A Stochastic Model of Macroscopic Traffic Flow: Theoretical Foundations

A DISSERTATION

SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL

OF THE UNIVERSITY OF MINNESOTA

BY

Saif Eddin Jabari

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

Doctor of Philosophy

Adviser: Henry X. Liu

August, 2012

Acknowledgements

I owe my gratitude to many people who have made this dissertation possible.

I would first like to express my deepest gratitude to my adviser, Professor Henry Liu, for his guidance, patience, and support throughout my graduate studies at the University of Minnesota. He has truly served as a mentor to me; over the years, he has presented me with challenging problems that helped me develop my technical skills, but also consistently welcomed and encouraged me to develop my own research questions.

I would like to thank Professor David Levinson for the many thought provoking questions and his constant willingness over the years to provide me with guidance whenever I asked him for it. I would like to thank Professor William Cooper for the technical help that he has provided me. The literature that he has recommended has substantially helped develop my understanding of the fluid and diffusion approximations of queueing systems. I wish to also thank Professor Gary Davis for his technical guidance throughout the years. My interest in the probabilistic modeling of traffic flow was inspired by classes that I took with him.

Finally, and most importantly, I would like to thank my wife Andrea. Without her support, encouragement, patience, and unwavering love, this dissertation would not have been possible. I thank my parents, Ghazi and Laila for believing in me and instilling in me the desire to learn.

Dedication

I dedicate this thesis to my father, Ghazi Jabari, for inspiring this journey and to my son, Adam Jabari, whom I hope my journey will inspire.

Abstract

In this thesis, a new stochastic extension of Godunov scheme based traffic flow dynamics is developed using a queuing theoretic approach. In contrast to the common approach of adding noise to deterministic models of traffic flow, the present approach considers probabilistic vehicle inter-crossing times (time headways) at various positions along the road as the source of randomness. Subsequently, time headways are used to describe stochastic vehicle counting processes. These counting processes represent the boundary flows in stochastic conservation equations of traffic flow. The advantage of this approach is that (i) non-negativity of time varying traffic variables (namely, traffic densities) is implicitly ensured, and (ii) the mean dynamic of the stochastic model is the Godunov scheme itself. Neither issue has been addressed in previous stochastic modeling approaches which extend the Godunov scheme and its special case, the cell transmission model. A Gaussian approximation of the queueing model is also proposed for purposes of model tractability. The Gaussian approximation is characterized by deterministic mean and covariance dynamics; the mean dynamics are those of the Godunov scheme. By deriving the Gaussian model, as opposed to assuming Gaussian noise arbitrarily, covariance matrices of traffic variables follow from the physics of traffic flow and can be computed using only few parameters, regardless of system size or how finely the system is discretized. Stationary behavior of the covariance function is analyzed and it is shown that the covariance matrices are bounded. Consequently, estimated covariance matrices are also bounded. As a result, Kalman filters that use the proposed model are stochastically observable, which is a critical issue in real time estimation of traffic dynamics. Model validation was carried out in a real-world signalized arterial setting, where cycle-by-cycle maximum queue sizes were estimated using the Gaussian model as a description of state dynamics in a Kalman filter. The estimated queue sizes were compared to observed maximum queue sizes and the results indicate very good agreement between estimated and observed queue sizes.

Contents

Acknowledgements i

Dedication ii

Abstract iii

List of Tables viii

List of Figures ix

Notation xi

1 Introduction 1

1.1 Background and Problem Statement 1

1.2 Research Scope and Contributions 3

1.3 Thesis Organization 5

2 Motivation and Illustration 7

2.1 Introduction 7

2.2 Motivation 7

2.3 Fluid Limits and Gaussian Approximation: An Illustration 9

3 Background: Deterministic Macroscopic Traffic Flow Modeling 12

3.1 Introduction 12

3.2 Macroscopic Variables of Traffic Flow . . . 13
3.3 Derivation of Conservation Laws . . . 15
3.4 Constitutive Relations: The Fundamental Diagram . . . 18
3.4.1 Empirical Fundamental Relations . . . 19
3.4.2 Fundamental Relations Derived from Car-Following . . . 23
3.5 Shockwaves, Rarefaction, and the Riemann Problem . . . 27
3.6 Numerical Solutions: Discrete Dynamics . . . 37

4 A Stochastic Model of Traffic Flow and its Fluid Limit 42
4.1 Introduction . . . 42
4.2 Preliminaries: properties of the fluid model . . . 43
4.3 The Stochastic Model . . . 45
4.4 Time Headways . . . 46
4.4.1 Example: Exponential Time Headways . . . 46
4.4.2 Example: Mixed Time Headways . . . 47
4.4.3 Example: State Dependent Time Headways . . . 47
4.5 The Fluid Limit: A Simplified Setting . . . 49
4.5.1 Scaling and the Strong Law of Large Numbers . . . 50
4.5.2 Example: scaling . . . 52
4.6 Fluid Limits for State-Dependent Processes . . . 54
4.6.1 Lipschitz Continuity of Flux Functions . . . 57
4.6.2 Derivatives of Flux Functions . . . 58
4.6.3 Boundedness of Traffic Densities . . . 60
4.6.4 Derivation of the Fluid Limit . . . 61
4.7 Numerical Example . . . 65

5 Gaussian Approximation of the Stochastic Traffic Flow Model 69
5.1 Introduction . . . 69
5.2 Derivation of the Gaussian Approximation for the General Case . . . 71
5.2.1 State Dependent Flow Rates . . . 72

5.2.2 Solution of the SDE . . . 74
5.2.3 Generalized Counting Processes . . . 77
5.3 Stationary Behavior of the Covariance Function . . . 78
5.3.1 Free-flow (Sub-Critical) Mean Traffic Conditions . . . 79
5.3.2 Capacity (Critical) Mean Traffic Condition . . . 80
5.3.3 Congested (Super-Critical) Mean Traffic Conditions . . . 81
5.4 Numerical Examples . . . 82

6 Traffic State Estimation and Model Validation **88**
6.1 Introduction . . . 88
6.2 The Estimation Problem . . . 89
6.3 Model Specification and Filtering Algorithm . . . 90
6.4 Observability . . . 93
6.5 Model Testing and Validation . . . 97
6.5.1 Model Parameters . . . 98
6.5.2 Estimated Cycle-by-Cycle Maximum Queue Sizes . . . 101
6.5.3 Discussion . . . 103

7 Conclusion and Future Research **104**
7.1 Research Summary . . . 104
7.2 Future Research . . . 105

References **107**

Appendix A. Mathematical Background **118**
A.1 Linear Algebra Background . . . 118
A.2 Calculus Background . . . 118
A.3 Probability Background . . . 120

Appendix B. Traffic Flow Theory Background **124**
B.1 Linear Advection Equation . . . 124

B.2 Approximation of Discontinuous Initial Data 125

Appendix C. Glossary and Acronyms **127**

C.1 Glossary 127

C.2 Acronyms 129

List of Tables

3.1 Numerical flux 40
6.1 Fitted flow-density relation parameters 98
6.2 Fitted free-flow headway parameters 100
6.3 Fitted congested headway parameters 101
C.1 Acronyms 129

List of Figures

3.1 Vehicle time headways and vehicle spacings 14
3.2 Arbitrary road segment 15
3.3 Greenshields' fundamental relations 20
3.4 Greenberg's fundamental relations 21
3.5 Underwood's fundamental relations 22
3.6 Newell-Franklin fundamental relations 23
3.7 Newell's simplified car-following model 25
3.8 Newell's simplified fundamental relations 27
3.9 Trapezoidal flow-density relation 27
3.10 Example linear advection problem solution 28
3.11 Characteristic lines for the linear advection equation 29
3.12 The characteristic speed vs. traffic speed 29
3.13 Intersecting characteristic lines 30
3.14 Traffic densities when characteristics intersect 30
3.15 Discontinuous initial data 33
3.16 Riemann problem shockwave solutions 35
3.17 Approximation of discontinuous initial data 36
3.18 Riemann problem rarefaction fan solutions 37
3.19 Traffic densities implied by intersecting characteristics 38
3.20 Daganzo's sending and receiving functions 41
4.1 Example fundamental relation 48

4.2 Example probability density plots 49
4.3 Example sample path for counting process 52
4.4 Example sample path for scaled counting process with $n = 3$ 53
4.5 Example sample path comparisons; thin: $n = 1$, dense: $n = 10$ 53
4.6 Example sample path comparisons; thin: $n = 1$, dense: $n = 100$ 54
4.7 Example sample path comparisons; thin: $n = 1$, dense: $n = 1000$ 54
4.8 Cell 1 traffic density comparisons 67
4.9 Cell 2 traffic density comparisons 68
5.1 Stationary traffic densities in cell 1 with 95% confidence intervals 85
5.2 Traffic densities with 95% confidence intervals; (a) cell 1, (b) cell 2. 86
5.3 Simulated traffic densities with 95% confidence intervals; (a) cell 1, (b) cell 2. 86
6.1 Example unobservable scenario 94
6.2 Example shockwave observability 95
6.3 Data collection site and detector locations 97
6.4 Fitted flow-density relationship 98
6.5 Probability plot of free-flow time headways 100
6.6 Probability plot of congested time headways 101
6.7 Comparison between estimated and measured queue sizes 102

Notation

The notation is organized below into two lists. The first lists the notation used to denote mathematical operations, such as distance, equivalence, inversion, and transposition, in addition to general notation used to denote spaces of variables. The second lists the variables used in this thesis along with their physical meaning; this list is meant to serve as a thorogh glossary of all variables used in this thesis. The items in the first list are organized in alphabetical order according to their meaning (since some of the symbols do not belong to a particular alphabet), while the second list is organized alphabetically according to the symbol, with Latin characters appearing first followed by the Greek characters.

General Notation

$\lvert \cdot \rvert$	the absolute value
$\approx$	approximately equal to
$\overset{\mathcal{D}}{\approx}$	approximately equal to in distribution
$\lvert\mathcal{C}\rvert$	the cardinality of $\mathcal{C}$; used here to denote the index of the last cell
$\longrightarrow$	converges to
$\overset{\mathcal{D}}{\longrightarrow}$	converges to in distribution
$\text{diag}(a_1, ..., a_m)$	a diagonal matrix with diagonal elements $\{a_1, ..., a_m\}$
$\nabla_{\mathbf{z}} g$	the directional derivative of the function g along the vector $\mathbf{z}$
$\overset{\mathcal{D}}{=}$	equal to in distribution
$\equiv$	is equivalent to (by definition), or is defined as
$\mathbb{E}\xi$	expectation of the random variable ξ
$\lfloor u \rfloor$	the largest integer less than or equal to u

∇g	the gradient of the function g
$\mathbf{I}$	the identity matrix
$\Longleftrightarrow$	if and only if
$1_{\{\cdot\}}$	the indicator function, $1_{\{A\}} = 1$ if the condition A is true and $1_{\{A\}} = 0$, otherwise
$\mathbf{A}^{-1}$	matrix inversion
d	a metric (e.g., $d(g_1(\cdot), g_2(\cdot)) \equiv \|\|g_1(\cdot) - g_2(\cdot)\|\|_U$)
Θ	a metric space
$\mathbb{N}$	the natural numbers
$\mathbb{P}(A)$	probability of event A
$\mathbb{R}$	the real line
$\mathbb{R}_+$	the non-negative real numbers
$\|\| \cdot \|\|_U$	the uniform norm
$\|\| \cdot \|\|_p$	the L_p norm
$\mathcal{F}$	a sigma-field
ϵ, δ	tolerance parameters
$\mathbf{A}^{\mathsf{T}}$	vector or matrix transposition
$g(\cdot)$	an entire trajectory of the function g
$g(u)$	the value assumed by function g (a point) at the point u
$x-$	the point immediately preceding x; if $g(\cdot)$ is continuous at x, then $g(x-) = g(x)$
$x+$	the point immediately succeeding x; if $g(\cdot)$ is continuous at x, then $g(x+) = g(x)$
$\mathbf{0}$	the zero matrix (a matrix will all of its elements equal to zero)

Variables Used and Their Physical Meaning

$a_j(t)$	the acceleration/deceleration of vehicle j at time t
$\mathbf{B}$	routing matrix

$\mathcal{C}$	a set of cell indices, $\mathcal{C} \subset \mathbb{N}$
$\bar{c}$	coefficient of variation
$\mathbf{D}(t)$	a matrix which captures the dependence of traffic densities between cells
$\hat{\mathbf{D}}(t)$	a matrix which captures the dependence of cumulative flows between cells
$G_h(\cdot\|y)$	the conditional distribution function of time headways
$G_t(\cdot)$	the distribution function of tracking headways
$\mathbf{H}$	measurement matrix
$h_j(x)$	the time headway between vehicles $j-1$ and j measured at position x
$h_j^f(x)$	the free component of time headway j at position x
$h_j^t(x)$	the tracking component of time headway j at position x
$\mathbf{K}(t_i)$	filtering (Kalman) gain matrix
l_x	length of cell x
l_{min}	the minimum cell length
$\mathbf{m}(t_i)$	measurement residual at time t_i
$N\big([x_1, x_2], [t_1, t_2]\big)$	the number of vehicles in road section $[x_1, x_2]$ during $[t_1, t_2]$
$N\big([x_1, x_2], t\big)$	the number of vehicles in road section $[x_1, x_2]$ at time instance t
$\mathcal{N}(t)$	a unit rate homogeneous Poisson process
$\widehat{\mathcal{N}}(t)$	a counting process with generally distributed i.i.d. time headways with mean 1
n	a scaling parameter
$\bar{O}_i^{(obs)}$	the ith observed average occupancy (% of time a detector is occupied)
$O_j^{(obs)}$	the occupancy time of the jth vehicle in seconds
$o^n(x,t)$	the deviation (or refinement) process for cumulative flows

$\tilde{o}(x,t)$	the approximated (Gaussian) refinement process for cumulative flows
$p\big(\boldsymbol{\eta}(t), t \mid \mathbf{Z}(t)\big)$	the conditional probability density function of the traffic state vector at time t given the measurement sequence
$p_x(k,t)$	the probability that the traffic density in cell x at time t is equal to k; i.e., $p_x(k,t) = \mathbb{P}\left(\rho(x,t) = k\right)$
$\tilde{\mathbf{Q}}(t)$	vector of approximated cumulative cell boundary flows
$\mathcal{Q}(x,t)$	the cumulative number of vehicles that have crossed the downstream boundary of cell x by time t
$Q_e\big(\bar{\rho}(x,t)\big)$	a fundamental flow-density relation (the subscript "e" means empirical or equilibrium)
q_{max}	the maximum rate of flow
$q(x,t)$	the rate of flow at x at time t
$\bar{q}(x,t)$	the mean rate of flow at x at time t
$R_e\big(\bar{\rho}(x,t)\big)$	a receiving function
$\tilde{\mathbf{r}}(t)$	vector refinement process
$r^n(x,t)$	the deviation (or refinement) process for traffic densities
$\tilde{r}(x,t)$	the approximated (Gaussian) refinement process for traffic densities
$S_e\big(\bar{\rho}(x,t)\big)$	a sending function
$s_j(t)$	the spacing between vehicles $j-1$ and j at time t
t	a time instance, $t \in \mathbb{R}_+$
$T_i^{(obs)}$	observation time used to calculate the ith observed flow rate
$t_j(x)$	the time at which vehicle j crosses position x
$\tilde{t}$	scaled time
U	a finite horizon time
$V_e\big(\bar{\rho}(x,t)\big)$	a fundamental speed-density relation (the subscript "e" means empirical or equilibrium)

v_f	the free-flow speed
$v_j(t)$	the speed of vehicle j at time t
v_s	the speed of a shockwave
$\bar{v}(x,t)$	the mean speed of vehicles at x at time t
w	the backward wave speed
$W_x(t)$	standard Brownian motion associated with outflows from cell x
$\mathbf{W}(t)$	vector standard Brownian motion
x	position along the road; either $x \in \mathbb{R}$ or $x \in \mathcal{C}$
$\hat{x}(t)$	the position of a discontinuity at time t
$x_j(t)$	the position of vehicle j at time t
$y(x,t)$	the vector of relevant traffic conditions at the downstream boundary of cell x; $y(x,t) = [\rho(x,t),\ \rho(x+1,t)]^{\mathsf{T}}$
$\bar{y}(x,t)$	the vector of relevant mean traffic conditions at the downstream boundary of cell x; $\bar{y}(x,t) = [\bar{\rho}(x,t),\ \bar{\rho}(x+1,t)]^{\mathsf{T}}$
$\mathbf{Z}(t)$	a sequence of traffic measurements available at time t
$\mathbf{z}(t_j)$	traffic flow measurements at time t_j
$\alpha^n(x,t)$	the vector $[r^n(x,t),\ r^n(x+1,t)]^{\mathsf{T}}$
$\tilde{\alpha}(x,t)$	the limit of $\alpha^n(x,t)$
β, β_j	arbitrary parameters of a statistical model
$\mathbf{\Gamma}(t)$	a diagonal matrix containing the standard deviations of cell boundary flows (the Itô integrands)
γ	a constant of integration
Δt	discrete time interval length
Δx	discrete space interval length
$\{\boldsymbol{\zeta}(t_i)\}$	measurement noise sequence
$\boldsymbol{\eta}(t)$	the state-space vector of the system at time t
$\mathbb{E}\boldsymbol{\eta}(t_i\|t_i)$	estimated (updated) mean state vector
$\mathbb{E}\boldsymbol{\eta}(t_i\|t_{i-1})$	predicted mean state vector

$\boldsymbol{\Theta}(t)$	covariance of traffic densities and cumulative flows at time t
θ	proportion of vehicles, which are tracking
λ_f	rate of free-flowing time headways
$\lambda\big(\rho^k(x-\Delta x), \rho^k(x)\big)$	the numerical flux at position $x - \frac{\Delta t}{2}$
$\lambda\big(\bar{y}(x,t)\big)$	the flux at the downstream boundary of cell x
$\bar{\lambda}(x,t)$	deterministic time-varying flow rates
$\boldsymbol{\lambda}(t)$	vector of mean fluxes at the cell boundary
$\lambda_i^{(obs)}$	the ith observed flow rate
$\{\boldsymbol{\Xi}(t_i)\}$	measurement noise covariance sequence
$\xi_1, \xi_2, ...$	a sequence of random variables
$\boldsymbol{\Pi}(t_i)$	residual covariance matrix at time t_i
$\rho(x,t)$	the traffic density at position x at time t; when $x \in \mathcal{C}$, $\rho(x,t)$ is the traffic density in cell x
$\boldsymbol{\rho}(t)$	vector of traffic densities
$\tilde{\rho}(x,t)$	the approximated Gaussian traffic density
$\tilde{\boldsymbol{\rho}}(t)$	approximated (Gaussian) vector of traffic densities
$\bar{\rho}^k(x)$	the traffic density in $\left[x - \frac{\Delta x}{2}, x + \frac{\Delta x}{2}\right]$ at time $k\Delta t$
$\bar{\rho}^{k,*}(x)$	an intermediate traffic density at position x during the time interval $[k\Delta t,\ (k+1)\Delta t]$
$\bar{\rho}(x,t)$	the mean traffic density (or traffic concentration) at position x at time t; when $x \in \mathcal{C}$, $\bar{\rho}(x,t)$ is the mean traffic density in cell x
$\bar{\boldsymbol{\rho}}(t)$	mean vector of traffic densities
$\rho_0(x)$	the prescribed initial traffic densities: $\rho_0(x) = \rho(x,0)$
$\rho_0^\delta(x)$	a continuous function that approximates $\rho_0(x)$
$\bar{\rho}_0(x)$	a step function which represents approximated initial data
$\rho^n(x,t)$	traffic density associated with scaled counting processes
ρ_{crit}	the critical traffic density, below which traffic is free and above which traffic is congested

ρ_{jam}	the maximum (or jam) density
ρ_l	the traffic density to the left of a discontinuity in $\rho_0(x)$
ρ_r	the traffic density to the right of a discontinuity in $\rho_0(x)$
$\mathbf{\Sigma}(t)$	covariance matrix of state-space vector
τ_j	the traversal time of vehicle j through an arbitrary road section
$\mathbf{\Sigma}(t_i\|t_{i-1})$	estimated (updated) state covariance matrix
$\mathbf{\Sigma}(t_i\|t_{i-1})$	predicted state covariance matrix
$\mathbf{\Psi}(t)$	covariance matrix of traffic densities at time t
$\hat{\mathbf{\Psi}}(t)$	covariance matrix of cumulative flows at time t

Chapter 1

Introduction

1.1 Background and Problem Statement

A variety of traffic management applications require probabilistic models of traffic flow. These include traffic simulation, real-time estimation of traffic conditions along freeways and signalized arterials, and applications that involve short-term traffic prediction such as adaptive traffic signal control. One of the main challenges pertaining to the use and analysis of probabilistic models of traffic flow is tractability. In a deterministic context, it is well known that prominent models, namely the model of Lighthill and Whitham [63] and Richards [95] (LWR) and its higher order extensions are ill-posed; adding randomness only exacerbates analytical complications.

The most widely used deterministic models of traffic flow are those that constitute discrete time and discrete space numerical solution schemes for the LWR model, such as Godunov scheme based traffic flow models [39, 58], the cell transmission model (CTM) [22, 23], and its finite difference generalization [24]; both the CTM and its fintite difference generalization are special cases of the Godunov scheme. These numerical methods are forward in time recursive schemes. For this reason, they now constitute traffic flow models in their own right. Their appeal stems from their simplicity and their ability to capture queue build-up and dissipation dynamics in both space and time.

In developing stochastic models of traffic flow, the most common approach is to add Gaussian noise to discrete deterministic models of traffic flow, such as the CTM. Examples include [10, 19, 36, 37, 45, 73, 75, 102–104, 106, 107], most of which are quite recent and were all developed for purposes of estimation and prediction of traffic flow conditions along freeways. This new interest in the subject emerged from the more widespread availability of real-time traffic sensor data. The rationale behind the use of Gaussian noise is that Gaussian models are characterized by their first two moments alone (the mean and the variance), a desirable feature from a tractability point of view. However, the arbitrary addition of noise could lead to two problems: (i) the possibility of producing negative sample paths (i.e., predicting negative traffic states) and (ii) mean dynamics that do not coincide with the original deterministic dynamics to which noise was added, due to the nonlinearity of the dynamic equations. In fact, non-linear functions of Gaussian random variables are generally non-Gaussian. These issues were also overlooked in recent theoretical developments in this area, such as in [54] and have yet to be addressed in the theory of stochastic conservation laws (see, for example, [46] for a brief discussion and numerical illustration). Other approaches to stochastic modeling of traffic flow in the literature include Botlzmann-like models of traffic flow (e.g., [86, 90]), Markovian/queueing network approaches (e.g., [29, 31, 49, 53, 85]), and cellular automata (e.g. [40, 74, 98, 99]). In general, these approaches do not suffer the two problems cited above (in some cases minor modification may be needed), but are generally intractable. Specific to problems of real-time traffic state estimation, successive Monte-Carlo or particle filter based methods are required for most of these models, which could be computationally prohibitive. On the other hand, extended Kalman filters, which are fast by comparison to particle filters, rely on first-order Taylor series approximations of the nonlinear dynamics. This limits the applicability of extended Kalman filtering to differentiable dynamics (such as those proposed in [29, 31, 53]) and precludes some of the most prominent traffic flow dynamics, such as the Godunov scheme and the CTM due to disjunctive flux functions (i.e., flux functions that involve extrema of traffic variables or conditional statements).

Another problem that arises in this context is related to computing covariance matrices

of traffic state variables, a crucial component of Kalman filters. To accurately capture queue build-up and dissipation dynamics, finer discretization of space and time are required. This results in large numbers of traffic flow variables and, consequently, large system covariance matrices. Furthermore, since flows across boundaries of cells (discrete space intervals) depend on traffic states on either side of the boundaries, dependencies between traffic state variables arise and assumptions of diagonal covariance matrices (typically made in the traffic state estimation literature) are not valid. Estimation of such large time varying covariance matrices may be prohibitive. More generally, for traffic estimation problems, finer discretization results in larger numbers of variables and increases the sparsity of the available measurements. This could lead to an observability issue; that is, introducing more variables, we have fewer observations. This type of observability pertains to the mean dynamics of the system. A critical issue related to the stochastic features of the system, from an estimation point of view, is *stochastic observability*. Stochastic observability is related to the behavior of the estimated covariance matrices of the system, which provide "a statistical description of the errors associated with the estimated state mean vector" [5]. Since the objective of any filter is to compute a minimum variance estimate, estimated covariance matrices must be bounded. This boundedness property is what defines stochastic observability.

1.2 Research Scope and Contributions

This research develops a stochastic extension of Godunov scheme based traffic flow dynamics as a stochastic queueing model of traffic flow. The source of randomness in the model is the randomness in vehicle time headways, which are used to develop stochastic counting processes that describe cumulative numbers of vehicles crossing cell boundaries over time. The counting processes are then used to develop (stochastic) conservation equations of traffic flow. The main advantage of using queueing models, in contrast to existing methods that extend Godunov scheme dynamics, is twofold: (i) non-negativity of traffic variables is

implicitly ensured and (ii) it is shown, via application of functional strong laws of large numbers (FSLLN), that the Godunov scheme dynamics arise as mean dynamics of the queueing model. That is, the qualitative behavior of the proposed queueing model is consistent with well-established principles of traffic flow theory. Furthermore, by application of a functional central limit theorem (FCLT), a tractable Gaussian model of traffic flow is obtained. By deriving a Gaussian model, rather than assuming Gaussian noise arbitrarily, the covariance matrices of the system can be computed from only few parameters.

The derived covariance matrices of the Gaussian model vary depending on traffic state and are shown to be bounded. The dependence on traffic state can be illustrated as follows: under free-flow traffic conditions, there is little variability in the model; in the extreme case of zero traffic density in a section of road, the variance in flows out of the section is also zero, which means, in this extreme case, that the model predicts zero outflow with certainty. Likewise, in the extreme case of jam traffic density, the model predicts zero inflows with certainty. The variance is largest around capacity traffic flow conditions, which may be interpreted as being less certain about traffic flow conditions when large numbers of vehicles that interact frequently are present on the road. It is notable that such dependence on traffic state has been overlooked in all but few research efforts on Gaussian macroscopic traffic flow models.

The derived mean and covariance dynamics of the Gaussian model are first-order deterministic differential equations that depend on the *expected values* of traffic flow variables, not the traffic flow variables themselves. This allows for implementation of a standard Kalman filter for purposes of traffic state estimation and prediction, which results in computational tractability and permits real-time implementation of the proposed model. The continuous time setting in which the model is derived offers the flexibility of using different computational time scales for the state and the measurement equations of the Kalman filter. That is, the availability of measurements at regular time intervals is not required to run the filter. In general, due to the sparsity of measurements, observability is difficult to establish. However, under certain traffic flow conditions, such as free-flow conditions, the presence of traffic sensors on either end of a road section will allow for reconstruction of initial mean

traffic densities within the road section. To overcome observability issues, this study uses a warm-up period where the initial conditions are observable free-flow traffic conditions. Furthermore, the number of cells used has no impact on whether mean traffic conditions are observable or not. In terms of stochastic observability, a crucial contribution of this research is that the covariance matrices of the Gaussian model are bounded. Thus, Kalman filters built using the proposed model are stochastically observable.

1.3 Thesis Organization

This thesis is organized as follows: Chapter 2 motivates the methods applied in this thesis and highlights the difficulties that would arise in applying "classical" Markovian techniques. The chapter also demonstrates application of the FSLLN and the FCLT to a simple queueing model of traffic flow so as to develop the intuition behind the detailed treatment in subsequent chapters.

Chapter 3 provides a self contained introduction to deterministic macroscopic models of traffic flow, in which traffic flow variables are defined and detailed derivations of the conservation laws of traffic flow are given. A discussion of constitutive relations of traffic flow (namely, the fundamental diagram) is given along with illustrations of how some of the more popular relations are derived from microscopic traffic flow considerations. The chapter culminates in a derivation of the Godunov scheme for concave fundamental relations and presents a particular manifestation of the scheme, the CTM.

Chapter 4 develops the stochastic queueing model proposed in this thesis and discusses, in detail, the properties of the model which allow the development of fluid and Gaussian approximations. A more detailed derivation of the fluid limit of the simplified system presented above is given so as to motivate some of the technical details needed to derive the fluid limit in the general case when flow rates depend on the stochastic traffic densities. The generalization utilizes the continuous mapping approach presented in [112]. It is shown that the fluid limit of the general process is a continuous time version of the Godunov scheme and a numerical example is given to illustrate the convergence.

Chapter 5 develops the Gaussian approximation of the stochastic process starting from a simplified setting and continuing on to the general setting capitalizing again on the continuous mapping approach. It is found that the approximation converges to the solution of a stochastic differential equation (SDE) which is linear in the narrow sense. Thus, an explicit solution of the SDE is established. The solution is shown to be a Gaussian process with mean dynamics equivalent to the Godunov scheme and covariance function which is easy to solve numerically. This chapter also discusses the stationarity and boundedness properties of the covariance matrices of the Gaussian approximation, which are a crucial ingredient in establishing that the proposed model is stochastically observable. Numerical examples are given to illustrate the stationary behavior of the covariance function.

Chapter 6 discusses use of the Gaussian model in traffic state estimation with fixed traffic sensor data (inductance loop detectors). The state-space and measurement models are given and the discrete-continuous Kalman filtering algorithm is presented. The chapter also discusses observability issues and impact of dividing the road section into a large number of cells. Finally, a real-world estimation example is presented as a validation test of the proposed model. The findings indicate a good match between observed traffic conditions and estimated traffic conditions. Chapter 7 concludes the thesis.

Chapter 2

Motivation and Illustration

2.1 Introduction

Derivation of the fluid and Gaussian approximations of a queueing model constitute a major portion of this research. The purpose of this chapter is to motivate the use of these approximations as opposed to the queueing model itself. An illustration is then given using the simplified setting of processes with deterministic time-varying rates. The intuition developed here carries over to the more sophisticated (state-dependent) setting, which is the main interest in this thesis. Detailed treatments of both the simplified setting and the state dependent setting are presented in subsequent chapters.

2.2 Motivation

Consider a homogeneous roadway without sources or sinks, which is divided into cells. Let $\mathcal{Q}(x,t)$ denote a stochastic counting process describing the cumulative number of vehicles that have crossed the downstream boundary of cell $x \in \mathcal{C}$ at time t, where $\mathcal{C}$ is a set of cell indices. The conservation of traffic density in cell x is written as:

$$\rho(x,t) = \rho(x,0) + \frac{1}{l_x}\big(\mathcal{Q}(x-1,t) - \mathcal{Q}(x,t)\big), \tag{2.1}$$

where $\rho(x,t)$ is the (random) traffic density in cell x at time t and l_x is the length (or size) of cell x. We wish to characterize $\rho(x,t)$ probabilistically. For the sake of illustration, let's assume that $\rho(x,0) = 0$ almost surely for all x and that the cell lengths are equal and normalized to 1; that is $l_x = 1$. Then the traffic density in x is characterized by the two counting processes $\mathcal{Q}(x-1,t)$ and $\mathcal{Q}(x,t)$. Again, for the sake of illustration, let's assume that $\{\mathcal{Q}(x,t)\}_{x\in\mathcal{C}}$ may be represented by independent (non-homogeneous) Poisson processes with time-varying deterministic rates $\{\bar{\lambda}(x,t)\}_{x\in\mathcal{C}}$; that is, we assume here that the instantaneous flow rates are given time varying constants. We then obtain a non-homogeneous birth and death process for each of the cells, or an $M_t/M_t/1$ queueing system for each cell. Then the Markovian approach to characterizing the probabilistic nature of $\{\rho(x,t)\}_{x\in\mathcal{C}}$ consists of solving the following system of (birth and death) equations for each x:

$$\frac{\partial}{\partial t}p_x(k,t) = \bar{\lambda}(x-1,t)p_x(k-1,t) + \bar{\lambda}(x,t)p_x(k+1,t) - \big(\bar{\lambda}(x-1,t) + \bar{\lambda}(x,t)\big)p_x(k,t), \quad (2.2)$$

for $k = 1, 2, ...$ and

$$\frac{\partial}{\partial t}p_x(0,t) = \bar{\lambda}(x,t)p_x(1,t) - \bar{\lambda}(x-1,t)p_x(0,t), \quad (2.3)$$

for $k = 0$, where $p_x(k,t) \equiv \mathbb{P}\big(\rho(x,t) = k\big)$.

The Markovian approach, which considers non-stationary dynamics (i.e., solving (2.2) and (2.3)), can be rather difficult, and a closed form solution is rarely available explicitly [66]. In fact, even when closed form expressions are available, their complexity could render their use and analysis preventative. As an example, when the rates $\bar{\lambda}(x,t)$ do not vary with time, the solution involves modified Bessel functions [69]. For this reason, one resorts to asymptotic analysis of the probabilities; that is, the probabilities which arise when $t \to \infty$. In this case, the left-hand sides of (2.2) and (2.3) are zero (i.e., $\frac{\partial}{\partial t}p_x(k,t) = 0$) and the differential equations become difference equations, which are easier to solve and typically deliver simpler solutions.

However, while it may not be difficult to obtain/analyze the stationary probabilities corresponding to the long-run probabilistic behavior of the system, this is of little use to a

traffic engineer concerned with the transient features of traffic flow (e.g., changes in traffic characteristics from one traffic light cycle to the next). To quote Mandlebaum and Massey [66]: "approximating the behavior of the system in the here and now by its behavior at time infinity is typically futile". Then, instead of attempting to explicitly solve equations (2.2) and (2.3), an alternative approach is to approximate (2.1) by a more tractable stochastic model. Specifically, if one could approximate (2.1) by a Gaussian process for which the mean (deterministic) behavior of the system and the time-varying covariance are easy to compute, then the problem becomes much simpler. Indeed, this approach first appeared in [79, 80] for the $M_t/M/1$ queue and made rigorous in [69, 70] for the $M_t/M_t/1$ queue; in particular, the latter introduced the idea of "uniform acceleration" as an approximation method that preserves the transient features of the queueing process. Using "uniform acceleration", [66] extended the work of [70] to the asymptotic analysis of the sample paths of the $M_t/M_t/1$ queue. This is the approach taken in this thesis, which is illustrated next for the simple $M_t/M_t/1$ queue given above. It is crucial to note here that the difficulties are substantially exacerbated when the flow rates $\bar{\lambda}(x,t)$ depend on the (random) traffic densities, which is the case for macroscopic traffic flow.

2.3 Fluid Limits and Gaussian Approximation: An Illustration

The main idea behind uniform acceleration of queueing processes is to consider, instead of the original process, a sequence of processes where all of the rates are multiplied by the scaling factor n; i.e., processes which, on average, count n times as many vehicles per unit time. The counting processes are then divided by n. Then letting $n \to \infty$, we obtain a fluid process. This is analogous to counting fractions of size n of vehicles rather than whole vehicles and letting the fractions get smaller and smaller while counting larger and larger numbers of fractions.

The scaled process is denoted by $\frac{1}{n}\mathcal{Q}(x,nt)$. When $\mathcal{Q}(x,t)$ is a non-homogeneous Poisson

process the FSLLN says that:

$$\frac{1}{n}\mathcal{Q}(x, n\cdot) \underset{n\to\infty}{\longrightarrow} \int_0^{\bullet} \bar{\lambda}(x, u)du \quad \text{almost surely (a.s.)}, \tag{2.4}$$

where the notation "$\cdot$" is used to indicate that we are talking about the convergence of a process (i.e., a function of time) to another process, and the notation "$\bullet$" is used when this appears as the upper limit of integration (for typographic purposes). It is to be understood, in the sequel, unless otherwise stated, that by $\mathcal{Q}(x, nt)$ is meant $\mathcal{Q}(x, n\cdot)$.

The fluid limit in (2.4) (in conjunction with the continuous mapping theorem [112] – see Appendix A.3 Theorem A.3.7) can the be used to determine the fluid limit of the simple traffic density process $\rho(x, t) = \mathcal{Q}(x-1, t) - \mathcal{Q}(x, t)$, which is a deterministic process; that is, as $n \to \infty$:

$$\rho^n(x, t) \longrightarrow \bar{\rho}(x, t) = \int_0^t \bar{\lambda}(x-1, u)du - \int_0^t \bar{\lambda}(x, u)du \quad \text{almost surely (a.s.)}, \tag{2.5}$$

where $\rho^n(x, t)$ is the traffic density associated with the scaled counting processes $\frac{1}{n}\mathcal{Q}(x-1, nt)$ and $\frac{1}{n}\mathcal{Q}(x, nt)$, and $\bar{\rho}(x, t)$ is the deterministic traffic density process.

The fluid limit, in essence, captures the mean (qualitative) behavior of the process. The Gaussian approximation *refines*[1] the fluid limit by adding a stochastic refinement process. The refinement process can simply be thought of as a measure of the deviation of the stochastic model from its fluid limit. The refinement process is:

$$r^n(x, t) = \sqrt{n}\left(\frac{1}{n}\mathcal{Q}(x, nt) - \int_0^t \bar{\lambda}(x, u)du\right), \tag{2.6}$$

where the re-scaling factor $\sqrt{n}$ in essence amplifies the deviation between the two processes, which would otherwise converge to zero in accordance with the fluid limit.

Letting $n \to \infty$ in (2.6), we have, by the FCLT, that the refinement process converges, in distribution, to a Gaussian process. In particular, we have that

$$r^n(x, \cdot) \xrightarrow{\mathcal{D}} \int_0^{\bullet} \sqrt{\bar{\lambda}(x, u)}\, dW_x(u), \tag{2.7}$$

[1] For a clarification of the nomenclature, see for example [17].

where $\xrightarrow{\mathcal{D}}$ means "converges in distribution to", $W_x(\cdot)$ is standard Brownian motion, and $\int_0^\bullet \sqrt{\bar{\lambda}(x,u)}\, dW_x(u)$ is an Itô integral with deterministic integrand, which may be understood for, for some fixed t to be a Gaussian random variable with mean zero and variance $\int_0^t \bar{\lambda}(x,u)du$.

Returning to the traffic density process, we see that $\rho(x,t)$, with $\mathcal{Q}(x-1,t)$ and $\mathcal{Q}(x,t)$ represented by independent non-homogeneous Poisson processes, may be approximated by:

$$\tilde{\rho}(x,t) = \int_0^t \bar{\lambda}(x-1,u)du - \int_0^t \bar{\lambda}(x,u)du + \int_0^t \sqrt{\bar{\lambda}(x-1,u)}dW_{x-1}(u) - \int_0^t \sqrt{\bar{\lambda}(x,u)}dW_x(u), \tag{2.8}$$

where $\tilde{\rho}(x,t)$ denotes approximated traffic density. This process has mean

$$\mathbb{E}\tilde{\rho}(x,t) = \int_0^t \big(\bar{\lambda}(x-1,u) - \bar{\lambda}(x,u)\big)du, \tag{2.9}$$

which coincides with the fluid limit of the process. The process variance is:

$$\mathbb{E}\big(\tilde{\rho}(x,t) - \mathbb{E}\tilde{\rho}(x,t)\big)^2 = \int_0^t \left(\sqrt{\bar{\lambda}(x-1,u)} + \sqrt{\bar{\lambda}(x,u)}\right) du \tag{2.10}$$

The variance given by (2.10) grows without bound so as long as vehicles continue to flow through the cells. This is an undesirable feature from both physical and application standpoints. That the variance grows without bound means that the probability of negative traffic states and traffic densities that exceed jam density increase with time, even under free-flow traffic conditions. Moreover, from the perspective of traffic state estimation applications, the unboundedness of the covariance matrices could easily lead to stochastic unobservability, particularly in cases where road sections are divided into many cells. The Gaussian approximations associated with the state-dependent counting processes studied in this thesis do not suffer this limitation as will be discussed in subsequent chapters.

Chapter 3

Background: Deterministic Macroscopic Traffic Flow Modeling

3.1 Introduction

The purpose of this chapter is to provide a self-contained background on first-order macroscopic modeling of traffic flow. Specifically, this chapter provides complete derivations of conservation laws in various integral forms and the more popular differential form, which was first studied by Lighthill and Whitham [63] and independently by Richards [95]; the theory is commonly referred to as LWR theory, due to the three authors. The different forms of the conservation law serve different purposes; the physical meaning behind them is most easily seen in the integral forms, while the differential form is used to derive methods to solve the conservation equation. This chapter culminates in a presentation of the Godunov scheme [39, 58], a forward in time numerical method for solving the LWR model, and a special kind of Godunov scheme that is popular in the traffic flow literature: the cell transmission model [22, 23]. The stochastic model proposed in this thesis is a probabilistic extension to Godunov scheme based dynamics.

