rho_x (U'(\rho))^2 .
\] (6.6)

Therefore, the total FC for an LWR model can be calculated via

\[
TFC_{\text{LWR}} = \int_0^T \int_0^Z \rho F_c\left(U(\rho), -\rho \rho_x (U'(\rho))^2\right) \, dx \, dt ,
\] (6.7)

and for the ARZ model, where \( u(x,t) \) is part of the solution output, the total FC can be calculated using the formula

\[
TFC_{\text{ARZ}} = \int_0^T \int_0^Z \rho F_c\left(u, u_t + uu_x\right) \, dx \, dt .
\] (6.8)

Since it is plausible that the solutions to both the LWR model and the ARZ model possess shocks/discontinuities — which manifests as sharp jumps in the data — this makes functions \( u_t, u_x \) and \( \rho_x \) undefined/unrealistic at some points. To avoid this problem, smooth the data via a Gaussian Kernel before
calculating any of the derivatives or applying any FC formulas. This technique has the effect of making shocks vary smoothly over a 40–50m interval, instead of having the shock appear as a sharp jump in a 5–10m interval in the original data.

6.3 Methods for quantifying ARZ cumulative effects using aARZ-1

In this section, the aim is to demonstrate that the aARZ-1 model can be used to capture the non-linear effects (jamitons) of the ARZ model without having to solve (3.1) explicitly. The focus here will be three different metrics: fuel consumption, total flow, and braking time.

6.3.1 Fuel consumption

Even though the structure of the aARZ-1 model is that of a first order LWR model, it would be naive to use the formula in (6.7) to estimate the total FC of all vehicles in a simulation. Because we understand that aARZ-1 solves for $\bar{\rho}$ instead of $\rho$, extra information can be extracted. In particular, whenever the solution of the averaged model contains a point $\bar{\rho}(\tilde{x}, \tilde{t})$ such that $\bar{Q}(\bar{\rho}(\tilde{x}, \tilde{t})) < \bar{\rho}(\tilde{x}, \tilde{t})U(\bar{\rho}(\tilde{x}, \tilde{t}))$ (i.e. that point lies on the lower jamiton envelope portion of the fundamental diagram of aARZ-1), we can infer that this $\bar{\rho}$ is
the average of a jamiton profile.

As discussed in §3.3, for any point \((\bar{\rho}_0, \bar{Q}_0)\) that lies on the lower jamiton envelope, there is only one jamiton with an average density of \(\bar{\rho}_0\) and an average flow of \(\bar{Q}_0\). For that specific jamiton, we can calculate the FC rate (Liter/hr) after applying the smoothing kernel. For a jamiton \(J_0 = [\rho_0(x), u_0(x)]\) with traveling speed \(s_0\) we have \(u_t = -s_0u_x\). The formula to get the FC rate of \(J_0\) per vehicle is

\[
FC_{J_0} = \frac{1}{N_0} \int_0^{L_0} \rho_0 F_c \left( u_0, u_0u_{0x} - s_0u_{0x} \right) dx.
\]  

(6.9)

\(L_0\) is the length of the jamiton, and \(N_0\) is the number of vehicles in the jamiton. The units of \(FC_{J_0}\) are in Liters/hr/veh.

Given \(\bar{\rho}(x, t)\) as the solution of the aARZ-1 model, we propose to calculate the total FC according to the formula

\[
TFC_{aARZ1} = \int \int_{\{x(t)|\bar{Q}(\bar{\rho}) = \bar{\rho}U(\bar{\rho})\}} \bar{\rho} F_c \left( U(\bar{\rho}), -\bar{\rho}\bar{p}_x(U'(\bar{\rho}))^2 \right) dx dt
\]

\[
+ \int \int_{\{x(t)|\bar{Q}(\bar{\rho}) < \bar{\rho}U(\bar{\rho})\}} \bar{\rho} FC_{J\bar{\rho}} dx dt.
\]  

(6.10)

This formula means that if \(\bar{\rho}\) is not on the lower jamiton envelope, we calculate the \(TFC\) using the same protocol followed in LWR models. Otherwise, we approximate the \(TFC\) using \(FC_{J\bar{\rho}}\), the average FC rate per vehicle in the jamiton with average density \(\bar{\rho}\) and average flow rate \(\bar{Q}(\bar{\rho})\).

