KATHOLIEKE UNIVERSITEIT LEUVEN
FACULTEIT INGENIEURSWETENSCHAPPEN
DEPARTEMENT BURGERLIJKE BOUWKUNDE
AFDELING VERKEER EN INFRASTRUCTUUR
Kasteelpark Arenberg 40 - B-3001 Leuven

The Link Transmission Model for Dynamic Network Loading

Promotor:
Prof. ir. L.H. Immers

Proefschrift voorgedragen tot
het behalen van het doctoraat
in de ingenieurswetenschappen

door

Isaak YPERMAN

June 2007

KATHOLIEKE UNIVERSITEIT LEUVEN
FACULTEIT INGENIEURSWETENSCHAPPEN
DEPARTEMENT BURGERLIJKE BOUWKUNDE
AFDELING VERKEER EN INFRASTRUCTUUR
Kasteelpark Arenberg 40 - B-3001 Leuven

The Link Transmission Model for Dynamic Network Loading

Jury:

Prof. dr. ir. D. Vandermeulen, voorzitter
Prof. ir. L.H. Immers, promotor
Prof. dr. S. Proost, co-promotor
Prof. dr. ir. E.C. van Berkum, Universiteit Twente
Prof. dr. ir. B. De Moor
Prof. dr. ir. J. Vandewalle
Dr. M.C.J. Bliemer, Technische Universiteit Delft
Prof. ir. J.-B. Lesort, INRETS/ENTPE Lyon

Proefschrift voorgedragen tot
het behalen van het doctoraat
in de ingenieurswetenschappen

door

Isaak YPERMAN

U.D.C. 656.021

June 2007

© Katholieke Universiteit Leuven
Faculteit Ingenieurwetenschappen
Kasteelpark Arenberg 1
B-3001 Leuven-Heverlee (Belgium)

Alle rechten voorbehouden. Niets uit deze uitgave mag worden vermenigvuldigd en/of openbaar gemaakt worden door middel van druk, fotokopie, microfilm, elektronisch of op welke andere wijze ook zonder voorafgaande schriftelijke toestemming van de uitgever.

All rights reserved. No part of this publication may be reproduced in any form by print, photoprint, microfilm or any other means without written permission from the publisher.

D/2007/7515/71
ISBN 978-90-5682-837-0

ACKNOWLEDGEMENTS

I would like to thank my promotor, Ben Immers, for creating this exceptional environment in which I could perform my own research, the way I wanted to do it myself. His open-mind, his enthusiasm, energy, generosity and brightness have been inspiring to me. I am grateful for the trust he put in me.

I owe a lot to Steven Logghe and Chris Tampère.
Steven's contagious passion for first order traffic flow theory triggered off this thesis research. The Link Transmission Model originated from one of his crucial insights.
Chris provided me with an invaluable support and guidance during the whole research process. Our numerous brainstorm sessions and his insightful ideas and comments put me on track toward completing this PhD. Knowing that I could always count on him gave me the necessary mental peace.

Jim Stada has been a great colleague as well. I thank him for his support, for his mildness, and for the many interesting and illuminating discussions we had.
Ben, Chris, Steven, Jim: working with you has been a privilege.

Furthermore, I would like to thank Eric van Berkum, Stef Proost and Bart De Moor for their feedback in seminars, and Dirk Vandermeulen, Joos Vandewalle, Michiel Bliemer and Jean-Baptiste Lesort for being part of the jury.

Finally, my deepest gratitude goes to my family for their constant support and for the faith they have in me. I am very fortunate to be surrounded by such caring and encouraging people.

Isaak Yperman
June 2007

SUMMARY

Dynamic Traffic Assignment (DTA) models are widely used tools to forecast the impact of proposed transportation investments and to estimate the effects of Advanced Traveler Information Systems and Advanced Traveler Management Systems. At the heart of the DTA model is the Dynamic Network Loading (DNL) model. This thesis presents a DNL model that realistically describes traffic propagation on both motorways and urban roads in practical large scale traffic networks.

The Link Transmission Model provides substantially more realism in the representation of queue-propagation (blocking back) and queue-dissipation compared to existing approaches in state-of-the-art macroscopic DTA models. Urban node models account for a detailed description of local flow restrictions and intersection delays. The Link Transmission Model is based on a computationally efficient algorithm, which allows for the modeling of traffic flows in large scale networks in a reasonable amount of time.