3.2 Macroscopic Variables of Traffic Flow

Macroscopic traffic flow modeling looks at the interplay between the following three dynamic variables of traffic flow (or any subset of two thereof): mean traffic density, mean flow rate, and mean speed, denoted respectively by $\bar{\rho}(x,t)$, $\bar{q}(x,t)$, and $\bar{v}(x,t)$, where $x \in \mathbb{R}$ is a position along the road and $t \in \mathbb{R}_+$ is a time instance. The macroscopic traffic variables are defined broadly as the average number of vehicles per unit length of the road at position x at time t, the average number of vehicles to cross position x per unit time at time t, and the mean speed of vehicles at position x at time t, respectively. How they arise as averages, and in what sense, may differ based on how the variables are defined.

The classical definitions of the three variables is due to Wardrop [108] and Lighthill and Whitham [63], and constitute the most widely used definitions of the three variables in traffic engineering since they are based on how one would measure the variables in practice. From the broad definitions above, we have the following basic relation:

$$\bar{q}(x,t) = \bar{\rho}(x,t)\bar{v}(x,t), \tag{3.1}$$

which restricts us to a particular definition of mean speed, namely, the "space mean speed" [108]. To see this, proceed as follows: fix a section of the road of length Δx, centered at position x, and a time interval of length Δt, centered at time instance t. Then record the traversal times of N vehicles through the road section during the specified time interval. Denote these traversal times by τ_1, τ_2, ..., τ_j, ..., τ_N. The space mean speed is defined as the ratio of the length of the road segment (Δx) to the average traversal time:

$$\bar{v}(x,t) \equiv \frac{\Delta x}{\frac{1}{N(x,t)} \sum_{j=1}^{N(x,t)} \tau_j}, \tag{3.2}$$

where $N \equiv N(x,t)$, short for $N \equiv N\left(\left[x - \frac{\Delta x}{2}, x + \frac{\Delta x}{2}\right], t\right)$, is used to remind us that N varies with x and t. Define the mean rate of flow through our road segment over the time interval of interest as:

$$\bar{q}(x,t) \equiv \frac{N(x,t)}{\Delta t} \tag{3.3}$$

Since some of the vehicles may have left the section during the interval $[t - \frac{\Delta t}{2}, t + \frac{\Delta t}{2}]$, $\frac{N(x,t)}{\Delta x}$ is not the correct expression for the traffic density. Instead, Lighthill and Whitham

[63] suggest using the average number of vehicles divided by the section length:

$$\bar{\rho}(x,t) \equiv \frac{\frac{1}{\Delta t}\sum_{j=1}^{N(x,t)} \tau_j}{\Delta x} \tag{3.4}$$

Then, from equations (3.2) - (3.4), it can be seen that the interpretation of $\bar{v}(x,t)$ as the space mean speed delivers the basic relation (3.1).

Alternatively, one could define traffic density and flow rate as the reciprocals of microscopic variables of traffic flow. Denote by $t_j(x)$ the time vehicle j crosses position x and let $h_j(x)$ denote the j^{th} inter-arrival (or inter-departure) time at position x, which shall be referred to hereafter as the j^{th} time headway. Consequently, $h_j(x) = t_j(x) - t_{j-1}(x)$. Also, let $x_j(t)$ denote the position of vehicle j at time t and $s_j(t)$ the spacing between vehicles $j-1$ and j; that is $s_j(t) = x_{j-1}(t) - x_j(t)$. These variables are illustrated in Figure 3.1.

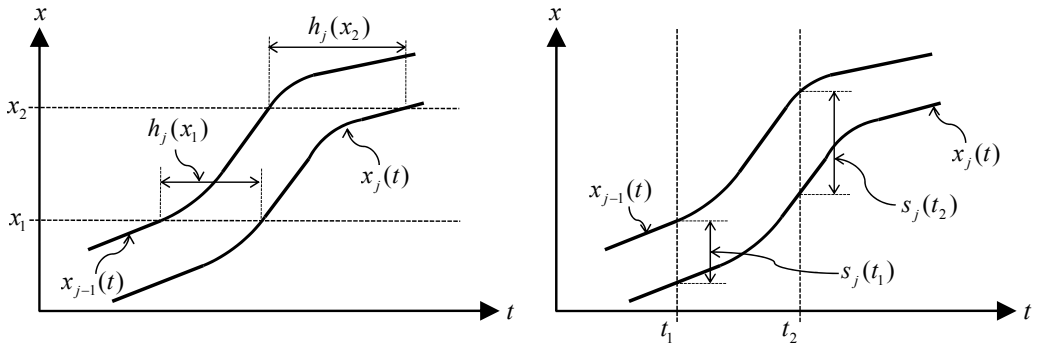


Figure 3.1: Trajectories of two successive vehicles; (a) vehicle time headways; (b) vehicle spacings

From the microscopic variables of traffic flow, traffic density, flow rate, and speed can be defined as follows:

$$\rho(x,t) = \frac{1}{s_j(t)}, \quad \text{for } x \in [x_j(t), x_{j-1}(t)) \tag{3.5}$$

$$q(x,t) = \frac{1}{h_j(x)}, \quad \text{for } t \in (t_{j-1}(x), t_j(x)] \tag{3.6}$$

$$v(x,t) = \frac{q(x,t)}{\rho(x,t)} = \frac{s_j(t)}{h_j(x)}, \quad \text{for } x \in [x_j(t), x_{j-1}(t)), t \in (t_{j-1}(x), t_j(x)] \tag{3.7}$$

Note that the overhead bars (e.g., $\bar{\rho}$) have been omitted in equations (3.5) - (3.7), since the quantities in these equations are no longer mean quantities. For instance, for a fixed position x, $q(x,t)$ varies for *every pair* of vehicles that cross over time and likewise, for a fixed time instance t, $\rho(x,t)$ varies with x depending on the spacing between *pairs* of consecutive vehicles.

3.3 Derivation of Conservation Laws

This section provides full derivations of integral and differential conservation laws of traffic flow. They are derived here as relations amongst the three mean variables of traffic flow ($\bar{\rho}(x,t)$, $\bar{q}(x,t)$, and $\bar{v}(x,t)$) although the derivations and the laws still hold for the macroscopic variables defined in equations (3.5) - (3.7). The derivations given here are universal, just as conservation laws are universal and apply in a variety of disciplines including fluid dynamics, geophysics, and biomechanics. Further information can be found in [59, 60, 62, 110].

Let $[x_1, x_2]$ denote an arbitrary road segment with instantaneous rates of flow through the boundaries $\bar{q}(x_1,t)$ and $\bar{q}(x_2,t)$. See Figure 3.2.

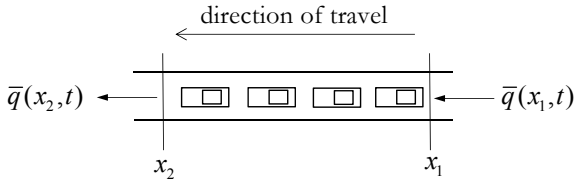


Figure 3.2: Arbitrary road segment

At any time instance t, the number of vehicles present in the road segment is:

$$N\big([x_1,x_2],t\big) = \int_{x_1}^{x_2} \bar{\rho}(x,t)dx \tag{3.8}$$

The rate at which $N\big([x_1,x_2],t\big)$ changes with time is written as:

$$\begin{aligned} \frac{d}{dt}\int_{x_1}^{x_2} \bar{\rho}(x,t)dx &= \bar{q}(x_1,t) - \bar{q}(x_2,t) \\ &= \bar{v}(x_1,t)\bar{\rho}(x_1,t) - \bar{v}(x_2,t)\bar{\rho}(x_2,t) \end{aligned} \tag{3.9}$$

Equation (3.9) will be referred to as *the first integral form of the conservation law*, which will be used in deriving shockwave speeds in the sequel. Suppose now we are interested in change in the number of vehicles in our road segment over a time interval $[t_1, t_2]$. To write this, one simply integrates both sides of equation (3.9) over $[t_1, t_2]$:

$$\int_{t_1}^{t_2} \left(\frac{d}{dt} \int_{x_1}^{x_2} \bar{\rho}(x,t)dx \right) dt = \int_{t_1}^{t_2} \Big(\bar{v}(x_1,t)\bar{\rho}(x_1,t) - \bar{v}(x_2,t)\bar{\rho}(x_2,t) \Big) dt \tag{3.10}$$

By applying (a corollary to) the fundamental theorem of calculus, (see Appendix A.2, Corollary A.2.1), to the left-hand side (LHS) of (3.10), we get:

$$\int_{t_1}^{t_2} \left(\frac{d}{dt} \int_{x_1}^{x_2} \bar{\rho}(x,t)dx \right) dt = \int_{x_1}^{x_2} \bar{\rho}(x,t_2)dx - \int_{x_1}^{x_2} \bar{\rho}(x,t_1)dx, \tag{3.11}$$

and separating the right-hand side (RHS) of (3.10) into two integrals, we get the *second integral form of the conservation law*:

$$\begin{aligned} \int_{x_1}^{x_2} \bar{\rho}(x,t_2)dx - \int_{x_1}^{x_2} \bar{\rho}(x,t_1)dx &= \int_{t_1}^{t_2} \bar{v}(x_1,t)\bar{\rho}(x_1,t)dt - \int_{t_1}^{t_2} \bar{v}(x_2,t)\bar{\rho}(x_2,t)dt \\ &= \int_{t_1}^{t_2} \bar{q}(x_1,t)dt - \int_{t_1}^{t_2} \bar{q}(x_2,t)dt \end{aligned} \tag{3.12}$$

The second integral form of the conservation law has the following interpretation: the first integral on the LHS represents the number of vehicles present in the road section at (future) time t_2, while the second integral on the LHS represents the number of vehicles present in the road segment at (earlier) time t_1. The first integral on the RHS represents the total number of vehicles that have entered the road segment (through x_1) over the time interval $[t_1, t_2]$, while the second integral on the RHS represents the number of vehicles that have left the road segment (through x_2) over the time interval $[t_1, t_2]$. That is, equation (3.12) says that number of vehicles in the road segment at any future time, t_2, is equivalent to the number of vehicles that were present at an earlier time, t_1, plus the total inflow during the time interval $[t_1, t_2]$, less the total outflow during the time interval $[t_1, t_2]$. **Note:** for conservation of vehicles to be honored, (3.12) must hold for all quadruples $\{x_1, x_2, t_1, t_2\}$, which is difficult to check both in practice and in theory, particularly, since these appear as limits in the integrals. However, the second integral form of the conservation law is useful

for designing numerical solution schemes, where space and intervals are fixed a priori; this will be revisited later in this chapter.

To derive the (classical) differential form of the conservation law (i.e., the LWR equation), one needs to make the following critical assumption:

Assumption. The quantities $\bar{\rho}(x,t)$, $\bar{q}(x,t)$, and $\bar{v}(x,t)$ are continuous and possess continuous partial derivatives on $\mathbb{R} \times \mathbb{R}_+$.

The importance of this assumption lies in the fact that –as we shall see later– it does not hold in general, which renders the LWR partial differential equation (PDE) *ill-posed* and further analysis is required to characterize a physically acceptable solution. Nonetheless, the differential form remains useful for purposes of solving conservation laws. **Note:** this assumption was not made in our derivation of the integral forms (3.9) and (3.12), which consequently do not suffer these limitations of the LWR PDE.

Let us now return to the derivation of the differential form. Start with equation (3.10): the assumption above allows us to interchange the order of differentiation and integration on the LHS (see Appendix A.2, Theorem A.2.2). That is,

$$\int_{t_1}^{t_2} \left(\frac{d}{dt} \int_{x_1}^{x_2} \bar{\rho}(x,t) dx \right) dt = \int_{t_1}^{t_2} \int_{x_1}^{x_2} \frac{\partial}{\partial t} \bar{\rho}(x,t) dx dt, \tag{3.13}$$

and from (a corollary to) the fundamental theorem of calculus (Appendix A.2, Corollary A.2.1), the RHS of (3.10) may be written as:

$$\begin{aligned} \int_{t_1}^{t_2} \Big(\bar{v}(x_1,t)\bar{\rho}(x_1,t) - \bar{v}(x_2,t)\bar{\rho}(x_2,t) \Big) dt &= \int_{t_1}^{t_2} \int_{x_2}^{x_1} \frac{\partial}{\partial x} \Big(\bar{v}(x,t)\bar{\rho}(x,t) \Big) dx dt \\ &= -\int_{t_1}^{t_2} \int_{x_1}^{x_2} \frac{\partial}{\partial x} \Big(\bar{v}(x,t)\bar{\rho}(x,t) \Big) dx dt \end{aligned} \tag{3.14}$$

Assembling (3.13) and (3.14), we get:

$$\int_{t_1}^{t_2} \int_{x_1}^{x_2} \left(\frac{\partial}{\partial t} \bar{\rho}(x,t) + \frac{\partial}{\partial x} \Big(\bar{v}(x,t)\bar{\rho}(x,t) \Big) \right) dx dt = 0 \tag{3.15}$$

which implies the differential form of the conservation law:

$$\frac{\partial}{\partial t} \bar{\rho}(x,t) + \frac{\partial}{\partial x} \Big(\bar{v}(x,t)\bar{\rho}(x,t) \Big) = 0, \tag{3.16}$$

or, equivalently,

$$\frac{\partial}{\partial t}\bar{\rho}(x,t) + \frac{\partial}{\partial x}\bar{q}(x,t) = 0 \tag{3.17}$$

Despite its limitations, the differential form of the conservation law has the advantage of giving insight into the solution of the conservation equation. Techniques for establishing good solution properties (e.g., existence and uniqueness) are more widely available for PDEs than integral equations, the latter possessing multiple solutions in the case of conservation laws; heuristically speaking, it can be easily seen, for example, in (3.9) or (3.12) that one could propose various paths $\bar{\rho}(\cdot, t)$ over $[x_1, x_2]$ that can be integrated to produce the same value. For this reason, one prescribes additional ("entropy") conditions to pick out physically relevant solutions.

3.4 Constitutive Relations: The Fundamental Diagram

While conservation laws are universal, constitutive relations are problem-specific and capture the idiosyncrasies of the system at hand. Here, a constitutive relation shall also *close* the conservation law: that is, one needs an additional relation in order to solve for the two unknowns that appear in the conservation equation (regardless of form); this additional relation is what is commonly referred to as the constitutive relation. In the context of macroscopic traffic flow modeling, there exist two classes of constitutive relations: the first are *stationary* relations between any two macroscopic variables of traffic flow, referred to as "fundamental relations" of traffic flow, due to Haight [43], who first introduced the terminology. The second class describe speed evolution as solutions to partial differential equations and are intended to capture both stationary and non-stationary traffic conditions; classical versions can be found in [87, 110] and more contemporary versions can be found in [4, 115]. The first class combined with the conservation equation constitute what is referred to as *first-order* macroscopic models of traffic flow, while the conservation equation combined with the second class of constitutive relations are referred to as *second-order* models. This thesis shall focus on first-order models, and no further discussion of the second class of constitutive relations will be given.

Fundamental relations of traffic flow come in two flavors: speed-density relations and flow-density relations, which express the mean speed and the mean flow rate as functions of traffic density, respectively. These relations, in essence, say the following: (i) the mean speed along the road decreases from a maximum value (the free-flow speed) as the mean traffic density on the road increases and until a maximum density is reached (the jam density) where vehicles can no longer move and the speed is zero; (ii) the mean rate of flow increases as traffic density increases (free-flow conditions) and then reaches a maximum value (congestion sets in) beyond which the flow decreases as the density increases and reaches zero when at jam density. The common notation for these relations is, respectively, $V_e\big(\bar{\rho}(x,t)\big)$ and $Q_e\big(\bar{\rho}(x,t)\big)$; the subscript "$e$" stands for *empirical* or *equilibrium*, as the relation has been commonly established by means of statistical fitting or derived from microscopic traffic flow considerations under stationarity (or equilibrium) assumptions.

Some of the more prominent fundamental relations established empirically are presented below; this is followed by a brief discussion of derivations of fundamental relations from car-following considerations. It is notable that, in recent years, researchers have proposed several new fundamental relations so as to encompass observed phenomena such as capacity drops and traffic hysteresis phenomena in addition to extensions to multi-lane traffic. The reader is referred to [25, 55, 56, 116] (and references therein) for further information on these recent extensions. It is notable that, to date, there exists no consensus as to what the "correct" shape of the fundamental relation should be. In this review, only "classical" fundamental relations shall be presented and no further discussion of recent extensions will be given.

3.4.1 Empirical Fundamental Relations

The first fundamental relation of traffic flow was fitted by Greenshields [42] as a linear relationship between mean speed and traffic density, which from (3.1) yields a quadratic relation between flow and density. These relations are expressed as follows: let v_f denote the free-flow speed of vehicles (the largest mean speed), q_{max} the maximum rate of flow, ρ_{crit} the critical density (below which we have free-flow conditions and beyond which we

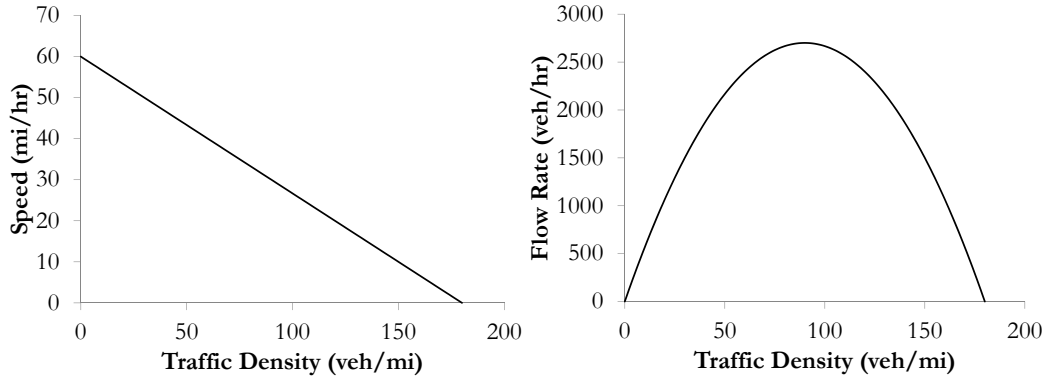
have congested traffic conditions), and ρ_{jam} the jam density (the maximum traffic density). The relations are given by:

$$V_e\big(\bar{\rho}(x,t)\big) = v_f\left(1 - \frac{\bar{\rho}(x,t)}{\rho_{jam}}\right) \tag{3.18}$$

and

$$Q_e\big(\bar{\rho}(x,t)\big) = v_f\bar{\rho}(x,t)\left(1 - \frac{\bar{\rho}(x,t)}{\rho_{jam}}\right) \tag{3.19}$$

Greenshields' fundamental relations are illustrated in Figure 3.3 for the case where $v_f = 60$ mi/hr (96.56 km/hr) and $\rho_{jam} = 180$ veh/mi (111.85 veh/km).



(a) (b)

Figure 3.3: Greenshields' fundamental relations; (a) speed-density; (b) flow-density

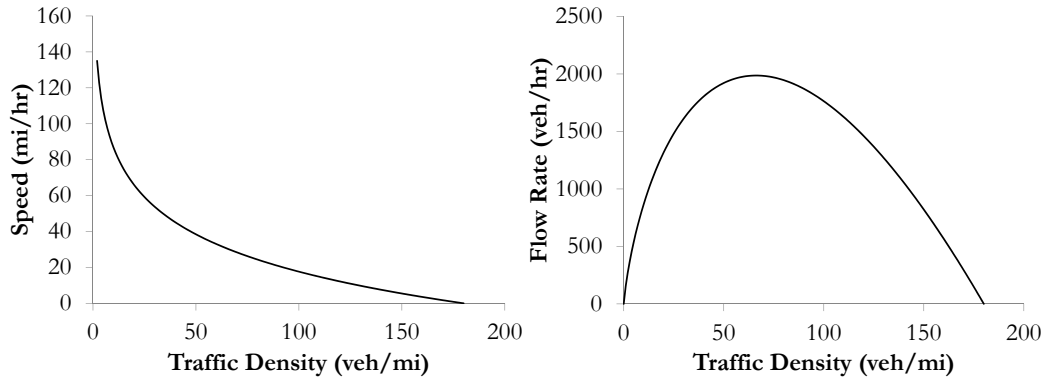
Greenshields' fundamental relations were developed using limited data (seven data points) and further empirical investigation was carried out in the late 1950's and early 1960's; Greenberg [41] proposed non-linear speed-density and flow-density models based on the fluid-like behavior assumptions of traffic flow, which he then fitted using the Lincoln Tunnel data. His proposed relations are given by:

$$V_e\big(\bar{\rho}(x,t)\big) = \beta\log\left(\frac{\rho_{jam}}{\bar{\rho}(x,t)}\right) \tag{3.20}$$

and

$$Q_e\big(\bar{\rho}(x,t)\big) = \beta\bar{\rho}(x,t)\log\left(\frac{\rho_{jam}}{\bar{\rho}(x,t)}\right), \tag{3.21}$$

where β is a parameter of the model, which determines the maximum flow rate, q_{max}. When log denotes the natural logarithm (as in Greenberg's original work), $\beta = \dfrac{q_{max}}{\rho_{crit}}$ and $\rho_{crit} = \frac{\rho_{jam}}{e}$. Greenberg's fundamental relations are illustrated in Figure 3.4 for the case where $\beta = 30$ mi/hr (48.28 km/hr) and $\rho_{jam} = 180$ veh/mi (111.85 veh/km).



(a) (b)

Figure 3.4: Greenberg's fundamental relations; (a) speed-density; (b) flow-density

Noticing the unreasonable behavior in Greenberg's speed-density model at low densities (the presence of an asymptote), Underwood [105] proposed the following relations:[1]

$$V_e\big(\bar{\rho}(x,t)\big) = \beta_1 \exp\left(\frac{-\bar{\rho}(x,t)}{\rho_{crit}}\right) - \beta_2 \tag{3.22}$$

and

$$Q_e\big(\bar{\rho}(x,t)\big) = \beta_1 \bar{\rho}(x,t) \exp\left(\frac{-\bar{\rho}(x,t)}{\rho_{crit}}\right) - \beta_2 \bar{\rho}(x,t), \tag{3.23}$$

where β_1 and β_2 are fitting parameters that are related to traffic flow parameters as follows: $\beta_2 = \beta_1 - v_f$ and $\beta_1 = \dfrac{v_f}{1 - \exp(-\rho_{jam}/\rho_{crit})}$.

Underwood's fundamental relations are illustrated in Figure 3.5 for the case where $v_f = 60$ mi/hr (96.56 km/hr), $\rho_{crit} = 30$ veh/mi (55.92 veh/km), and $\rho_{jam} = 180$ veh/mi (111.85 veh/km).

More recently, Del Castillo and Benitez [30] proposed a family of fundamental relations based on a set of constraints that ensure: (i) free-flow speed at zero density, (ii) zero speed

[1] In fact, this is the modified version of Underwood's model [105, pg. 149]; the modification ensures that $V_e(\rho_{jam}) = 0$ and $Q_e(\rho_{jam}) = 0$, where an asymptote appears in the unmodified version.

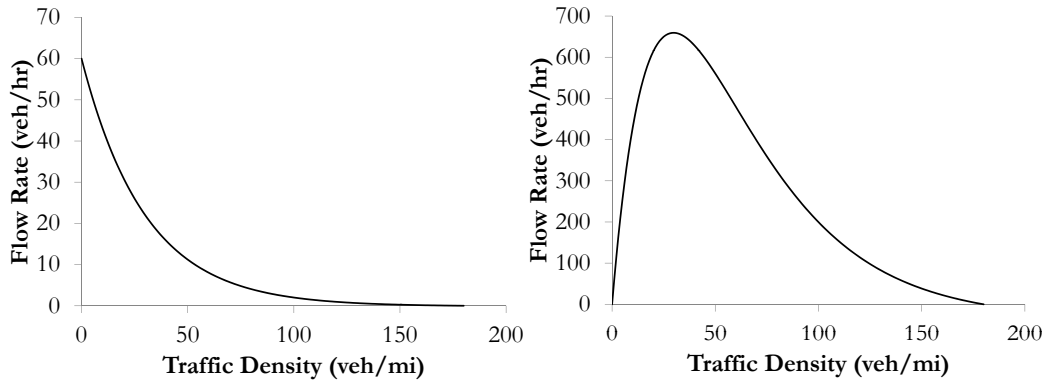


Figure 3.5: Underwood's fundamental relations; (a) speed-density; (b) flow-density

at jam density, (iii) a decreasing speed-density relation except near zero and jam densities, and (iv) a concave flow-density relation. (It is notable that the strict concavity assumption can be relaxed; allowing for non-concave flow-density relations, such as Underwood's, have been studied in [61] and their numerical solutions in [84].) As a special case, which was of particular interest to the authors, their relations include the relations previously proposed independently by Newell [76] and Franklin [35]. Hereafter, these relation shall be referred to as the Newell-Franklin relations, written as:

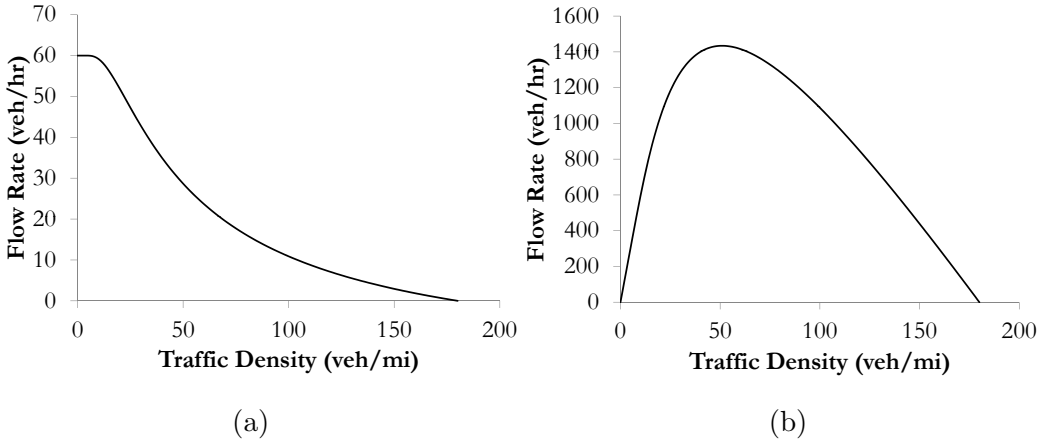
$$V_e\big(\bar{\rho}(x,t)\big) = v_f\left(1-\exp\left[\frac{|w|}{v_f}\left(1-\frac{\rho_{jam}}{\bar{\rho}(x,t)}\right)\right]\right) \tag{3.24}$$

and

$$Q_e\big(\bar{\rho}(x,t)\big) = v_f\bar{\rho}(x,t)\left(1-\exp\left[\frac{|w|}{v_f}\left(1-\frac{\rho_{jam}}{\bar{\rho}(x,t)}\right)\right]\right), \tag{3.25}$$

where w is wave speed at jam density; that is $w = \dfrac{dQ_e(\rho_{jam})}{d\bar{\rho}(x,t)}$.

The Newell-Franklin relations are depicted in Figure 3.6 for the case where $v_f = 60$ mi/hr (96.56 km/hr), $\rho_{jam} = 180$ veh/mi (111.85 veh/km), and $w = -15$ mi/hr (-24.14 km/hr).



(a) (b)

Figure 3.6: Newell-Franklin fundamental relations; (a) speed-density; (b) flow-density

3.4.2 Fundamental Relations Derived from Car-Following

Early derivations of the fundamental diagram from car-following dynamics were carried out in order to investigate the aggregate-level behavior of proposed car-following rules and for purposes of validation against real world data [32, 38, 76, 88]. It wasn't long, however, before the fundamental diagram induced by car-following theory itself became an object of interest, and researchers proposed car-following models that reproduce known empirical relations. Pipes [89] proposed the general car-following equation:

$$a_j(t) = \beta_1 \frac{v_{j-1}(t) - v_j(t)}{\big(x_{j-1}(t) - x_j(t)\big)^{\beta_2}}, \tag{3.26}$$

where $j-1$ and j are the indices of a leading vehicle and a following vehicle, respectively; (see Figure 3.1). β_1 and β_2 are model parameters, $a_j(t)$ is the acceleration/deceleration of the follower vehicle, j, at time t, and $v_{j-1}(t)$ and $v_j(t)$ are the speeds of the leader and the follower at time t. Roughly, the model says that the followers acceleration/deceleration is directly proportional to the speed difference between the two vehicles and inversely proportional to the spacing between them with β_2 representing the follower's sensitivity to spacing. Different choices for β_2 lead to different fundamental relations. For example, suppose we

choose $\beta_2 = 2$, then[2]

$$\begin{aligned}\frac{dv_j(t)}{dt} &= \beta_1 \frac{v_{j-1}(t) - v_j(t)}{\big(x_{j-1}(t) - x_j(t)\big)^2} \\ &= \beta_1 \Big(x_{j-1}(t) - x_j(t)\Big)^{-2} \frac{d}{dt}\Big(x_{j-1}(t) - x_j(t)\Big) \\ &= \beta_1 \Big(s_j(t)\Big)^{-2} \frac{ds_j(t)}{dt} \end{aligned} \tag{3.27}$$

Multiplying both sides by dt and integrating:

$$\int dv_j(t) = \beta_1 \int \frac{ds_j(t)}{\big(s_j(t)\big)^2}, \tag{3.28}$$

we get the following relation:

$$v_j(t) = \beta_1 \left(\frac{-1}{s_j(t)} + \gamma\right), \tag{3.29}$$

where γ is a constant of integration. Assuming $v_j(t)$ depends neither on j nor on t, but on $\bar{\rho}(x,t) = \big(s_j(t)\big)^{-1}$ and model parameters, (3.29) may be interpreted as the stationary speed-density relation:

$$V_e\big(\bar{\rho}(x,t)\big) = \beta_1 \left(-\bar{\rho}(x,t) + \gamma\right) \tag{3.30}$$

Using the boundary conditions: $V_e\big(0\big) = v_f$ and $V_e\big(\rho_{jam}\big) = 0$, β_1 and γ are such that we get the following relation:

$$V_e\big(\bar{\rho}(x,t)\big) = v_f \left(1 - \frac{\bar{\rho}(x,t)}{\rho_{jam}}\right), \tag{3.31}$$

which is Greenshields' speed-density relation.

Likewise, setting $\beta_2 = 1$, we get the following:

$$\begin{aligned}\frac{dv_j(t)}{dt} &= \beta_1 \frac{v_{j-1}(t) - v_j(t)}{x_{j-1}(t) - x_j(t)} \\ &= \beta_1 \big(s_j(t)\big)^{-1} \frac{ds_j(t)}{dt}, \end{aligned} \tag{3.32}$$

and upon multiplying by dt, integrating, and assuming stationary conditions, we get:

$$V_e\big(\bar{\rho}(x,t)\big) = \beta_1 \log\left(\frac{1}{\bar{\rho}(x,t)}\right) + \gamma, \tag{3.33}$$

---

[2] Recall that $a_j(t) = \dfrac{dv_j(t)}{dt}$, $v_j(t) = \dfrac{ds_j(t)}{dt}$, and $s_j(t) = x_{j-1}(t) - x_j(t)$.

which with the boundary condition $V_e(\rho_{jam}) = 0$ yields Greenberg's speed-density relation.

A recent, more widely used, fundamental relation is based on the simplified car-following model of Newell [82], which hypothesizes a linear relationship between speed and spacing, written as:

$$s_j(t) = \beta_{j,2} + \beta_{j,1} v_j(t) \qquad 0 \le v_j(t) < v_f, \tag{3.34}$$

The linear relation is illustrated in Figure 3.7a.

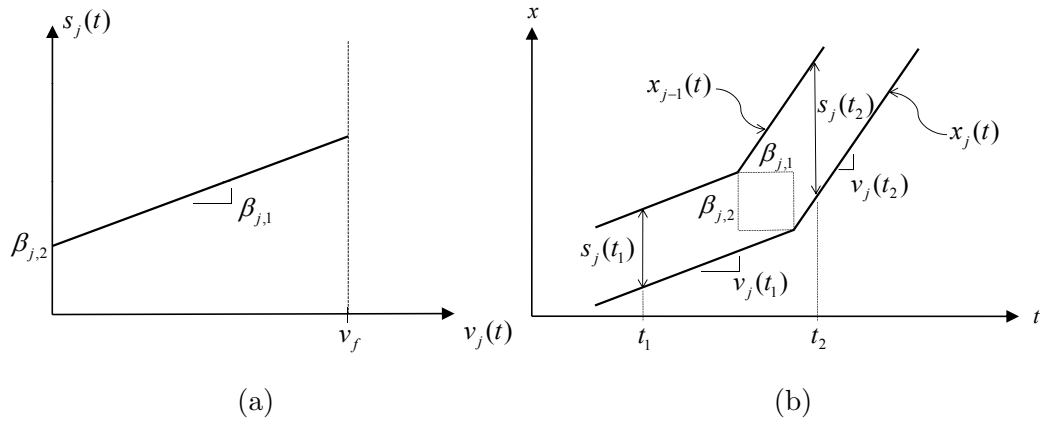


Figure 3.7: Newell's simplified car-following model [82] (reproduced); (a) the linear speed-spacing relation for a single vehicle, (b) linear trajectories

In Newell's simplified model, $\beta_{j,1}$ can be loosely interpreted as a minimum time headway between the follower j and their leader, $j-1$, while $\beta_{j,2}$ may be interpreted as the minimum spacing between vehicles j and $j-1$ when the vehicles are not moving. The parameters would vary from one driver to the other. Furthermore, the theory only applies when the speeds are less than the free-flow speed, since otherwise vehicles move unrestricted by their leaders. The linear relationship also implies piecewise linear trajectories of vehicles, where the follower's trajectory is simply a translation of the leader's trajectory an amount $\beta_{j,1}$ in time and $\beta_{j,2}$ in space. It is notable that the simplified model was empirically validated in [2], while the resulting piecewise linear flow-density relation was empirically verified in [20].

To derive a fundamental relation from the simplified car-following model, one assumes stationary conditions (i.e., independence of j and t), so that $\beta_{2,j} \equiv \beta_2$ and $\beta_{1,j} \equiv \beta_1$ (for

all j). Then, from (3.34):

$$V_e\big(\bar{\rho}(x,t)\big) = \frac{1}{\beta_1}\left(\frac{1}{\bar{\rho}(x,t)} - \beta_2\right) \qquad \rho_{crit} \le \bar{\rho}(x,t) \le \rho_{jam}, \tag{3.35}$$

where the critical density is $\rho_{crit} = \dfrac{q_{max}}{v_f}$. Furthermore, $\beta_2 = \dfrac{1}{\rho_{jam}}$ since $V_e(\rho_{jam}) = 0$ and

$$\frac{1}{\beta_1} = \left(\frac{q_{max}}{\rho_{jam} - \rho_{crit}}\right)\rho_{jam} \equiv |w|\rho_{jam}, \tag{3.36}$$

since $Q_e(\rho_{crit}) = q_{max}$. Here, as in the Newell-Franklin model, w is interpreted as a constant backward wave speed.

For sub-critical traffic densities (i.e., free-flow conditions), we have that $V_e\big(\bar{\rho}(x,t)\big) = v_f$ and, consequently, $Q_e\big(\bar{\rho}(x,t)\big) = v_f\bar{\rho}(x,t)$. Finally, we have the following fundamental relations due to Newell:

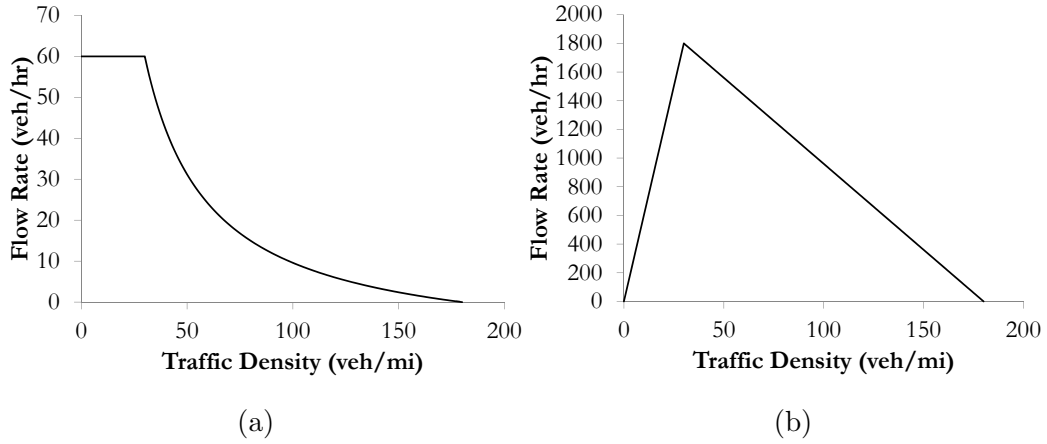
$$V_e\big(\bar{\rho}(x,t)\big) = \begin{cases} v_f & \text{if } 0 \le \bar{\rho}(x,t) < \rho_{crit} \\ |w|\left(\dfrac{\rho_{jam}}{\bar{\rho}(x,t)} - 1\right) & \text{if } \rho_{crit} \le \bar{\rho}(x,t) \le \rho_{jam} \end{cases}, \tag{3.37}$$

and

$$Q_e\big(\bar{\rho}(x,t)\big) = \begin{cases} v_f\bar{\rho}(x,t) & \text{if } 0 \le \bar{\rho}(x,t) < \rho_{crit} \\ |w|\big(\rho_{jam} - \bar{\rho}(x,t)\big) & \text{if } \rho_{crit} \le \bar{\rho}(x,t) \le \rho_{jam} \end{cases} \tag{3.38}$$

Newell's simplified relations are depicted in Figure 3.8 with parameters $v_f = 60$ mi/hr (96.56 km/hr), $\rho_{crit} = 30$ veh/mi (18.64 veh/km), $\rho_{jam} = 180$ veh/mi (111.85 veh/km), and $w = -12$ mi/hr (-19.31 km/hr).

Newell's simplified flow-density relation has also been extended to a trapezoidal shape to accommodate discharge capacities which are associated with a range of traffic densities. Such discharge capacities typically arise along urban arterials, where queued vehicles, regardless of queue size, tend to discharge at saturation flow rates. The trapezoidal flow-density relation is depicted in Figure 3.9 with parameters $v_f = 60$ mi/hr (96.56 km/hr), $\rho_{jam} = 180$ veh/mi (111.85 veh/km), $w = -12$ mi/hr (-19.31 km/hr), and $q_{max} = 1400$ veh/hr.



(a) (b)

Figure 3.8: Newell's simplified fundamental relations; (a) speed-density; (b) flow-density

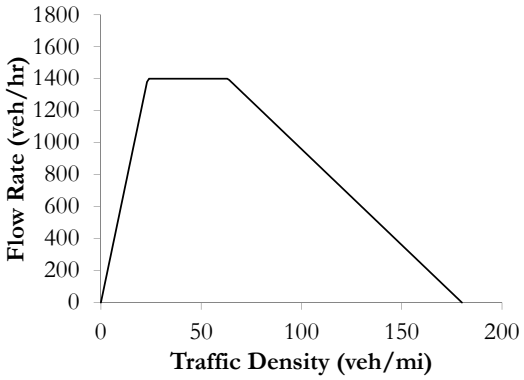


Figure 3.9: Trapezoidal flow-density relation; (a) speed-density; (b) flow-density

3.5 Shockwaves, Rarefaction, and the Riemann Problem

A solution to the LWR model can, in essence, be thought of as a formula $\bar{\rho}(x,t)$ where one plugs in x and t and the formula gives the value of the mean traffic density at position x and time t. As the LWR model is ill-posed, such a formula cannot be obtained in general, but may be obtained by piecing together solutions to simpler problems. The nature of the solution depends on two things: the fundamental relation and the traffic densities at an initial time, hereafter referred to as the initial data. The latter is a prescribed relation, written as: $\bar{\rho}(x,0) \equiv \rho_0(x)$, where $\rho_0(x)$ is a given function of x. To illustrate, suppose

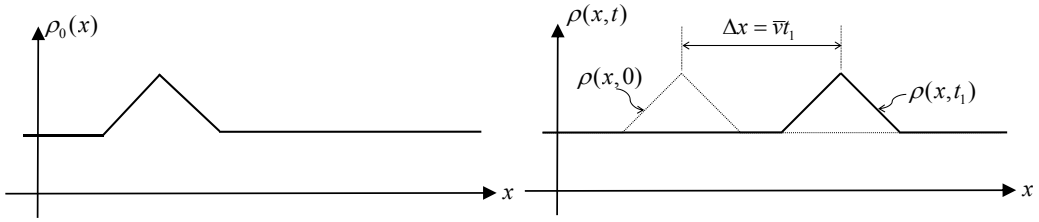
$\rho_0(x)$ is some continuous function of x and assume that the flow-density relation is given by $Q_e\big(\bar{\rho}(x,t)\big) = \bar{v}\bar{\rho}(x,t)$, where $\bar{v} > 0$ is a constant. Such a scenario would describe free-flow traffic. The conservation equation (3.16) is then written as:

$$\frac{\partial}{\partial t}\bar{\rho}(x,t) + \bar{v}\frac{\partial}{\partial x}\bar{\rho}(x,t) = 0, \tag{3.39}$$

This version of the problem, known as the *linear advection equation*, is well-posed and has the following solution:

$$\bar{\rho}(x,t) = \rho_0(x - \bar{v}t) \tag{3.40}$$

That is, the initial traffic densities remain unchanged, but shift in space at a speed of $\bar{v}$ over time. Derivation of this solution is given in Appendix B, Section B.1. To illustrate the properties of the solution, suppose the initial traffic densities are given by Figure 3.10a, then the traffic density profile at some time $t_1 > 0$ in the future is illustrated in Figure 3.10b.