One of the challenges here is finding \(J_{\bar{\rho}}\) given \(\bar{\rho}\). The procedure outlined in §3.3 constructs the jamiton profiles \(\rho(x)\) and \(u(x)\) starting from a given \(\rho_S\) and
\(\rho^+\). In this case, we are given neither of these quantities. However, knowing that \((\bar{\rho}, \bar{Q}(\bar{\rho}))\) lies on the lower jamiton envelope, we can use equation (3.29) to find \(\rho_s\). And knowing that for a given \(\rho_s\), \(\bar{\rho}\) is a decreasing function of \(\rho^+\), it is straightforward to search for the jamiton \(J_{\bar{\rho}}\) by iterating over the jamiton construction procedure in §3.3, until we land on the jamiton with average density \(\bar{\rho}\), with any desired tolerance. In this work, we use a tolerance \(\epsilon = 0.01\) veh/km, i.e. we accept a jamiton if its average density is in the interval \([\bar{\rho} - \epsilon, \bar{\rho} + \epsilon]\).

We will again make a comparison between the LWR model with the flux function \(Q(\rho) = \rho U(\rho)\), the full ARZ model (3.1), and the \texttt{aARZ-1} model (6.2). This time, the comparison will focus on the TFC on a ring road (periodic domain) of length 8 km. The starting initial conditions are constant uniform flow with \(\rho(x, 0) = \rho_0 = 59\) km/hr, and in the case of the ARZ model, \(u(x, 0) = U(\rho_0)\). We run the simulations until \(t = T_f = 1200\) sec using the three different models. We then calculate the TFC using the formulas and techniques discussed above in equations (6.7), (6.8) and (6.10).

Table 6.1 below summarizes the results. The total fuel consumed in the ARZ simulation is 477.6 Liters. In the commonly used LWR model, that quantity is 401.1 Liters, and in the \texttt{aARZ-1} model it is 466.7 Liters. If we assume that true traffic dynamics behave similarly to the ARZ model, then using a simple LWR model produces a 16\% underestimate in the TFC, while

\[\text{\texttt{aARZ-1}}\]
Table 6.1: A summary of total fuel consumption (Liters) resulting from the ARZ, LWR and aARZ-1 models. The three simulations were run until the final time $T_f = 1200$ sec, and are for a ring road of length 8 km. All three simulations share the same initial conditions.

<table>
<thead>
<tr>
<th>Model</th>
<th>Role</th>
<th>Total Fuel Consumption ($TFC$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARZ</td>
<td>Ground truth</td>
<td>477.6 Liters</td>
</tr>
<tr>
<td>LWR</td>
<td>Commonly used</td>
<td>401.1 Liters</td>
</tr>
<tr>
<td>aARZ-1</td>
<td>Proposed model</td>
<td>466.7 Liters</td>
</tr>
</tbody>
</table>

the aARZ-1 model does a better job in estimating the $TFC$, limiting the mismatch to a 2.2% underestimate.

In this section, we considered the ARZ model to be the ground truth. Even though it captures the phenomenon of phantom traffic jams, it is important to note that there is still a lot of work to be done in the calibration of this model to real data.

6.3.2 Total flow

Another comparison that can be made is related to the traffic flow itself. Following the same example discussed in §6.3.1, we can quantify the total flow of vehicles at any fixed position. Since this example uses a ring road, the
number of vehicles in the domain is fixed at all times. However, the flow rates might differ. To quantify differences between the three models with regard to the total flow, we pick the fixed location $x = 0$, and calculate the total number of vehicles that pass at $x = 0$ between $t = 0$ and $t = T_f$. We call this quantity $TF_{x=0}$. Higher values of $TF_{x=0}$, given a sufficiently large $T_f$, correspond to a situation where vehicles have travelled further.