SAMENVATTING

Dynamische verkeerstoedelingsmodellen worden gebruikt om de impact van infrastructuuraanpassingen in verkeersnetwerken te voorspellen en om de effecten van informatieverstrekking en verkeersbeheersingsmaatregelen in te schatten. Een verkeerssimulatiemodel is een basiscomponent van het dynamisch toedelingsmodel. In dit eindwerk wordt een verkeerssimulatiemodel ontwikkeld.
Het Link Transmissie Model simuleert verkeersstromen in grote praktische netwerken die zowel snelwegen als stedelijke regio's omvatten. Het gemodelleerde file-opbouw en file-afbouw proces sluit nauwer aan bij de realiteit dan in state-of-the-art macroscopische verkeerstoedelingsmodellen. Op kruispunten worden lokale capaciteitsbeperkingen en knoopvertragingen gedetailleerd in rekening gebracht. Het Link Transmissie Model stoelt op een rekenefficiënt algoritme waardoor verkeersstromen in grote netwerken gesimuleerd kunnen worden in een beperkte rekentijd.

Een uitgebreide Nederlandstalige samenvatting is opgnenomen als appendix.

CONTENTS

Notation xi

1 Introduction 1
1.1. Background and Context 1
1.1.1. General framework of the DTA model 2
1.1.2. Types of assignment 3
1.2. Objectives and scope 4
1.2.1. Objective 4
1.2.2. Model validation 4
1.2.3. Higher order traffic phenomena 5
1.2.4. Multiple user classes 5
1.2.5. Time-dependent delay at intersections 5
1.2.6. DTA components 5
1.3. Thesis contributions 6
1.4. Thesis outline 7

2 Overview of DTA and DNL approaches 9
2.1. Analytical approach 9
2.2. Simulation-based approach 12
2.2.1. Equilibrium versus en-route DTA models 12
2.2.2. Micro-, meso- and macroscopic DTA models 13
2.2.2.1. Microscopic simulation-based DTA models 13
2.2.2.2. Mesoscopic simulation-based DTA models 15
2.2.2.3. Macroscopic simulation-based DTA models 16
2.3. Discussion 17

3 Framework of the Link Transmission Model 19
3.1. Definitions 19
3.2. LTM solution algorithm 23

4 Link model **27**

4.1. Macroscopic variables 27
4.2. The fundamental diagram 29
4.3. The conservation law 30
4.4. Kinematic wave theory 32
4.5. Newell’s simplified theory of kinematic waves 36
4.6. Sending and Receiving flows in LTM 40
4.7. Extension to a piecewise linear fundamental diagram 43
4.8. Computational efficiency and accuracy in comparison with CTM 46
4.9. Conclusions 50

5 Node models **51**

5.1. Node models for motorway intersections 52
  5.1.1. Inhomogeneous node 52
  5.1.2. Origin node 52
  5.1.3. Destination node 53
  5.1.4. Diverge node 53
  5.1.5. Merge node 56
5.2. Node models for urban intersections 57
  5.2.1. Basic signalized urban cross node 59
    5.2.1.1. Prototype signalized intersection 59
    5.2.1.2. Adjustments for other types of signalized intersections 63
  5.2.2. Basic unsignalized urban cross node 63
    5.2.2.1. Prototype unsignalized intersection 64
    5.2.2.2. Adjustments for other types of unsignalized intersections . . . 70
5.3. Conclusions 72

6 Intersection delay model **73**

6.1. Introduction of Point-Queues to implicitly realize average intersection delays 73
  6.1.1 Objective 73
  6.1.2. Implicit versus explicit consideration of intersection delays 74
  6.1.3. Implicit consideration of intersection delays: flow based versus travel-time based models 75
  6.1.4. Point-Queues versus kinematic wave queues 76
  6.1.5. Point-Queues: definitions and characteristics 77
6.2. Extended LTM solution algorithm including intersection delays 79
  6.2.1. Target value of intersection delay (algorithm step 4.1) 81
    6.2.1.1. Deterministic intersection delay for signalized intersections . .82
    6.2.1.2. Stochastic intersection delay for signalized intersections 83
    6.2.1.3. Total time-dependent intersection delay for signalized intersections 85
    6.2.1.4. Total time-dependent intersection delay for unsignalized intersections 85
  6.2.2. Target number of vehicles in the P-Q (algorithm step 4.2) 86