(a) (b)

Figure 3.10: Example linear advection problem solution; (a) initial traffic density profile $\rho_0(x)$; (b) traffic density profile at time $t_1 > 0$

An alternative way of visualizing the solution is by looking at the characteristic lines of the problem. In the context of conservation of traffic flow, characteristic lines are lines in the time-space diagram along which traffic densities are constant. Figure 3.11 illustrates how one would obtain the traffic density at position x_1 at time t_1 by following the characteristic line -on which the point (x_1, t_1) lies- to the origin, where the traffic density is known (i.e., determine the corresponding $x_0 = x_1 - \bar{v}t_1$). The slopes of the characteristic lines in Figure 3.11 are $1/\bar{v}$ rather than $\bar{v}$ due to the orientation of the axes.

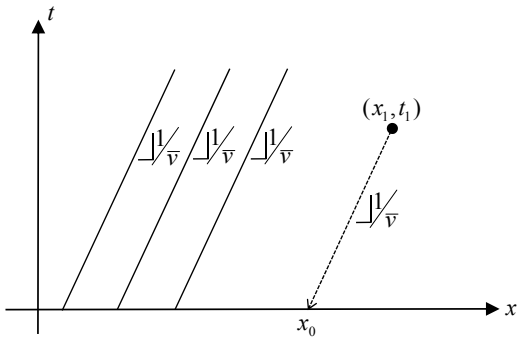


Figure 3.11: Characteristic lines for the linear advection equation

In general the constant slopes of the characteristic lines throughout the time-space diagram in Figure 3.11 indicate that traffic densities shift at a constant speed over time, preserving their initial shape. When instead of a constant speed we have speeds that depend on traffic density, the slopes of the characteristic lines are no longer equal. For any flow-density relation, $Q_e\big(\bar{\rho}(x,t)\big)$, the conservation equation to be solved is:

$$\frac{\partial}{\partial t}\bar{\rho}(x,t) + Q_e'\big(\bar{\rho}(x,t)\big)\frac{\partial}{\partial x}\bar{\rho}(x,t) = 0, \tag{3.41}$$

where $Q_e' = \frac{dQ_e}{d\bar{\rho}}$, the slope of the flow-density relation, is referred to as *the characteristic speed*. The slopes of the characteristic lines are $1/Q_e'\big(\bar{\rho}(x,t)\big)$ and now depend on the traffic density. Note that the characteristic speeds, while they may coincide with the speed of traffic, depending on traffic density and the shape of the flow-density relation, generally differ from the speed of traffic. This is illustrated in Figure 3.12.

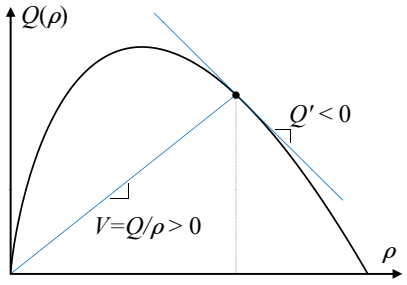


Figure 3.12: The characteristic speed (Q') vs. traffic speed (V) for the same traffic density

For this non-linear case, proceed in the same way as in the linear case. Take, for

instance, the initial data shown in Figure 3.10a and assume that the traffic densities climb to a value that exceeds the critical density, ρ_{crit}, and then descends as shown in Figure 3.14a. As the traffic density increases, the characteristic speed decreases, the slope of the corresponding characteristic line increases, and vice versa. What we then have is the intersecting characteristic lines depicted in Figure 3.13.

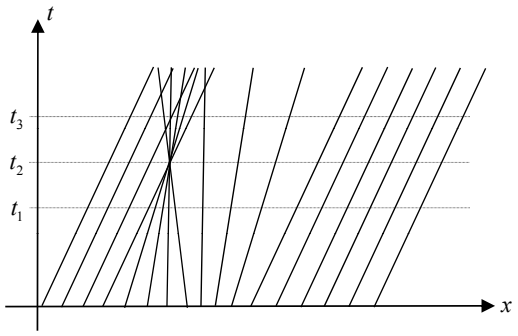


Figure 3.13: Intersecting characteristic lines

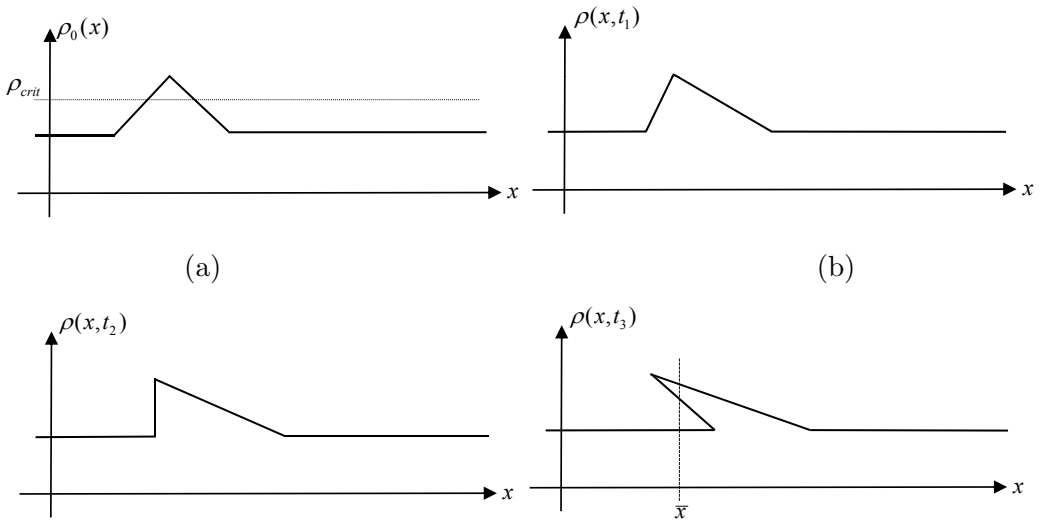


Figure 3.14: Traffic densities when characteristics intersect; (a) $t = 0$: initial data; (b) $t = t_1$: before the characteristics intersect; (c) $t = t_2$: time of intersection; (d) $t = t_3$: after intersection of characteristic lines

The traffic density profile implied by the characteristic lines in Figure 3.13 at times $t = 0$, t_1, t_2, and t_3 are shown in Figure 3.14. Figure 3.14b illustrates the traffic density profile at time t_1, before the characteristic lines intersect. In this figure, we see that increasing traffic densities tend to "compress" with time, while decreasing traffic densities tend to "rarefy" (or spread). That is, on the compression side, the traffic densities increase faster in space as faster vehicles in the upstream will have joined the slower vehicles in the downstream. On the other hand, on the rarefaction side at time t_1, slower vehicles accelerate to the speeds of the faster vehicles further downstream. At time t_2, the time at which the characteristic lines intersect, Figure 3.14c illustrates how traffic densities will have compressed to a point where we see an instantaneous jump in the traffic density, followed by a wider rarefaction. Beyond time t_2, we see the deficiency of the theory. When characteristic lines intersect, we obtain multi-valued traffic densities at position $\bar{x}$ as illustrated in Figure 3.14d for time t_3. From a physical standpoint, one would never observe different traffic densities at the same position in space. From a theoretical standpoint, this implies that the derivatives of traffic density with respect to time do not exist when we have multi-valued traffic densities in space, which violates the assumption that was made in deriving the differential form of the conservation law, equations (3.16) and (3.17). However, the integral forms, (3.9) and (3.12) remain valid.

It is notable that the approach used to solve the linear advection equation, when applied to the non-linear problem (3.41), delivers a reasonable solution up to time t_2, the time at which we have a jump (a discontinuity) in the density profile. At time t_2, to the left and right of the discontinuity, the solution is also reasonable. After time t_2, the solution given by the method of characteristics is no longer useful. The question, then, is: what happens **at** the discontinuity at time t_2? To answer this question, we need to use an integral form of the conservation equation. Let $\hat{x}(t)$ denote the position of a discontinuity at time t. Using the first integral form of the conservation law, for any (short) section of road $[x_1, x_2]$ which contains $\hat{x}(t)$, the conservation equation may be written as:

$$\frac{d}{dt}\int_{x_1}^{\hat{x}(t)-} \bar{\rho}(x,t)dx + \frac{d}{dt}\int_{\hat{x}(t)+}^{x_2} \bar{\rho}(x,t)dx = Q_e\big(\bar{\rho}(x_1,t)\big) - Q_e\big(\bar{\rho}(x_2,t)\big) \tag{3.42}$$

Applying the Leibniz rule (see Appendix A, Theorem A.2.3) to the LHS of (3.42), we get:

$$\bar{\rho}(\hat{x}(t)-,t)\frac{d\hat{x}(t)}{dt}+\int_{x_1}^{\hat{x}(t)-}\frac{\partial}{\partial t}\bar{\rho}(x,t)dx-\bar{\rho}(\hat{x}(t)+,t)\frac{d\hat{x}(t)}{dt}+\int_{\hat{x}(t)+}^{x_2}\frac{\partial}{\partial t}\bar{\rho}(x,t)dx \tag{3.43}$$

Now, to investigate what happens **at** the discontinuity, simply let the points x_1 and x_2 approach $\hat{x}(t)$ from below and from above, respectively. Since $\bar{\rho}(x,t)$ is differentiable away from the discontinuity, as $x_1 \nearrow \hat{x}(t)-$ and $x_2 \searrow \hat{x}(t)+$, the integrals in (3.43) go to zero. Combining this with (3.42), we get the following equation for the behavior of the discontinuity:

$$\bar{\rho}(\hat{x}(t)-,t)\frac{d\hat{x}(t)}{dt}-\bar{\rho}(\hat{x}(t)+,t)\frac{d\hat{x}(t)}{dt}=Q_e\big(\bar{\rho}(\hat{x}(t)-,t)\big)-Q_e\big(\bar{\rho}(\hat{x}(t)+,t)\big), \tag{3.44}$$

and, consequently,

$$\frac{d\hat{x}(t)}{dt}=\frac{Q_e\big(\bar{\rho}(\hat{x}(t)-,t)\big)-Q_e\big(\bar{\rho}(\hat{x}(t)+,t)\big)}{\bar{\rho}(\hat{x}(t)-,t)-\bar{\rho}(\hat{x}(t)+,t)}, \tag{3.45}$$

so that as long as one can track the location of the discontinuity, the solution to the left and to the right of it are known. The moving discontinuity is referred to as a *shockwave*. The speed of the shockwave at time t is given by $\frac{d\hat{x}(t)}{dt}$. Physically speaking, a shockwave represents the build-up of a queue of vehicles.

Since any scenario that involves compression of traffic densities will eventually develop such a discontinuity, it is informative to consider problems where the initial data includes a discontinuity. The position of a discontinuity in traffic flow represents the position of the end of a queue along the road. Such initial data are depicted in Figure 3.15, where ρ_l denotes the traffic density to the left of the discontinuity and ρ_r denotes the traffic density to the right of the discontinuity. Note that we have the two separate cases: (i) $\rho_l < \rho_r$ and (ii) $\rho_l > \rho_r$.

Despite the simplicity of such a problem, solutions of more sophisticated problems can be obtained by analyzing this case. A nonlinear conservation equation, such as (3.41), combined with either of the two initial density profiles shown in Figure 3.15 is referred to as *the Riemann problem*. In the case of a strictly concave flow-density relation, discontinuities that compress always occur when $\rho_l < \rho_r$, while rarefaction always occurs when $\rho_l > \rho_r$.

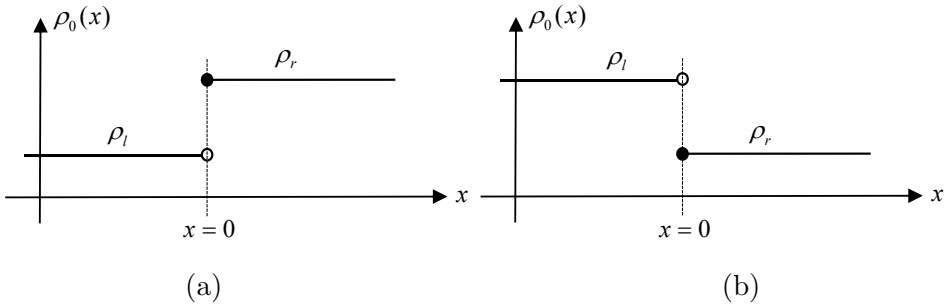


Figure 3.15: Discontinuous initial data; (a) $\rho_l < \rho_r$; (b) $\rho_l > \rho_r$

Nonetheless, the theory derived for the case of a concave flow-density relation is only slightly modified (by only changing the cases) in the more general setting. We shall, therefore, focus on the case of strictly concave flow-density relations in what follows.

For the initial data depicted in Figure 3.15a, since the traffic densities are constant on either side of the discontinuity, the slopes of the characteristic lines on either side of the discontinuity are constant and the shockwave speed is also constant and calculated, using (3.46), as:

$$\frac{d\hat{x}(t)}{dt} = \frac{Q_e(\rho_l) - Q_e(\rho_r)}{\rho_l - \rho_r} \equiv v_s \tag{3.46}$$

This formula for the speed of a shockwave is known as the *Rankine-Hugoniot jump condition* (R-H), which is a well-known formula in the traffic flow literature. It is notable, however, that the R-H formula was not originally developed for traffic flow applications. Now, note that the RHS, v_s, is constant; integrating both sides of (3.46), one obtains the position of the shockwave at any time $t > 0$ via:

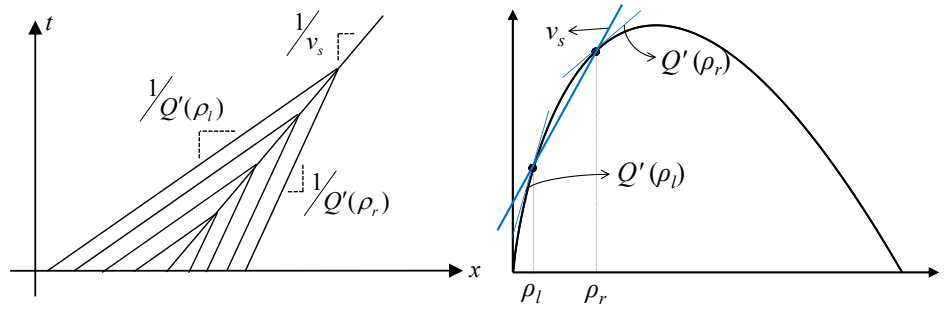
$$\hat{x}(t) = v_s t \tag{3.47}$$

The solution is, thus, obtained by observing that at any time t, for any position x that lies to the left of $\hat{x}(t)$ (upstream of the shock front), the traffic density is $\bar{\rho}(x,t) = \rho_l$. Likewise, for any position x that lies to the right of $\hat{x}(t)$ (downstream of the shock front), the traffic density is $\bar{\rho}(x,t) = \rho_r$. We have, thus, obtained the following formula for the

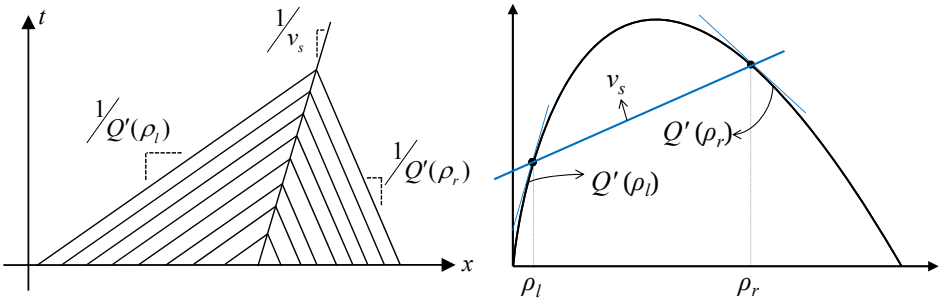
solution of the Riemann problem with initial data, $\rho_l < \rho_r$, shown in Figure 3.15a:

$$\bar{\rho}(x,t) = \begin{cases} \rho_l & \text{if } x < v_s t \\ \rho_r & \text{if } x \geq v_s t \end{cases} \tag{3.48}$$

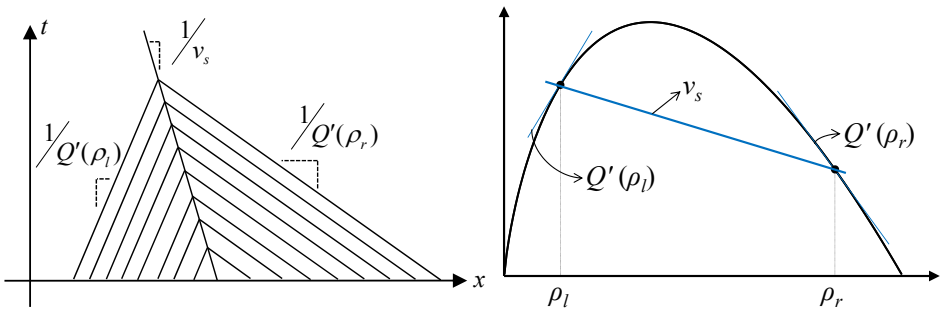
For any concave flow-density relation, the solution (3.48) involves a shockwave, which can be thought of as a disturbance that propagates through traffic at a speed (possibly negative) that depends on the values of ρ_l and ρ_r. Figure 3.16 illustrates the characteristic lines for the four cases: (a) $\rho_l < \rho_r < \rho_{crit}$, (b) $\rho_l < \rho_{crit} < \rho_r$ and $Q_e(\rho_l) < Q_e(\rho_r)$, (c) $\rho_l < \rho_{crit} < \rho_r$ and $Q_e(\rho_l) \geq Q_e(\rho_r)$, and (d) $\rho_{crit} < \rho_l < \rho_r$. Included in the figure are the shockwave speeds and the characteristic slopes obtained from the fundamental diagram. These shockwave solutions of the Riemann problem, in essence, say that the shape of the initial traffic density profile is preserved over time (as in the linear advection equation case) with the discontinuity traveling at the speed v_s.



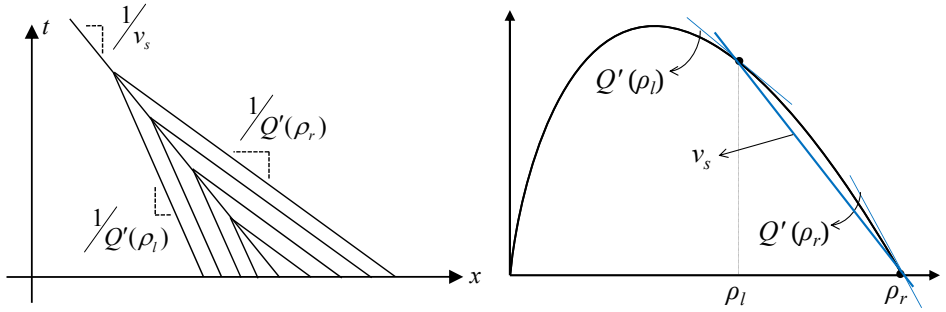
(a)



(b)



(c)



(d)

Figure 3.16: Riemann problem shockwave solutions; (a) $\rho_l < \rho_r < \rho_{crit}$; (b) $\rho_l < \rho_{crit} < \rho_r$ and $Q_e(\rho_l) < Q_e(\rho_r)$; (c) $\rho_l < \rho_{crit} < \rho_r$ and $Q_e(\rho_l) \geq Q_e(\rho_r)$; (d) $\rho_{crit} < \rho_l < \rho_r$

For the initial data $\rho_l > \rho_r$ depicted in Figure 3.15b, the solution is obtained by perturbing the shape of the initial profile a negligible amount and then investigating the behavior of the problem with the perturbed initial data. The perturbation is, in essence, an approximation of the discontinuous initial data by a continuous function, $\rho_0^\delta(x)$ as depicted, for instance, in Figure 3.17.

An example form for the function $\rho_0^\delta(x)$, which would approximated $\rho_0(x)$ well is:

$$\rho_0^\delta(x) = (\rho_l - \rho_r)\frac{e^{-\delta x}}{1 + e^{-\delta x}} + \rho_r \tag{3.49}$$

As the parameter δ in (3.49) is increased, $\rho_0^\delta(x)$ begins to look more and more like a step function. In fact, it can be easily shown that for any tolerance $\epsilon > 0$, one can find a value for the parameter δ in (3.49) so that the distance between the two functions $\rho_0(x)$

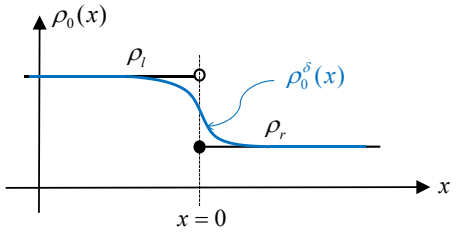


Figure 3.17: Approximation of discontinuous initial data

and $\rho_0^\delta(x)$ is less than ϵ (see Appendix B, Section B.2). With the perturbed initial data, the solution is simply given by the rarefying traffic densities presented above, which is written as

$$\bar{\rho}(x,t) = \begin{cases} \rho_l & \text{if } \frac{x}{t} < Q_e'(\rho_l) \\ \rho_{int} & \text{if } Q_e'(\rho_l) \leq \frac{x}{t} \leq Q_e'(\rho_r) \\ \rho_r & \text{if } \frac{x}{t} > Q_e'(\rho_r) \end{cases}, \tag{3.50}$$

where ρ_{int} is an intermediate traffic density obtained by solving the equation:

$$Q_e'(\rho_{int}) = \frac{x}{t} \tag{3.51}$$

It is notable that perturbing the initial data for the shockwave case, $\rho_l < \rho_r$, the problem will quickly develop a discontinuity as was shown in Figure 3.14. Therefore, an analysis based on approximating the initial data is not appropriate in this case.

Examples.

(i) If $Q_e(\bar{\rho})$ is Greenshields' fundamental relationship, we have that

$$\rho_{int} = \frac{\rho_{jam}}{2v_f}\left(v_f - \frac{x}{t}\right)$$

(ii) If $Q_e(\bar{\rho})$ is Greenberg's fundamental relationship, we have that

$$\rho_{int} = \rho_{jam} \exp\left(\frac{-\rho_{crit}}{q_{max}}\frac{x}{t} - 1\right)$$

Figure 3.18 depicts the characteristic lines for the three cases: (a) $\rho_r < \rho_l < \rho_{crit}$, (b) $\rho_r < \rho_{crit} < \rho_l$, and (c) $\rho_{crit} < \rho_r < \rho_l$. These solutions are referred to as *rarefaction fans* due to their shapes.

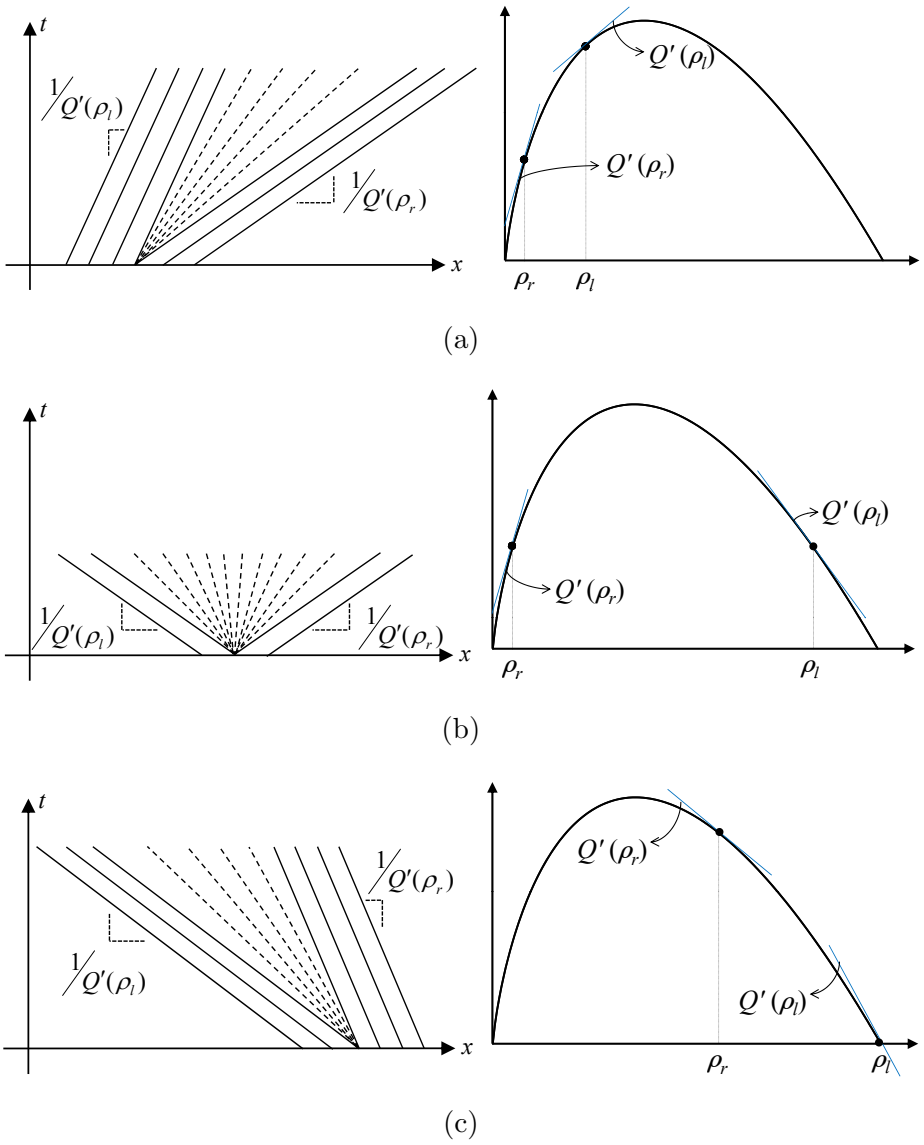
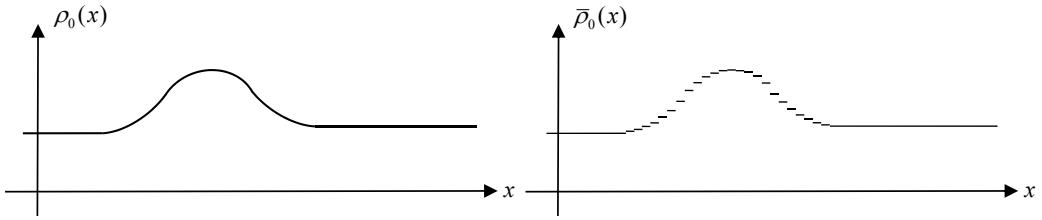


Figure 3.18: Riemann problem rarefaction fan solutions; (a) $\rho_r < \rho_l < \rho_{crit}$; (b) $\rho_r < \rho_{crit} < \rho_l$; (c) $\rho_{crit} < \rho_r < \rho_l$

3.6 Numerical Solutions: Discrete Dynamics

In the previous section, we obtained formulas for solving the Riemann problem. While limited to initial traffic densities that are far from being general enough to apply to real-world traffic flow problems, the Riemann problem can be effectively used to design numerical

schemes that apply in general settings. The method described in this section is referred to as the *Godunov scheme* due to Sergei Godunov, [39]. It's applications in traffic flow were investigated in [58]. The method solves the conservation equation over discrete space and time intervals of length Δx and Δt, respectively, by computing averages of traffic densities in the discrete space intervals in discrete time steps. One begins by approximating the initial data, $\rho_0(x)$, over the discrete space intervals by a step function, $\bar{\rho}_0(x)$, as illustrated in Figure 3.19.



(a) (b)

Figure 3.19: Approximating the initial data; (a) $\rho_0(x)$; (b) $\bar{\rho}_0(x)$

Let $\bar{\rho}^k(x)$ denote the mean traffic density in *cell* $\left[x-\frac{\Delta x}{2}, x+\frac{\Delta x}{2}\right]$ at time instance $k\Delta t$; that is:

$$\bar{\rho}^k(x) = \frac{1}{\Delta x}\int_{x-\frac{\Delta x}{2}}^{x+\frac{\Delta x}{2}} \bar{\rho}(u, k\Delta t)du \tag{3.52}$$

For each discrete space-time interval of size $\Delta x \times \Delta t$, we have the following conservation equation (second integral form):

$$\begin{aligned}\int_{x-\frac{\Delta x}{2}}^{x+\frac{\Delta x}{2}} \bar{\rho}\big(u, (k+1)\Delta t\big)du \;\; &= \;\; \int_{x-\frac{\Delta x}{2}}^{x+\frac{\Delta x}{2}} \bar{\rho}(u, k\Delta t)du + \int_{k\Delta t}^{(k+1)\Delta t} Q_e\left(\bar{\rho}\left(x-\frac{\Delta x}{2}, u\right)\right)du \\ &\quad - \int_{k\Delta t}^{(k+1)\Delta t} Q_e\left(\bar{\rho}\left(x+\frac{\Delta x}{2}, u\right)\right)du \end{aligned} \tag{3.53}$$

Dividing both sides of (3.53) by Δx, the LHS becomes $\bar{\rho}^{k+1}(x)$ and the first term on the RHS is $\bar{\rho}^k(x)$. It remains to obtain a formula for the second and third integrals on the RHS. If Δt is chosen carefully (sufficiently small), so that shockwaves and rarefaction fans that are initiated at a cell boundary $x-\frac{\Delta x}{2}$ at time $k\Delta t$ do not reach the following cell boundary $x+\frac{\Delta x}{2}$ (or the preceding cell boundary $x-\Delta t$) before time $(k+1)\Delta t$, then the second and

third integrals on the RHS of (3.53) are easy to calculate. The reason for this is as follows: take the second integral on the RHS; since neither a shockwave nor a rarefaction fan from the upstream (or the downstream) will cross the cell boundary at position $x-\frac{\Delta x}{2}$ during the time interval $[k\Delta t,(k+1)\Delta t]$, the traffic density at this position remains constant over the time interval. Call this constant traffic density $\bar{\rho}^{k,*}\left(x-\frac{\Delta x}{2}\right)$. We have, since $\bar{\rho}^{k,*}\left(x-\frac{\Delta x}{2}\right)$ is constant, that

$$\int_{k\Delta t}^{(k+1)\Delta t} Q_e\left(\bar{\rho}\left(x-\frac{\Delta x}{2},u\right)\right)du = Q_e\left(\bar{\rho}^{k,*}\left(x-\frac{\Delta x}{2}\right)\right)\Delta t, \tag{3.54}$$

and all that remains is to determine $\bar{\rho}^{k,*}\left(x-\frac{\Delta x}{2}\right)$. A rule that is typically used to determine an appropriate Δt is the Courant, Friedrichs, and Lewy (CFL) condition [59]:

$$\frac{\Delta t}{\Delta x}\max_{0\le\bar{\rho}\le\rho_{jam}}\left|Q_e'(\bar{\rho})\right| \le 1 \tag{3.55}$$

Since this ensures that shockwaves and rarefaction fans do not cross cell boundaries during discrete time intervals, we actually have Riemann problems with their (potential) discontinuities at each of the cell boundaries, which do not interfere with one another; that is, they may be solved independently at each time step with each solution serving as the initial data for the next time step. Take an arbitrary cell centered at x and let's take the upstream cell boundary located at position $x-\frac{\Delta x}{2}$; if $\rho^k(x-\Delta x)<\rho^k(x)$, a shockwave would be initiated at the boundary and we would have one of the four cases depicted in Figure 3.17:

(a) $\rho^k(x-\Delta x)<\rho^k(x)<\rho_{crit}$: $\bar{\rho}^{k,*}\left(x-\frac{\Delta x}{2}\right)=\rho^k(x-\Delta x)$

(b) $\rho^k(x-\Delta x)<\rho_{crit}<\rho^k(x)$ and $Q_e\big(\rho^k(x-\Delta x)\big)<Q_e\big(\rho^k(x)\big)$: $\bar{\rho}^{k,*}\left(x-\frac{\Delta x}{2}\right)=\rho^k(x-\Delta x)$

(c) $\rho^k(x-\Delta x)<\rho_{crit}<\rho^k(x)$ and $Q_e\big(\rho^k(x-\Delta x)\big)\ge Q_e\big(\rho^k(x)\big)$: $\bar{\rho}^{k,*}\left(x-\frac{\Delta x}{2}\right)=\rho^k(x)$

(d) $\rho_{crit}<\rho^k(x-\Delta x)<\rho^k(x)$: $\bar{\rho}^{k,*}\left(x-\frac{\Delta x}{2}\right)=\rho^k(x)$

Likewise, if $\rho^k(x-\Delta x)>\rho^k(x)$, a rarefaction fan would be initiated at the boundary and we would have one of the three cases depicted in Figure 3.18:

(a) $\rho^k(x) < \rho^k(x-\Delta x) < \rho_{crit} : \bar{\rho}^{k,*}\left(x-\frac{\Delta x}{2}\right) = \rho^k(x-\Delta x)$

(b) $\rho^k(x) < \rho_{crit} < \rho^k(x-\Delta x) : \bar{\rho}^{k,*}\left(x-\frac{\Delta x}{2}\right) = \rho_{crit}$

(c) $\rho_{crit} < \rho^k(x) < \rho^k(x-\Delta x) : \bar{\rho}^{k,*}\left(x-\frac{\Delta x}{2}\right) = \rho^k(x)$

We have just determined $\bar{\rho}^{k,*}\left(x-\frac{\Delta x}{2}\right)$ from the values of $\rho^k(x-\Delta x)$ and $\rho^k(x)$. To emphasize this, we write:

$$Q_e\left(\bar{\rho}^{k,*}\left(x-\frac{\Delta x}{2}\right)\right) \equiv \lambda\big(\rho^k(x-\Delta x), \rho^k(x)\big), \tag{3.56}$$

where $\lambda\big(\rho^k(x-\Delta x), \rho^k(x)\big)$ is referred to as the *numerical flux function* and can be calculated from Table 3.1, which summarizes the cases listed above.

Table 3.1: Numerical flux

$\lambda\big(\rho^k(x-\Delta x), \rho^k(x)\big)$	$\rho^k(x) \le \rho_{crit}$	$\rho^k(x) > \rho_{crit}$
$\rho^k(x-\Delta x) \le \rho_{crit}$	$Q_e\big(\rho^k(x-\Delta x)\big)$	$\min\Big\{Q_e\big(\rho^k(x-\Delta x)\big), Q_e\big(\rho^k(x)\big)\Big\}$
$\rho^k(x-\Delta x) > \rho_{crit}$	q_{max}	$Q_e\big(\rho^k(x)\big)$

Now, the conservation equation, (3.53), may simply be written (after dividing by Δx) as:

$$\bar{\rho}^{k+1}(x) = \bar{\rho}^k(x) + \frac{\Delta t}{\Delta x}\Big(\lambda\big(\rho^k(x-\Delta x), \rho^k(x)\big) - \lambda\big(\rho^k(x), \rho^k(x+\Delta x)\big)\Big), \tag{3.57}$$

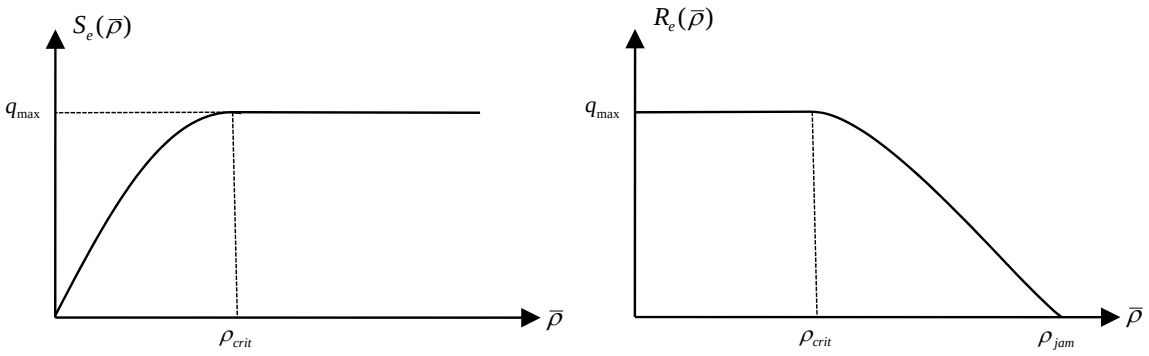
and starting from $\bar{\rho}_0(x)$, using (3.57) along with Table 3.1, traffic densities are computed recursively, moving forward in time.

Alternative choices for the numerical flux function equivalent to the one shown in Table 3.1 exist [16, 58–60, 84]. The most general numerical flux function, which applies to concave and non-concave flow-density relations alike was proposed by Osher [84]

$$\lambda\big(\rho^k(x-\Delta x), \rho^k(x)\big) = \begin{cases} \min\limits_{\rho^k(x-\Delta x)\le\rho\le\rho^k(x)} Q_e(\rho) & \text{if } \rho^k(x-\Delta x) \le \rho^k(x) \\ \max\limits_{\rho^k(x-\Delta x)\ge\rho\ge\rho^k(x)} Q_e(\rho) & \text{if } \rho^k(x-\Delta x) \ge \rho^k(x) \end{cases} \tag{3.58}$$

The most commonly used flux function in the traffic flow literature was proposed by Daganzo [24] and will, throughout this thesis be referred to as *Daganzo's flux*. Daganzo's

flux function was developed for piecewise linear flow-density relations, which are not strictly concave. This means that Daganzo's flux is more general than the flux function given in Table 3.1. It is obtained by splitting the fundamental diagram into a *sending function* and a *receiving function* denoted respectively by $S_e(\bar{\rho})$ and $R_e(\bar{\rho})$, depicted in Figure 3.20.



(a) (b)

Figure 3.20: Daganzo's sending and receiving functions; (a) the sending function; (b) the receiving function

Daganzo's flux function is given by:

$$\lambda\big(\rho^k(x-\Delta x),\rho^k(x)\big)=\min\Big\{S_e\big(\rho^k(x-\Delta x)\big),R_e\big(\rho^k(x)\big)\Big\}, \tag{3.59}$$

It can be easily checked that Daganzo's flux coincides with Osher's numerical flux function, (3.58). One of the most popular numerical methods, which is a special case of the Godunov scheme using Dganzo's flux is the cell transmission model (CTM) [22, 23]. The CTM is obtained by using Newell's simplified flow-density relation and choosing Δt and Δx so that $\frac{\Delta x}{\Delta t}=v_f$, which constitutes the upper bound of the CFL condition.

It is notable that, due to their forward in time nature, Godunov scheme based numerical methods, and particularly the CTM, have come to constitute models of traffic flow rather than just computational schemes. This is also adopted in this thesis, where the mean behavior of the stochastic models proposed are compared to Godunov scheme based dynamical equations.

Chapter 4

A Stochastic Model of Traffic Flow and its Fluid Limit

4.1 Introduction

The terminology *fluid approximation*[1] was first introduced by Gordon Newell in [77], in which a deterministic approximation was proposed for purposes of analyzing delays at signalized intersections. It is interesting to note that, while the first application of this idea (fluid approximation) appeared in the traffic flow literature, its most recent developments and applications appear in other areas of operations research[2] . In essence, a fluid limit of a (stochastic) queueing model is a deterministic approximation of the queueing model, which describes the long-run (qualitative) behavior of the stochastic model. The term *fluid* comes from the continuous nature of the fluid model used to approximate a discrete valued dynamic.

Most applications of fluid limits start with a stochastic queueing model and then proceed to derive and analyze its fluid limit. In this thesis, the question of interest is: what type of a queueing model delivers, as its fluid limit, the first-order macroscopic traffic flow dynamics

[1] Here, *fluid approximation* and *fluid limit* are used interchangeably.

[2] In [81], a 2002 article on the history of traffic flow research, Newell gives an account of the dwindling activity in traffic flow research as of the early 1970s.

presented in Chapter 3? That is, the problem is worked out in the reverse direction: a desired deterministic dynamic is known and the stochastic model is derived as an extension to the fluid model.

That the first order macroscopic dynamics described in Chapter 3 arise as long-run mean dynamics is not a new notion. As was presented in Chapter 3, the LWR model was developed by its authors as a relationship between averaged quantities. Furthermore, when deriving the fundamental diagram, assumptions of stationarity were made, which is a property that arises in the long-run. As another example, in the kinetic models of traffic flow, the LWR model arises as a long-run stationary dynamic or as a first-order (mean) approximation (see, for instance [90, Section 3.4] and [44]).