For the LWR model, this quantity can be calculated using the formula

$$TF_{x=0} = \int_0^{T_f} \rho(0,t) U\left(\rho(0,t)\right) \, dt.$$  \hspace{1cm} (6.11)

Similarly, for the aARZ-1, total flow at $x = 0$ can be calculated via

$$TF_{x=0} = \int_0^{T_f} \tilde{Q}\left(\tilde{\rho}(0,t)\right) \, dt.$$  \hspace{1cm} (6.12)

And finally, for the full ARZ model, the same quantity has the formula

$$TF_{x=0} = \int_0^{T_f} \rho(0,t) u(0,t) \, dt.$$  \hspace{1cm} (6.13)

The units associated with $TF_{x=0}$ is vehicles. Table 6.2 summarizes the results for the same simulation in the example presented in §6.3.1. In the ARZ model, 762.7 vehicles pass $x = 0$ in the 1200 seconds of the simulation. That number is 861.8 vehicles for the LWR model (13% overestimate from the ARZ model), and 749.4 vehicles for the aARZ-1 model (1.7% underestimate from the ARZ model). Again, it is evident that the simple LWR model, which is commonly used, does poorly in approximating the ARZ model, while the
aARZ-1 model performs much better. The results here suggest that vehicles in the LWR model travelled further than in the ARZ model. Note that for the same simulations, the total fuel consumed in the LWR model was 16% lower than the ARZ model. The fact that the LWR model has vehicles travelling longer distances suggests that the disparity in the fuel economy is even greater than 16%.

<table>
<thead>
<tr>
<th>Model</th>
<th>Role</th>
<th>Total Flow $TF_{x=0}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARZ</td>
<td>Ground truth</td>
<td>762.7 vehicles</td>
</tr>
<tr>
<td>LWR</td>
<td>Commonly used</td>
<td>861.8 vehicles</td>
</tr>
<tr>
<td>aARZ-1</td>
<td>Proposed model</td>
<td>749.4 vehicles</td>
</tr>
</tbody>
</table>

Table 6.2: A summary of total flow (vehicles) at $x = 0$ resulting from the ARZ, LWR and aARZ-1 models. The three simulations were run until the final time $T_f = 1200$ sec, and are for a ring road of length 8 km. All three simulations share the same initial conditions.

6.3.3 Braking

To quantify the “amount” of braking that happens during a given simulation, the quantity that will be used is braking time per vehicle, per hour. This is a quantity of interest because it could be used as a proxy for “safety” in
Here, a vehicle is assumed to be braking if its acceleration is below $-0.6 \text{ m/s}^2$. We call this quantity $BTR$, and calculate it as follows:

$$BTR = \frac{1}{N \times T_f} \int_0^{T_f} \int_0^Z \rho(x,t) \mathbb{H}(x,t) \, dx \, dt,$$

(6.14)

where

$$\mathbb{H}(x,t) = \begin{cases} 
1 & \text{if } a(x,t) < -0.6 \text{ m/s}^2 \\
0 & \text{if } a(x,t) \geq -0.6 \text{ m/s}^2
\end{cases}.$$  

Here $a(x,t)$ is vehicle acceleration at position $x$ and time $t$, defined for different models by (6.5) and (6.6). $N$ is the number of vehicles in the simulation, which remains constant on a ring road. $T_f$ is the final time and $Z$ is the length of the domain.

For the LWR and ARZ models, the $BTR$ quantity can be calculated directly. But for the aARZ-1 model, we take the same approach taken in §6.3.1: Whenever $\bar{\rho}$ is not on the equilibrium curve (defined by $\rho U(\rho)$) part of $\bar{Q}(\bar{\rho})$, we search for the jamitont with average density $\bar{\rho}$ and average flow rate $\bar{Q}(\bar{\rho})$, and then extract the desired quantities from that jamitont. Again, smoothing (smearing shocks) is applied before calculating any $BTR$ values.