6.2.3. P-Q Sending flow $S'_{ij}(t)$ (algorithm step 4.3) 86
6.2.3.1. Solution method for deterministic intersection delays at signalized intersections 87
6.2.3.2. Adjustments for time-dependent intersection delays at signalized and unsignalized intersections 96
6.2.4. P-Q transition flows $G'_{ij}(t)$ (algorithm step 5) 98
6.3. Conclusions 100

7 Case studies 103
7.1. Case 1: Simple diverge network 103
7.1.1. Aims 103
7.1.2. Model properties: LTM versus DQM 103
7.1.3. Inputs 105
7.1.4. Outputs 107
7.1.5. Conclusions of the diverge case study 111
7.2. Case 2: Ghent-Brussels network 112
7.2.1. Aims 112
7.2.2. Model description 112
7.2.3. Ghent-Brussels network: Inputs 114
7.2.4. Ghent-Brussels network: Outputs 115
7.2.5. Affligem-Brussels network: Inputs 122
7.2.6. Affligem-Brussels network: Outputs 123
7.2.7. Conclusions of the Ghent-Brussels case study 125

8 Conclusions 127
8.1. Properties of the developed model 127
8.2. General conclusions 128
8.2.1. Link model 128
8.2.2. Node models 128
8.2.3. Intersection delay model 129
8.3. Further research 130
8.3.1. Disaggregation of traffic flows to include Multiple User Classes (MUC) 130
8.3.2. Development of a solution method to consider stochastic intersection delays 130
8.3.3. Model validation 130
8.3.4. Evaluation of computational efforts 131

References 133

Nederlandstalige samenvatting 141

List of publications 155

About the author 157

NOTATION

Acronyms

ATIS	Advanced Traveler Information Systems
ATMS	Advanced Traveler Management Systems
BPR	Bureau of Public Roads
CTM	Cell Transmission Model
DNL	Dynamic Network Loading
DQM	Dynamic Queuing Model
DTA	Dynamic Traffic Assignment
DUE	Dynamic User Equilibrium
DUO	Dynamic User Optimum
FIFO	First-In-First-Out
LCTM	Lagged Cell Transmission Model
LTM	Link Transmission Model
MC	Multi-Commodity
MNL	Multinomial Logit
MP	Mathematical Programming
MSA	Method of Successive Averages
MUC	Multiple User Class
OCT	Optimal Control Theory
OD	Origin-Destination
P-Q	Point-Queue
SDUE	Stochastic Dynamic User Equilibrium
SO	System Optimum
VI	Variational Inequalities
VMS	Variable Message Sign

Sets

O	set of nodes
A	set of links
I_n	set of incoming links into node n
J_n	set of outgoing links out of node n
RS	set of origin-destination pairs
P	set of routes
P_{rs}	set of routes connecting origin-destination pairs rs, $P_{rs} \subset P$
X	total network space
T	total simulation time
Z	set of piecewise linear fundamental diagram branches

Indices

n	node, $n \in O$
r	origin node, $r \in R \subset O$
s	destination node, $s \in S \subset O$
a	link, $a \in A$
i	incoming link, $i \in I_n$
j	outgoing link, $j \in J_n$
ij	turning movement from incoming link $i \in I_n$ to outgoing link $j \in J_n$
rs	origin-destination pair, $rs \in RS$
p	route, $p \in P$
z	linear fundamental diagram branch, $z \in Z$

Space indices

x	point in network space, $x \in X$
x_a^0	entrance point (upstream end) of link a
x_a^L	exit point (downstream end) of link a
$x_{P-Q,ij}^0$	entrance point (upstream end) of the P-Q of turning movement ij
$x_{P-Q,ij}^L$	exit point (downstream end) of the P-Q of turning movement ij
Δx	space interval length

Time indices

t	point in time, $t \in T$
$t_x(N)$	time at which vehicle number N passed location x
Δt	time interval length, simulation time step

Link model variables

Macroscopic traffic flow variables

q	flow (*veh/h*)
k	density (*veh/km*)
v	average speed (*km/h*)
Q_e	equilibrium flow (*veh/h*)
κ	physically infeasible density (*veh/km*)

Given link variables

v_f	free flow speed, forward wave speed (*km/h*)
q_M	(link) capacity (*veh/h*)
k_M	critical density (*veh/km*)
k_{jam}	jam density (*veh/km*)
w	backward wave speed (*km/h*)
L	(link) length (*km*)
k_{queue}	queue density (*veh/km*)