This chapter develops a stochastic queueing model of traffic flow, for which the fluid limit is the Godunov scheme based dynamic presented in Chapter 3. The mathematical properties of the numerical flux functions, which are needed in order to derive the fluid limit, are discussed. In particular, Lipschitz continuity is shown and how to calculate their derivatives (needed in the following Chapter) are presented. It is also shown that the proposed stochastic model implicitly produces traffic densities that are non-negative and do not exceed the jam density. Numerical examples are given throughout the chapter to illustrate the stochastic model and its fluid limit.

4.2 Preliminaries: properties of the fluid model

The fluid model used in this thesis is based on the Godunov scheme. The physical properties of this dynamic and its connection to the LWR model were presented in Chapter 3. In this section, the objects of interest are the mathematical properties needed in order to establish the asymptotic limits derived in the sequel. Specifically, the setting is one where the road is divided into cells; let $\mathcal{C} \subset \mathbb{N}$ denote the set of cell indices with $|\mathcal{C}|$ representing the index of the last cell. To simplify notation, let $\bar{y}(x,t) = [\bar{\rho}(x,t),\ \bar{\rho}(x+1,t)]^{\mathsf{T}}$ denote the (mean) *vector of relevant traffic conditions* at the downstream boundary of cell x, which consists of the mean traffic densities in the two cells adjacent to the subject boundary. Here, with

a slight abuse of notation, $\bar{\rho}(x,t)$ is used to denote the mean traffic density in cell x at time t. Also, variable cell lengths (sizes) shall be allowed, where l_x denotes the length of cell x. The flux across the boundary, using this new notation, is written as: $\lambda\big(\bar{y}(x,t)\big)$. To establish the Godunov scheme as a deterministic queueing (network) model, it is written in terms of cumulative flows and in continuous time. This is carried out as follows: with the new notation convention, (3.57) may be written as:

$$\bar{\rho}\big(x,(k+1)\Delta t\big) = \bar{\rho}(x,k\Delta t) + \frac{\Delta t}{l_x}\Big(\lambda\big(\bar{y}(x-1,k\Delta t)\big) - \lambda\big(\bar{y}(x,k\Delta t)\big)\Big) \tag{4.1}$$

Starting from time zero, (4.1) is given by:

$$\bar{\rho}\big(x,(k+1)\Delta t\big) = \bar{\rho}(x,0) + \frac{\Delta t}{l_x}\left(\sum_{j=0}^{k}\lambda\big(\bar{y}(x-1,j\Delta t)\big) - \sum_{j=0}^{k}\lambda\big(\bar{y}(x,j\Delta t)\big)\right) \tag{4.2}$$

Upon letting $\Delta t \to 0$, we have the continuous time version of the Godunov scheme:

$$\bar{\rho}(x,t) = \bar{\rho}(x,0) + \frac{1}{l_x}\left(\int_0^t \lambda\big(\bar{y}(x-1,u)\big)du - \int_0^t \lambda\big(\bar{y}(x,u)\big)du\right) \tag{4.3}$$

Note that the CFL condition is still honored in this continuous time version of the Godunov scheme since Δt only gets smaller. Since variable cell lengths are allowed, for the sake of accuracy, the CFL condition should be re-written as:

$$\frac{\Delta t}{l_{min}} \max_{0\le\bar{\rho}\le\rho_{jam}} \big|Q_e'(\bar{\rho})\big| \le 1, \tag{4.4}$$

where $l_{min} \equiv \min_{x\in\mathcal{C}} l_x$ is the minimum cell length.

The mathematical properties of the conservation equation (4.3) are determined from the properties of the flux function $\lambda\big(\bar{y}(x,t)\big)$. Of particular interest to the developments in the sequel are: (i) boundedness, (ii) Lipschitz continuity (see Appendix A, Definition A.2.1), and (iii) differentiability. Formally, we have the following requirements:

(i) For all $\bar{y}(x,t) \in [0,\rho_{jam}] \times [0,\rho_{jam}]$, the flux functions are bounded and non-negative.

(ii) For all $\bar{y}(x,t) \in [0,\rho_{jam}] \times [0,\rho_{jam}]$, the flux functions are Lipschitz continuous.

(iii) The flux functions are differentiable in both elements of $\bar{y}(x,t)$.

That requirement (i) is honored is obvious from the definition of a fundamental diagram, regardless of shape (see Section 3.4), where the upper bound is q_{max} and the lower bound is zero; the corresponding sending and receiving functions inherit these properties and so does their minimum. Properties (ii) and (iii) on the other hand are not obvious. In fact, neither property holds in general, property (ii) holds for Daganzo's flux (but this is subtle) but not the general numerical flux given in Table 3.1, while property (iii) does not hold (everywhere) in the classical sense, which was also noted in [9, 113]. These two issues are addressed in Section 4.6.

4.3 The Stochastic Model

let $\mathcal{Q}(x,t)$ denote a stochastic counting process describing the cumulative number of vehicles that have crossed the downstream boundary of cell $x \in \mathcal{C}$ at time $t \in [0, U]$, where $U < \infty$ is a horizon time. The conservation of traffic density in x is written as:

$$\rho(x,t) = \rho(x,0) + \frac{1}{l_x}\big(\mathcal{Q}(x-1,t) - \mathcal{Q}(x,t)\big), \tag{4.5}$$

where $\rho(x,t)$ is the random traffic density in cell x at time t.

As in Chapter 3, $h_i(x)$ denotes the (random) time headway of the i^{th} vehicle departing cell x; then the crossing time of vehicle k at the downstream boundary of cell x is:

$$t_k(x) = \sum_{i=1}^{k} h_i(x), \tag{4.6}$$

where $t_0(x) \equiv 0$. Consequently, the cumulative flow is defined as

$$\mathcal{Q}(x,t) \equiv \max\{k : t_k(x) \leq t\}, \tag{4.7}$$

or, alternatively, the counting process may be characterized by the events:

$$\{\mathcal{Q}(x,t) = k\} \iff \{t_k(x) \leq t \cap t_{k+1}(x) > t\}, \tag{4.8}$$

where $\iff$ means "if and only if". Both cases, (4.7) and (4.8), are interpreted as: "$\mathcal{Q}(x,t)$ is the index of the most recent departure from cell x at time t".

It can, thus, be seen that the probabilistic properties of $\mathcal{Q}(x,t)$, which characterize our stochastic conservation equation, (4.5), are closely related to the properties of the random time headways at the downstream boundary of x. Properties of time headways are discussed next.

4.4 Time Headways

Time headways measured at location x depend on traffic conditions in the immediate upstream and immediate downstream of x; that is, they depend on the vector $y(x,t)$. More precisely, the i^{th} time headway shall be taken to depend on traffic conditions at the time of the most recent departure: $y\big(x,t_{i-1}(x)\big)$, i.e., past traffic densities; then consecutive time headways are conditionally independent. The mean time headway (time per vehicle) is the reciprocal of the mean rate of flow (vehicles per unit time). Then, dependence of time headways on traffic densities is obtained via the dependence of flows on traffic densities. In this thesis, microscopic assumptions that lead to a particular relationship between headways and traffic densities shall not be made. However, in allowing for any of the fundamental relations presented in Section 3.4, any of the (classical) microscopic assumptions that appear in the literature can be accommodated. Let $G_h(\cdot|y)$ denote the conditional distribution function of the time headways and $\mathbb{E}\Big(h_i(x)|y\big(x,t_{i-1}(x)\big)\Big) \equiv \bar{h}\Big(y\big(x,t_{i-1}(x)\big)\Big) = 1/\lambda\Big(y\big(x,t_{i-1}(x)\big)\Big)$ the conditional expectation of time headway i. The conditional variance of time headway i is given by $\bar{c}^2\bar{h}^2\Big(y\big(x,t_{i-1}(x)\big)\Big)$, where $\bar{c}$ is the coefficient of variation of time headways, taken here to be independent of traffic state. Next, two examples of headway distributions that appear in the traffic flow literature are given, followed by an example of a state-dependent headway.

4.4.1 Example: Exponential Time Headways

Let $G_h(\cdot|y) \equiv G_h(\cdot)$ be independent of traffic state (or simply take the same form for all possible values of y) and define the mean headway as: $\bar{h}(y) \equiv 1/\lambda$, where $\lambda > 0$ is a known

scalar. Then, writing the distribution function as:

$$G_h(a) = 1 - e^{-\lambda a}, \qquad a \geq 0, \tag{4.9}$$

one obtains a model with independent and identically distributed (i.i.d.) exponential headways with $\bar{c} = 1$, which is appropriate for modeling free-flow traffic conditions (see for example [12, 15, 43, 78, 92, 97, 109]). In this case $\mathcal{Q}(x,t)$ is a homogeneous Poisson process with rate λ.

4.4.2 Example: Mixed Time Headways

A common generalization of exponentially distributed time headways to situations where vehicles are not traveling freely is to consider the time headways as consisting of two independent components: a "free" component and a "tracking" component, denoted h_i^f and h_i^t, respectively [11, 21, 43, 47]. Let $\theta \in [0,1]$ be a given probability that the vehicle is tracking and $(1-\theta)$ the probability that the vehicle is moving freely. Then the time headway may be written as $h_i(x) = h_i^f(x) + h_i^t(x)$, where $h_i^f(x)$, the free headway, is zero with probability θ and exponentially distributed with parameter λ_f with probability $(1-\theta)$; $h_i^t(x)$, the tracking component, is a non-negative random variable with distribution function $G_t(\cdot)$. Then,

$$G_h(a) = \theta G_t(a) + (1-\theta)\int_0^a G_t(a-u)\lambda_f e^{-\lambda_f u} du, \qquad a \geq 0 \tag{4.10}$$

with expectation

$$\mathbb{E}h_i(x) = \bar{h} = \mathbb{E}h_i^t(x) + \frac{1-\theta}{\lambda_f} \tag{4.11}$$

and variance

$$\mathrm{Var}\big(h_i(x)\big) = \mathbb{E}\big(h_i(x) - \mathbb{E}h_i(x)\big)^2 = \mathrm{Var}\big(h_i^t(x)\big) + \frac{(1-\theta)(1+3\theta)}{(\lambda_f)^2} \tag{4.12}$$

In this case $\mathcal{Q}(x,t)$ is a renewal process with time headway distributions given by (4.10).

4.4.3 Example: State Dependent Time Headways

In this example, we examine the impact of state dependence on the probability distribution of time headways. The setting is a cell boundary with varying traffic densities on either

side. Suppose we have the triangular flow density relation (Newell's simplified relation), $Q_e(\rho)$, shown in Figure 4.1.

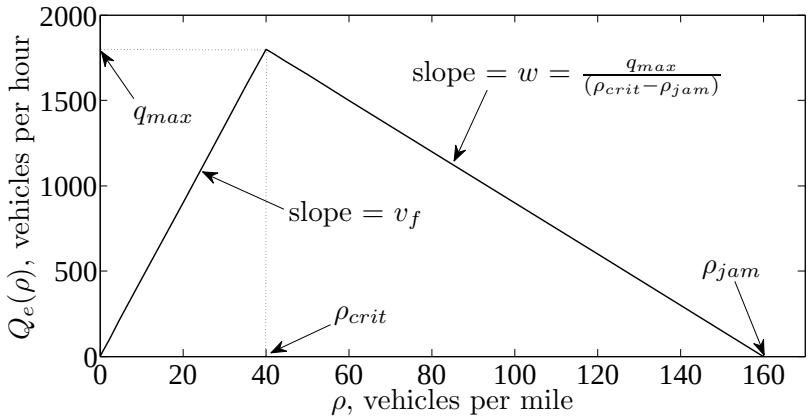


Figure 4.1: Example fundamental relation

The sending and receiving functions are written as $S_e(\rho) = \min\{v_f\rho, q_{max}\}$ and $R_e(\rho) = \min\{q_{max}, |w|(\rho_{jam} - \rho)\}$, respectively. Then, Daganzo's flux function is written as:

$$\lambda(y) = \min\{S_e(y_1), R_e(y_2)\} \tag{4.13}$$

Here the relevant traffic conditions are $y = [y_1,\ y_2]^\mathsf{T}$, where y_1 represents the traffic density upstream the cell boundary (i.e., pertaining to the subject cell) and y_2 is the traffic density downstream the cell boundary (i.e., that of the following cell). We now have that

$$\bar{h}(y) = \frac{1}{\lambda(y)} \tag{4.14}$$

where, as shown in Figure 4.1, $v_f = 45$ mi/hr (72.42 km/hr), $q_{max} = 1800$ veh/hr, $|w| = 15$ mi/hr (24.14 km/hr), $\rho_{jam} = 160$ veh/mi (99.42 veh/km), and, consequently, $\lambda(y) = \min\{45y_1, 1800, 15(160 - y_2)\}$ veh/hr (for y_1 and y_2 in veh/mi).

For $G_h(\cdot|y)$, suppose the time headways are gamma distributed with shape parameter 2 and mean $\bar{h}(y)$ (i.e., $\bar{c} = \sqrt{\frac{1}{2}}$). Then, we have the following conditional probability density function:

$$f_h(a|y) = \frac{dG_h(a|y)}{da} = a\lambda^2(y)e^{-a\lambda(y)}, \qquad a \geq 0 \tag{4.15}$$

To illustrate the impact of state dependence, Figure 4.2 is a plot of the probability density functions of the time headways, $f_h(\cdot|y)$, for three cases: (i) $y = [10,\ 10]^{\mathsf{T}}$ veh/mi ($[6.21,\ 6.21]^{\mathsf{T}}$ veh/km) representing free-flow conditions in both upstream and downstream cells, (ii) $y = [100,\ 10]^{\mathsf{T}}$ veh/mi ($[62.1,\ 6.21]^{\mathsf{T}}$ veh/km) representing queue discharge conditions, and (iii) $y = [100,\ 100]^{\mathsf{T}}$ veh/mi ($[62.1,\ 62.1]^{\mathsf{T}}$ veh/km) representing congested conditions in both upstream and downstream cells.

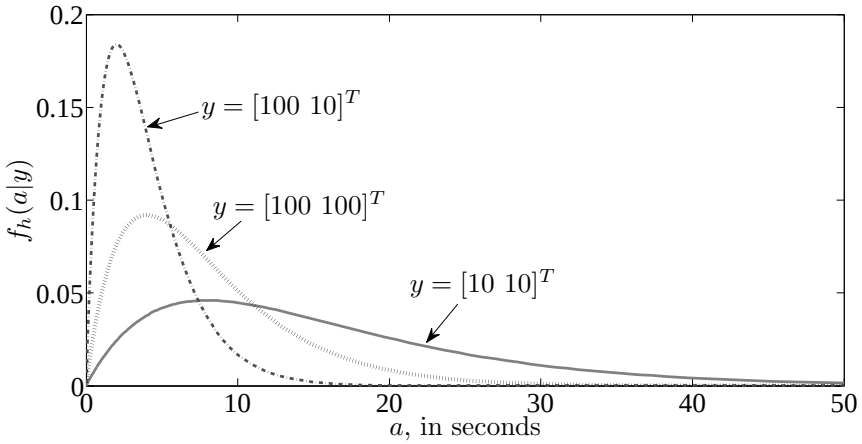


Figure 4.2: Example probability density plots for state dependent headways

We see in Figure 4.2 that under free-flow conditions ($y = [10,\ 10]^{\mathsf{T}}$ veh/mi), larger headways are "more likely" to arise with smaller headways occurring less frequently. This is in contrast to a queue discharge situation ($y = [100,\ 10]^{\mathsf{T}}$ veh/mi), where small headways tend to occur with high frequency and larger headways are "less likely" to occur. Under congested conditions ($y = [100,\ 100]^{\mathsf{T}}$ veh/mi), we see something that falls in between the first two cases; one can think of this as a dispersed version of case (ii) as vehicles tend to encounter slower moving traffic in the downstream.

4.5 The Fluid Limit: A Simplified Setting

The following notation will be used in the sequel: $\mathcal{N}(t)$ is a unit rate Poisson process (i.e., averaging one event per unit time); see Appendix A, Definition A.3.2 for a definition of the

Poisson process. For example, when $\mathcal{Q}(x,t)$ is a homogeneous Poisson process with constant rate λ, this shall be indicated by writing:

$$\mathcal{Q}(x,t) = \mathcal{N}_x(\lambda t), \tag{4.16}$$

where the subscript x is a reminder that this Poisson process represents vehicles departing cell x. The rate, λ, shall be thought of as a re-scaling of the time axis. In other words, a homogeneous Poisson process with rate λ counts (on average) λ vehicles per unit time. This is equivalent to a process that counts one vehicle λ times per unit time and can be interpreted as stretching (or contracting) the time axis a constant rate of λ and counting (on average) one vehicle per unit "scaled" time. With time-varying rates (e.g., a non-homogeneous Poisson process), the "stretching/contracting" of the time axis is not constant.

More generally, a counting process with generally distributed i.i.d. time headways with mean 1 shall be denoted by $\widehat{\mathcal{N}}(t)$. (The Poisson process is the special case when the time headways are exponentially distributed.) If the mean time headways are $\bar{h} \equiv \frac{1}{\lambda}$, the corresponding counting process can also be thought of as the *time-changed* counting process $\widehat{\mathcal{N}}(\lambda t)$. To see this, define $\tilde{h}_i(x) \equiv \frac{h_i(x)}{\bar{h}}$ and note that $\mathbb{E}\tilde{h}_i(x) = 1$, then, following (4.8), the counting process can be written as:

$$\max\left\{k : \sum_{i=1}^{k} h_i(x) \le t\right\} = \max\left\{k : \sum_{i=1}^{k} \tilde{h}_i(x) \le \frac{t}{\bar{h}}\right\} = \widehat{\mathcal{N}}\left(\frac{t}{\bar{h}}\right) = \widehat{\mathcal{N}}(\lambda t) \tag{4.17}$$

4.5.1 Scaling and the Strong Law of Large Numbers

In order to perform an asymptotic analysis of $\mathcal{Q}(x,t)$, i.e., determine the stationary flow rate: $\frac{\mathcal{Q}(x,t)}{t}$ as $t \to \infty$, one counts a large number of time headways. As letting k in (4.17) go to infinity provides the future behavior of the process, this is of little use as discussed in Section 2.2. Instead, we may count fractions of time headways and as we let the fraction size get smaller we obtain a larger number of fractions to which we may apply asymptotic analysis without loosing the transient information in the process. This is illustrated as follows: suppose we wish to count k whole vehicles by looking at the "index of the most recent *fractional* arrival". That is, suppose we divide the headways uniformly into $n \in \mathbb{N}$

fractions, then the crossing time of the k^{th} whole vehicle is the crossing time of the $(nk)^{\text{th}}$ fraction:

$$\sum_{i=1}^{nk} \frac{h_i(x)}{n} = \frac{1}{n} t_{nk}(x), \tag{4.18}$$

where the right-hand side follows from (4.6). Suppose this occurs by some time t, but the crossing time of the next fraction occurs after time t; that is, the following event takes place:

$$\left\{ \frac{1}{n} t_{nk}(x) \leq t \cap \frac{1}{n} t_{nk+1}(x) > t \right\} = \{ t_{nk}(x) \leq nt \cap t_{nk+1}(x) > nt \} \tag{4.19}$$

From (4.8), this implies the following event:

$$\{ \mathcal{Q}(x, nt) = nk \} = \left\{ \frac{1}{n} \mathcal{Q}(x, nt) = k \right\} \tag{4.20}$$

As $n \to \infty$, we count a larger and larger number of smaller and smaller *jumps* across x, which in the limit resemble a *fluid* process. $\mathcal{Q}(x, nt)$ may then be interpreted as the process $\mathcal{Q}(x, t)$ with its flow rates accelerated uniformly by a factor of n (n times as many vehicles per unit time). Dividing $\mathcal{Q}(x, nt)$ by n serves as a reminder that we are counting fractions (of size $\frac{1}{n}$) of vehicles.

For the simple cases where $\mathcal{Q}(x, t)$ is either $\mathcal{N}(t)$ or $\widehat{\mathcal{N}}(t)$, uniformly accelerating $\mathcal{Q}(x, t)$ a rate n and "aggregating", we write (for the latter):

$$\frac{1}{n} \mathcal{Q}(x, nt) = \frac{1}{n} \widehat{\mathcal{N}}_x(nt) \tag{4.21}$$

The asymptotic behavior of this process is obtained as an application of the strong law of large numbers (SLLN) [18, 93]; see Appendix A, Theorem A.3.1 for statement of the classical SLLN. First note that $\widehat{\mathcal{N}}_x(nt) \to \infty$ as $n \to \infty$ and that $\frac{1}{n} t_n(x) \to \mathbb{E} h_1(x)$ as $n \to \infty$ by the SLLN. Then, from the RHS of (4.19), we have that:

$$t_{\widehat{\mathcal{N}}_x(nt)}(x) \leq nt < t_{\widehat{\mathcal{N}}_x(nt)+1}(x) \tag{4.22}$$

Taking the reciprocal and multiplying by $\widehat{\mathcal{N}}_x(nt)$:

$$\frac{\widehat{\mathcal{N}}_x(nt)}{t_{\widehat{\mathcal{N}}_x(nt)}(x)} \geq \frac{\widehat{\mathcal{N}}_x(nt)}{nt} > \frac{\widehat{\mathcal{N}}_x(nt) + 1}{t_{\widehat{\mathcal{N}}_x(nt)+1}(x)} \frac{\widehat{\mathcal{N}}_x(nt)}{\widehat{\mathcal{N}}_x(nt) + 1} \tag{4.23}$$

In the limit, it follows immediately that:

$$\frac{1}{n}\widehat{\mathcal{N}}_x\left(nt\right) \underset{n\to\infty}{\longrightarrow} t \quad \text{almost surely (a.s.)} \tag{4.24}$$

In fact, this holds for any t. For example, let $\tilde{t} \equiv \lambda t$; then $\frac{1}{n}\widehat{\mathcal{N}}_x(n\lambda t) = \frac{1}{n}\widehat{\mathcal{N}}_x(n\tilde{t})$ converges to $\tilde{t} = \lambda t$. As another example, take the non-homogeneous Poisson process with the time varying deterministic rates $\lambda(x,t)$. Define $\tilde{t} \equiv \int_0^t \lambda(x,u)du$, we have, from (4.24), that:

$$\frac{1}{n}\mathcal{N}_x\left(n\int_0^t \lambda(x,u)du\right) \underset{n\to\infty}{\longrightarrow} \int_0^t \lambda(x,u)du \quad \text{a.s.} \tag{4.25}$$

This latter example shows, by applying uniform acceleration, how the transient information in the process is preserved in the mean process.

4.5.2 Example: scaling

Let $\{h_i\}$ be i.i.d. gamma distributed time headways, with shape parameter 2, mean headway $\bar{h} = 10$ sec/veh, and $\mathcal{Q}(t)$ is the associated counting process. Figure 4.3 below plots a single sample path of the counting process for the first 10 arrivals.

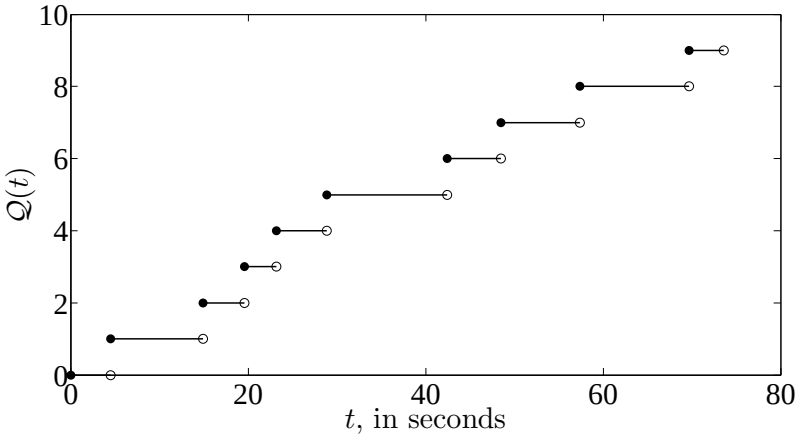


Figure 4.3: Example sample path for counting process ($n = 1$)

The scaled version of the process, $\frac{1}{n}\mathcal{Q}(nt)$, in essence, involves a larger number of smaller jumps. To see this, Figure 4.4 plots a sample path of the scaled counting process for $n = 3$.

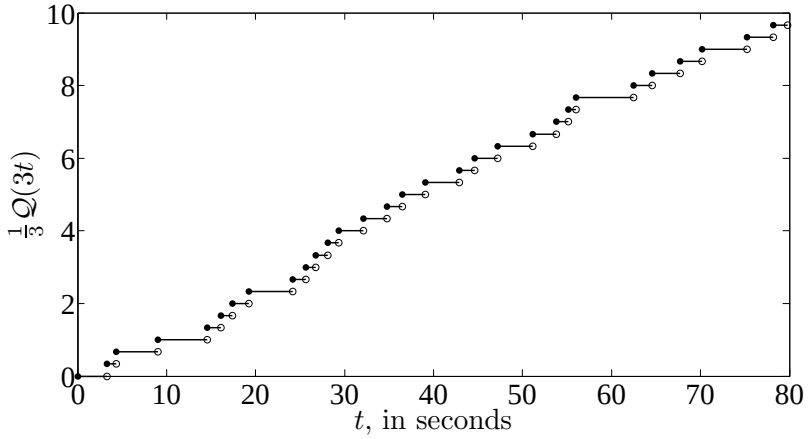


Figure 4.4: Example sample path for scaled counting process with $n = 3$

As n gets larger, the variance of the process shrinks, and in the limit the process degenerates to a deterministic process. To illustrate this, the following figures compare plots of the jump points of 500 sample paths of $\mathcal{Q}(t)$ (i.e., $n = 1$) with 500 sample paths of $\frac{1}{n}\mathcal{Q}(nt)$ for $n = 10$, $n = 100$, and $n = 1000$, in Figures 4.5, 4.6, and 4.7, respectively, for a time horizon $U = 400$ seconds. In these figures, the more thinly dispersed points belong the sample paths of $\mathcal{Q}(t)$, while the darker more densely dispersed points are the sample paths of the scaled processes, $\frac{1}{n}\mathcal{Q}(nt)$.

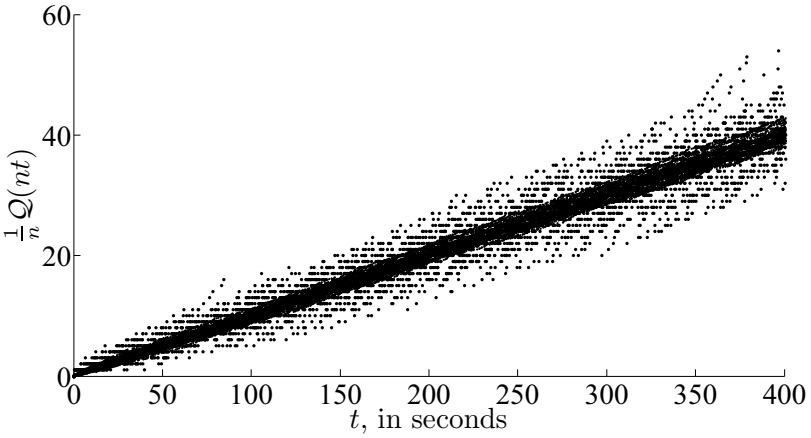


Figure 4.5: Example sample path comparisons; thin: $n = 1$, dense: $n = 10$

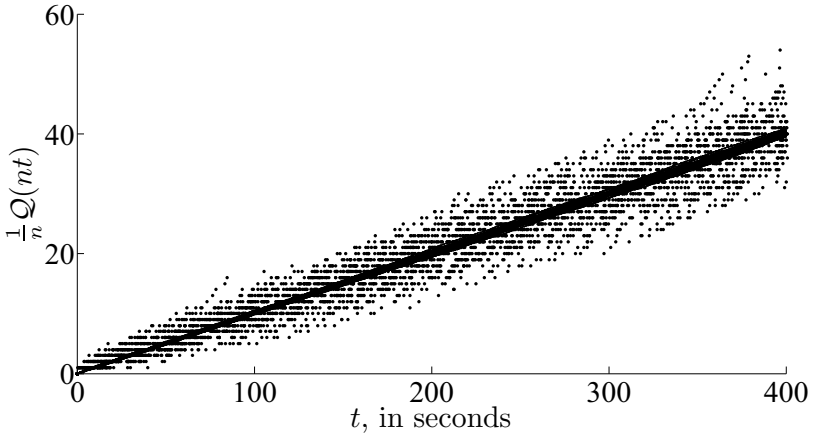


Figure 4.6: Example sample path comparisons; thin: $n = 1$, dense: $n = 100$

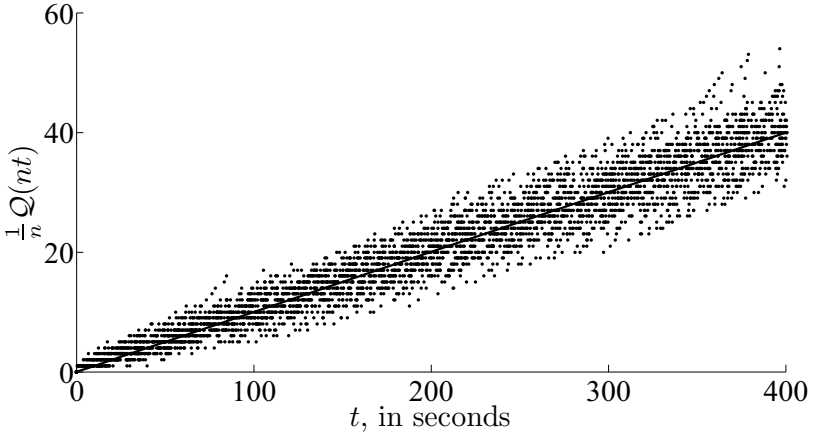


Figure 4.7: Example sample path comparisons; thin: $n = 1$, dense: $n = 1000$

Clearly, as n gets larger, process variation gets smaller. In Figure 4.7, it is difficult to distinguish between the scaled process and the (continuous) deterministic process $\bar{\mathcal{Q}}(t) = \lambda t = 0.1t$. This is the fluid limit of $\mathcal{Q}(t)$.

4.6 Fluid Limits for State-Dependent Processes

In this section, the general case where instantaneous flow rates depend on traffic state, via appropriate flux functions, is considered. The same intuition behind scaling applies here as

well; however, due to state dependence, the flow rates are random. Thus, the time change is also random. Specifically, for each $x \in \mathcal{C}$, $\mathcal{Q}(x,t)$ is a point process with rates defined by the flux function $\lambda\big(y(x,t)\big)$. In the language of probability theory, we say that the point process $\mathcal{Q}(x,t)$ is measurable with respect to the sigma-field $\mathcal{F}_t$ while its intensity, $\int_0^t \lambda\big(y(x,u)\big)du$, is measurable with respect to $\mathcal{F}_{t-}$, where $\mathcal{F}_{t-} \subseteq \mathcal{F}_t$. The sigma field is, in essence, an event space for a random variable and when considering a stochastic process, a sequence of sigma-fields (a filtration) represents the history of the process and therefore grows with time. This technical detail, referred to as *predictability* in the stochastic calculus literature, is needed to ensure that the stochastic integrals we construct in the sequel possess the desired mathematical properties (namely, that they are well defined).

The best way to illustrate this predictability condition in our traffic flow model is to consider how one would simulate the process: suppose one devises a discrete time approximation of our processes and computes traffic variables over small discrete time intervals of length Δt, which for the sake of illustration may be assumed fixed. One begins with initial traffic densities, $\{\rho(x,0)\}$, which are possibly random. These traffic densities are used to compute/simulate cumulative flows over the time interval $[0,\Delta t)$. These in turn are used to compute the traffic densities at time Δt, $\{\rho(x,\Delta t)\}$. The procedure is then repeated for the interval $[\Delta t, 2\Delta t)$, and so on. We note that this is precisely a stochastic version of how both the Godunov scheme (and the CTM) operates; that is flows are computed using the most recent past values of traffic density and the new traffic densities are computed using these flows.

These conditions allows us to represent $Q(x,t)$ as a unit rate Poisson process with the stochastic time change $\int_0^t \lambda\big(y(x,u)\big)du$; that is,

$$\mathcal{Q}(x,t) = \mathcal{N}_x\left(\int_0^t \lambda\big(y(x,u)\big)du\right) \tag{4.26}$$

That any point process can be transformed via a time change to a Poisson process is an important result in point processes theory; a detailed exposition of the result can be found in [13, 26, 27] and a simpler exposition is given in [14]. The Poisson process with the random time change is a doubly stochastic Poisson process and can be interpreted in the

same manner as the time-changed non-homogeneous Poisson process, except that the time axis is stretched/contracted in accordance with the crossing times $\{t_i(x)\}_{i\geq 0}$. We may now write the scaled conservation equation as:

$$\rho^n(x,t) = \rho^n(x,0) + \frac{1}{l_x}\left(\frac{1}{n}\mathcal{Q}(x-1,nt) - \frac{1}{n}\mathcal{Q}(x,nt)\right) \tag{4.27}$$

The scaled counting processes are written as:

$$\frac{1}{n}\mathcal{Q}(x,nt) = \frac{1}{n}\mathcal{N}_x\left(n\int_0^t \lambda\big(y^n(x,u)\big)du\right) \tag{4.28}$$

Note that flow rates are functions of the scaled traffic densities, $y^n(x,t) = [\rho^n(x,t)\ \rho^n(x+1,t)]^{\mathsf{T}}$. The interpretation of the scaled traffic density, $\rho^n(x,t)$, is: the traffic density in cell x which corresponds to scaled counting processes at the cell boundaries. That is, the traffic density itself is not scaled, but constructed from scaled processes.

The initial traffic density $\rho(x,0)$ is assumed here to be a random variable, independent of the future of the process (but not vice versa) and independent in x. This preserves causality in the model. Denote the mean and variance of $\rho(x,0)$ by $\bar{\rho}(x,0)$ and $\sigma^2(x,0)$, respectively. The scaled initial traffic densities shall be assumed to converge to their mean values as $n \to \infty$. Since we are only interested in the limiting behavior, this can be easily accomplished by construction. For example, let $\rho_1(x), \rho_2(x), ..., \rho_n(x)$ be a sequence of i.i.d. random variables, identically distributed as $\rho(x,0)$, then

$$\rho^n(x,0) \equiv \frac{1}{n}\sum_{j=1}^{n}\rho_j(x) \tag{4.29}$$

achieves the desired limit, as $n \to \infty$ by the classical SLLN. Furthermore, under re-scaling and centering, we have that $\rho^n(x,0)$ converges in distribution to a Normal random variable with mean $\bar{\rho}(x,0)$ and variance $\sigma^2(x,0)$ by the classical central limit theorem (CLT); see Appendix A, Theorem A.3.2 for statement of the classical CLT.

Before discussing the limiting behavior of the scaled traffic density process, (4.27), Lipschitz continuity and differentiability of the flux function (mentioned in Section 4.2) must be addressed. This is done next.

4.6.1 Lipschitz Continuity of Flux Functions

The Godunov scheme flux function in Table 3.1 may be written as:

$$\begin{aligned}\lambda\big(\bar{y}(x,t)\big) \;=\; & Q_e\big(\bar{\rho}(x,t)\big)1_{\{\bar{\rho}(x,t)\le\rho_{crit}\,\cap\,\bar{\rho}(x+1,t)\le\rho_{crit}\}} \\ & +q_{max}1_{\{\bar{\rho}(x+1,t)\le\rho_{crit}\le\bar{\rho}(x,t)\}} \\ & +Q_e\big(\bar{\rho}(x+1,t)\big)1_{\{\bar{\rho}(x,t)\ge\rho_{crit}\,\cap\,\bar{\rho}(x+1,t)\ge\rho_{crit}\}} \\ & +\min\{Q_e\big(\bar{\rho}(x,t)\big),Q_e\big(\bar{\rho}(x+1,t)\big)\}1_{\{\bar{\rho}(x,t)\le\rho_{crit}\le\bar{\rho}(x+1,t)\}} \qquad (4.30)\end{aligned}$$

where $1_{\{\cdot\}}$ is the indicator function: $1_{\{A\}}=1$ if the condition A is true and $1_{\{A\}}=0$, otherwise. The indicator function is discontinuous and hence does not possess the Lipschitz property. However, it may be closely approximated by a logistic function (see, for example, Appendix B, Section B.2), which does possess the Lipschitz property. While this takes care of the first three components on the RHS of (4.30), it is not clear whether the fourth component is Lipschitz, due to the minimum function. Then Lipschitz continuity is established (after approximating the indicator functions), if it can be shown that the minimum function in (4.30) is Lipschitz continuous. Alternatively, this can be established for Daganzo's flux function:

$$\lambda\big(\bar{y}(x,t)\big)=\min\Big\{S_e\big(\bar{\rho}(x,t)\big),R_e\big(\bar{\rho}(x+1,t)\big)\Big\}, \qquad (4.31)$$

where both the sending and the receiving function are Lipschitz continuous. Note that the Lipschitz property holds for any sending and receiving function obtained from any of the fundamental relations presented in Section 3.4. To proceed, note that (4.31) can alternatively be written as:

$$\min\{S_e(a),R_e(b)\}=\frac{1}{2}\Big(S_e(a)+R_e(b)-\big|S_e(a)-R_e(b)\big|\Big), \qquad (4.32)$$

for any pair of traffic densities $(a,b)\in[0,\rho_{jam}]\times[0,\rho_{jam}]$.

Now, to see that (4.31) is Lipschitz, first note that $\lambda:\Theta_1\to\Theta_2$ is a map between two metric spaces (Θ_1,d_1) and (Θ_2,d_2) where $\Theta_1\subset\mathbb{R}^2_+$ and $\Theta_2\subset\mathbb{R}_+$, both bounded, and the metrics can be taken as $d_1([a_1\ \ a_2]^\mathsf{T},[b_1\ \ b_2]^\mathsf{T})=\max\{|a_1-b_1|,|a_2-b_2|\}$, and $d_2(\lambda_1,\lambda_2)=|\lambda_1-\lambda_2|$. Then, for λ to be Lipschitz continuous, a constant $0\le K<\infty$ must

exist such that the following holds:

$$d_2\big(\lambda([a_1\ a_2]^\mathsf{T}), \lambda([b_1\ b_2]^\mathsf{T})\big) \leq K d_1([a_1\ a_2]^\mathsf{T}, [b_1\ b_2]^\mathsf{T}) \tag{4.33}$$

for all $[a_1\ a_2]^\mathsf{T}, [b_1\ b_2]^\mathsf{T} \in \Theta_1$. We proceed as follows:

$$\begin{aligned}
&d_2\big(\lambda([a_1\ a_2]^\mathsf{T}), \lambda([b_1\ b_2]^\mathsf{T}) = \Big|\min\{S_e(a_1), R_e(a_2)\} - \min\{S_e(b_1), R_e(b_2)\}\Big| \\
&= \frac{1}{2}\Big|\big(S_e(a_1) + R_e(a_2)\big) - \big|S_e(a_1) - R_e(a_2)\big| - \big(S_e(b_1) + R_e(b_2)\big) + \big|S_e(b_1) - R_e(b_2)\big|\Big| \\
&\leq \frac{1}{2}\Big|S_e(a_1) - S_e(b_1)\Big| + \frac{1}{2}\Big|R_e(a_2) - R_e(b_2)\Big| + \frac{1}{2}\Big|\big|S_e(b_1) - R_e(b_2)\big| - \big|S_e(a_1) - R_e(a_2)\big|\Big| \\
&\leq \frac{1}{2}\Big|S_e(a_1) - S_e(b_1)\Big| + \frac{1}{2}\Big|R_e(a_2) - R_e(b_2)\Big| + \frac{1}{2}\Big|S_e(b_1) - R_e(b_2) - S_e(a_1) + R_e(a_2)\Big| \\
&\leq \frac{1}{2}\Big|S_e(a_1) - S_e(b_1)\Big| + \frac{1}{2}\Big|R_e(a_2) - R_e(b_2)\Big| + \frac{1}{2}\Big|S_e(b_1) - S_e(a_1)\Big| + \frac{1}{2}\Big|R_e(a_2) - R_e(b_2)\Big| \\
&= \Big|S_e(a_1) - S_e(b_1)\Big| + \Big|R_e(a_2) - R_e(b_2)\Big|
\end{aligned} \tag{4.34a}$$

where the second line in (4.34a) relies on (4.32) and all bounds in (4.34a) follow from the triangle inequality (see Appendix A, Definition A.2.2 for definition).

Now, from Lipschitz continuity of the sending and receiving functions, we have that

$$\begin{aligned}
\Big|S_e(a_1) - S_e(b_1)\Big| \ &+ \ \Big|R_e(a_2) - R_e(b_2)\Big| \\
&\leq \ K_S\big|a_1 - b_1\big| + K_R\big|a_2 - b_2\big| \\
&\leq \ 2\max\{K_S, K_R\}\max\left\{|a_1 - b_1|, |a_2 - b_2|\right\},
\end{aligned} \tag{4.35}$$

where K_S and K_R denote the Lipschitz constants of the sending function and the receiving function, respectively. This proves the Lipschitz continuity of the flux function given in (4.31), where the constant K may be taken as $2\max\{K_S, K_R\}$. Note that no assumptions of concavity were made, only that the sending and receiving functions are Lipschitz continuous. This means that a wide range of fundamental relations can be accommodated in the proposed theory.