The units we will use for $BTR$ is min/veh/hr. Table 6.3 below summarizes the results. Since the LWR model is stable with respect to any uniform base state, the solution to it starting with $\rho(x,0) = \rho_0$, is $\rho(x,t) = \rho_0$. Therefore, the LWR model assumes that the acceleration of all vehicles is zero at all times, which means that no vehicle is considered to be braking at any time. This
is of course not the case for the ARZ model, where jamitons can arise from uniform base states. Here, the average vehicle brakes for 8.85 minutes every hour. Using the aARZ-1 model, we get that the average vehicle is braking for 6.16 minutes every hour, which is a 30% underestimate when compared to the ARZ model. Despite that, this is still a much better estimate than the simple LWR model, which cannot capture most braking events.

<table>
<thead>
<tr>
<th>Model</th>
<th>Role</th>
<th>Braking Time/vehicle/hr (BTR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARZ</td>
<td>Ground truth</td>
<td>8.85 min/veh/hr</td>
</tr>
<tr>
<td>LWR</td>
<td>Commonly used</td>
<td>0 min/veh/hr</td>
</tr>
<tr>
<td>aARZ-1</td>
<td>Proposed model</td>
<td>6.16 min/veh/hr</td>
</tr>
</tbody>
</table>

Table 6.3: A summary of braking time per vehicle per hour (min/veh/hr) resulting from the ARZ, LWR and aARZ-1 models. The three simulations were run until the final time $T_f = 1200$ sec, and are for a ring road of length 8 km. All three simulations share the same initial conditions.

6.4 A more realistic example

To verify further the ability of the aARZ-1 model, we will present next comparisons similar to the ones displayed in §6.3, but for a more elaborate example. The setup of this example was discussed in §6.1, and snapshots of
the initial conditions, and solutions of the LWR, ARZ and aARZ-1 models at later times were plotted in Figure 6.3. This is an example of a ring road of length 20 km, where the initial density is set to be piecewise linear. The simulations were run until the final time was reached at 2400 seconds (40 minutes).

Then, the techniques discussed in §6.3 were used to calculate the total fuel consumption, the total flow at $x = 2$ km, and the braking time of the average vehicle per hour for all three models. The results are summarized in Table 6.4 below. This table contains an extra row for a naive version of the aARZ-1 model. This is only different from the aARZ-1 row in terms of how the traffic effects are quantified. In fact, the results of the naive version of the aARZ-1 model come from the exact same simulation of the aARZ-1 model, but where the $TFC$ and the $BTR$ are treated in a naive way. In particular, the only two differences between the aARZ-1 model and the naive version of the aARZ-1 model are:

- Equation (6.10) is used to calculate the $TFC$ of the aARZ-1 model, while equation (6.7) is used to calculate the $TFC$ of the naive version of the aARZ-1 model.

- The underlying jamitons are taken into account when calculating the $BTR$ of the aARZ-1 model, while equation (6.14) is used directly when calculating the $BTR$ of the naive version of the aARZ-1 model.
Table 6.4: A summary of the total fuel consumption, total vehicle flow at \( x = 2 \) km, and the braking time per hour of the average vehicle, resulting from the ARZ, LWR and aARZ-1 models. The three simulations were run until the final time \( T_f = 2400 \) sec, and are for a ring road of length 20 km. All three simulations share the same initial conditions. The naive aARZ-1 model is the same as the aARZ-1 model, but where it is treated as a simple LWR model when calculating \( TFC \) and \( BTR \).

Note that the total flow of vehicles is exactly the same in the two aARZ-1 rows. To summarize, this naive version of the aARZ-1 model captures the effective dynamics of the ARZ model, but not the cumulative effects it has in terms of fuel consumption and braking.