4.6.2 Derivatives of Flux Functions

The flux function (4.31) is non-differentiable in the classical sense. To demonstrate this, let a and b denote arbitrary traffic densities in two adjacent cells. We see from (4.32), and

using $|S_e(a) - R_e(b)| = \sqrt{\big(S_e(a) - R_e(b)\big)^2}$, that the partial derivatives may be written as:

$$\begin{aligned}\frac{\partial \lambda}{\partial a} &= \frac{dS_e(a)}{da}\left(\frac{1}{2} - \frac{S_e(a) - R_e(b)}{2\big|S_e(a) - R_e(b)\big|}\right)\\ \frac{\partial \lambda}{\partial b} &= \frac{dR_e(b)}{db}\left(\frac{1}{2} + \frac{S_e(a) - R_e(b)}{2\big|S_e(a) - R_e(b)\big|}\right)\end{aligned} \tag{4.36}$$

Now, note that when $S_e(a) = R_e(b)$, the partial derivatives in (4.36) do not exist. The set of points where this occurs is non-negligible: for any $a \in [0, \rho_{jam}]$ it is easy to see that there exists a point $b \in [0, \rho_{jam}]$ such that $S_e(a) = R_e(b)$ and vice versa. Furthermore, the derivatives are generally not equal for these cases.

To overcome this problem, we apply the following intuition: when $S_e(a) = R_e(b)$ we have that $|S_e(a) - R_e(b)| = 0$ and the formula for the minimum in (4.32), for such points of discontinuity, reduces to:

$$\min\{S_e(a), R_e(b)\} = \frac{1}{2}\big(S_e(a) + R_e(b)\big), \tag{4.37}$$

and the derivatives, when $S_e(a) = R_e(b)$, are then computed as:

$$\begin{aligned}\frac{\partial \lambda}{\partial a} &= \frac{1}{2}\frac{dS_e(a)}{da}\\ \frac{\partial \lambda}{\partial b} &= \frac{1}{2}\frac{dR_e(b)}{db}\end{aligned} \tag{4.38}$$

The derivatives above are in fact weak derivatives (see for example [48] for definition and further discussion). It is well known that the absolute value function $|z|$ possesses the following weak derivative[3] :

$$\frac{d|z|}{dz} = \begin{cases} 1 & \text{if } z > 0\\ 0 & \text{if } z = 0\\ -1 & \text{if } z < 0\end{cases} \tag{4.39}$$

[3] An alternative intuition is: the absolute value function has a V shape. To approximate the derivative at $z = 0$, one may approximate $|z|$ by a V with a very small parabola (of negligible size) at $z = 0$; the derivative at $z = 0$ is then 0.

We, hence, see that the weak derivatives of λ in (4.32) can be written as:

$$\frac{\partial \lambda}{\partial a} = \begin{cases} 0 & \text{if } S_e(a) > R_e(b) \\ \frac{1}{2}\frac{dS_e(a)}{da} & \text{if } S_e(a) = R_e(b) \\ \frac{dS_e(a)}{da} & \text{if } S_e(a) < R_e(b) \end{cases} \tag{4.40}$$

and

$$\frac{\partial \lambda}{\partial b} = \begin{cases} \frac{dR_e(b)}{db} & \text{if } S_e(a) > R_e(b) \\ \frac{1}{2}\frac{dR_e(b)}{db} & \text{if } S_e(a) = R_e(b) \\ 0 & \text{if } S_e(a) < R_e(b) \end{cases} \tag{4.41}$$

which is consistent with (4.36) when $S_e(a) \neq R_e(b)$ and (4.38) when $S_e(a) = R_e(b)$.

4.6.3 Boundedness of Traffic Densities

Another crucial property of the stochastic model is that traffic densities are implicitly non-negative and do not exceed the jam density ρ_{jam}.

Negative traffic densities arise when more vehicles artificially depart a cell than available in the cell, while traffic densities that exceed jam density arise when vehicles are allowed to artificially enter a cell for which the traffic density is at jam density.

To investigate the former case, suppose that at some time t the traffic density in cell x is zero. A negative traffic density arises if an instantaneous change in the number of vehicles by an amount less than zero is allowed. We are thus interested in the following probability:

$$\mathbb{P}\left(\mathcal{N}_{x-1}\left(\int_t^{t+dt} \lambda\big(y(x-1,u)\big)du\right) - \mathcal{N}_x\left(\int_t^{t+dt} \lambda\big(y(x,u)\big)du\right) < 0 \,\middle|\, \rho(x,t) = 0\right), \tag{4.42}$$

where dt is sufficiently small to ensure that at most one jump can occur at either cell boundary. Then, this probability is bounded from above by:

$$\mathbb{P}\left(\mathcal{N}_x\left(\int_t^{t+dt} \lambda\big(y(x,u)\big)du\right) = 1 \,\middle|\, \rho(x,t) = 0\right), \tag{4.43}$$

which is equal to $\lambda\big(y(x,t)\big)dt$ (see, for example [57, Section 9.4, pg. 256]). But since $\rho(x,t) = 0$, we have that $\lambda\big(y(x,t)\big) = S_e\big(\rho(x,t)\big) = 0$. This means that a negative traffic density occurs with probability zero.

Likewise, suppose at time t the traffic density in cell x is the jam density. A traffic density that exceeds the jam density occurs if an instantaneous change in the number of vehicles by an amount greater than zero is allowed. In this case, we are interested in the following probability:

$$\mathbb{P}\left(\mathcal{N}_{x-1}\left(\int_t^{t+dt}\lambda\big(y(x-1,u)\big)du\right)-\mathcal{N}_x\left(\int_t^{t+dt}\lambda\big(y(x,u)\big)du\right)>0\,\middle|\,\rho(x,t)=\rho_{jam}\right), \tag{4.44}$$

which is bounded from above by

$$\mathbb{P}\left(\mathcal{N}_{x-1}\left(\int_t^{t+dt}\lambda\big(y(x-1,u)\big)du\right)=1\,\middle|\,\rho(x,t)=\rho_{jam}\right) \tag{4.45}$$

This probability is equal to $\lambda\big(y(x-1,t)\big)dt$. Since $\rho(x,t)=\rho_{jam}$, we have that $\lambda\big(y(x-1,t)\big)=R_e\big(\rho(x,t)\big)=0$. Consequently, the probability that the traffic density will exceed the jam density is zero.

4.6.4 Derivation of the Fluid Limit

Before deriving the limit, we will need an appropriate way to measure distance between processes. Since the limiting processes throughout this thesis are continuous, it shall suffice to consider the uniform norm [18, 67]. For more general metrics and their applications, see [112]. The uniform norm, for the vector valued process $\mathbf{g}(t)\in\mathbb{R}^m$, is written as:

$$\big|\big|\mathbf{g}(\cdot)\big|\big|_U=\sup_{0\le t\le U}\max_{1\le j\le m}\big\{\big|g_j(t)\big|\big\}, \tag{4.46}$$

where $g_j(t)$ is the j^{th} element of the vector $\mathbf{g}(t)$. When $g(\cdot)$ is a scalar process, the uniform norm is simply written as:

$$\big|\big|g(\cdot)\big|\big|_U=\sup_{0\le t\le U}\big\{\big|g(t)\big|\big\} \tag{4.47}$$

Since the SLLN for the Poisson process, (4.24), holds for any $t\in[0,U]$[4] , we have that

$$\left|\left|\frac{1}{n}\mathcal{N}_x\left(nt\right)-t\right|\right|_U\underset{n\to\infty}{\longrightarrow}0\quad\text{a.s.} \tag{4.48}$$

[4] In fact, the more general counting process also.

This is known as the *functional* strong law of large numbers (FSLLN) for the unit rate Poisson process. Now consider

$$\left|\left|\frac{1}{n}\mathcal{N}_x\left(n\int_0^t \lambda\big(y^n(x,u)\big)du\right) - \int_0^t \lambda\big(y^n(x,u)\big)du\right|\right|_U \tag{4.49}$$

In order to obtain the limiting behavior of (4.49), we will need the following random time change theorem [18, Theorem 5.3], which is proved in [8, Section 14].

Theorem 4.6.1 (random time change theorem)**.** *Let $\mathcal{X}_n(\cdot)$ and $\mathcal{Y}_n(\cdot)$ be two sequences of stochastic processes which possess right-continuous sample paths with left-hand limits (RCLL). Suppose that $\mathcal{Y}_n(\cdot)$ is non-decreasing with $\mathcal{Y}_n(0) = 0$ and that as $n \to \infty$, $\mathcal{X}_n(\cdot)$ and $\mathcal{Y}_n(\cdot)$ converge uniformly on compact sets to processes $\mathcal{X}(\cdot)$ and $\mathcal{Y}(\cdot)$ which possess continuous sample paths. Then, the time-changed process $\mathcal{X}_n(\mathcal{Y}_n)$ converges uniformly on compact sets to $\mathcal{X}(\mathcal{Y})$.*

Explanation: RCLL processes generalize continuous processes[5] to processes with sample paths which may contain jumps, such as the point processes $\mathcal{N}(\cdot)$ and $\widehat{\mathcal{N}}(\cdot)$ (see Figures 4.3 and 4.4 for example). The theorem, in essence, says that if two such processes which take values in $[0, U]$ (a compact set) and converge to the processes $\mathcal{X}(\cdot)$ and $\mathcal{Y}(\cdot)$ almost surely (i.e., "converge uniformly on compact sets") and $\mathcal{Y}(\cdot)$ is a continuous non-decreasing process with $\mathcal{Y}(0) = 0$, then $\mathcal{X}_n(\mathcal{Y}_n)$ converges to $\mathcal{X}(\mathcal{Y})$. That is, knowing how the two processes converge individually, we can characterize the convergence of their composition, provided the conditions of the theorem hold.

For (4.49), the Lipschitz continuity and boundedness of the flux functions, $\lambda(\cdot)$, the continuity of integration, and that $\lambda(\cdot)$ in (4.49) does not depend on n (except in its argument) ensures that $\int_0^\bullet \lambda\big(y^n(x,u)\big)du$ converges to a continuous process. Furthermore, it is easy to see that $\mathcal{Y}_n(t) \equiv \int_0^t \lambda\big(y^n(x,u)\big)du$ is non-decreasing in t, since $\lambda(\cdot)$ is non-negative, and that $\mathcal{Y}_n(0) = 0$. Consequently, by the FSLLN for a unit rate Poisson process, we have that

$$\left|\left|\frac{1}{n}\mathcal{N}_x\left(n\int_0^t \lambda\big(y^n(x,u)\big)du\right) - \int_0^t \lambda\big(y^n(x,u)\big)du\right|\right|_U \underset{n\to\infty}{\longrightarrow} 0 \quad \text{a.s.} \tag{4.50}$$

[5] More accurately, processes with continuous sample paths.

In fact, from (4.24), and by the same arguments above, this can also be established for the more general point processes, $\widehat{\mathcal{N}}_x(\cdot)$; that is,

$$\left|\left|\frac{1}{n}\widehat{\mathcal{N}}_x\left(n\int_0^t \lambda\big(y^n(x,u)\big)du\right)-\int_0^t \lambda\big(y^n(x,u)\big)du\right|\right|_U \underset{n\to\infty}{\longrightarrow} 0 \quad \text{a.s.} \tag{4.51}$$

Now, let's return to the scaled conservation equation:

$$\begin{aligned}\rho^n(x,t)=\rho^n(x,0) \quad &+\frac{1}{l_x}\left(\frac{1}{n}\mathcal{N}_{x-1}\left(n\textstyle\int_0^t \lambda\big(y^n(x-1,u)\big)du\right)\right.\\ &\left.-\frac{1}{n}\mathcal{N}_x\left(n\textstyle\int_0^t \lambda\big(y^n(x,u)\big)du\right)\right)\end{aligned} \tag{4.52}$$

It will be shown that this process converges almost surely (or uniformly on compact sets) to the deterministic process $\bar{\rho}(x,\cdot)$ given by (4.3); that is, the fluid limit of this stochastic process is a Godunov scheme based dynamic. That is, it will be shown that the following holds:

$$\left|\left|\rho^n(x,\cdot)-\bar{\rho}(x,\cdot)\right|\right|_t \underset{n\to\infty}{\longrightarrow} 0 \quad \text{a.s} \tag{4.53}$$

This is shown as follows: for any $t\in[0,U]$,

$$\begin{aligned}\left|\left|\rho^n(x,\cdot)-\bar{\rho}(x,\cdot)\right|\right|_t \;\leq\; &\left|\left|\rho^n(x,0)-\bar{\rho}(x,0)\right|\right|_t\\ &+\frac{1}{l_x}\left|\left|\frac{1}{n}\mathcal{N}_{x-1}\left(n\int_0^\bullet \lambda\big(y^n(x-1,u)\big)du\right)-\int_0^\bullet \lambda\big(y^n(x-1,u)\big)du\right|\right|_t\\ &+\frac{1}{l_x}\left|\left|\int_0^\bullet \lambda\big(y^n(x-1,u)\big)du-\int_0^\bullet \lambda\big(\bar{y}(x-1,u)\big)du\right|\right|_t\\ &+\frac{1}{l_x}\left|\left|\frac{1}{n}\mathcal{N}_x\left(n\int_0^\bullet \lambda\big(y^n(x,u)\big)du\right)-\int_0^\bullet \lambda\big(y^n(x,u)\big)du\right|\right|_t\\ &+\frac{1}{l_x}\left|\left|\int_0^\bullet \lambda\big(y^n(x,u)\big)du-\int_0^\bullet \lambda\big(\bar{y}(x,u)\big)du\right|\right|_t,\end{aligned} \tag{4.54}$$

where $\int_0^t \lambda\big(y^n(x-1,u)\big)du$ and $\int_0^t \lambda\big(y^n(x,u)\big)du$ have been added to and subtracted from the equations and the inequality follows from the triangle inequality.

As $n\to\infty$, by the classical law of large numbers, the first element on the RHS converges to zero almost surely. In accordance with (4.50), both the second and the fourth elements converge to zero, a.s., as well. Now, by Lipschitz continuity of the flux functions combined

with the triangle inequality, we have that

$$\frac{1}{l_x}\left|\left|\int_0^{\bullet}\lambda\big(y^n(x-1,u)\big)du-\int_0^{\bullet}\lambda\big(\bar{y}(x-1,u)\big)du\right|\right|_t\leq K_x\int_0^t\left|\left|y^n(x-1,\cdot)-\bar{y}(x-1,\cdot)\right|\right|_u du \tag{4.55}$$

and likewise

$$\frac{1}{l_x}\left|\left|\int_0^{\bullet}\lambda\big(y^n(x,u)\big)du-\int_0^{\bullet}\lambda\big(\bar{y}(x,u)\big)du\right|\right|_t\leq K_x\int_0^t\left|\left|y^n(x,\cdot)-\bar{y}(x,\cdot)\right|\right|_u du, \tag{4.56}$$

where $K_x\equiv\frac{1}{l_x}K$ and K is the Lipschitz constant.

Since the traffic densities are bounded, we have that

$$\left|\left|y^n(x-1,\cdot)-\bar{y}(x-1,\cdot)\right|\right|_t\leq H\left|\left|\rho^n(x,\cdot)-\bar{\rho}(x,\cdot)\right|\right|_t, \tag{4.57}$$

and

$$\left|\left|y^n(x,\cdot)-\bar{y}(x,\cdot)\right|\right|_t\leq H\left|\left|\rho^n(x,\cdot)-\bar{\rho}(x,\cdot)\right|\right|_t, \tag{4.58}$$

where $1\leq H<\infty$ is a constant. Note that such a constant does not exist when $\left|\left|\rho^n(x,\cdot)-\bar{\rho}(x,\cdot)\right|\right|_t=0$ while $\left|\left|\rho^n(x+1,\cdot)-\bar{\rho}(x+1,\cdot)\right|\right|_t>0$ and $\left|\left|\rho^n(x-1,\cdot)-\bar{\rho}(x-1,\cdot)\right|\right|_t>0$, but this is the trivial case which delivers our desired result immediately.

We now have the following upper bound for the sum of the LHSs of (4.55) and (4.56):

$$2HK_x\int_0^t\left|\left|\rho^n(x,\cdot)-\bar{\rho}(x,\cdot)\right|\right|_u du \tag{4.59}$$

Denote by κ^n the sum of the first, second, and fourth elements on the RHS of (4.54), we then have:

$$\left|\left|\rho^n(x,\cdot)-\bar{\rho}(x,\cdot)\right|\right|_t\leq\kappa^n+2HK_x\int_0^t\left|\left|\rho^n(x,\cdot)-\bar{\rho}(x,\cdot)\right|\right|_u du \tag{4.60}$$

By the Bellman-Gronwall inequality (see Appendix A, Theorem A.3.4), this implies that

$$\left|\left|\rho^n(x,\cdot)-\bar{\rho}(x,\cdot)\right|\right|_t\leq\kappa^n e^{2HK_x t} \tag{4.61}$$

Since $\kappa^n\underset{n\to\infty}{\longrightarrow}0$ almost surely, we have the desired result. That is, the scaled stochastic process converges almost surely to the Godunov scheme based dynamic. In is notable that this strengthens the result derived in [49], in which weak convergence to the fluid limit was established.

4.7 Numerical Example

The purpose of this example is to illustrate the convergence of the scaled stochastic model to the fluid limit. Consider a simple two cell setting, where both cells are of length 264 ft (80.47 m). Assume a triangular fundamental relationship (i.e., a CTM fluid model) with the following parameters: free-flow speed $v_f = 60$ mi/hr (96.56 km/hr), capacity $q_{max} = 1800$ veh/hr, backward wave propagation speed $w = 12$ mi/hr (19.31 km/hr), a jam density of $\rho_{jam} = 180$ veh/mi (111.85 veh/km), and coefficient of variation of time headways of $\bar{c} = 1$; that is, it is assumed that the headways are conditionally exponentially distributed. The simulation time horizon is $U = 200$ seconds. The boundary flow rates (in veh/hr)) are in accordance with the following:

$$\lambda\big(y(0,t)\big) = \min\{1800, 12\big(180 - \rho(1,t)\big)\} \tag{4.62a}$$

$$\lambda\big(y(1,t)\big) = \min\{60\rho(1,t), 1800, 12\big(180 - \rho(2,t)\big)\} \tag{4.62b}$$

$$\lambda\big(y(2,t)\big) = \min\{60\rho(2,t), 1800g(t)\}, \tag{4.62c}$$

where $\rho(1,t)$ and $\rho(1,t)$ are in units of veh/mi. The function $g(t)$ captures downstream capacity restrictions due to the traffic light, which turns red during the time interval [50,70) and green during the remainder of the 200 second time period. Thus, $g(t) = 0$ if $t \in [50, 70)$ and $g(t) = 1$, otherwise. A sample path of the scaled process is simulated using Algorithm 3.1 below.

Algorithm 3.1: Example scaled model simulation

Initialization:
$\rho^n(x,0) := \rho_0(x)$, for all $x \in \mathcal{C}$
Iteration:
while $t < U$ **do**
  for $x = 0 \to |\mathcal{C}|$ **do**
    $\lambda_x := \lambda\big(y^n(x,t)\big)/3600$ (flow in vehicles per second)
    if $\lambda_x == 0$ **then**
      $t_x := \infty$
    else

10 $v_x := uniform(0,1)$ (uniform pseudo-random number in $[0,1]$)
11 $t_x := -\log(1-v_x)/\lambda_x$ (exponential pseudo-random number)
12 $t_x := t_x/n$ (scaled time headways)
13 **end if**
14 **end for**
15 $i := y : t_y == \min_{x \in \mathcal{C}}(t_x)$ ($i :=$ index of the smallest headway)
16 $t := t + t_i$ (update system clock)
17 **Compute (scaled) numbers of vehicles leaving cells:**
18 **for** $x = 0 \rightarrow |\mathcal{C}|$ **do**
19 **if** $x == i$ **then**
20 $q_x^n := 1/n$
21 **else**
22 $q_x^n := 0$
23 **end if**
24 **end for**
25 **Compute cell densities:**
26 **for** $x = 1 \rightarrow |\mathcal{C}|$ **do**
27 $\rho^n(x,t) = \rho^n(x, t - t_i) + (1/l_x)(q_{x-1}^n - q_x^n)$
28 **end for**
29 **end while**

In essence, the algorithm proceeds as follows: (i) compute the fluxes at the cell boundaries, (ii) use the fluxes to generate $|\mathcal{C}|$ exponential pseudo-random numbers (headways), (iii) find the minimum headway and scale it, t_x, (iv) set the flow across the cell boundary with the smallest headway to $1/n$ and all other boundary flows to 0, (v) compute the new densities, and (vi) repeat the process until $t == U$. Figures 4.8 and 4.9 show sample paths of traffic densities for $n = 1$, $n = 10$, $n = 100$, and $n = 1000$, for cell 1 and cell 2, respectively. The figures also show the (mean) time varying traffic densities using the CTM, for comparison.

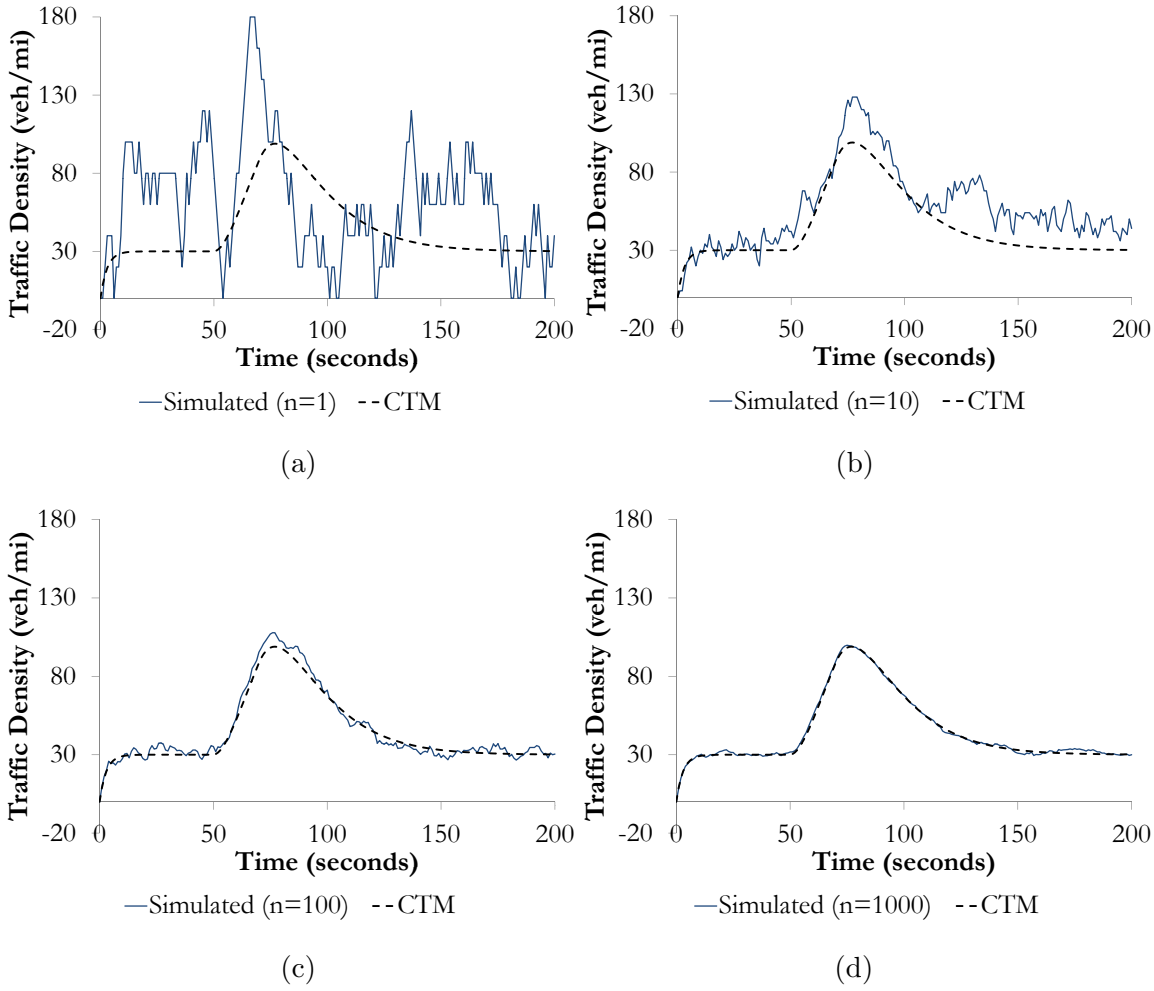
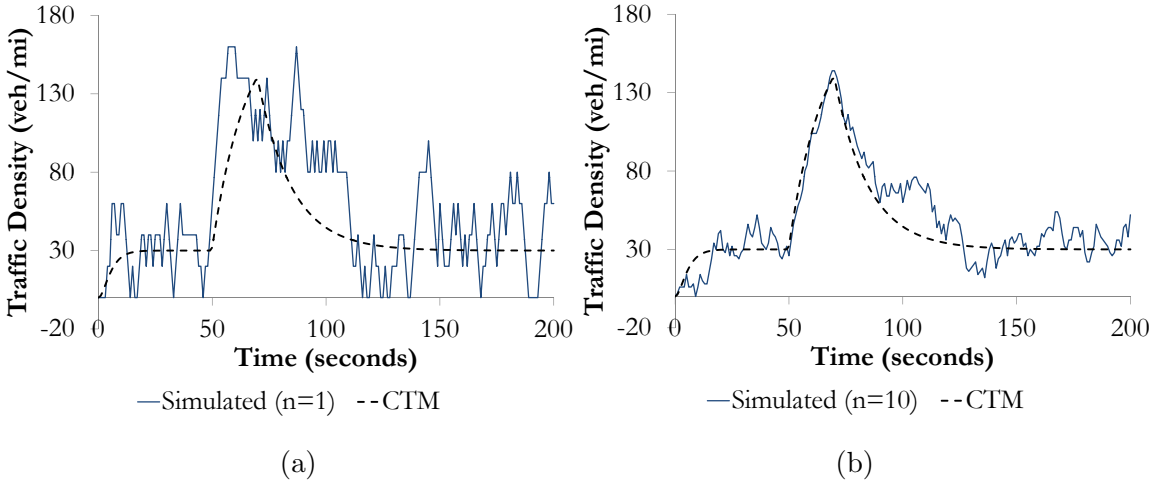


Figure 4.8: Cell 1 traffic densities; (a) $n = 1$; (b) $n = 10$; (c) $n = 100$; (d) $n = 1000$



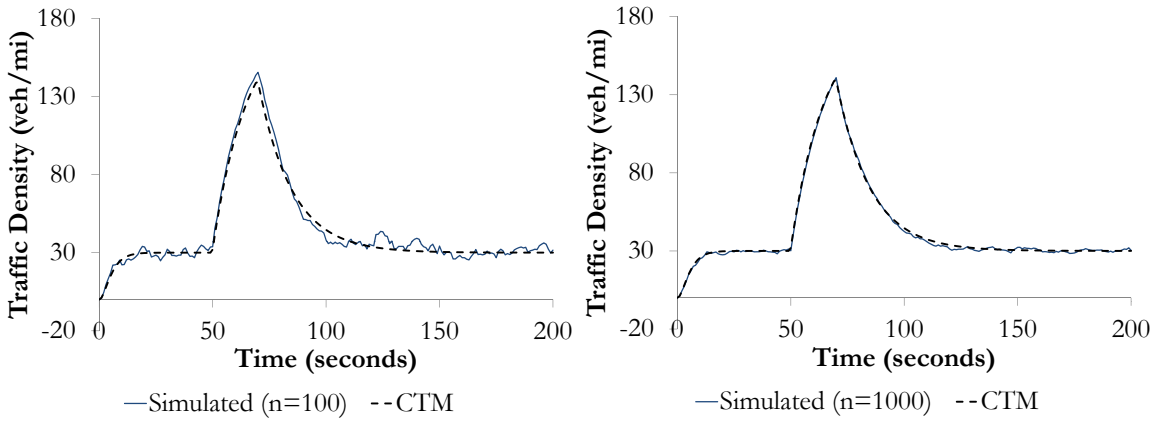


Figure 4.9: Cell 2 traffic densities; (a) $n = 1$; (b) $n = 10$; (c) $n = 100$; (d) $n = 1000$

The sample paths in the two figures illustrate that, as n gets larger and larger, the sample paths get closer and closer to that of the CTM, where the variability in the sample paths decrease with increasing n, just as in Figures 4.5, 4.6, and 4.7. The figures also illustrate how uniform acceleration scaling preserves the transient features of the traffic flow dynamics.

Chapter 5

Gaussian Approximation of the Stochastic Traffic Flow Model

5.1 Introduction

The Gaussian approximation is a second order (stochastic) refinement of the fluid process. The refinement component captures the deviation of the stochastic model from its fluid limit. Then, considering both the mean behavior and the probabilistic deviation from the mean, we obtain a second order (stochastic) approximation of the process. Consider the scaled unit rate Poisson process, $\frac{1}{n}\mathcal{N}_x(nt)$: when n is large, the random variable $\mathcal{N}_x(nt)$ is approximately equal to the sum of $\lfloor nt \rfloor$ independent and identically distributed (i.i.d.) Poisson random variables with mean 1, where $\lfloor nt \rfloor$ is the largest integer less than or equal to nt; that is, let $\{\xi_j\}_{j=1}^{n} \sim$ i.i.d. Poisson(1). Then,

$$\mathcal{N}_x(nt) \approx \sum_{j=1}^{\lfloor nt \rfloor} \xi_j \tag{5.1}$$

Consider the re-scaled centered process (centered at its fluid limit):

$$\sqrt{n}\left(\frac{1}{n}\mathcal{N}_x(nt) - t\right) \tag{5.2}$$

Using (5.1), this is approximated by

$$\frac{1}{\sqrt{n}}\sum_{j=1}^{\lfloor nt\rfloor}(\xi_j - 1) \tag{5.3}$$

By Donsker's theorem (see Appendix A, Theorem A.3.3), (5.3) converges in distribution to standard Brownian motion, denoted by $W_x(t)$, for vehicles leaving cell x (see Appendix A, Definition A.3.3). That is:

$$\sqrt{n}\left(\frac{1}{n}\mathcal{N}_x(nt) - t\right) \overset{\mathcal{D}}{\approx} W_x(t), \tag{5.4}$$

where $\overset{\mathcal{D}}{\approx}$ means "approximately equal, in distribution, to". Defining the scaled time $\tilde{t} \equiv nt$ and multiplying both sides of (5.4) by $\sqrt{n}$, we get:

$$\mathcal{N}_x(\tilde{t}) - \tilde{t} \overset{\mathcal{D}}{\approx} \sqrt{n}W_x\left(\frac{\tilde{t}}{n}\right) \overset{\mathcal{D}}{=} W_x(\tilde{t}), \tag{5.5}$$

where equivalence in distribution ($\overset{\mathcal{D}}{=}$) follows from the scale invariance property[1] of Brownian motion.

This may also be extended, via random time change (see Chapter 4, Theorem 4.6.1), to the more general counting process $\widehat{\mathcal{N}}_x(t)$ (see [18, Theorem 5.11] or [112, Corollary 7.3.1]). That is, let $\zeta_1, \zeta_2, ...$ be sequence of i.i.d. inter-arrival times with mean $\mathbb{E}\zeta_j = 1$ and variance $\mathrm{Var}(\zeta_j) = \sigma^2$ for all j and denote by $\bar{c} = \frac{\sqrt{\mathrm{Var}(\zeta_j)}}{\mathbb{E}\zeta_j}$ the coefficient of variation. Then,

$$\widehat{\mathcal{N}}_x(t) - t \overset{\mathcal{D}}{\approx} \bar{c}W_x(t) \tag{5.6}$$

We have thus obtained that the deviation of the scaled Poisson process from its mean, $\mathcal{N}_x(t) - t$, behaves, in terms of distribution, like standard Brownian motion, $W_x(t)$. Since this holds for any time t, this result may immediately be extended, via time scaling, to a non-homogeneous Poisson process with deterministic time varying rates:

$$\mathcal{N}_x\left(\int_0^t \bar{\lambda}(x,u)du\right) - \int_0^t \bar{\lambda}(x,u)du \overset{\mathcal{D}}{\approx} W_x\left(\int_0^t \bar{\lambda}(x,u)du\right), \tag{5.7}$$

[1] It is this property which motivates the choice of $\sqrt{n}$ as a re-scaling factor. In essence, multiplying the difference in (5.2) by this re-scaling factor plays the role of an amplification of the deviation, which otherwise converges to zero.

where $\bar{\lambda}(x,t)$ are the deterministic time varying rates.

The time changed Brownian motion on the right-hand side of (5.7) can be understood, for a fixed time t, as a normal random variable with mean zero and variance $\int_0^t \bar{\lambda}(x,u)du$. For varying t, $W_x\left(\int_0^t \bar{\lambda}(x,u)du\right)$ may also be written as an Itô integral with deterministic integrand (i.e., a Gaussian process):

$$\int_0^t \sqrt{\bar{\lambda}(x,u)}\, dW_x(u) \tag{5.8}$$

Notice that the variance of (5.8) grows without bound as time progresses. Despite the fact that the derivation of this Gaussian approximation was relatively painless, the approximation has this undesirable feature and cannot be trusted in applications that involve, for instance, traffic prediction as one cannot preclude negative traffic densities and traffic densities which may exceed the jam density. However, as will be shown in the next section, this is not the case when instantaneous flow rates depend on traffic density. The subject of this chapter is the derivation of the Gaussian approximation for the case where instantaneous flow rates depend on traffic density. This derivation is carried out in Section 5.2, while stationary properties of the covariance function are discussed in Section 5.3.

5.2 Derivation of the Gaussian Approximation for the General Case

In this section, the Gaussian approximation of the stochastic model is derived for the general case where the instantaneous flow rates depend on traffic density. The derivation capitalizes on the results given in the previous section for the simple unit rate counting processes. The line of attack here is to consider the general case as a random time change applied to the simple case. Then, the continuous mapping approach given in [112] (see Appendix A, Theorem A.3.7) is used to obtain new convergence results from old.

5.2.1 State Dependent Flow Rates

We first recall the state-dependent conservation equation:

$$\rho^n(x,t) = \rho^n(x,0) \quad + \frac{1}{l_x}\left(\frac{1}{n}\mathcal{N}_{x-1}\left(n\int_0^t \lambda\big(y^n(x-1,u)\big)du\right) - \frac{1}{n}\mathcal{N}_x\left(n\int_0^t \lambda\big(y^n(x,u)\big)du\right)\right), \tag{5.9}$$

and its fluid limit

$$\bar{\rho}(x,t) = \bar{\rho}(x,0) + \frac{1}{l_x}\left(\int_0^t \lambda\big(\bar{y}(x-1,u)\big)du - \int_0^t \lambda\big(\bar{y}(x,u)\big)du\right) \tag{5.10}$$

We are interested in the limiting behavior of the re-scaled process:

$$r^n(x,t) \equiv \sqrt{n}\big(\rho^n(x,t) - \bar{\rho}(x,t)\big) \tag{5.11}$$

Expanding (5.11), we have:

$$\begin{aligned} r^n(x,t) &= \sqrt{n}\big(\rho^n(x,0) - \bar{\rho}(x,0)\big) && \text{(5.12a)}\\ &+ \frac{\sqrt{n}}{l_x}\left(\frac{1}{n}\mathcal{N}_{x-1}\left(n\int_0^t \lambda\big(y^n(x-1,u)\big)du\right) - \int_0^t \lambda\big(\bar{y}(x-1,u)\big)du\right) && \text{(5.12b)}\\ &- \frac{\sqrt{n}}{l_x}\left(\frac{1}{n}\mathcal{N}_x\left(n\int_0^t \lambda\big(y^n(x,u)\big)du\right) - \int_0^t \lambda\big(\bar{y}(x,u)\big)du\right) && \text{(5.12c)} \end{aligned}$$

As discussed in Chapter 4, the first difference in (5.12a) converges in distribution to a zero mean Normal random variable with variance $\sigma^2(x,0)$. Since (5.12b) and (5.12c) are similar, their limiting behavior can be obtained in the same manner. Take the following process:

$$\sqrt{n}\left(\frac{1}{n}\mathcal{N}_x\left(n\int_0^t \lambda\big(y^n(x,u)\big)du\right) - \int_0^t \lambda\big(\bar{y}(x,u)\big)du\right) \tag{5.13}$$

Adding and subtracting $\int_0^t \lambda\big(y^n(x,u)\big)du$, (5.13) can be expressed as the sum of the two differences:

$$\sqrt{n}\left(\frac{1}{n}\mathcal{N}_x\left(n\int_0^t \lambda\big(y^n(x,u)\big)du\right) - \int_0^t \lambda\big(y^n(x,u)\big)du\right) \tag{5.14}$$

and

$$\sqrt{n}\left(\int_0^t \lambda\big(y^n(x,u)\big)du - \int_0^t \lambda\big(\bar{y}(x,u)\big)du\right) \tag{5.15}$$

For (5.14), since the flux functions $\lambda(\cdot)$ are Lipshitz continuous, we immediately obtain the following convergence from the fluid limit:

$$\int_0^t \lambda\big(y^n(x,u)\big)du \underset{n\to\infty}{\longrightarrow} \int_0^t \lambda\big(\bar{y}(x,u)\big)du \quad \text{a.s.} \tag{5.16}$$

by the continuous mapping theorem (see Appendix A, Theorem A.3.7).

Then, again by the continuous mapping theorem, combined with Donsker's theorem (see Appendix A, Theorem A.3.3) and the Skorohod representation theorem (see Appendix A, Theorem A.3.6), we obtain the following for (5.14):

$$\sqrt{n}\left(\frac{1}{n}\mathcal{N}_x\left(n\int_0^t \lambda\big(y^n(x,u)\big)du\right) - \int_0^t \lambda\big(y^n(x,u)\big)du\right) \underset{n\to\infty}{\overset{\mathcal{D}}{\longrightarrow}} W_x\left(\int_0^t \lambda\big(\bar{y}(x,u)\big)du\right), \tag{5.17}$$

where the Brownian processes $W_x(\cdot)$ are independent in x. This is precisely the limit obtained for the non-homogeneous Poisson process presented in Section 5.1. The dependence on traffic states between cells is captured by the limit of (5.15), which is derived next.

Let $\alpha^n(x,t) \equiv [r^n(x,t),\ r^n(x+1,t)]^\mathsf{T}$; then $y^n(x,t) = \bar{y}(x,t) + \frac{1}{\sqrt{n}}\alpha^n(x,t)$, in accordance with (5.11). Now, (5.15) can be written as:

$$\int_0^t \left(\frac{\lambda\left(\bar{y}(x,u) + \frac{1}{\sqrt{n}}\alpha^n(x,u)\right) - \lambda\big(\bar{y}(x,u)\big)}{\frac{1}{\sqrt{n}}}\right)du \tag{5.18}$$

and upon multiplying and dividing by $||\alpha^n(x,t)||$, we get

$$\int_0^t ||\alpha^n(x,u)|| \left(\frac{\lambda\left(\bar{y}(x,u) + \frac{1}{\sqrt{n}}\alpha^n(x,u)\right) - \lambda\big(\bar{y}(x,u)\big)}{\frac{1}{\sqrt{n}}||\alpha^n(x,u)||}\right)du \tag{5.19}$$

Here, $||\alpha^n(x,t)||$ is the "length" of the vector $\alpha^n(x,t)$. The term inside the brackets in (5.19) converges to the directional derivative of $\lambda\big(\bar{y}(x,t)\big)$ along $\tilde{\alpha}(x,t)$, which is the (weak) limit of $\alpha^n(x,t)$. Consequently, (5.19) converges to:[2]

$$\int_0^t \Big(\tilde{\alpha}(x,u)^\mathsf{T} \cdot \nabla\lambda\big(\bar{y}(x,u)\big)\Big)du \tag{5.20}$$

[2] Here, the relationship between the directional derivative and the gradient is used: for two vectors $\mathbf{u}$ and $\mathbf{z}$, and the function $g(\mathbf{u})$, this relationship is:

$$\nabla_{\mathbf{z}} g(\mathbf{u}) = \frac{\mathbf{z}^\mathsf{T}}{||\mathbf{z}||}\nabla g(\mathbf{u})$$

That the product of random variables converges to the product of their limits is established by appeal to Cramér's theorem (see Appendix A, Theorem A.3.5), which is allowed since $\nabla\lambda\big(\bar{y}(x,u)\big)$ is deterministic. Convergence under the integral sign is allowed by the continuous mapping theorem.

Now, expanding the dot product in (5.20), we get:

$$\int_0^t \left(\frac{\partial\lambda\big(\bar{y}(x,u)\big)}{\partial\bar{\rho}(x,u)}\tilde{r}(x,u) + \frac{\partial\lambda\big(\bar{y}(x,u)\big)}{\partial\bar{\rho}(x+1,u)}\tilde{r}(x+1,u) \right) du, \tag{5.21}$$

where $\tilde{r}(x,t)$ and $\tilde{r}(x+1,t)$ are the (weak) limits of $r^n(x,t)$ and $r^n(x+1,t)$, respectively.

We have, thus, found that the deviation of the stochastic traffic flow model from its fluid limit, for cell x, is obtained as the solution of the following stochastic differential equation (written in integral form):

$$\begin{aligned}
\tilde{r}(x,t) = \tilde{r}(x,0) \quad &+ \quad \frac{1}{l_x}\int_0^t \left(\frac{\partial\lambda\big(\bar{y}(x-1,u)\big)}{\partial\bar{\rho}(x-1,u)}\tilde{r}(x-1,u) + \frac{\partial\lambda\big(\bar{y}(x-1,u)\big)}{\partial\bar{\rho}(x,u)}\tilde{r}(x,u) \right) du \\
&- \quad \frac{1}{l_x}\int_0^t \left(\frac{\partial\lambda\big(\bar{y}(x,u)\big)}{\partial\bar{\rho}(x,u)}\tilde{r}(x,u) + \frac{\partial\lambda\big(\bar{y}(x,u)\big)}{\partial\bar{\rho}(x+1,u)}\tilde{r}(x+1,u) \right) du \\
&+ \quad \frac{1}{l_x}\int_0^t \sqrt{\lambda\big(\bar{y}(x-1,u)\big)}\, dW_{x-1}(u) - \frac{1}{l_x}\int_0^t \sqrt{\lambda\big(\bar{y}(x,u)\big)}\, dW_x(u)
\end{aligned} \tag{5.22}$$

The only random terms in this equation are the $\tilde{r}(\cdot,\cdot)$'s and the standard Brownian motions $W_{x-1}(\cdot)$ and $W_x(\cdot)$. Since the stochastic differential equation (SDE) (5.22) involves terms that appear in SDEs for cells $x-1$ and $x+1$, we have a system of SDEs, which should be solved simultaneously. In the next section, the SDE is written in vector form and the solution is derived.

5.2.2 Solution of the SDE

The system of equations (5.22) may be written compactly as:

$$\tilde{\mathbf{r}}(t) = \tilde{\mathbf{r}}(0) + \int_0^t \mathbf{D}(u)\tilde{\mathbf{r}}(u)du + \int_0^t \mathbf{B}\boldsymbol{\Gamma}(u)d\mathbf{W}(u), \tag{5.23}$$

where $\tilde{\mathbf{r}}(t)$ and $\mathbf{W}(t)$ are, respectively, the $|\mathcal{C}|$ and $|\mathcal{C}|+1$ dimensional vector valued processes $\big[\tilde{r}(1,t) \;\cdots\; \tilde{r}(|\mathcal{C}|,t)\big]^\mathsf{T}$ and $\big[W_0(t) \;\cdots\; W_{|\mathcal{C}|}(t)\big]^\mathsf{T}$; here, $W_0(t)$ represents flows into the first

cell. $|\mathcal{C}|$ denotes the number of cells and is also used here to denote the index of the last cell.

The $|\mathcal{C}| \times |\mathcal{C}|$ matrix $\mathbf{D}(t)$ captures the dependence between cells (the second and third components on the right-hand side of (5.22)). For arbitrary cell x, the corresponding row in $\mathbf{D}(t)$ is:

$$\frac{1}{l_x}\left[\begin{array}{ccccccc} \cdots & 0 & \dfrac{\partial\lambda\big(\bar{y}(x-1,t)\big)}{\partial\bar{\rho}(x-1,t)} & \left(\dfrac{\partial\lambda\big(\bar{y}(x-1,t)\big)}{\partial\bar{\rho}(x,t)} - \dfrac{\partial\lambda\big(\bar{y}(x,t)\big)}{\partial\bar{\rho}(x,t)}\right) & -\dfrac{\partial\lambda\big(\bar{y}(x,t)\big)}{\partial\bar{\rho}(x+1,t)} & 0 & \cdots \end{array}\right], \tag{5.24}$$

where the middle element lies in the x^{th} column of the row. The constant $|\mathcal{C}|\times(|\mathcal{C}|+1)$ matrix $\mathbf{B}$ may be interpreted as a normalized routing matrix, since each row contains two non-zero elements which are equal in value but opposite in sign. $\mathbf{B}$ is given by:

$$\mathbf{B} \equiv \left[\begin{array}{ccccccc} 1/l_1 & -1/l_1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1/l_2 & -1/l_2 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1/l_3 & -1/l_3 & \cdots & 0 & 0 \\ & & & & \ddots & & \\ 0 & 0 & 0 & 0 & \cdots & 1/l_{|\mathcal{C}|} & -1/l_{|\mathcal{C}|} \end{array}\right] \tag{5.25}$$

$\mathbf{\Gamma}(t)$ is the $(|\mathcal{C}|+1)\times(|\mathcal{C}|+1)$ matrix of Itô integrands

$$\mathbf{\Gamma}(t) \equiv \left[\begin{array}{ccccc} \sqrt{\lambda\big(\bar{y}(0,t)\big)} & 0 & \cdots & 0 & 0 \\ & & \ddots & & \\ 0 & 0 & \cdots & 0 & \sqrt{\lambda\big(\bar{y}(|\mathcal{C}|,t)\big)} \end{array}\right] \tag{5.26}$$

where $\bar{y}(0,t)$ can be taken to be equal to $\bar{\rho}(1,t)$. Note that in this vector notation dependence on $\bar{y}(\cdot,\cdot)$ has been omitted. This is both to simplify notation and to emphasize that these are deterministic (albeit time-varying) quantities.

The vector stochastic integral equation (5.23) can also be written (symbolically) in differential form as:

$$d\tilde{\mathbf{r}}(t) = \mathbf{D}(t)\tilde{\mathbf{r}}(t)dt + \mathbf{B}\mathbf{\Gamma}(t)d\mathbf{W}(t) \tag{5.27}$$

with the initial condition $\tilde{\mathbf{r}}(0)$ (a zero mean Gaussian random vector); this is a *narrow sense linear* SDE, since all coefficient matrices are (time varying) constant matrices that do not

depend on $\tilde{\mathbf{r}}(t)$ (see Appendix A, Definition A.3.4). The solution is written as:

$$\tilde{\mathbf{r}}(t) = \mathbf{\Phi}(t)\left(\tilde{\mathbf{r}}(0) + \int_0^t \mathbf{\Phi}^{-1}(u)\mathbf{B}\mathbf{\Gamma}(u)d\mathbf{W}(u)\right), \tag{5.28}$$

where $\mathbf{\Phi}(\cdot)$ is a $|\mathcal{C}| \times |\mathcal{C}|$ *fundamental matrix*. That is, $\mathbf{\Phi}(\cdot)$ solves the matrix differential equation:

$$\frac{d\mathbf{\Phi}(t)}{dt} = \mathbf{D}(t)\mathbf{\Phi}(t), \quad \mathbf{\Phi}(0) = \mathbf{I} \tag{5.29}$$

where $\mathbf{I}$ is a $|\mathcal{C}| \times |\mathcal{C}|$ identity matrix. Note that the fundamental matrix $\mathbf{\Phi}(t)$, except in specific situations, is not easy to obtain explicitly. However, as will be shown shortly, this is not an obstacle.

The solution (5.28) is an Itô integral with deterministic integrand; thus, $\tilde{\mathbf{r}}(t)$ is a Gaussian process.

Let $\boldsymbol{\rho}(t)$ denote the $|\mathcal{C}|$-dimensional vector of traffic densities, $\bar{\boldsymbol{\rho}}(t)$, $\tilde{\boldsymbol{\rho}}(t)$ its second-order approximation, and $\boldsymbol{\lambda}(t) \equiv \left[\lambda\big(\bar{y}(0,t)\big) \cdots \lambda\big(\bar{y}(|\mathcal{C}|,t)\big)\right]^{\mathsf{T}}$. From (5.28), we have the approximated traffic density process

$$\tilde{\boldsymbol{\rho}}(t) = \tilde{\boldsymbol{\rho}}(0) + \int_0^t \mathbf{B}\boldsymbol{\lambda}(u)du + \mathbf{\Phi}(t)\left(\tilde{\mathbf{r}}(0) + \int_0^t \mathbf{\Phi}^{-1}(u)\mathbf{B}\mathbf{\Gamma}(u)d\mathbf{W}(u)\right) \tag{5.30}$$

Since $\tilde{\mathbf{r}}(t)$ is Gaussian, then so is $\tilde{\boldsymbol{\rho}}(t)$, which is characterized by it's first two moments. The mean dynamic is

$$\mathbb{E}\tilde{\boldsymbol{\rho}}(t) = \bar{\boldsymbol{\rho}}(t) = \bar{\boldsymbol{\rho}}(0) + \int_0^t \mathbf{B}\boldsymbol{\lambda}(u)du, \tag{5.31}$$

where $\bar{\boldsymbol{\rho}}(0)$ is the mean traffic density at time $t = 0$. Equation (5.31) is simply the fluid limit written in vector form (the Godunov scheme)

The covariance matrix, denoted by $\mathbf{\Psi}(t)$, is

$$\begin{aligned}\mathbf{\Psi}(t) &= \mathbb{E}\Big(\big(\tilde{\boldsymbol{\rho}}(t) - \mathbb{E}\tilde{\boldsymbol{\rho}}(t)\big)\big(\tilde{\boldsymbol{\rho}}(t) - \mathbb{E}\tilde{\boldsymbol{\rho}}(t)\big)^{\mathsf{T}}\Big) = \mathbb{E}\big(\tilde{\mathbf{r}}(t)\tilde{\mathbf{r}}(t)^{\mathsf{T}}\big) \\ &= \mathbf{\Phi}(t)\left(\mathbf{\Psi}(0) + \int_0^t \mathbf{\Phi}^{-1}(u)\mathbf{B}\mathbf{\Gamma}(u)\mathbf{\Gamma}(u)^{\mathsf{T}}\mathbf{B}^{\mathsf{T}}\mathbf{\Phi}^{-1}(u)^{\mathsf{T}}du\right)\mathbf{\Phi}(t)^{\mathsf{T}}\end{aligned} \tag{5.32}$$

Taking the first order derivative of $\mathbf{\Psi}(t)$ with respect to time and using (5.29), we see that $\mathbf{\Psi}(t)$ is obtained by solving the linear matrix differential equation

$$\frac{d\mathbf{\Psi}(t)}{dt} = \mathbf{D}(t)\mathbf{\Psi}(t) + \mathbf{\Psi}(t)\mathbf{D}(t)^{\mathsf{T}} + \mathbf{B}\mathbf{\Gamma}(t)\mathbf{\Gamma}(t)^{\mathsf{T}}\mathbf{B}^{\mathsf{T}} \tag{5.33}$$

with initial covariance matrix $\mathbf{\Psi}(0)$ given by the variances of the initial cell traffic densities. Note that the fundamental matrix $\mathbf{\Phi}(t)$ does not appear in (5.33); more importantly, its inverse does not appear in (5.33) either. This simplifies computation of the covariance matrix, which is discussed in Section 5.3 below.

We have thus obtained two deterministic equations, which can be solved to give the mean and covariance of the approximated process. Since the approximated process is Gaussian, it is fully characterized by these two equations. The dependence of traffic densities in cell x on traffic densities in adjacent cells $x-1$ and $x+1$ is preserved in the Gaussian model (in both the mean and the covariance matrix). Furthermore, the only parameters needed to compute both the mean and the covariance matrix (given initial conditions) are those pertaining to the fundamental diagram.

5.2.3 Generalized Counting Processes

The weak convergence of the doubly stochastic Poisson process was established by applying the continuous mapping theorem to the sequence of scaled homogeneous Poisson processes. That is, the FCLT for the simple case was used to establish the FCLT for the general case. Since we have a FCLT for the simple counting process $\widehat{\mathcal{N}}(t)$ (5.6), the same approach presented above can be applied to generalized counting processes constructed by applying a random time change. That is, we may consider cell boundary cumulative flows of the form:

$$\mathcal{Q}(x,t) \equiv \widehat{\mathcal{N}}_x\left(\int_0^t \lambda\big(y(x,u)\big)du\right) \tag{5.34}$$

The only difference between this case and the doubly stochastic Poisson is inclusion of the coefficient of variation, $\bar{c}$, to the matrix $\mathbf{\Gamma}(t)$[3] , which is now given by:

$$\mathbf{\Gamma}(t) \equiv \begin{bmatrix} \bar{c}\sqrt{\lambda\big(\bar{y}(0,t)\big)} & 0 & \cdots & 0 & 0 \\ & & \ddots & & \\ 0 & 0 & \cdots & 0 & \bar{c}\sqrt{\lambda\big(\bar{y}(|\mathcal{C}|,t)\big)} \end{bmatrix} \tag{5.35}$$

The solution of the SDE remains the same Gaussian process. That is, with $\mathbf{\Gamma}(t)$ given by (5.35), the time varying means and covariances are obtained from (5.31) and (5.33).

[3] For the case of the doubly stochastic Poisson process $\bar{c}=1$.

5.3 Stationary Behavior of the Covariance Function

The stationary behavior of the covariance matrix gives an indication of what the covariance matrix tends to with time. This sheds light on whether or not the covariance grows without bound or is bounded. When the covariance matrix is bounded, the stationary behavior aids in designing numerical schemes for computing the covariance matrix; that is, it helps with choosing appropriate Δt.

As discussed in Section 5.1, for the simple non-stationary Poisson model, variances grow without bound. This is not the case for the approximated state-dependent process. To see this, we first point out that, for stationary mean traffic conditions, we have that the matrices $\mathbf{D}(t) = \mathbf{D}$ and $\mathbf{\Gamma}(t) = \mathbf{\Gamma}$ do not vary with time. Consequently, the fundamental matrix, $\mathbf{\Phi}(t)$ can be written explicitly as $\mathbf{\Phi}(t) = e^{\mathbf{D}t}$ (see for example [1]) and the solution of the stochastic differential equation for the deviation becomes[4] :

$$\tilde{\mathbf{r}}(t) = e^{\mathbf{D}t}\tilde{\mathbf{r}}(0) + \int_0^t e^{(t-u)\mathbf{D}}\mathbf{B}\mathbf{\Gamma} d\mathbf{W}(u), \tag{5.36}$$

for which the covariance matrix is:

$$\mathbf{\Psi}(t) = e^{\mathbf{D}t}\mathbf{\Psi}(0)e^{\mathbf{D}^{\mathsf{T}}t} + \int_0^t e^{(t-u)\mathbf{D}}\mathbf{B}\mathbf{\Gamma}\mathbf{\Gamma}^{\mathsf{T}}\mathbf{B}^{\mathsf{T}}e^{(t-u)\mathbf{D}^{\mathsf{T}}}du \tag{5.37}$$

The behavior of the covariance matrix depends on the traffic conditions. Next, the following three cases are investigated: (i) free-flow traffic conditions (sub-critical mean traffic densities), (ii) capacity conditions (critical mean densities), and (iii) congested traffic conditions (super-critical mean traffic densities).

[4] The matrix exponential is defined by:

$$e^{\mathbf{A}} = \sum_{k=0}^{\infty} \frac{1}{k!}\mathbf{A}^k$$

When $\mathbf{A}$ is diagonal of size j, we have that

$$\mathbf{A} = \begin{bmatrix} a_1 & \cdots & 0 \\ & \ddots & \\ 0 & \cdots & a_j \end{bmatrix} \Longrightarrow e^{\mathbf{A}} = \begin{bmatrix} e^{a_1} & \cdots & 0 \\ & \ddots & \\ 0 & \cdots & e^{a_j} \end{bmatrix}$$

5.3.1 Free-flow (Sub-Critical) Mean Traffic Conditions

Under free-flow traffic conditions, we have, for each cell boundary, that $S_e\big(\bar{\rho}(x,t)\big) < R_e\big(\bar{\rho}(x+1,t)\big)$, where $\bar{\rho}$ is the (sub-critical) mean traffic density in the cells. Consequently,

$$\frac{\partial\lambda\big(\bar{y}(x,t)\big)}{\partial\bar{\rho}(x,t)} = \frac{dS_e\big(\bar{\rho}(x,t)\big)}{d\bar{\rho}(x,t)} \tag{5.38}$$

and

$$\frac{\partial\lambda\big(\bar{y}(x,t)\big)}{\partial\bar{\rho}(x+1,t)} = 0 \tag{5.39}$$

See Chapter 4.1, Section 4.6.2 for formulas of these derivatives. Let

$$\omega_x \equiv \frac{1}{l_x}\frac{dS_e\big(\bar{\rho}(x,t)\big)}{d\bar{\rho}(x,t)}$$

We then see that $\mathbf{D}$ has the following structure:

$$\mathbf{D} = \begin{bmatrix} -\omega_1 & 0 & 0 & 0 & \cdots & 0 & 0 \\ \omega_2 & -\omega_2 & 0 & 0 & \cdots & 0 & 0 \\ 0 & \omega_3 & -\omega_3 & 0 & \cdots & 0 & 0 \\ & & & & \ddots & & \\ 0 & 0 & 0 & 0 & \cdots & \omega_{|\mathcal{C}|} & -\omega_{|\mathcal{C}|} \end{bmatrix} \tag{5.40}$$

This can be written as the sum of two matrices $\mathbf{D} = \mathbf{D}_1 + \mathbf{D}_2$, where $\mathbf{D}_1$ is a diagonal matrix and $\mathbf{D}_2$ is a nilpotent matrix (see Appendix A, Definition A.1.1). Assuming cell lengths are all equal[5] , we see that $\mathbf{D}_1\mathbf{D}_2 = \mathbf{D}_2\mathbf{D}_1$ and, thus, $e^{(\mathbf{D}_1+\mathbf{D}_2)t} = \big(e^{\mathbf{D}_1}e^{\mathbf{D}_2}\big)^t$. Since $\mathbf{D}_2$ is nilpotent, $e^{\mathbf{D}_2}$ is a lower triangular matrix with 1s along the diagonal, while $e^{\mathbf{D}_1}$ is a diagonal matrix with diagonal elements $e^{-\omega_x}$. Consequently, $\big(e^{\mathbf{D}_1}e^{\mathbf{D}_2}\big)$ has spectral radius strictly less than 1 (see Appendix A, Definition A.1.2). Then as $t \to \infty$, $\big(e^{\mathbf{D}_1}e^{\mathbf{D}_2}\big)^t \to \mathbf{0}$ (see Appendix A, Theorem A.1.1). Furthermore, for any $u < t$, $e^{(t-u)\mathbf{D}} \to 0$, and when $u = t$, $e^{(t-u)\mathbf{D}} = \mathbf{I}$. We have just established that:

$$\mathbf{\Psi}(t) \underset{t\to\infty}{\longrightarrow} \mathbf{B}\mathbf{\Gamma}\mathbf{\Gamma}^{\mathsf{T}}\mathbf{B}^{\mathsf{T}}dt, \tag{5.41}$$

which is bounded since $\mathbf{B}$ and $\mathbf{\Gamma}$ are bounded.

[5] This is easily generalized by considering upper and lower bounds established by substituting all cell lengths with the minimum and maximum cell lengths and seeing that both bounds converges to the same limit.

5.3.2 Capacity (Critical) Mean Traffic Condition

When mean traffic conditions are critical, we have that $S_e\big(\bar{\rho}(x,t)\big) = R_e\big(\bar{\rho}(x+1,t)\big)$. Consequently,

$$\frac{\partial\lambda\big(\bar{y}(x,t)\big)}{\partial\bar{\rho}(x,t)} = \frac{1}{2}\frac{dS_e\big(\bar{\rho}(x,t)\big)}{d\bar{\rho}(x,t)} \tag{5.42}$$

and

$$\frac{\partial\lambda\big(\bar{y}(x,t)\big)}{\partial\bar{\rho}(x+1,t)} = \frac{1}{2}\frac{dR_e\big(\bar{\rho}(x+1,t)\big)}{d\bar{\rho}(x+1,t)} \tag{5.43}$$

Assuming again equal cell lengths, let

$$\omega_l \equiv \frac{1}{2l_x}\frac{dS_e\big(\bar{\rho}(x-1,t)\big)}{d\bar{\rho}(x-1,t)},$$

$$\omega_m \equiv \frac{1}{2l_x}\left(\frac{dS_e\big(\bar{\rho}(x,t)\big)}{d\bar{\rho}(x,t)} - \frac{dR_e\big(\bar{\rho}(x,t)\big)}{d\bar{\rho}(x,t)}\right),$$

and

$$\omega_r \equiv \frac{1}{2l_x}\frac{dR_e\big(\bar{\rho}(x+1,t)\big)}{d\bar{\rho}(x+1,t)}$$

Then, $\mathbf{D}$ has the following structure:

$$\mathbf{D} = \begin{bmatrix} -\omega_m & \omega_r & 0 & 0 & \cdots & 0 & 0 \\ \omega_l & -\omega_m & \omega_r & 0 & \cdots & 0 & 0 \\ 0 & \omega_l & -\omega_m & \omega_r & \cdots & 0 & 0 \\ & & & & \ddots & & \\ 0 & 0 & 0 & 0 & \cdots & \omega_l & -\omega_m \end{bmatrix} \tag{5.44}$$

The matrix $\mathbf{D}$ in (5.44) is said to be of *Toeplitz type*, which is any tridiagonal matrix with equal elements along its diagonals. Let $\mathbf{P}$ be a matrix with its columns consisting of eigenvectors of the matrix $\mathbf{D}$. It can be shown that (i) $\mathbf{D}$ has distinct eigenvalues, denoted $\{a_1, ..., a_{|\mathcal{C}|}\}$; (ii) $\mathbf{P}$ diagonalizes $\mathbf{D}$, i.e., $\mathbf{D} = \mathbf{P}\text{diag}(a_1, ..., a_{|\mathcal{C}|})\mathbf{P}^{-1}$; and (iii) the eigenvalues are calculated as $a_j = -\omega_m + 2\omega_r\sqrt{\frac{\omega_l}{\omega_r}}\cos\left(\frac{j\pi}{|\mathcal{C}|+1}\right)$ (see, for instance, [71, Example 7.2.5, page 514] for detailed derivation). Consequently,

$$e^{\mathbf{D}} = \mathbf{P}\text{diag}\left(e^{a_1}, ..., e^{a_{|\mathcal{C}|}}\right)\mathbf{P}^{-1} \tag{5.45}$$

and

$$e^{\mathbf{D}t} = \mathbf{P}\text{diag}\left(e^{a_1 t}, ..., e^{a_{|\mathcal{C}|} t}\right)\mathbf{P}^{-1} \tag{5.46}$$

From the definition of sending and receiving functions, we have that $-\omega_m < 0$, $\omega_r < 0$, and $\cos\left(\frac{j\pi}{|\mathcal{C}|+1}\right) \in (0,1)$ for all $j \in \{1, ..., |\mathcal{C}|\}$. Hence, $\text{diag}\left(e^{a_1 t}, ..., e^{a_{|\mathcal{C}|} t}\right) \to \mathbf{0}$, as $t \to \infty$, then from (5.46), we have that $e^{\mathbf{D}t} \to \mathbf{0}$ as $t \to \infty$. Likewise, for $u < t$, we have that, as $t \to \infty$, $e^{(t-u)\mathbf{D}} \to \mathbf{0}$ and when $t = u$, $e^{(t-u)\mathbf{D}} \to \mathbf{PIP}^{-1} = \mathbf{I}$.

We have just established that the covariance matrix converges to the same limit as the free-flow case, (5.41), for the case of critical mean traffic densities.

5.3.3 Congested (Super-Critical) Mean Traffic Conditions

When mean traffic conditions are super-critical, we have that $S_e\big(\bar{\rho}(x,t)\big) > R_e\big(\bar{\rho}(x+1,t)\big)$ and, consequently,

$$\frac{\partial\lambda\big(\bar{y}(x,t)\big)}{\partial\bar{\rho}(x,t)} = 0 \tag{5.47}$$

and

$$\frac{\partial\lambda\big(\bar{y}(x,t)\big)}{\partial\bar{\rho}(x+1,t)} = \frac{dR_e\big(\bar{\rho}(x+1,t)\big)}{d\bar{\rho}(x+1,t)} \tag{5.48}$$

Now, let

$$\alpha \equiv \frac{1}{l_x}\frac{dR_e}{d\bar{\rho}},$$

then $\mathbf{D}$ has the following structure:

$$\mathbf{D} = \begin{bmatrix} \alpha & -\alpha & 0 & 0 & \cdots & 0 & 0 \\ 0 & \alpha & -\alpha & 0 & \cdots & 0 & 0 \\ 0 & 0 & \alpha & -\alpha & \cdots & 0 & 0 \\ & & & & \ddots & & \\ 0 & 0 & 0 & 0 & \cdots & 0 & \alpha \end{bmatrix} \tag{5.49}$$

Since $\bar{\rho}$ is super-critical, $dR_e/d\bar{\rho} < 0$, which implies that $\alpha < 0$. Consequently, the limiting behavior of $\mathbf{\Psi}(t)$ is also given by (5.41), by the same argument given for the free-flow case.

5.4 Numerical Examples

Consider a simple two cell setting, where both cells are of length 264 ft (80.47 m). Assume a triangular fundamental relationship (i.e., a CTM fluid model) with the following parameters: free-flow speed $v_f = 60$ mi/hr, capacity $q_{max} = 1800$ veh/hr, backward wave propagation speed $w = -12$ mi/hr (-19.31 km/hr), a jam density of $\rho_{jam} = 180$ veh/mi (111.85 veh/km), and coefficient of variation of time headways of $\bar{c} = 1$. The simulation time horizon is $U = 200$ seconds. The boundary flow rates (in veh/hr) are in accordance with the following:

$$\lambda\big(y(0,t)\big) = \min\{\lambda_0, 12\big(180 - \rho(1,t)\big)\} \tag{5.50a}$$

$$\lambda\big(y(1,t)\big) = \min\{60\rho(1,t), 1800, 12\big(180 - \rho(2,t)\big)\} \tag{5.50b}$$

$$\lambda\big(y(2,t)\big) = \min\{60\rho(2,t), 1800g(t)\}, \tag{5.50c}$$

where $\rho(1,t)$ and $\rho(2,t)$ are in units of veh/mi. λ_0 is a constant mean inflow rate into cell 1 and $g(t)$ captures downstream capacity restrictions (e.g., a traffic signal).

The mean traffic density is computed using the CTM and the variance is computed using

$$\mathbf{\Psi}(t + \Delta t) = \mathbf{\Psi}(t) + \mathbf{D}(t)\mathbf{\Psi}(t)\Delta t + \mathbf{\Psi}(t)\mathbf{D}(t)^{\mathsf{T}}\Delta t + \mathbf{B}\mathbf{\Gamma}(t)\mathbf{\Gamma}(t)^{\mathsf{T}}\mathbf{B}^{\mathsf{T}}\Delta t \tag{5.51}$$

with $\Delta t = 0.2$ seconds (for both the mean and the covariance). The matrices $\mathbf{B}$, $\mathbf{\Gamma}(t)$, and $\mathbf{D}(t)$ are computed according to (5.25), (5.26), and (5.24):

$$\mathbf{B} = \begin{bmatrix} 20 & -20 & 0 \\ 0 & 20 & -20 \end{bmatrix}, \tag{5.52}$$

$$\mathbf{\Gamma}(t) = \begin{bmatrix} \sqrt{\lambda\big(\bar{y}(0,t)\big)} & 0 & 0 \\ 0 & \sqrt{\lambda\big(\bar{y}(1,t)\big)} & 0 \\ 0 & 0 & \sqrt{\lambda\big(\bar{y}(2,t)\big)} \end{bmatrix}, \tag{5.53}$$

and

$$\mathbf{D}(t) = \begin{bmatrix} 10\left(\dfrac{\partial\lambda\big(\bar{y}(0,t)\big)}{\partial\bar{\rho}(1,t)} - \dfrac{\partial\lambda\big(\bar{y}(1,t)\big)}{\partial\bar{\rho}(1,t)}\right) & -10\left(\dfrac{\partial\lambda\big(\bar{y}(1,t)\big)}{\partial\bar{\rho}(2,t)}\right) \\ 10\left(\dfrac{\partial\lambda\big(\bar{y}(1,t)\big)}{\partial\bar{\rho}(1,t)}\right) & 10\left(\dfrac{\partial\lambda\big(\bar{y}(1,t)\big)}{\partial\bar{\rho}(2,t)} - \dfrac{\partial\lambda\big(\bar{y}(2,t)\big)}{\partial\bar{\rho}(2,t)}\right) \end{bmatrix}, \tag{5.54}$$

where all derivatives in (5.54) are to be interpreted in the weak sense given in Chapter 4, Section 4.6.2. This is illustrated for the derivative in the first row, in the second column. We may re-write (5.50b) as:

$$\lambda\big(\bar{y}(1,t)\big) = \min\{S_e\big(\bar{\rho}(1,t)\big), R_e\big(\bar{\rho}(2,t)\big)\}, \tag{5.55}$$

where

$$S_e\big(\bar{\rho}(1,t)\big) = \min\{60\bar{\rho}(1,t), 1800\} \tag{5.56a}$$

$$R_e\big(\bar{\rho}(2,t)\big) = \min\{1800, 20\big(180 - \bar{\rho}(2,t)\big)\} \tag{5.56b}$$

This gives

$$\frac{dR_e\big(\bar{\rho}(2,t)\big)}{d\bar{\rho}(2,t)} = \begin{cases} -12 & \text{if } 1800 > 12\big(180 - \bar{\rho}(2,t)\big) \\ -6 & \text{if } 1800 = 12\big(180 - \bar{\rho}(2,t)\big) \\ 0 & \text{if } 1800 < 12\big(180 - \bar{\rho}(2,t)\big) \end{cases}, \tag{5.57}$$

and, consequently,

$$\frac{\partial\lambda\big(\bar{y}(1,t),\beta\big)}{\partial\bar{\rho}(2,t)} = \begin{cases} \dfrac{dR_e\big(\bar{\rho}(2,t)\big)}{d\bar{\rho}(2,t)} & \text{if } S_e\big(\bar{\rho}(1,t)\big) > R_e\big(\bar{\rho}(2,t)\big) \\ \dfrac{1}{2}\dfrac{dR_e\big(\bar{\rho}(2,t)\big)}{d\bar{\rho}(2,t)} & \text{if } S_e\big(\bar{\rho}(1,t)\big) = R_e\big(\bar{\rho}(2,t)\big) \\ 0 & \text{if } S_e\big(\bar{\rho}(1,t)\big) < R_e\big(\bar{\rho}(2,t)\big) \end{cases} \tag{5.58}$$

To illustrate the behavior of the covariance in the model, we look at five scenarios, in which the mean traffic densities converge to different stationary traffic densities:

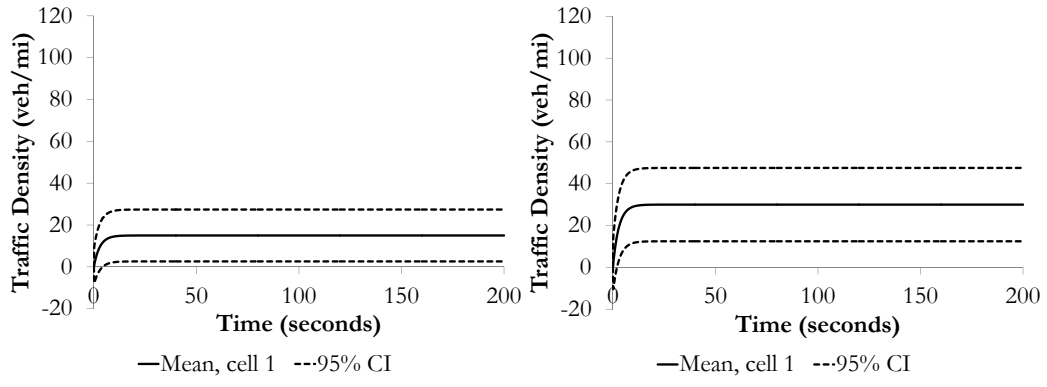
1. Convergence to free-flow traffic conditions: $\lambda_0 = 900$ veh/hr, $g(t) = 1$ for all $t \in [0, 200]$, and zero initial traffic densities and variances.

2. Convergence to critical traffic conditions: $\lambda_0 = 1800$ veh/hr, $g(t) = 1$ for all $t \in [0, 200]$, and zero initial traffic densities and variances.

3. Convergence to congested traffic conditions: $\lambda_0 = 1800$ veh/hr, $g(t) = 0.5$ for all $t \in [0, 200]$, and zero initial traffic densities and variances.

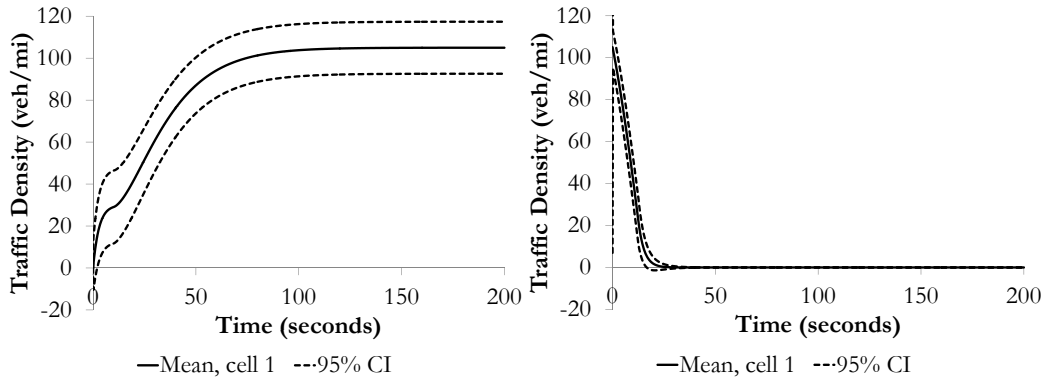
4. Convergence to zero traffic densities: $\lambda_0 = 0$ veh/hr, $g(t) = 1$ for all $t \in [0, 200]$, and mean initial traffic densities of 105 veh/mi with standard deviations of 50 veh/mi.

5. Convergence to jam traffic densities: $\lambda_0 = 1800$ veh/hr, $g(t) = 0$ for all $t \in [0, 200]$, and zero initial traffic densities and variances.



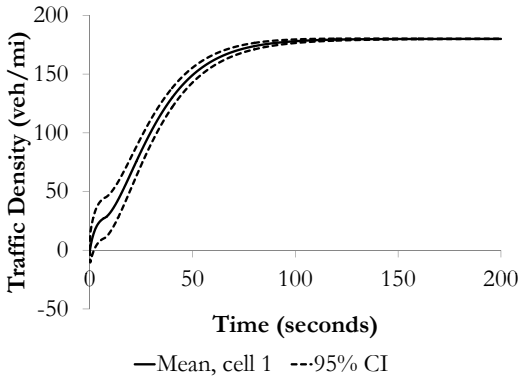
(a)

(b)



(c)

(d)



(e)

Figure 5.1: Stationary traffic densities in cell 1 with 95% confidence intervals; (a) free-flow; (b) capacity flow; (c) congested; (d) zero traffic density; (e) jam traffic density

The traffic densities for the five scenarios along with 95% confidence intervals are shown in Figures 5.1a - 5.1e, for cell 1; confidence intervals of traffic densities in cell 2 behaves in a similar fashion. The figures illustrate both the stationary behavior of the variance of the traffic densities and the speed with which it adapts to changing mean densities. Notice that the variance converges to zero in Figures 5.1d and 5.1e, corresponding to zero mean traffic density and mean jam density. This is interpreted as certainty of zero flows, which indicates that the model will not predict negative traffic densities or traffic densities greater than jam density. In both these cases, the process degenerates and the dynamics become deterministic. It is also notable that carrying out the same calculations using the stationary covariance function given by

$$\mathbf{\Psi}(t+\Delta t) = \mathbf{B}\mathbf{\Gamma}(t)\mathbf{\Gamma}^{\mathsf{T}}(t)\mathbf{B}^{\mathsf{T}}\Delta t \tag{5.59}$$

yields variances that cannot be distinguished from those shown in Figure 5.1 for all five scenarios. The reason for this is that the rate of convergence of (5.51) to (5.59) is proportional to the spectral radius of the matrix $\mathbf{D}(t)$ during each time step (as was discussed in Sections 5.3.1 - 5.3.3), which is much faster than the rate of change in traffic densities.

As an illustration of the transient behavior of the Gaussian model, consider another scenario where a traffic light is introduced at the downstream boundary of cell 2. As in

the numerical example in Section 4.7, suppose the signal turns red during the time interval [50,70) and green during the remainder of the 200 second time period. Thus, $g(t) = 0$ if $t \in [50, 70)$ and $g(t) = 1$, otherwise. The behavior of the traffic densities are shown in Figure 5.2 for both cells 1 and 2, along with 95% confidence intervals. Here, again, we shall assume capacity inflows: $\lambda_0 = 1800$ veh/hr in order to create traffic congestion.

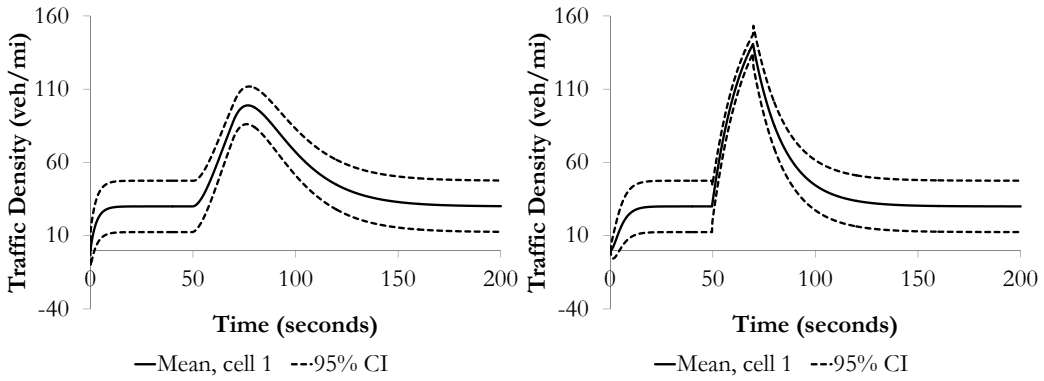


Figure 5.2: Traffic densities with 95% confidence intervals; (a) cell 1, (b) cell 2.

It is also easy to simulate sample paths of the Gaussian process, which is done by simulating normal random variables for each of the time intervals using the computed means and variances. Two such sample paths are illustrated in Figure 5.3.

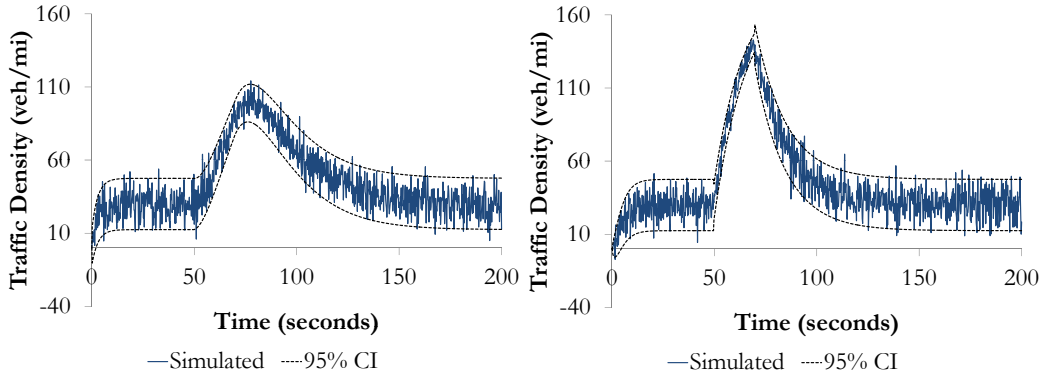


Figure 5.3: Simulated traffic densities with 95% confidence intervals; (a) cell 1, (b) cell 2.

As expected, we see in Figure 5.3 that the two sample paths fall mostly within the computed 95% confidence intervals. A larger confidence interval (e.g., 99%) would entirely encapsulate the sample paths. The variances in these figures are smaller than those for the unscaled queueing process simulated in Section 4.7. The reason for this is the discrete time nature of the simulated process presented here. An alternative method of simulating the queueing process of Section 4.7 would be to simulate a sequence of Poisson random variables over intervals of length $\Delta t = 0.2$ sec. In this case, it would be difficult to distinguish between the sample paths of the Gaussian process and the unscaled queueing process, and the latter would fall mostly within the confidence interval computed using the Gaussian model.