The results displayed in Table 6.4 are in line with the results of the simpler example in §6.3: The aARZ-1 model, when treated carefully, is able to capture cumulative jamiton effects while being a simple first order model. Table 6.5

<table>
<thead>
<tr>
<th>Model</th>
<th>Role</th>
<th>( TFC )</th>
<th>( T_{F_{x=2}} )</th>
<th>( BTR )</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARZ</td>
<td>Ground truth</td>
<td>2312 L</td>
<td>1491 veh</td>
<td>5.66 min/veh/hr</td>
</tr>
<tr>
<td>LWR</td>
<td>Commonly used</td>
<td>2026 L</td>
<td>1602 veh</td>
<td>0.25 min/veh/hr</td>
</tr>
<tr>
<td>aARZ-1</td>
<td>Proposed model</td>
<td>2284 L</td>
<td>1477 veh</td>
<td>4.57 min/veh/hr</td>
</tr>
<tr>
<td>naive aARZ-1</td>
<td>Naive version</td>
<td>1934 L</td>
<td>1477 veh</td>
<td>0.11 min/veh/hr</td>
</tr>
</tbody>
</table>
below summarizes the relative difference between the ARZ model (considered to be the ground truth here) and the LWR model, the aARZ-1 model, and the naive version of the aARZ-1 model.

<table>
<thead>
<tr>
<th>Model</th>
<th>% difference in $TFC$ compared to ARZ model</th>
<th>% difference in $TF_{x=2}$ compared to ARZ model</th>
<th>% difference in $BTR$ compared to ARZ model</th>
</tr>
</thead>
<tbody>
<tr>
<td>LWR</td>
<td>−13%</td>
<td>+7.4%</td>
<td>−96%</td>
</tr>
<tr>
<td>aARZ-1</td>
<td>−1.2%</td>
<td>−0.9%</td>
<td>−19%</td>
</tr>
<tr>
<td>naive aARZ-1</td>
<td>−16%</td>
<td>−0.9%</td>
<td>−98%</td>
</tr>
</tbody>
</table>

**Table 6.5:** A summary of the relative difference in fuel consumption, total vehicle flow at $x = 2$ km, and the braking time per hour of the average vehicle, between the ARZ model (considered to be the ground truth here) and the LWR model, the aARZ-1 model, and the naive version of the aARZ-1 model. These relative differences are for the results displayed in Table 6.4.

Much like the LWR model, the aARZ-1 is easy to solve analytically and numerically. One of the computational advantages it has is that it is only one hyperbolic conservation law with no relaxation term, which makes solving every step in the Godunov scheme faster. The main advantage, however, is that the aARZ-1 model is stable everywhere, which means that perturbations never grow. This implies that it is unnecessary to have a fine resolution when
simulating the aARZ-1 model. Whereas a cell size smaller than 4 meters is recommended when solving the ARZ model to capture jamitons with some level of detail, a cell size of 40 meters is enough to solve the aARZ-1. On average, it was observed that solving the ARZ model with a cell size of 4 meters takes 30x more time than solving the aARZ-1 with a cell size of 40 meters.

6.5 The shortcomings of the aARZ-1 model and possible directions

The aARZ-1 model assumes that the surviving jamitons (which minimize the flow rate for a given density) are established instantly. This is in fact not the case. Starting with uniform flow, if the SCC is violated, perturbations grow into short jamitons that merge together forming longer jamitons, which survive if stable. Also, if the initial conditions resemble the profile of a long jamiton, then a long jamiton will form before it splits into shorter stable jamitons. This process is not captured by the aARZ-1 model.

In addition, it was observed from simulations similar to the ones in Figure 6.2 that it takes a longer period of time for the stable jamitons to be established when the average density $\bar{\rho}$ is lower. We will not go into detail about how to solve this lack of detail in the jamiton formation mechanisms, but we will note
that an averaged second order model can be proposed to capture this process. This model would have the following form:

\[
\bar{\rho}_t + (\bar{Q}(\bar{\rho}, \eta))_x = 0, \\
\eta_t + (\bar{Q}/\bar{\rho})\eta_x = -b(\bar{\rho})\eta.
\]

(6.16)

We call the model in (6.16) aARZ-2. The flux function here \( \bar{Q}(\bar{\rho}, \eta) \) represents a family of curves in the fundamental diagram, parametrized by \( \eta \). One possible choice is

\[
\bar{Q}(\bar{\rho}, \eta) = \eta \bar{\rho}U(\bar{\rho}) + (1 - \eta) \bar{Q}(\bar{\rho})
\]

where \( \bar{Q} \) is the same function that appears in the aARZ-1 model. A plot of \( \bar{Q}(\bar{\rho}, \eta) \) is displayed in Figure 6.6 below.

The variable \( \eta \) here models how far \( \bar{\rho} \) is from the equilibrium average density of the jamiton that eventually survives, and it pushes \( \bar{\rho} \) in that direction via \( b(\bar{\rho}) \). Given that stable jamitons take different periods of time to arise, depending on the average density \( \bar{\rho} \), the function \( b(\bar{\rho}) > 0 \) can be calibrated to simulated data from the original ARZ model.

Even though the aARZ-2 model is a second order model, it is simpler than the ARZ model. In fact, aARZ-2 model is always linearly stable with respect to uniform base states. To verify that, we perform a linear stability analysis similar to the one in §3.2.

Start with base state solutions of (6.16) in which \( \bar{\rho} = \rho_0 \) and \( \eta = 0 \) are constant in space and time. Consider now infinitesimal wave perturbations of
Figure 6.6: A plot of $\bar{Q}(\bar{\rho}, \eta)$ for $\eta = 0, \eta = 0.25, \eta = 0.5, \eta = 0.75, \text{ and } \eta = 1$. 

the base state,

$$\hat{\rho} = \hat{R}e^{ikx+\sigma t} \quad \text{and} \quad \hat{\eta} = \hat{N}e^{ikx+\sigma t},$$

where $k$ is the wave number and $\sigma$ the complex growth rate. Substitute the perturbed solution $\rho = \rho_0 + \hat{\rho}$ and $\eta = 0 + \hat{\eta}$ into (6.16), and consider only constant and linear terms. This leads to the system

$$\begin{bmatrix}
\sigma + i k \bar{Q}'(\rho_0) & ik(\rho_0 U'(\rho_0) - \bar{Q}(\rho_0)) \\
0 & \sigma + ik \frac{\bar{Q}(\rho_0)}{\rho_0} + b(\rho_0)
\end{bmatrix}
\begin{bmatrix}
\hat{R} \\
\hat{N}
\end{bmatrix} =
\begin{bmatrix}
0 \\
0
\end{bmatrix}, \quad (6.17)$$

for the perturbation amplitudes. Nontrivial solutions can only exist if the
matrix in (6.17) has a vanishing determinant, which requires

\[
\left( \sigma + ik\tilde{Q}'(\rho_0) \right) \left( \sigma + i k \frac{\tilde{Q}(\rho_0)}{\rho_0} + b(\rho_0) \right) = 0.
\]

This means that \( \sigma \) is either purely imaginary, or \( \text{Re}(\sigma) = -b(\rho_0) \). Choosing \( b(\rho) \) to be a positive function implies that \( \text{Re}(\sigma) \leq 0, \forall \rho_0 \). This is equivalent to linear stability, which implies that this model is stable with respect to any uniform flow. Therefore, much like the \( \textbf{aARZ-1} \) model, the \( \textbf{aARZ-2} \) model can be simulated using a grid that is much more coarse than what is needed for the ARZ model. The main advantage to the \( \textbf{aARZ-2} \) model is that, similarly to the \( \textbf{aARZ-1} \) model, it contains information about the underlying jamiton of any pair \( (\bar{\rho}, \eta) \) in its solution, which can be extracted to estimate, for example, the fuel consumption of the ARZ model without having to solve it.
CHAPTER 7

CONCLUSION

7.1 Summary

This thesis is devoted to the study of non-equilibrium dynamics of second order macroscopic traffic models. In particular, we focus on analyzing the traveling wave solutions (jamitons) of the ARZ model when the sub-characteristic condition (SCC) is violated.

After going through the mathematical foundations of macroscopic traffic models in general, second order models in particular, and solving the ARZ model specifically in Chapters 2 and 3, we turned to the key research question: which jamiton solutions are dynamically stable?