Chapter 6

Traffic State Estimation and Model Validation

6.1 Introduction

The purpose of this chapter is to present the use of the Gaussian model developed in Chapter 5 as a state-space model in a Kalman filter. The estimation problem along with the state-space model and the Kalman filtering algorithm are first presented. As the proposed model requires a large number of cells to give an accurate representation of traffic flow dynamics, observability issues arise, which are also discussed. In general, when the system is observable, further space discretization has no impact on observability. The final section of this chapter presents a real-world estimation example, in which queue sizes are estimated on a cycle by cycle basis along a signalized arterial. Here, a queue size is defined as the number of consecutive vehicles queued behind the stop-line of the intersection which are moving at a speed less than 5 mi/hr (8.05 km/hr). Thus, while other vehicles may be present in the road section, they may not be part of the queue.

6.2 The Estimation Problem

Consider a countable sequence of noisy measurements of traffic flows (and/or traffic densities) at discrete time instants $t_1, t_2, ..., t_u \in [0, U]$. Without loss of generality, it will be assumed that only traffic count information is available. Let $\mathbf{z}(t_j)$ be a vector of such measurements made at time t_j taken at point sensor locations (e.g., loop detectors) that constitute a subset of cell boundaries; that is, for each t_j, $\mathbf{z}(t_j)$ is a vector of dimension $|\mathcal{M}| \leq |\mathcal{C}|$. Let $\boldsymbol{\eta}(t)$ denote an augmented state vector at time $t \in [0, U]$ of dimension $2|\mathcal{C}| + 1$ with the first $|\mathcal{C}|$ elements consisting of the traffic densities in the cells and the remaining elements consisting of the cell boundary traffic flows.

The stochastic filtering problem is one that seeks to determine a best estimate of the state vector $\boldsymbol{\eta}(t)$ given the measurement sequence $\mathbf{Z}(t) = \{\mathbf{z}(t_j), ...\mathbf{z}(t_j)\}$, where $t_j = \max_{1 \leq k \leq u} \{t_k : t_k \leq t\}$. It is well known that the optimal estimate (in the minimal mean squares sense) of the state vector given the measurement sequence is the conditional expectation $\mathbb{E}\big(\boldsymbol{\eta}(t) \big| \mathbf{Z}(t)\big)$. The problem is then one of determining the (time-varying) conditional probability density function $p\big(\boldsymbol{\eta}(t), t \;\big|\; \mathbf{Z}(t)\big)$ [6, 50, 114]. The Kalman filter [51, 52] is an efficient recursive algorithm for computing the first two moments of $\mathbb{E}\big(\boldsymbol{\eta}(t) \;\big|\; \mathbf{Z}(t)\big)$ when $p\big(\boldsymbol{\eta}(t), t \;\big|\; \mathbf{Z}(t)\big)$ is a Gaussian probability density. In general, the Kalman filter consists of iteratively carrying out the following steps:

1. Prediction of the mean state vector at time t_{k+1} from the most recent estimate calculated at time t_k.

2. Prediction of the covariance matrix of the state vector at time t_{k+1} from the most recent estimate calculated at time t_k.

3. Calculation of the filter gain matrix.

4. Calculation of the mean state estimate at time t_{k+1}.

5. Calculation of the estimated covariance matrix of the state vector at time t_{k+1}.

The first two steps are carried out using the Gaussian model derived in Chapter 5. The third step uses the measurements available at time t_{k+1} to compute optimal weighting factors (filter gain) that combine the measurements and the predictions. Then, using the filter gain, the optimal estimates of the mean and the covariance matrix at time t_{k+1} are calculated and the procedure is repeated for the next time step. Next, the components of the filter (namely, the state vector equations and the measurement equations) are specified.

6.3 Model Specification and Filtering Algorithm

Let $\tilde{\mathbf{Q}}(t)$ denote a vector of approximated cumulative cell boundary flows, then the state-space vector is $\boldsymbol{\eta}(t) = [\tilde{\boldsymbol{\rho}}(t)\ \tilde{\mathbf{Q}}(t)]^{\mathsf{T}}$. Let $\hat{\boldsymbol{\Psi}}(t)$ denote the covariance matrix of $\tilde{\mathbf{Q}}(t)$, which is derived in a similar way to $\boldsymbol{\Psi}(t)$. That is, for cell x, start with the centered difference

$$o^n(x,t) = \sqrt{n}\left(\frac{1}{n}\widehat{\mathcal{N}}_x\left(n\int_0^t \lambda\big(y^n(x,u)\big)du\right) - \int_0^t \lambda\big(\bar{y}(x,u)\big)du\right) \tag{6.1}$$

Then, as was shown in Section 5.2.1 that this converges to the solution of the following SDE:

$$\tilde{o}(x,t) = \int_0^t \left(\frac{\partial\lambda\big(\bar{y}(x,u)\big)}{\partial\bar{\rho}(x,u)}\tilde{o}(x,u) + \frac{\partial\lambda\big(\bar{y}(x,u)\big)}{\partial\bar{\rho}(x+1,u)}\tilde{o}(x+1,u)\right)du + \int_0^t \bar{c}\sqrt{\lambda\big(\bar{y}(x,u)\big)}dW_x(u), \tag{6.2}$$

where $\tilde{o}(x,t)$ is the weak limit of $o^n(x,t)$. The system of equations is written (symbolically) in vector form as:

$$d\tilde{\mathbf{o}}(t) = \hat{\mathbf{D}}(t)\tilde{\mathbf{o}}(t)dt + \boldsymbol{\Gamma}(t)d\mathbf{W}(t), \tag{6.3}$$

where $\tilde{\mathbf{o}}(t)$ is a vector of dimension $|\mathcal{C}|+1$, $\boldsymbol{\Gamma}(t)$ and $\mathbf{W}(t)$ retain their meaning given in Section 5.2.2. $\hat{\mathbf{D}}(t)$ is the following $(|\mathcal{C}|+1)\times(|\mathcal{C}|+1)$ matrix:

$$\hat{\mathbf{D}}(t) \equiv \begin{bmatrix} 0 & \dfrac{\partial\lambda\big(\bar{y}(0,t)\big)}{\partial\bar{\rho}(1,t)} & 0 & \cdots & 0 & 0 \\ 0 & \dfrac{\partial\lambda\big(\bar{y}(1,t)\big)}{\partial\bar{\rho}(1,t)} & \dfrac{\partial\lambda\big(\bar{y}(1,t)\big)}{\partial\bar{\rho}(2,t)} & \cdots & 0 & 0 \\ & & & \ddots & & \\ 0 & 0 & 0 & \cdots & 0 & \dfrac{\partial\lambda\big(\bar{y}(|\mathcal{C}|,t)\big)}{\partial\bar{\rho}(|\mathcal{C}|,t)} \end{bmatrix} \tag{6.4}$$

Following the same steps in Section 5.2.2, we have that

$$\mathbb{E}\tilde{\mathbf{Q}}(t) = \int_0^t \boldsymbol{\lambda}(u)du, \tag{6.5}$$

and the covariance matrix is obtained by solving:

$$\frac{d\hat{\boldsymbol{\Psi}}(t)}{dt} = \hat{\mathbf{D}}(t)\hat{\boldsymbol{\Psi}}(t) + \hat{\boldsymbol{\Psi}}(t)\hat{\mathbf{D}}(t)^{\mathsf{T}} + \boldsymbol{\Gamma}(t)\boldsymbol{\Gamma}(t)^{\mathsf{T}} \tag{6.6}$$

Let $\boldsymbol{\Theta}(t) = \mathbb{E}\Big(\big(\tilde{\boldsymbol{\rho}}(t) - \mathbb{E}\tilde{\boldsymbol{\rho}}(t)\big)\big(\tilde{\mathbf{Q}}(t) - \mathbb{E}\tilde{\mathbf{Q}}(t)\big)^{\mathsf{T}}\Big)$ denote the covariance of $\tilde{\boldsymbol{\rho}}(t)$ and $\tilde{\mathbf{Q}}(t)$. Following the same steps, we find that $\boldsymbol{\Theta}(t)$ is obtained by solving:

$$\frac{d\boldsymbol{\Theta}(t)}{dt} = \mathbf{D}(t)\boldsymbol{\Theta}(t) + \boldsymbol{\Theta}(t)\hat{\mathbf{D}}(t)^{\mathsf{T}} + \mathbf{B}\boldsymbol{\Gamma}(t)\boldsymbol{\Gamma}(t)^{\mathsf{T}} \tag{6.7}$$

Then, the state vector, a Gaussian vector, is characterized as follows:

$$\boldsymbol{\eta}(t) = \begin{bmatrix} \tilde{\boldsymbol{\rho}}(t) \\ \tilde{\mathbf{Q}}(t) \end{bmatrix},\ \mathbb{E}\boldsymbol{\eta}(t) = \begin{bmatrix} \bar{\boldsymbol{\rho}}(0) + \int_0^t \mathbf{B}\boldsymbol{\lambda}(u)du \\ \int_0^t \boldsymbol{\lambda}(u)du \end{bmatrix},\ \boldsymbol{\Sigma}(t) = \begin{bmatrix} \boldsymbol{\Psi}(t) & \boldsymbol{\Theta}(t) \\ \boldsymbol{\Theta}(t)^{\mathsf{T}} & \hat{\boldsymbol{\Psi}}(t) \end{bmatrix} \tag{6.8}$$

To characterize the measurement sequence $\{\mathbf{z}(t_i)\}$, let $\mathbf{H}$ be an $|\mathcal{M}| \times (2|\mathcal{C}|+1)$ incidence matrix with a 1 in row i column j if the ith element in the measurement vector represents a measurement of the jth state variable, and a 0 in row i column j otherwise. Here, it will be assumed that measurement error can be represented by Gaussian $|\mathcal{M}|$-dimensional noise vectors $\{\boldsymbol{\zeta}(t_i)\}$ with zero mean and covariance matrices $\{\boldsymbol{\Xi}(t_i)\}$. The true nature of measurement errors vary by sensor type, time of day, weather conditions, and controller sensitivity settings. A thorough investigation of measurement error, while enlightening, is beyond the scope of this thesis and the assumption of Gaussian noise is in accord with the convention in the literature on traffic state estimation [31, 36, 37, 53, 73, 75, 103, 104, 106, 107]. Thus, the measurement equation is given by:

$$\mathbf{z}(t_i) = \mathbf{H}\boldsymbol{\eta}(t_i) + \boldsymbol{\zeta}(t_i), \qquad i \in \{1, 2, ..., u\} \tag{6.9}$$

The state equations (6.8) are continuous time equations, while the measurement equations (6.9) are discrete. Then, traffic state estimation is carried out using a discrete-continuous Kalman filter [50]. With the discrete-continuous Kalman filter, the state equation is solved over intervals defined by the measurement sequence; i.e., the prediction intervals are defined by the measurement times. At the measurement times, the estimates

are updated in the same way this is done with a discrete Kalman filter. The algorithm is included below.

Algorithm 5.1: Discrete-continuous Kalman filter

1: **Initialization:**

2: $\mathbb{E}\boldsymbol{\eta}(0) = \begin{bmatrix} \mathbb{E}\tilde{\boldsymbol{\rho}}(0) \\ \mathbb{E}\tilde{\mathbf{Q}}(0) \end{bmatrix} = \begin{bmatrix} \bar{\boldsymbol{\rho}}(0) \\ \mathbf{0} \end{bmatrix}$

3: $\boldsymbol{\Sigma}(0) = \begin{bmatrix} \boldsymbol{\Psi}(0) & \boldsymbol{\Theta}(0) \\ \boldsymbol{\Theta}(0)^{\mathsf{T}} & \hat{\boldsymbol{\Psi}}(0) \end{bmatrix}$

4: **Iteration:**

5: **for** $i = 1 \to u$ **do**

6: **Predict:**

7: Predict mean $\mathbb{E}\boldsymbol{\eta}(t_i|t_{i-1}) := \begin{bmatrix} \bar{\boldsymbol{\rho}}(t_{i-1}) + \int_{t_{i-1}}^{t_i} \mathbf{B}\boldsymbol{\lambda}(u)du \\ \int_{t_{i-1}}^{t_i} \boldsymbol{\lambda}(u)du \end{bmatrix}$

8: Predict covariance $\boldsymbol{\Sigma}(t_i|t_{i-1}) := \begin{bmatrix} \boldsymbol{\Psi}(t_i|t_{i-1}) & \boldsymbol{\Theta}(t_i|t_{i-1}) \\ \boldsymbol{\Theta}(t_i|t_{i-1})^{\mathsf{T}} & \hat{\boldsymbol{\Psi}}(t_i|t_{i-1}) \end{bmatrix}$

9: **Update:**

10: Measurement residual $\mathbf{m}(t_i) := \mathbf{z}(t_i) - \mathbf{H} \cdot \mathbb{E}\boldsymbol{\eta}(t_i|t_{i-1})$

11: Residual covariance $\boldsymbol{\Pi}(t_i) := \mathbf{H}\boldsymbol{\Sigma}(t_i|t_{i-1})\mathbf{H}^{\mathsf{T}} + \boldsymbol{\Xi}(t_i)$

12: Kalman gain $\mathbf{K}(t_i) := \boldsymbol{\Sigma}(t_i|t_{i-1})\mathbf{H}^{\mathsf{T}}\boldsymbol{\Pi}(t_i)^{-1}$

13: Updated (posterior) mean $\mathbb{E}\boldsymbol{\eta}(t_i|t_i) := \mathbb{E}\boldsymbol{\eta}(t_i|t_{i-1}) + \mathbf{K}(t_i)\mathbf{p}(t_i)$

14: Updated (posterior) covariance $\boldsymbol{\Sigma}(t_i|t_i) := \big(\mathbf{I} - \mathbf{K}(t_i)\mathbf{H}\big)\boldsymbol{\Sigma}(t_i|t_{i-1})$

15: **end for**

The ability to use a classical (as opposed to an extended) Kalman filter stems from the linearity of the stochastic traffic flow model. This is due to dependence of the mean and the covariance on mean traffic variables rather than the stochastic traffic variables. The continuous time setting in which the models were derived offer the flexibility of using different computational time scales for the state equation and the measurement equation. That is, the availability of measurements at regular time intervals is not required in order to run the filter. An important contribution of the model to the filtering problem is the ability

to compute state covariance matrices using few parameters (parameters of the fundamental diagram and coefficients of variation) and the ability to capture correlations between traffic variables in different cells.

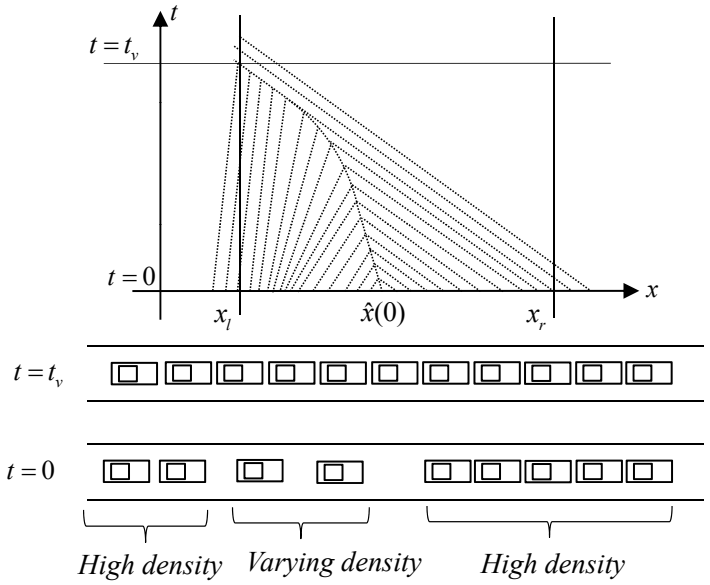
6.4 Observability

A system is said to be *observable* if the available measurement sequence, $\mathbf{Z}(t)$, is adequate to construct the initial conditions of the (mean) state vector, $\mathbb{E}\boldsymbol{\eta}(0)$ [50, 101]. Since the proposed research involves dividing a roadway into (potentially many) cells, observability becomes an important question, as this discretization results in larger numbers of state variables and increases the sparsity of the measurements.

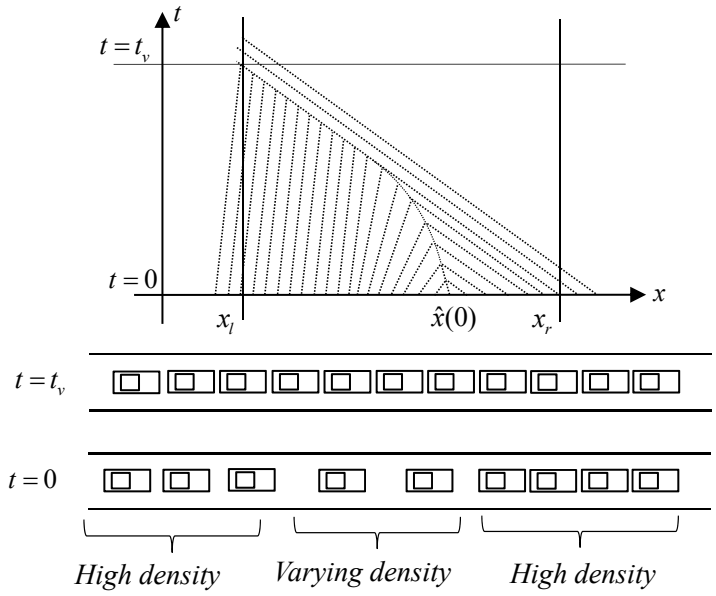
Consider the scenario where a road section is instrumented at either end, so that traffic densities at the upstream and downstream ends of the section and flow rates into the section and out of the section are observed. In the literature, several articles have noted that scenarios where, at time $t = 0$, the traffic density at the upstream end of the road section is sub-critical, while the traffic density at the downstream end is super-critical, result in unobservable traffic densities within the road section [72, 73, 100, 103]. These unobservable scenarios occur, for instance, when a rarefaction fan meets a shockwave within the road section, creating a wave front with varying slope. Such scenarios occur frequently along congested signalized arterials, where queues can start to build-up in the downstream (creating a shockwave that propagates upstream) before a queue in the upstream from a previous cycle has fully dissipated. This is illustrated in Figure 6.1.

While in both Figures 6.1a and 6.1b traffic conditions throughout the road section at time t_v are the same and despite seeing the same traffic conditions at positions x_l and x_r, the initial traffic conditions (at time t_0) are different. In Figure 6.1a, the shockwave meets the rarefaction fan farther upstream at time $t = 0$ than in Figure 6.1b. The position $\hat{x}(0)$ represents the back of a queue, which cannot be observed using sensors at positions x_l and x_r.

In general, the ability to observe traffic conditions within the road section using only



(a)

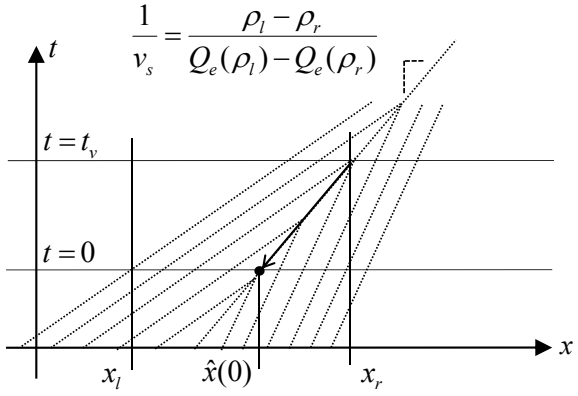


(b)

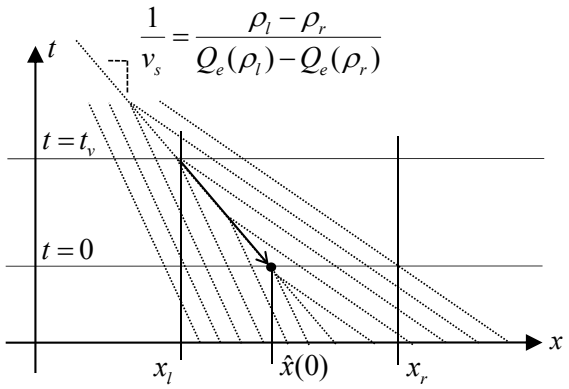
Figure 6.1: Example unobservable scenario; (a) larger queue; (b) smaller queue

observations at its ends is only possible when there is a single wavefront/rarefaction fan within the road section at time $t = 0$. As an example, suppose, the observed traffic densities

at the upstream and downstream positions, denoted, respectively, by ρ_l and ρ_r, are such that $\rho_l < \rho_r < \rho_{crit}$. Then there is a shockwave that was initiated at a position $\hat{x}(0)$ at time $t = 0$ and has a speed of $v_s = \frac{\rho_l - \rho_r}{Q_e(\rho_l) - Q_e(\rho_r)}$. If the system is observable then at some time t_v, the shockwave will be detected in the downstream (position x_r). In essence, this is the time that the traffic density at x_r changes from ρ_r to ρ_l. With t_v and v_s known, $\hat{x}(0)$ can be calculated. This is illustrated in Figure 6.2a.



(a)



(b)

Figure 6.2: Example shockwave observability; (a) $\rho_l < \rho_r < \rho_{crit}$; (b) $\rho_{crit} < \rho_l < \rho_r$

The arrows in Figures 6.2a and 6.2b indicate how the position of the shockwave at time 0, $\hat{x}(0)$, indicated by the solid dots in the figure, are determined from the position of the shockwave at a later time. The directions of the arrows are not to be interpreted as the

direction of the shockwave.

In both scenarios, the initial traffic densities throughout the road section are those observed at the ends; then, estimating the initial traffic densities is a question of determining how far downstream of the entrance of the road section are the traffic densities equal to those at the entrance. The same is true for the case where $\rho_{crit} < \rho_l < \rho_r$, except now the wave is detected at the upstream end of the road section (i.e., position x_r), which is depicted in Figure 6.2b. In both situations, using a larger number of cells results in a more accurate estimate of the position $\hat{x}(0)$ and, hence, more accurate initial traffic densities.

To overcome observability issues in this research, the start time of the estimation period is chosen such that free-flow traffic conditions prevail. Then, at some finite time in the future, traffic conditions observed at the upstream end of the road section will have propagated through the road section, while vehicles that were present within the section at time $t = 0$ will have left. In essence, this constitutes a warm-up period for the estimation process. If estimation of peak period is desired, the start time of the warm-up period should be chosen well before the onset of traffic congestion.

Observability, thus far, has been discussed in terms of the mean dynamics. This is the classical way defining of observability. For stochastic systems, a more relevant issue is *stochastic observability*, which is defined in terms of the behavior of the covariance matrix. A system is said to be stochastically observable if there exists a time, $t_o < \infty$ such that the largest element of the estimated covariance matrix $\mathbf{\Sigma}(t_i|t_i)$ is bounded for all $t_i \geq t_0$ (see [5] and references therein). The boundedness of $\mathbf{\Psi}(\cdot)$ was established in Section 5.3; following the same steps, it can be shown that $\hat{\mathbf{\Psi}}(\cdot)$ and $\mathbf{\Theta}(\cdot)$ are also bounded, since these two matrices are very similar in structure to $\mathbf{\Psi}(\cdot)$. Since the residual covariance matrix, $\mathbf{\Pi}(t_i)$, and the Kalman gain matrix, $\mathbf{K}(t_i)$, are both calculated directly from the predicted covariance matrix, $\mathbf{\Sigma}(t_i|t_{i-1})$, which is bounded, the estimated covariance matrix, $\mathbf{\Sigma}(t_i|t_i)$, is also bounded (see step 14 of Algorithm 5.1). This establishes stochastic observability of the system.

6.5 Model Testing and Validation

In this section, validation of the proposed model is carried out using real-world traffic data. In this test, maximum queue sizes are estimated along the westbound direction of Trunk Highway 55, a high-speed signalized arterial in Minnesota. The maximum queues sizes were estimated on a cycle by cycle basis for the (actuated) signalized intersection of Rhode Island Avenue (the west-most intersection in Figure 6.3) and during the morning peak period of 7:00AM - 9:00AM on December 10th, 2008. Inductance loop detector data and signal timing information were obtained using the SMART-Signal system described in [64]. A 30 minute warm-up period starting at 6:30AM was used, where initial traffic densities were assumed to be zero throughout the study road section.

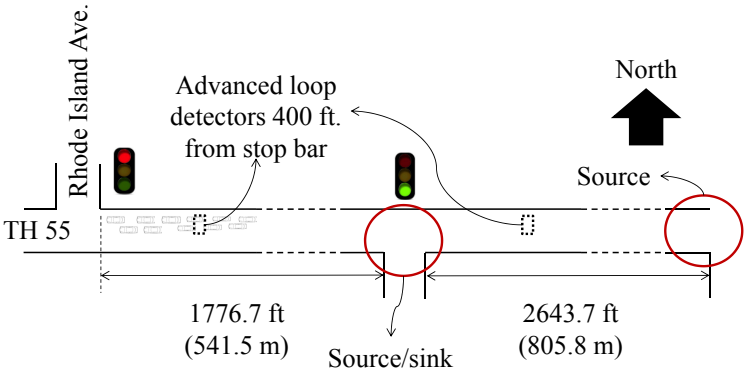


Figure 6.3: Data collection site and detector locations

Data that was used for estimation consists primarily of source and sink counts and advanced loop detector counts. Queues typically build up beyond the locations of the advanced detector and thus cannot be observed using the loop detector counts. For this reason, a manual queue size data collection effort was carried out by a Minneapolis-based transportation consulting firm, Alliant Engineering Inc., during the morning peak period on December 10th, 2008 [65]. This data set was used for comparison with the model estimates.

6.5.1 Model Parameters

Model parameters were fitted prior to estimation. The model parameters include fundamental diagram parameters and the coefficient of variation of time headways. For the fundamental diagram, a triangular *mean* relationship was assumed. The parameters are the free-flow speed v_f, the capacity q_{max}, and the jam density ρ_{jam}. These parameters were fitted using a minimum least squares estimate. The results are illustrated in Figure 6.4, and summarized in Table 6.1.

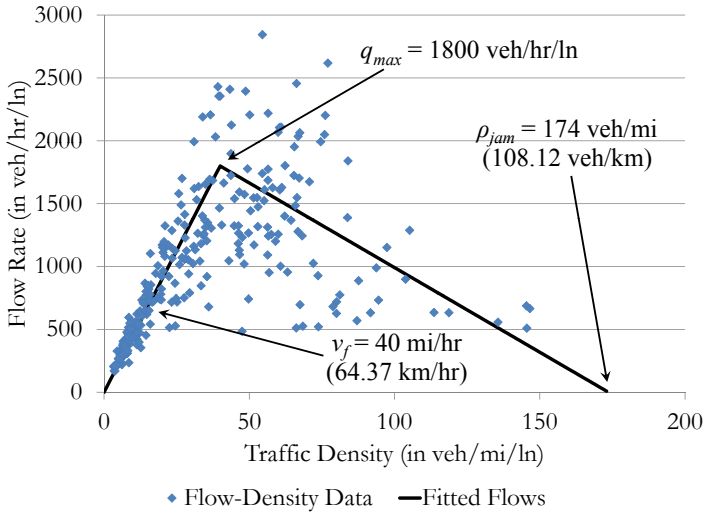


Figure 6.4: Fitted flow-density relationship

Table 6.1: Fitted flow-density relation parameters

v_f	q_{max}	ρ_{jam}
40 mi/hr (64.37 km/hr)	1800 veh/hr/ln	174 veh/mi (108.12 veh/km)

The dataset used to fit the fundamental diagram parameters consists of individual vehicle arrival times to the detector station and detector occupation times. Flow rates and occupancies in the scatter-plot in Figure 6.4 were averaged over 10-vehicle batches as follows:

$$\lambda_i^{(obs)} = \frac{10}{T_i^{(obs)}}, \tag{6.10}$$

where $\lambda_i^{(obs)}$ is the observed flow rate for the ith 10-vehicle batch and $T_i^{(obs)}$ is the difference between the departure time of the first vehicle in the batch and the last vehicle in the batch. Observed occupancies were computed as:

$$\bar{O}_i^{(obs)} = \frac{1}{T_i^{(obs)}} \sum_{j=1}^{10} O_j^{(obs)}, \tag{6.11}$$

where $O_j^{(obs)}$ is the occupancy time of the jth vehicle in the batch and $\bar{O}_i^{(obs)}$ is the average occupancy for the ith 10-vehicle batch (i.e., the percentage of time the detector is occupied).

The reason for aggregating was to average out the effect of unreasonable time headways that are due to measurement error (e.g., over-counting or under-counting when vehicles change lanes at the detector locations). The choice of ten as the batch size was arbitrary. TH-55 is high-speed arterial with a speed limit of 50 mph; the small fitted value of 40 mi/hr is a result of lower speeds due to snow accumulation on the road on the day the data was collected.

Note that the heavy scatter seen in Figure 6.4 is characteristic of traffic flow data that is averaged over short time intervals. Consequently, choosing a different shape can only provide slight improvements to the overall fit and only negligible improvements to the queue size estimates.

To fit the coefficient of variation parameter, $\bar{c}$, sets of consecutive vehicle time headways were collected and binned according to their associated vehicle occupancy times (to check dependency of time headways on traffic density). Probability plots were developed using the statistical software package Minitab in order to determine the appropriate time headway distributions. Figure 6.5 is a probability plot for time headways associated with small occupancies (free-flow traffic conditions) fitted to a 3-parameter log-normal distribution with $\bar{c} = 0.593$. The fitted distribution parameters are given in Table 6.2 along with the Anderson-Darling goodness of fit statistic.[1] Figure 6.6 is a probability plot for time headways associated with large occupancies (congested traffic conditions) also fitted to a 3-parameter log-normal with $\bar{c} = 0.544$ and Table 6.3 summarizes the fitted parameters. In

[1] The Anderson-Darling statistic provides a measure of the difference between the hypothesized distribution and the empirical distribution of the data while assigning relatively higher weights to the tails.

general, the 3-parameter log-normal distribution was found to give good fits for a variety of time headway datasets; the data points mostly fall within the 95% confidence region as illustrated in the probability plots. The Anderson-Darling statistics in both cases indicate a good fit as well.[2] Despite requiring different distribution parameters, a coefficient of variation of 0.5 - 0.6 seemed reasonable for a variety of traffic conditions.

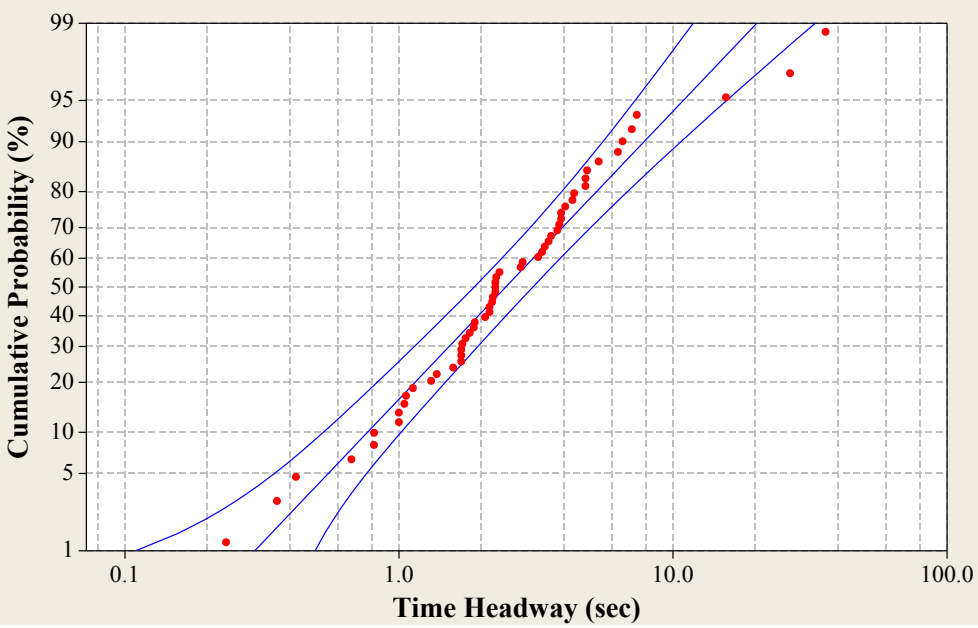


Figure 6.5: Probability plot of free-flow time headways with 95% confidence intervals

Table 6.2: Fitted free-flow headway parameters

Fitted parameters			Goodness of fit
Location	Scale	Shift	Anderson-Darling
0.8971	0.9069	4.001	0.626

[2] That is, we cannot reject the hypotheses that the data points are distributed according to 3-parameter log-normals.

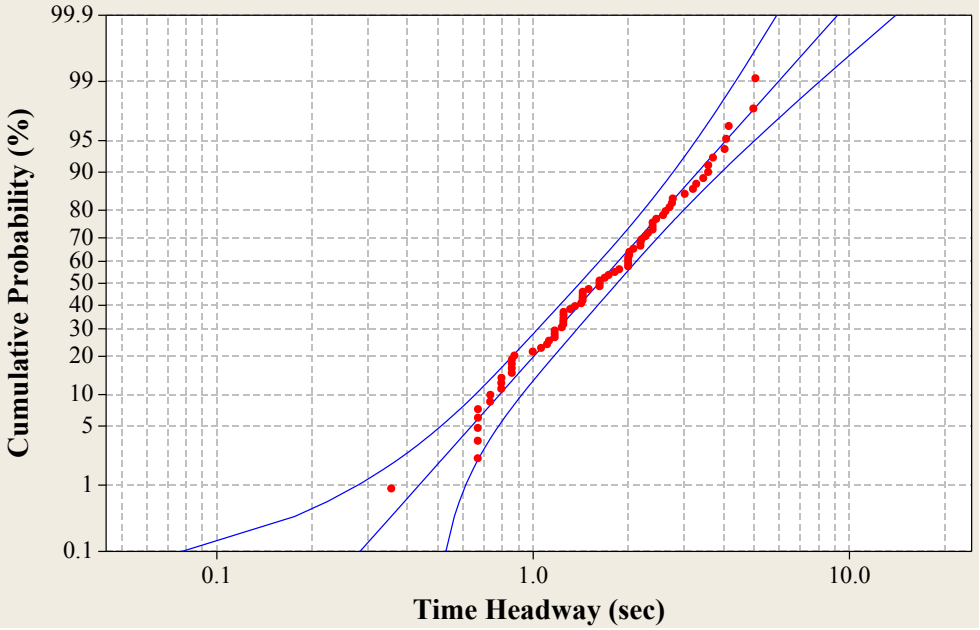


Figure 6.6: Probability plot of congested time headways with 95% confidence intervals

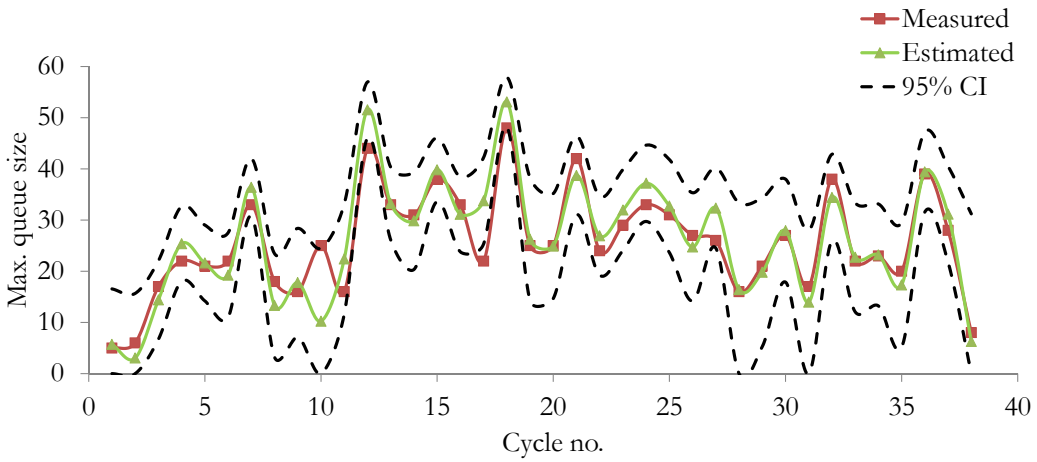
Table 6.3: Fitted congested headway parameters

Fitted parameters			Goodness of fit
Location	Scale	Shift	Anderson-Darling
0.4772	0.5624	0.9723	0.348

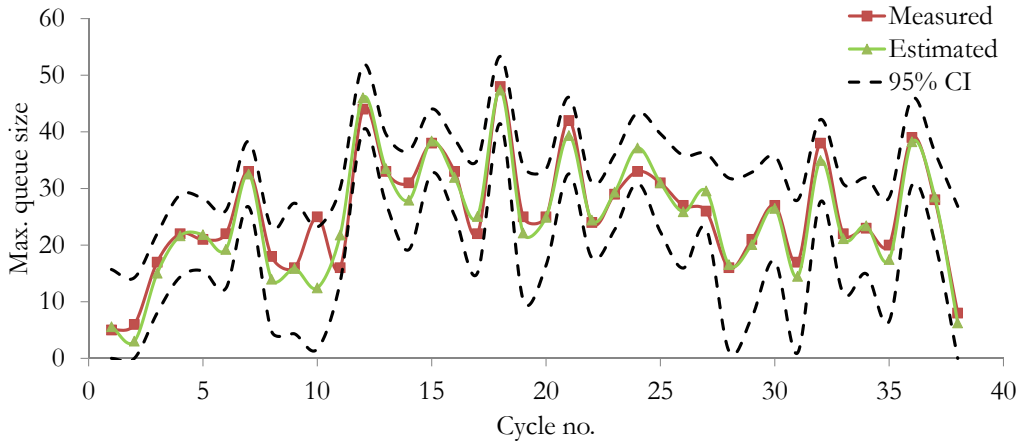
6.5.2 Estimated Cycle-by-Cycle Maximum Queue Sizes

With the parameters of the fundamental diagram and the coefficient of variation estimated, the traffic state dynamics (6.8) are fully characterized. For the measurement equation, it was assumed that measurement errors are uncorrelated and assumed a standard deviation of 5% of the measured value to describe the measurement covariance. This value is based on studies carried out in Minnesota and other parts of the U.S., which indicate that loop detectors, when operational, provide accurate counts and speeds (ranging 1 - 9% depending on facility type) [68]. The two links were divided into 90 cells of length 50 ft (15.24 m), except boundary cells to ensure a correct total length. The Kalman filtering algorithm presented

in Section 6.3 was then used to estimate cell densities and state covariance matrices.



(a)



(b)

Figure 6.7: Comparison between estimated and measured queue sizes; (a) with 5 mi/hr criterion (b) with 5 ±2 mi/hr (8.05 ±3.22 km/hr) criterion

Queue sizes are not state parameters, so the estimated queues were computed from estimated cell densities. Cells with average speeds less than or equal to 5 mi/hr (8.05 km/hr) were considered to be part of the queue and the estimated queue size was simply treated as the average density in consecutive cells that meet the criterion multiplied by the total length of the cells. This is, hence, an average over random variables with known

means and covariance matrix, so that a corresponding mean queue size and variance can be readily calculated. A comparison between the estimated queue sizes along with a 95% confidence interval and the measured queue sizes is shown in Figure 6.7a.

6.5.3 Discussion

The results shown in Figure 6.7a reveal a very good match between estimated and observed queue sizes. All but two observed queue sizes fall within the confidence region. For the queue in cycle no. 10, the model seems to have trapped vehicles in the upstream intersection, which is revealed by the larger estimated queue size in cycle no. 11. The discrepancy in cycle no. 17 however cannot be explained in the same way; here, observed measurement error is a plausible cause. Observers were instructed to follow the 5 mi/hr (8.05 km/hr) criterion, but did not possess an accurate speed measurement apparatus. If the criterion is changed slightly, by ± 2 mi/hr (± 3.22 km/hr), in the estimated queue calculation, the discrepancy disappears as shown in Figure 6.7b. Allowing for possibly larger error (e.g. ± 3 mi/hr) could also take care of the discrepancy in cycle no. 10 in the sense that the observations would fall within the confidence interval, but the qualitative difference between estimated and observed maximum queue sizes in cycles 10 and 11 (in terms of mean) would remain.

An interesting exercise is estimation of both traffic states and parameters simultaneously via an adaptive filtering algorithm. This would shed light on whether the proposed estimation framework would be capable of detecting the drop in estimated free-flow speed due to snowy conditions. Since the purpose of this validation exercise was to test the Gaussian model, not the filtering algorithm, this was not carried out as part of the present research, but should be considered in any future research concerned with the estimation problem itself.