A systematic computational study on the stability of jamitons was conducted in Chapter 4. The Godunov Method with an HLL approximate Rie-
mann solver was used to solve the ARZ model numerically. Starting with jamiton profile initial conditions, the system was evolved in time, and the solutions were checked at a later final time to distinguish between stable jamitons and unstable jamitons. Several patterns were observed in the results:

1. Very short jamitons are dynamically unstable. The mechanism of instability here is called a “merging” instability where a sequence of short jamitons merge together forming a longer jamiton.

2. Very long jamitons are dynamically unstable. The mechanism of instability here is called a “splitting” instability where a long jamiton splits into two or more shorter jamitons.

3. The length of a jamiton alone cannot identify whether it is stable or unstable.

4. In the \((\rho_S, L), (\bar{\rho}, s), (\rho_S, \rho^+)\) and \((\rho_S, \bar{\rho})\) jamiton phase planes, the region of stability separates two regions of instability, where one corresponds to the “merging” instability and the other corresponds to the “splitting” instability.

5. For a given average density \(\bar{\rho}\) of a jamiton, stable jamitons tend to have average flow rates \(\bar{Q}\) that are very close to the smallest possible average flow rate.
After that, a more rigorous analysis of the stability of jamitons was performed. This was started by deriving a linear model for infinitesimal perturbations of jamiton solutions of the ARZ model. This is different from deriving a linear model for perturbations of uniform states because a special treatment has to be applied to shocks. This perturbation analysis, presented in Chapter 5, led to a variable-coefficient linear advection-reaction system whose solutions characterize jamiton stability. As shown in the chapter, this system exhibits extremely complex dynamics that may not be suspected at first glance, given its simple fundamental structure.

A key reason for those complex dynamics is the zero-transition of one characteristic field, which corresponds to the sonic point in the nonlinear jamiton. The qualitative characterization presented in §5.2 reveals that there are two key mechanisms for instability that must be considered: first, asymptotic stability that captures the net amplification or decay of infinitesimal perturbations that traverse through periodic jamiton patterns; and second, the transient growth of small perturbations as they travel from near the sonic point down the jamiton profile until they eventually hit the next shock. The quantitative study in §5.5 reveals that for some jamitons, such transient amplifications may yield noise amplification by many orders of magnitude, which for many practical situations will definitely push the solutions into the fully nonlinear regime.
Based on those stability concepts, two criteria have been developed that are directly verifiable in terms of the model functions rather than requiring non-linear hyperbolic system simulations. Asymptotic stability reduces to finding the spectral radius of a sparse matrix, which in itself is a non-trivial problem as well, but it is an established standard task in numerical linear algebra. For the transient growth, a proxy criterion has been devised that boils down to a straightforward quadrature of two model functions. When compared with the “brute force” nonlinear stability results (Chapter 4), those two criteria capture the key qualitative essence of the stability boundary for long jamitons; but to reproduce the precise shape there is still room for improvement via more refined stability criteria.

One possible application of the results in Chapters 4 and 5 was explored in Chapter 6. In particular, an averaged model of the ARZ model was proposed. This model, named the \texttt{aARZ-1} model, is a first order model that is a hyperbolic conservation law that solves for the average of the traffic density in space $\bar{\rho}(x,t)$, where the averaging happens over a window of several kilometers. Even though the \texttt{aARZ-1} model cannot resolve individual jamitons, it carries information about the underlying jamitons. The advantage of this averaged model is that it is much easier and faster to solve numerically than the ARZ model, while at the same time containing information about the cumulative effects of jamitons. To that end, three metrics were investigated: fuel
consumption, total vehicular flow at a specific location, and braking time per hour of the average vehicle. Through several examples, it was demonstrated that the \textbf{aARZ-1} model can approximate the total fuel consumption of the ARZ model within around 2\%, total flow of vehicles within around 1\%, and braking time of the average vehicle within around 25\%. In contrast, using a simple first order LWR model produces an error (relative to the ARZ model) of around 15\% in terms of fuel consumption, 10\% in terms of total flow of vehicles, and 90–100\% in terms of braking time.