Chapter 7

Conclusion and Future Research

7.1 Research Summary

Godunov scheme based dynamic equations of traffic flow and, particularly, the cell transmission model (CTM) have gained a great deal of attention since the late 1990s amongst traffic flow researchers and modelers. This is due to their simplicity and ability to capture queue build-up dynamics, queue dissipation dynamics, and spill-over dynamics, both spatially and temporally. As real-time traffic flow data is becoming more widely available, the real-time estimation and short-term prediction of traffic conditions along roads has also gained recent attention. Consequently, many of the new stochastic models of traffic flow, developed for purposes of real-time estimation and short-term prediction, are stochastic extensions of the CTM. Most of these models are developed by adding noise to the CTM, which results in potentially large covariance matrices which need to be estimated. In this thesis, a new stochastic extension of Godunov scheme based dynamics is developed using a queuing theoretic approach. The advantage of this approach is that (i) non-negativity of the sample paths of the stochastic model is implicitly ensured and (ii) the mean dynamics (fluid limit) of the stochastic model is the Godunov scheme itself. Neither issue has been addressed in previous research. A Gaussian approximation of the queueing model is also proposed for purposes of model tractability. The Gaussian approximation proposed is

characterized by its mean and covariance dynamics; the mean dynamics are those of the Godunov scheme, while the covariance matrices can be computed using few parameters. The covariance function is derived from the queueing model as opposed to being estimated from data. Stationary behavior of the covariance function is analyzed and it is shown that the covariance matrices are bounded. Consequently, estimated covariance matrices are also bounded. As a result, Kalman filters that use the proposed model are stochastically observable, which is a critical issue in real time estimation of traffic dynamics. Model validation was carried out in a real-world signalized arterial setting where cycle-by-cycle maximum queue sizes were estimated using the Gaussian model as a state-space model in a Kalman filter and the estimated queue sizes were compared to observed maximum queue sizes. The results indicated good agreement between estimated and observed queue sizes.

Before discussing future research potential, it is important to note some of the limitations of the present research. First, while introducing randomness yields oscillatory traffic behavior, it is not clear whether this is consistent with observed stop-and-go traffic behavior in the real world. Nonetheless, as driver choice is the primary player in all traffic flow dynamics, probabilistic approaches to modeling such phenomena are the most well suited for in-depth analysis. Another limitation lies in the classical way in which the fundamental diagram was introduced into the stochastic model. As described in Section 3.4.2, fundamental relations can be reconciled with car-following behavior. This is leaves out a crucial component of traffic flow: lane-changing behavior, which could have serious implications on how macroscopic variables are related to one-another. This is a major limitation of the classical macroscopic modeling approaches presented in Chapter 3 and inherited by the proposed stochastic modeling framework.

7.2 Future Research

Future research could be carried out in various directions. From an application standpoint, the proposed model could be used in a variety of traffic control contexts, including ramp metering, adaptive signal control, and real-time traveler information systems. The proposed

model can, in effect, be used in any application that uses the CTM. The probabilistic nature of the model lends itself to a variety of extensions that incorporate measures of uncertainty, reliability, and robustness, all measures that are usually established on probabilistic grounds. Another direction for future research is the investigation of stochastic models of traffic flow in general network settings. As in deterministic models, this would be accomplished by first investigating merge and diverge dynamics. Applications of the latter include real-time estimation of turning counts at network junctions. A stochastic model which includes merge and diverge dynamics could also be used to develop stochastic dynamic traffic assignment models, which allows for incorporating both the detailed dynamics and travel time reliability measures in the investigation of route choice dynamics at the network level.

As another example of future research, a more thorough investigation into observability could shed light on questions related to sensor placement along road networks and how mobile sensors could be used to overcome observability issues. As a parallel line of research, issues related to controllability and stochastic controllability using the proposed model, in a variety of traffic management applications can be pursued; these include ramp metering and adaptive signal timing.

From a theoretical standpoint, stochastic conservation laws of traffic flow could be developed as continuous time and continuous space models (stochastic partial differential equations) to investigate wave propagation dynamics in a probabilistic context. Such models could have the potential to unify the various macroscopic traffic flow phenomena currently modeled using disparate deterministic theories, such as stop-and-go waves and traffic hysteresis. These different phenomena would arise as sample paths of the unifying stochastic framework and may be studied probabilistically.

References

[1] R. Agarwal and D. O'Regan. *An introduction to ordinary differential equations.* Springer, New York, 2008.

[2] S. Ahn, M. Cassidy, and J. Laval. Verification of a simplified car-following theory. *Transportation Research Part B*, 38(5):431–440, 2004.

[3] L. Arnold. *Stochastic differential equations: Theory and applications.* Wiley, New York, 1974.

[4] A. Aw and M. Rascle. Resurrection of second order models of traffic flow. *SIAM Journal on Applied Mathematics*, 60(3):916–938, 2000.

[5] V. Bageshwar, D. Gebre-Egziabher, W. Garrarc, and T. Georgiou. Stochastic observability test for discrete-time kalman filters. *Journal of Guidance, Control, and Dynamics*, 32(4):1356–1370, 2009.

[6] A. Bain and D. Crişan. *Fundamentals of stochastic filtering.* Springer, New York, 2008.

[7] R. Bellman. The stability of solutions of linear differential equations. *Duke Mathematical Journal*, 10(4):643–647, 1943.

[8] P. Billingsley. *Convergence of probability measures, 2nd edition.* Wiley, New York, 1999.

[9] S. Blandin, A. Couque, A. Bayen, and D. Work. On sequential data assimilation for scalar macroscopic traffic flow models. *Physica D: Nonlinear Phenomena*, (In Press), 2012.

[10] R. Boel and L. Mihaylova. A compositional stochastic model for real-time freeway traffic simulation. *Transportation Research Part B*, 40(4):319–334, 2006.

[11] D. Branston. Models of single lane time headway distributions. *Transportation Science*, 10(2):125–148, 1976.

[12] L. Breiman. The Poisson tendency in traffic distribution. *The Annals of Mathematical Statistics*, 34(1):308–311, 1963.

[13] P. Brémaud. *Point processes and queues: martingale dynamics*. Springer, New York, 1981.

[14] E. Brown, R. Barbieri, V. Ventura, R. Kass, and L. Frank. The time-rescaling theorem and its application to neural spike train data analysis. *Neural Computation*, 14(2): 325–346, 2002.

[15] M. Brown. Some results on a traffic model of Rényi. *Journal of Applied Probability*, 6(2):293–300, 1969.

[16] D. Bui, P. Nelson, and S. Narasimhan. Computational realizations of the entropy condition in modeling congested traffic flow. *Federal Highway Administration Report*, FHWA/TX-92/1232-7, 1992.

[17] H. Chen and A. Mandelbaum. Hierarchical modeling of stochastic networks. Part I: Fluid models, Part II: Strong approximations. In *Yao, D. (ed), Stochastic modeling and analysis of manufacturing systems*, pages 47–131. Springer, New York, 1994.

[18] H. Chen and D. Yao. *Fundamentals of queueing networks*. Springer, New York, 2001.

[19] H. Chen, H. Rakha, S. Sadek, and B. Katz. A particle filter approach for real-time freeway traffic state prediction. In *Proceedings of the 91st Annual Meeting of the Transportation Research Board, (paper no. 12-2605)*, 2012.

[20] N. Chiabaut, C. Buisson, and L. Leclercq. Fundamental diagram estimation through passing rate measurements in congestion. *IEEE Transactions on Intelligent Transportation Systems*, 10(2):355–359, 2009.

[21] R. Cowan. Useful headway models. *Transportation Research*, 9(6):371–375, 1975.

[22] C. Daganzo. The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory. *Transportation Research Part B*, 28 (4):269–287, 1994.

[23] C. Daganzo. The cell transmission model, part II: Network traffic. *Transportation Research Part B*, 29(2):79–93, 1995.

[24] C. Daganzo. A finite difference approximation of the kinematic wave model of traffic flow. *Transportation Research Part B*, 29(4):261–276, 1995.

[25] C. Daganzo. A behavioral theory of multi-lane traffic flow. Part I: Long homogeneous freeway sections. Part II: Merges and the onset of congestion. *Transportation Research Part B: Methodological*, 36(2):131–169, 2002.

[26] D. Daley and D. Vere-Jones. *An introduction to the theory of point processes, 2nd edition. Volume I: Elementary theory and methods.* Springer, New York, 2003.

[27] D. Daley and D. Vere-Jones. *An introduction to the theory of point processes, 2nd edition. Volume II: General theory and structure.* Springer, New York, 2008.

[28] J. Davidson. *Stochastic limit theory: An introduction for econometricians.* Oxford University Press, Oxford, 1994.

[29] G. Davis and J. Kang. Estimating destination-specific traffic densities on urban freeways for advanced traffic management. *Transportation Research Record: Journal of the Transportation Research Board*, 1457:143–148, 1994.

[30] J. Del Castillo and F. Benitez. On the functional form of the speed-density relationship. I: General theory, II: Empirical investigation. *Transportation Research Part B*, 29(5):373–406, 1995.

[31] X. Di, H. Liu, and G. Davis. Hybrid Extended Kalman filtering approach for traffic density estimation along signalized arterials. *Transportation Research Record: Journal of the Transportation Research Board*, 2188:165–173, 2010.

[32] L. Edie. Car-following and steady-state theory for noncongested traffic. *Operations Research*, 9(1):66–76, 1961.

[33] L. Evans. *Partial differential equations, 2nd edition.* American Mathematical Society, Providence, 2010.

[34] H. Flanders. Differentiation under the integral sign. *The American Mathematical Monthly*, 80(6):615–627, 1973.

[35] R. Franklin. The structure of a traffic shock wave. *Civil Engineering Public Works Review*, 56:1186–1188, 1961.

[36] D. Gazis and C. Knapp. On-line estimation of traffic densities from time-series of flow and speed data. *Transportation Science*, 5(3):283–301, 1971.

[37] D. Gazis and C. Liu. Kalman filtering estimation of traffic counts for two network links in tandem. *Transportation Research Part B*, 37(8):737–745, 2003.

[38] D. Gazis, R. Herman, and R. Potts. Car-following theory of steady-state traffic flow. *Operations Research*, 7(4):499–505, 1959.

[39] S. Godunov. A difference scheme for numerical computation of discontinuous solutions of hydrodynamic equations. *Math. Sbornik*, 47(3):271–306, 1959.

[40] L. Gray and D. Griffeath. The ergodic theory of traffic jams. *Journal of Statistical Physics*, 105(3):413–452, 2001.

[41] H. Greenberg. An analysis of traffic flow. *Operations Research*, 7(1):79–85, 1959.

[42] B. Greenshields. A study of traffic capacity. In *Proceedings of the 14th Annual Meeting of the Highway Research Board*, pages 448–477, 1935.

[43] F. Haight. *Mathematical theories of traffic flow.* Academic Press, New York, 1963.

[44] D. Helbing. Gas-kinetic derivation of Navier-Stokes-like traffic equations. *Physical Review E*, 53(3):2366–2381, 1996.

[45] J. Herrera and A. Bayen. Incorporation of Lagrangian measurements in freeway traffic state estimation. *Transportation Research Part B*, 44(4):460–481, 2010.

[46] H. Holden and N. Risebro. Conservation laws with a random source. *Applied Mathematics and Optimization*, 36(2):229–241, 1997.

[47] S. Hoogendoorn and P. Bovy. New estimation technique for vehicle-type-specific headway distributions. *Transportation Research Record: Journal of the Transportation Research Board*, 1646:18–28, 1998.

[48] S. Howison. *Practical applied mathematics: Modelling, analysis, approximation.* Cambridge University Press, Cambridge, 2005.

[49] S. Jabari and H. Liu. A stochastic model of traffic flow: Theoretical foundations. *Transportation Research Part B*, 46(1):156–174, 2012.

[50] A. Jazwinski. *Stochastic processes and filtering theory.* Academic Press, New York, 1970.

[51] R. Kalman. A new approach to linear filtering and prediction problems. *Transactions of the ASME. Series D, Journal of Basic Engineering*, 82:35–45, 1960.

[52] R. Kalman and R. Bucy. New results in linear filtering and prediction theory. *Transactions of the ASME. Series D, Journal of Basic Engineering*, 83:95–107, 1961.

[53] J. Kang. *Estimation of destination-specific traffic densities and identification of parameters on urban freeways using Markov models of traffic flow.* PhD thesis, University of Minnesota, 1995.

[54] M. Khoshyaran and J. Lebacque. A stochastic macroscopic traffic model devoid of diffusion. In *Appert-Rolland, C., Chevoir, F., Gondret, P., Lassarre, S., Lebacque, J., Schreckenberg, M. (eds.), Traffic and Granular Flow '07*, pages 139–150. Springer, New York, 2009.

[55] T. Kim and H. Zhang. An empirical study on gap time and its relation to the fundamental diagram of traffic flow. In *Proceedings of the 7th International IEEE Conference on Intelligent Transportation Systems*, pages 94–99. IEEE. Washington, D.C., 2004.

[56] T. Kim and H. Zhang. A stochastic wave propagation model. *Transportation Research Part B*, 42(7):619–634, 2008.

[57] F. Klebaner. *Introduction to stochastic calculus with applications, 2nd edition.* Imperial College Press, London, 2005.

[58] J. Lebacque. The Godunov scheme and what it means for first order traffic flow models. In *Lesort, J. (ed), Proceedings of the 13th Internaional Symposium on Transportation and Traffic Theory*, pages 647–677. Elsevier. Lyon, France, 1996.

[59] R. LeVeque. *Numerical methods for conservation laws, 2nd edition.* Birkhäuser, Berlin, 1992.

[60] R. LeVeque. *Finite volume methods for hyperbolic problems.* Cambridge University Press, Cambridge, 2002.

[61] T. Li. Global solutions of nonconcave hyperbolic conservation laws with relaxation arising from traffic flow. *Journal of Differential Equations*, 190(1):131–149, 2003.

[62] H. Lieu. *Revised monograph on traffic flow theory.* US Department of Transportation Federal Highway Administration, Washington, D.C., 2003.

[63] M. Lighthill and G. Whitham. On kinematic waves. I: Flood movement in long rivers, II: A theory of traffic flow on long crowded roads. In *Proceedings of the Royal Society (London) A229*, pages 281–345, 1955.

[64] H. Liu and W. Ma. A virtual vehicle probe model for time-dependent travel time estimation on signalized arterials. *Transportation Research Part C*, 17(1):11–26, 2009.

[65] H. Liu, X. Wu, W. Ma, and H. Hu. Real-time queue length estimation for congested signalized intersections. *Transportation Research Part C*, 17(4):412–427, 2009.

[66] A. Mandelbaum and W. Massey. Strong approximations for time-dependent queues. *Mathematics of Operations Research*, 20(1):33–64, 1995.

[67] A. Mandelbaum and G. Pats. State-dependent queues: Approximations and applications. In *Kelly, F. and Williams, R. (eds), IMA Volumes in Mathematics and Its Applications*, pages 239–282. Springer, Berlin, 1995.

[68] P. Martin, Y. Feng, and X. Wang. Detector technology evaluation. *Mountain-Plains Consortium*, 2003.

[69] W. Masey. *Non-stationary queues.* PhD thesis, Stanford University, 1981.

[70] W. Massey. Asymptotic analysis of the time dependent M/M/1 queue. *Mathematics of Operations Research*, 10(2):305–327, 1985.

[71] C. Meyer. Matrix Analysis and Applied Linear Algebra. *SIAM, Philadelphia*, 2000.

[72] L. Muñoz, X. Sun, R. Horowitz, and L. Alvarez. Piecewise-linearized cell transmission model and parameter calibration methodology. *Transportation Research Record: Journal of the Transportation Research Board*, 1965.

[73] L. Muñoz, X. Sun, R. Horowitz, and L. Alvarez. Traffic density estimation with the cell transmission model. In *Proceedings of the 2003 American Control Conference*, pages 3750–3755, 2003.

[74] K. Nagel and M. Schreckenberg. A cellular automaton model for freeway traffic. *Journal de Physique I*, 2(12):2221–2229, 1992.

[75] C. Nanthawichit, T. Nakatsuji, and H. Suzuki. Application of probe-vehicle data for real-time traffic-state estimation and short-term travel-time prediction on a freeway. *Transportation Research Record: Journal of the Transportation Research Board*, 1855: 49–59, 2003.

[76] G. Newell. Nonlinear effects in the dynamics of car following. *Operations Research*, 9 (2):209–229, 1961.

[77] G. Newell. Approximation methods for queues with application to the fixed-cycle traffic light. *Siam Review*, 7(2):223–240, 1965.

[78] G. Newell. Equilibrium probability distributions for low density highway traffic. *Journal of Applied Probability*, 3(1):247–260, 1966.

[79] G. Newell. Queues with time-dependent arrival rates. I: The transition through saturation. *Journal of Applied Probability*, 5(2):436–451, 1968.

[80] G. Newell. Queues with time-dependent arrival rates. II: The maximum queue and the return to equilibrium, III: A mild rush hour. *Journal of Applied Probability*, 5(3): 579–606, 1968.

[81] G. Newell. Memoirs on highway traffic flow theory in the 1950s. *Operations Research*, 50(1):173–178, 2002.

[82] G. Newell. A simplified car-following theory: a lower order model. *Transportation Research Part B*, 36(3):195–205, 2002.

[83] B. Oksendal. *Stochastic differential equations, 6th edition.* Springer, Berlin, 2007.

[84] S. Osher. Riemann solvers, the entropy condition, and difference approximations. *SIAM Journal on Numerical Analysis*, 21(2):217–235, 1984.

[85] C. Osorio, G. Flötteröd, and M. Bierlaire. Dynamic network loading: A stochastic differentiable model that derives link state distributions. *Transportation Research Part B*, 45(9):1410–1423, 2011.

[86] S. Paveri-Fontana. On Boltzmann-like treatments for traffic flow: A critical review of the basic model and an alternative proposal for dilute traffic analysis. *Transportation Research*, 9(4):225–235, 1975.

[87] H. Payne. Models of freeway traffic and control. In *Bekey, G. (ed.) Mathematical Models of Public Systems. Proceedings of the Simulation Council*, pages 51–61, 1971.

[88] L. Pipes. Wave theories of traffic flow. *Journal of the Franklin Institute*, 280(1):23–41, 1965.

[89] L. Pipes. Car following models and the fundamental diagram of road traffic. *Transportation Research*, 1(1):21–29, 1967.

[90] I. Prigogine and R. Herman. *Kinetic theory of traffic flow*. Elsevier, New York, 1971.

[91] C. Pugh. *Real mathematical analysis*. Springer, New York, 2002.

[92] A. Rényi. On two mathematical models of the traffic on a divided highway. *Journal of Applied Probability*, 1(2):311–320, 1964.

[93] S. Resnick. *Adventures in stochastic processes*. Birkhäuser, Boston, 1992.

[94] S. Resnick. *A probability path*. Birkhäuser, Boston, 1999.

[95] P. Richards. Shock waves on the highway. *Operations Research*, 4(1):42–51, 1956.

[96] M. Rosenlicht. *Introduction to analysis*. Dover, New York, 1986.

[97] H. Solomon and P. Wang. Nonhomogeneous Poisson fields of random lines with applications to traffic flow. In *Lecam, L., Neyman, J., Scott, E. (eds.) Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability. Volum III: Probability Theory.*

[98] A. Sopasakis. Lattice free stochastic dynamics. *Communications in Computational Physics*, 12(3):691–702, 2012.

[99] A. Sopasakis and M. Katsoulakis. Stochastic modeling and simulation of traffic flow: Asymmetric single exclusion process with Arrhenius look-ahead dynamics. *SIAM Journal on Applied Mathematics*, 66(3):921–944, 2006.

[100] K. Stanková and B. De Schutter. On freeway traffic density estimation for a jump Markov linear model based on Daganzo's cell transmission model. In *Proceedings of the 13th International IEEE Conference on Intelligent Transportation Systems*, pages 13–18, 2010.

[101] R. Stengel. *Optimal control and estimation.* Dover, New York, 1994.

[102] A. Sumalee, R. Zhong, T. Pan, and W. Szeto. Stochastic cell transmission model (SCTM): A stochastic dynamic traffic model for traffic state surveillance and assignment . *Transportation Research Part B*, 45(3):507–533, 2011.

[103] X. Sun, L. Muñoz, and R. Horowitz. Mixture Kalman filter based highway congestion mode and vehicle density estimator and its application. In *Proceedings of the 2004 American Control Conference*, pages 2098–2103. IEEE, 2004.

[104] M. Szeto and D. Gazis. Application of Kalman filtering to the surveillance and control of traffic systems. *Transportation Science*, 6(4):419–439, 1972.

[105] R. Underwood. Speed, volume, and density relationships. In *Proceedings of the Quality and Theory of Traffic Flow Symposium*, pages 141–188, 1961.

[106] Y. Wang and M. Papageorgiou. Real-time freeway traffic state estimation based on extended Kalman filter: A general approach. *Transportation Research Part B*, 39(2): 141–167, 2005.

[107] Y. Wang, M. Papageorgiou, and A. Messmer. Real-time freeway traffic state estimation based on extended Kalman filter: A case study. *Transportation Science*, 41(2): 167–181, 2007.

[108] J. Wardrop. Some theoretical aspects of road traffic research. In *Proceedings of the Institution of Civil Engineers (London), pt. II*, volume 1, pages 325–362, 1952.

[109] G. Weiss and R. Herman. Statistical properties of low-density traffic. *Quarterly of Applied Mathematics*, 20:121–130, 1962.

[110] G. Whitham. *Linear and nonlinear waves*. Wiley, New York, 1974.

[111] W. Whitt. *Internet supplement to stochastic-process limits*. Available at: http://www.columbia.edu/~ww2040/supplement.html, 2002.

[112] W. Whitt. *Stochastic process limits*. Springer, New York, 2002.

[113] D. Work, O. Tossavainen, S. Blandin, A. Bayen, T. Iwuchukwu, and K. Tracton. An ensemble Kalman filtering approach to highway traffic estimation using GPS enabled mobile devices. In *Proceedings of the 47th IEEE Conference on Decision and Control*, pages 5062–5068, 2008.

[114] J. Xiong. *An introduction to stochastic filtering theory*. Oxford University Press, New York, 2008.

[115] H. Zhang. A theory of nonequilibrium traffic flow. *Transportation Research Part B*, 32(7):485–498, 1998.

[116] H. Zhang and T. Kim. A car-following theory for multiphase vehicular traffic flow. *Transportation Research Part B*, 39(5):385–399, 2005.

Appendix A

Mathematical Background

A.1 Linear Algebra Background

Definition A.1.1 (nilpotent matrix)**.** *A square matrix* $\mathbf{A}$ *is said to be nilpotent if there exists a finite number* $k \in \mathbb{N}$ *such that* $\mathbf{A}^k = \mathbf{0}$*, where* $\mathbf{0}$ *is the zero matrix. Any triangular matrix with zeros along the diagonal is nilpotent.*

Definition A.1.2 (spectral radius)**.** *Let* $\mathbf{A} \in \mathbb{R}^{m \times m}$ *be a square matrix and let* $a_1, ..., a_m$ *be its eigenvalues. The spectral radius of* $\mathbf{A}$*, denoted by* $\mathcal{R}(\mathbf{A})$*, is defined as:*

$$\mathcal{R}(\mathbf{A}) = \max_{j=1,...,m} |a_j| \tag{A.1}$$

Theorem A.1.1 (matrix convergence)**.** *Let* $\mathbf{A} \in \mathbb{R}^{m \times m}$ *be a square matrix, then*

$$\mathbf{A}^k \underset{k \to \infty}{\longrightarrow} \mathbf{0} \iff \mathcal{R}(\mathbf{A}) < 1 \tag{A.2}$$

Proof. See for example [71, Theorem 7.10.5, page 617]. □

A.2 Calculus Background

Definition A.2.1 (Lipschitz continuity)**.** *A real-valued function* $f : \Theta_1 \to \Theta_2$ *is said to be Lipschitz continuous if there exists a constant* $K \in \mathbb{R}$ *such that for any pair* $a, b \in \Theta_1$*, we*

have that:

$$|f(a) - f(b)| \leq K|a - b| \tag{A.3}$$

Definition A.2.2 (triangle inequality)**.** *For any metric $d : \Theta \to \mathbb{R}_+$ on the metric space Θ and any points $a, b, c \in \Theta$ the triangle inequality says that the following hold:*

- $d(a, b) \leq d(a, c) + d(c, b)$
- $|d(a, c) + d(c, b)| \leq d(a, b)$

In particular, in a normed vector space, we have, from the latter, that $\Big|\,||a|| - ||b||\,\Big| \leq ||a - b||$.

Theorem A.2.1 (the fundamental theorem of calculus)**.** *Let f be a continuous real-valued function on the interval $[a, b]$ and let the function F be defined by:*

$$F(u) = \int_a^u f(z)dz, \quad \text{for all } u \in [a, b] \tag{A.4}$$

Then, F is differentiable and

$$\frac{dF(u)}{du} = f(u) \tag{A.5}$$

Proof. This is a classical result, the proof of which can be found in a variety of texts. See for example [96, page 127] or [91, page 171]. □

Corollary A.2.1 (corollary to the fundamental theorem of calculus)**.** *Let f be a continuous function on the interval $[a, b]$ and suppose there exists a function F such that f is the anti-derivative of F. That is,*

$$\frac{dF(u)}{du} = f(u) \tag{A.6}$$

Then,

$$\int_a^b f(u)du = F(b) - F(a) \tag{A.7}$$

Proof. See for example [96, Corollary 2, page 127]. □

Theorem A.2.2 (differentiation under the integral sign)**.** *Let $f : [a, b] \times [c, d] \to \mathbb{R}$ be a real-valued function that is continuous for all pairs $(u, z) \in [a, b] \times [c, d]$. Further suppose that $\partial f(u, z)/\partial u$ exists and is also continuous for all pairs $(u, z) \in [a, b] \times [c, d]$. Then,*

$$\frac{d}{du}\int_c^d f(u, z)dz = \int_c^d \frac{\partial f(u, z)}{\partial u}dz \tag{A.8}$$

Proof. See for example [96, page 159]. □

Theorem A.2.3 (the Leibniz integration rule)**.** *Define the function F by*

$$F(u) = \int_{g_1(u)}^{g_2(u)} f(u,z)dz, \tag{A.9}$$

where $g_1(u)$ and $g_2(u)$ are continuously differentiable on $[a,b]$ and both $f(u,z)$ and its partial derivative $\partial f(u,z)/\partial u$ are continuous for all pairs $(u,z) \in [a,b] \times [g_1(u), g_2(u)]$. Then, for $u \in [a,b]$,

$$\frac{dF(u)}{du} = f(g_2(u),u)\frac{dg_2(u)}{du} - f(g_1(u),u)\frac{dg_1(u)}{du} + \int_{g_1(u)}^{g_2(u)} \frac{\partial f(u,z)}{\partial u}dz \tag{A.10}$$

Proof. See for example [34]. □

A.3 Probability Background

Definition A.3.1 (counting process)**.** *A process $\mathcal{N}(\cdot)$ is said to be a counting process, if*

(i) $\mathcal{N}(t) \in \mathbb{Z}_+$ for all t.

(ii) $\mathcal{N}(\cdot)$ has non-decreasing sample paths.

(iii) For $t_1 < t_2$, $\mathcal{N}(t_2) - \mathcal{N}(t_1)$ is the number of events that occur in the interval $(t_1, t_2]$.

Definition A.3.2 (the homogeneous Poisson process)**.** *The counting process $\mathcal{N}(\cdot)$ is said to be a homogeneous Poisson process with rate $\lambda > 0$, if*

(i) $\mathcal{N}(0) = 0$ with probability one.

(ii) $\mathcal{N}(\cdot)$ has stationary independent increments.

(iii) The number of events in any interval of length Δt is a Poisson distributed random variable with mean $\lambda\Delta t$.

Definition A.3.3 (standard Brownian motion)**.** *The stochastic process $W(\cdot)$ is said to be a standard Brownian motion (or a Wiener process), if*

(i) $W(0) = 0$ *with probability one.*

(ii) $W(\cdot)$ *has stationary independent increments, which are normally distributed with mean 0 and variance* Δt*, the increment length.*

(iii) $W(\cdot)$ *has continuous sample paths.*

Definition A.3.4 (narrow-sense linear stochastic differential equation)**.** *Let* $\mathbf{a}(t)$ *be a deterministic time varying vector, and let* $\mathbf{A}_1(t)$ *and* $\mathbf{A}_2(t)$ *be time varying deterministic matrices. Then, the SDE*

$$d\mathbf{X}(t) = \mathbf{a}(t)dt + \mathbf{A}_1(t)\mathbf{X}(t)dt + \mathbf{A}_2(t)d\mathbf{W}(t) \tag{A.11}$$

is called a narrow-sense linear SDE, for which the solution may be written explicitly as [3]:

$$\mathbf{X}(t) = \mathbf{\Phi}(t)\left(\mathbf{X}(0) + \int_0^t \mathbf{\Phi}(u)^{-1}\big(\mathbf{a}(u)du + \mathbf{A}_2(u)d\mathbf{W}(u)\big)\right), \tag{A.12}$$

where $\mathbf{\Phi}(\cdot)$ *is a fundamental matrix.*

Theorem A.3.1 (the strong law of large numbers (SLLN))**.** *Let* $\xi_1, \xi_2, \ldots$ *be a sequence of i.i.d. random variables with finite mean* μ *and finite variance* σ^2*. Then*

$$\frac{1}{n}\sum_{i=1}^{n}\xi_i \underset{n\to\infty}{\longrightarrow} \mu \quad a.s. \tag{A.13}$$

Proof. The proof can be found in most probability theory texts. See, for example, [94, Chapter 7]. □

Theorem A.3.2 (the central limit theorem (CLT))**.** *Let* $\xi_1, \xi_2, \ldots$ *be a sequence of i.i.d. random variables with finite mean* μ *and finite variance* σ^2 *and let* Z *denote a standard normal random variable. Then*

$$\frac{1}{\sqrt{n}\sigma}\left(\sum_{i=1}^{n}(\xi_i - \mu)\right) \underset{n\to\infty}{\overset{\mathcal{D}}{\longrightarrow}} Z, \tag{A.14}$$

where $\overset{\mathcal{D}}{\to}$ *means "converges in distribution to".*

Proof. See, for example, [94, Chapter 9]. □

Theorem A.3.3 (Donsker's theorem). *Let $\xi_1, \xi_2, \ldots$ be a sequence of i.i.d. random variables with finite mean μ and finite variance σ^2 and let $W(\cdot)$ be standard Brownian motion. Then,*

$$\frac{1}{\sqrt{n}\sigma}\left(\sum_{i=1}^{\lfloor n\cdot\rfloor}(\xi_i - \mu)\right) \xrightarrow[n\to\infty]{\mathcal{D}} W(\cdot), \tag{A.15}$$

where $\lfloor u \rfloor$ is the largest integer less than or equal to u.

Proof. See [8, Theorem 14.1] □

Theorem A.3.4 (Bellman-Gronwall inequality). *Let $g(t)$ be a non-negative function, such that, for $t \in [0, U]$,*

$$g(t) \leq \kappa + K\int_0^t g(u)du \tag{A.16}$$

for some constants κ and K. Then, for $t \in [0, U]$,

$$g(t) \leq \kappa e^{Kt} \tag{A.17}$$

Proof. See, for example, [7] and [83, Exercise 5.17]. □

Theorem A.3.5 (Cramér's theorem). *Let $\xi_1, \xi_2, \ldots$ and $\zeta_1, \zeta_2, \ldots$ be two sequence of random variables and ξ be another random variable such that as $n \to \infty$, $\xi_n \xrightarrow{\mathcal{D}} \xi$ and $\mathbb{P}(|\zeta_n - a| > \epsilon) \longrightarrow 0$ for some constant a and any $\epsilon > 0$. Then,*

(i) $(\xi_n + \zeta_n) \xrightarrow{\mathcal{D}} (\xi + a)$,

(ii) $\xi_n\zeta_n \xrightarrow{\mathcal{D}} a\xi$, *and*

(iii) $\frac{\xi_n}{\zeta_n} \xrightarrow{\mathcal{D}} \frac{\xi}{a}$, *when* $a \neq 0$.

Proof. See [28, Theorem 22.14]. □

Theorem A.3.6 (Skorohod representation theorem). *Let $\xi_1(\cdot), \xi_2(\cdot), \ldots$ be a sequence of RCLL stochastic processes and let $\xi(\cdot)$ be another RCLL stochastic process such that $\xi_n(\cdot) \xrightarrow{\mathcal{D}} \xi(\cdot)$ as $n \to \infty$. Then there exist a RCLL sequence $\widetilde{\xi}_1(\cdot), \widetilde{\xi}_2(\cdot), \ldots$ and a RCLL random process $\widetilde{\xi}(\cdot)$, all defined on a common probability space, such that $\xi_j(\cdot) \overset{\mathcal{D}}{=} \widetilde{\xi}_j(\cdot)$ for all j, $\xi(\cdot) \overset{\mathcal{D}}{=} \widetilde{\xi}(\cdot)$, and where $\widetilde{\xi}_n(\cdot) \longrightarrow \widetilde{\xi}(\cdot)$ almost surely as $n \to \infty$.*

Proof. See [111, Section 1.5]. □

Theorem A.3.7 (continuous mapping theorem)**.** *Let $\xi_1(\cdot), \xi_2(\cdot), \ldots$ be a sequence of stochastic processes and let $\xi(\cdot)$ be another stochastic process such that $\xi_n(\cdot) \xrightarrow{\mathcal{D}} \xi(\cdot)$ as $n \to \infty$. Then, the following hold.*

(i) For any continuous function $g(\cdot)$, we have that $g\big(\xi_n(\cdot)\big) \xrightarrow{\mathcal{D}} g\big(\xi(\cdot)\big)$ as $n \to \infty$.

(ii) Let $g_n(\cdot)$ be a sequence of functions and let A be a set of points such that $g_n(u) \to g(u)$ as $n \to \infty$ for all points $u \in A$. If $\mathbb{P}\big(\xi(\cdot) \in A\big) = 1$, then $g_n\big(\xi_n(\cdot)\big) \xrightarrow{\mathcal{D}} g\big(\xi(\cdot)\big)$ as $n \to \infty$.

Proof. See [111, Section 1.3]. □

Appendix B

Traffic Flow Theory Background

B.1 Linear Advection Equation

The linear advection problem is written as:

$$\begin{aligned} &\frac{\partial \bar{\rho}}{\partial t} + \bar{v}\frac{\partial \bar{\rho}}{\partial x} = 0 \\ &\bar{\rho}(x,0) = \rho_0(x), \end{aligned} \tag{B.1}$$

where $\rho_0(x)$ is a prescribed initial traffic density profile (the initial data). To solve the problem, first write the total derivative of $\rho(x,t)$ with respect to time:

$$\frac{d\bar{\rho}}{dt} = \frac{\partial \bar{\rho}}{\partial t} + \frac{dx(t)}{dt}\frac{\partial \bar{\rho}}{\partial x}, \tag{B.2}$$

which includes the effect of changing x on the derivative with respect to time. Now noticing the similarity between (B.1) and (B.2), we get that:

$$\frac{dx(t)}{dt} = \bar{v} \tag{B.3}$$

and

$$\frac{d\bar{\rho}}{dt} = 0, \tag{B.4}$$

which is interpreted as: traffic densities do not change along lines of slope $\bar{v}$. These lines are referred to as *characteristic lines* and equations (B.3) and (B.4) are sometimes referred

to as characteristic equations. Solving (B.3), we get the formula:

$$x(t) = \bar{v}t + x(0) \tag{B.5}$$

Then, the formula for the solution of (B.1) is obtained by answering the following question: for any position x and time t, what is the corresponding position at time $t = 0$, which has the same traffic density? The answer is obtained from (B.5) as: $x_0 = x - \bar{v}t$ and the solution is given simply by the formula:

$$\bar{\rho}(x,t) = \rho_0(x - \bar{v}t) \tag{B.6}$$

B.2 Approximation of Discontinuous Initial Data

It will be shown here that the step function

$$\rho_0(x) = \begin{cases} \rho_l & \text{if } x < 0 \\ \rho_r & \text{if } x \geq 0 \end{cases} \tag{B.7}$$

is "well" approximated by the continuous function

$$\rho_0^\delta(x) = (\rho_l - \rho_r)\frac{e^{-\delta x}}{1 + e^{-\delta x}} + \rho_r \tag{B.8}$$

Without loss of generality, ρ_l and ρ_r will be taken to be equal to 1 and 0, respectively. (This can be achieved with a simple change of units.) Now, (B.7) and (B.8) simplify, respectively, to

$$\rho_0(x) = \begin{cases} 1 & \text{if } x < 0 \\ 0 & \text{if } x \geq 0 \end{cases} \tag{B.9}$$

and

$$\rho_0^\delta(x) = \frac{e^{-\delta x}}{1 + e^{-\delta x}} \tag{B.10}$$

It can be shown that for any $\epsilon > 0$, there exists a $\delta > 0$, such that $\left|\left|\rho_0(\cdot) - \rho_0^\delta(\cdot)\right|\right|_1 < \epsilon$, where $||\cdot||_1$ is the L_1 distance between the two functions. That is,

$$\left|\left|\rho_0(\cdot) - \rho_0^\delta(\cdot)\right|\right|_1 = \int_{-\infty}^{\infty} \left|\rho_0(x) - \rho_0^\delta(x)\right| dx$$

$$
\begin{aligned}
&= \int_{-\infty}^{0} \frac{dx}{1+e^{-\delta x}} + \int_{0}^{\infty} \frac{e^{-\delta x}}{1+e^{-\delta x}} dx \\
&= \frac{2}{\delta} \log_e(2) \qquad \text{(B.11)}
\end{aligned}
$$

Consequently, any $\delta > \frac{2}{\epsilon} \log_e(2)$ delivers the desired approximation.

Appendix C

Glossary and Acronyms

Care has been taken in this thesis to minimize the use of jargon and acronyms, but this cannot always be achieved. This appendix defines jargon terms in a glossary, and contains a table of acronyms and their meaning.

C.1 Glossary

- **Characteristic lines** – In the context of conservation of traffic flow, characteristic lines are lines in the time-space diagram along which traffic densities are constant.
- **Characteristic speed** – The reciprocal of the slope of the characteristic line. The reciprocal is used since the orientation of time-space diagrams in this thesis place time on the y-axis and position on the x-axis.
- **Courant, Friedrichs, and Lewy condition** – A condition that related discrete space and time interval lengths, Δx and Δt, to the maximum wave speed implied by a flow-density relation, which is used to choose Δx and Δt so that shockwaves and rarefaction fans do not cross downstream or upstream cell boundaries.
- **Daganzo's flux** – A numerical flux function computed by constructing taking the minimum of a sending function constructed from the first half of the fundamental

diagram (flow-density) and a receiving function constructed from the secod half of the fundamental diagram.

- **Doubly stochastic Poisson process** – A Poisson process with random rates.
- **First-order traffc flow model** – A non-linear conservation law combined with a fundamental relation of traffic flow.
- **Fundamental relation of traffic flow** – A stationary relation between any two of the three macroscopic variables of traffic flow. Most common relations are flow-density relations and speed density relations.
- **Ill-posed problem** – A PDE problem that is *not* well-posed (see definition of *well-posed problems*)
- **Initial data** – In the context of conservation laws, this refers to the traffic densities along the road at some initial time t, which is usually taken to be time zero ($t = 0$), and is denoted by the prescribed relation $\rho_0(x)$.
- **Linear advection equation** – A PDE of the form $u_t + au_x = 0$, where a is a given constant. For continuous initial data, $u(x, 0) \equiv u_0(x)$, the solution is given by $u(x, t) = u_0(x - at)$.
- **Numerical flux function** – A simplified way of calculating the total number of vehicles that cross a cell boundary over a small discrete time interval.
- **Observability** – The ability to reconstruct the initial conditions of the system state from the available measurements.
- **Rankine-Hugoniot jump condition** – A formula for calculating the speed of a shockwave.
- **Rarefaction fan** – A gradual reduction of traffic densities in the time-space diagram.
- **Receiving function** – See Daganzo's flux.

- **Riemann problem** – A non-linear conservation equation combined with initial data in the form of a step function.
- **Sending function** – See Daganzo's flux.
- **Sigma-field** – An event space representing the domain of a random variable.
- **Shock front** – See *shockwave*.
- **Shockwave** – A traveling discontinuity in a characteristic plot. The position of a *shock front* represents the position of the back of a queue in traffic.
- **Spectral radius** – The maximum absolute value of the eigenvalues of a matrix.
- **Stationary** – Does not change with time. A stationary variable does not depend on time, a stationary relation is one that does not change with time, although its argument might.
- **Stochastic observability** – A boundedness condition on estimated system covariance matrices.
- **Toeplitz matrix** – a tridiagonal matrix with equal elements along its diagonals.
- **Well-posed problem** [33] – A PDE is said to be *well-posed* if: (i) a solution to the problem exists, (ii) the solution is unique, and (iii) the solution depends continuously on the data given in the problem (e.g., initial and boundary conditions).

C.2 Acronyms

Table C.1: Acronyms

Acronym	Meaning
a.s.	almost surely
Continued on next page	

Table C.1 – continued from previous page

Acronym	Meaning
CFL	Courant, Friedrichs, and Lewy
CLT	central limit theorem
CTM	cell transmission model
FCLT	functional central limit theorem
FSLLN	functional strong law of large numbers
i.i.d.	independent and identically distributed
LHS	left-hand side (of an equation or inequality)
LWR	Lighthill, Whitham, and Richards
PDE	partial differential equation
RCLL	right-continuous with left hand limits
R-H	the Rankine-Hugoniot jump condition
RHS	right-hand side (of an equation or inequality)
SDE	stochastic differential equation
SLLN	strong law of large numbers