### 7.2 Discussion

The study presented in this thesis highlights important structural properties of hyperbolic conservation law systems with relaxation terms, in the regime when the SCC is violated. Such PDE are of importance in the macroscopic modeling of vehicular traffic flow (the main focus here), but also for other applications, such as roll waves in open channels [71] and circular hydraulic jumps [41]. Furthermore, many of the issues are similar to those that appear in the context of the ZND theory for the stability of Chapman-Jouguet (CJ) detonations [26]. In fact, jamitons are mathematical analogs of detonation waves [27]. While for detonation waves the notion of an SCC does not seem to apply, CJ detonations do have a sonic point, which renders their stability analysis [91, 12] difficult. It is our hope that the relative simplicity of systems
such as the ARZ model will provide a route to advance in this challenging topic.

Mathematically, understanding the solution behavior of relaxation systems in which the SCC is violated is a crucial challenge [64, 59, 40], and this work provides some insight. In addition, the jamiton perturbation system (5.8) is full of challenging structure (see §5.3), and this thesis provides criteria to characterize its stability properties. Moreover, the non-normal structure of the system in (5.8), leading to the transient growth behavior it exhibits (§5.5), has interesting connections to the task of stabilizing traffic flow with a single autonomous vehicle [15].

The work done in this thesis provides a pathway to understanding important stability questions for the ARZ model (3.1). In the regime of violated SCC, this model can reproduce the practically relevant [90] phenomena of phantom traffic jams and stop-and-go traffic waves, while preserving the advantages of a macroscopic description (see Chapter 1). The dynamic stability of jamitons determines which of the many theoretically possible jamiton solutions of the model can/will be selected by the equations’ dynamics. The study in Chapter 4 reveals that short jamitons tend to merge, and long jamitons tend to split, resulting in a middle range of stable jamiton wave lengths. A remarkable aspect about this dynamic selection via stability/instability is that it selects a length scale (range), even though there is no length scale that is
explicitly inserted into the model.

\section*{7.3 Future directions}

The qualitative analysis of the behaviour of the solution to the perturbation system in Chapter 5 was enough to extract the main mechanisms of instability of the system. Despite that, a complete analysis of that behaviour remains to be conducted in future work (including a full WKB analysis \cite{7}).

For the key application of traffic flow, the understanding of which jamiton solutions are dynamically stable is a critical step towards determining which models reproduce real-world phenomena best. One possible research direction here is to use experimental data to systematically calibrate the ARZ model. The model functions picked in §3.1 are chosen to roughly match the jamiton lines with the spread in the fundamental diagram data. Instead, the model functions could be calibrated to ensure that only jamitons with realistic features (traveling speed, length, etc...) are stable.

Additionally, a clear progression of what was presented in Chapter 6, regarding the averaging of the ARZ model, is to study the \texttt{aARZ-2} model further and in more detail, and to compare it to the \texttt{aARZ-1} model. In principle, the same techniques used for the \texttt{aARZ-1} model to quantify fuel consumption, total flow and braking time should be applicable to the \texttt{aARZ-2} model.
Moreover, the use of autonomous vehicles as moving bottlenecks to reduce total fuel consumption has been demonstrated both experimentally [15] and theoretically [82, 62, 61, 78]. The theoretical work relies either on microscopic models or first order LWR models. In the latter case, such studies rely on solving a PDE-ODE coupled systems [20], where the PDE tracks the vehicle density on the road, and the ODE tracks the position of the controlled vehicle. Extending this framework to the ARZ model can be done [98], but calculations for long highways can be challenging computationally. A possible research direction is to verify that the simpler aARZ-1 model can be used in this setting instead of the ARZ model or other LWR models.

Finally, an obvious extension of the analysis presented in Chapter 5 is to tackle the merging instability. This would require deriving a linear perturbation system for a sequence of two or more jamitons. At least in principle, the methodology of this current work can be extended to include the merging instabilities by allowing multiple shock perturbation.
REFERENCES


[40] S. Jin and M. A. Katsoulakis. Hyperbolic systems with supercharac-


[72] Seri Oh, Stephen G Ritchie, and Cheol Oh. Real-time traffic measure-


[101] G. B. Whitham. Some comments on wave propagation and shock wave