Link model variables

$N\ (x,t)$	cumulative vehicle number at place x at time t (*veh*)
$S_{ij}\ (t)$	Sending flow (increment) from link i to link j during time interval $[t\ ,\ t+\Delta t]$ (*veh*)
$R_{ij}\ (t)$	Receiving flow (increment) of link j from link i during time interval $[t\ ,\ t+\Delta t]$ (*veh*)
$G_{ij}\ (t)$	transition flow (increment) from link i to link j during time interval $[t\ ,\ t+\Delta t]$ (*veh*)
δ	link-route incidence matrix, $\delta_{ap} = 1$ if route p contains link a and $\delta_{ap} = 0$ otherwise (-)

Node model variables

$q_{n,ij}\ (t)$	(node) capacity for turning movement ij, imposed by node n during time interval $[t\ ,\ t+\Delta t]$ (*veh/h*)
p_{ij}	priority parameter for turning movement ij (-)
$q_{M,ij}\ (t)$	saturation flow for turning movement ij during time interval $[t\ ,\ t+\Delta t]$ (*veh/h*)
g_{ij}	effective green time for turning movement ij (s)
c	cycle length (s)
$A_{\overline{ij},ij}$	conflict matrix to consider the priority rules (-)

$D_{\bar{ij}}$	set of turning movements that are conflicting with turning movement *ij* (-)
$t_{s,ij}$	occupation time caused by a vehicle from movement *ij* (*s*)
$t_{a,ij}$	time margin during which an approaching major stream vehicle from movement *ij* is blocking the intersection in advance of its arrival (*s*)

Intersection delay model variables

$D_{\text{int},ij}(t)$	intersection delay for turning movement *ij* at time *t* (*h*)
$D_{\widehat{\text{int},ij}}(t)$	target value of intersection delay for turning movement *ij* at time *t* (*h*)
$M_{P-Q,ij}(t)$	number of vehicles in the P-Q of turning movement *ij* at time *t* (*veh*)
$M_{\widehat{P-Q},ij}(t)$	target number of vehicles in the P-Q of turning movement *ij* at time *t* (*veh*)
$q_{P\text{-}Qin,ij}(t)$	inflow rate in the P-Q of turning movement *ij* during time interval [*t* , *t*+*Δt*] (*veh/h*)
$q_{P\text{-}Qout,ij}(t)$	outflow rate out of the P-Q of turning movement *ij* during time interval [*t* , *t*+*Δt*] (*veh/h*)
$S_{ij}'(t)$	P-Q Sending flow (increment) of link *i* to link *j* during time interval [*t* , *t*+*Δt*] (*veh*)
$G_{ij}'(t)$	P-Q transition flow (increment) from link *i* to link *j* during time interval [*t* , *t*+*Δt*] (*veh*)
$x_{ij}(t)$	degree of saturation for turning movement *ij* (flow to capacity ratio) during time interval [*t* , *t*+*Δt*] (-)

Logit model variables

μ	scale parameter of the MNL model (-)
$Pr(p(t))$	probability of choosing route *p* at time *t* (-)

DNL model variables

$D_{rs}(t)$	traffic demand for OD pair *rs* during time interval [*t* , *t*+*Δt*] (*veh*)
$f_{rs}^{p}(t)$	flow entering route $p \in P$ from origin *r* to destination *s* during time interval [*t* , *t*+*Δt*] (*veh/h*)
$\tau_{rs}^{p}(t)$	travel time on route *p* from origin *r* to destination *s* at time *t* (*h*)
$C_{rs}^{p}(t)$	travel cost on route *p* from origin *r* to destination *s* at time *t* (€)

1

INTRODUCTION

1.1. Background and Context

Dynamic Traffic Assignment (DTA) models are widely used tools for transportation planning and traffic management analysis. To forecast the impact of proposed transportation investments, and to estimate the effects of Advanced Traveler Information Systems (ATIS) and Advanced Traveler Management Systems (ATMS), traffic planners rely on DTA models.

The many applications of DTA models may be subdivided in two categories: equilibrium DTA applications and en-route DTA applications.

Equilibrium DTA applications include estimating ‘typical day’ traffic conditions, forecasting the impact of proposed changes to the transportation system, and testing and evaluating Advanced Traveler Management Systems (ATMS), such as ramp metering, signal coordination, tolling, and other control systems before they are implemented in practice.

En-route DTA applications include estimating the impact of incidents, special events, unusual weather, etc… , evaluating Advanced Traveler Information Systems (ATIS) strategies such as in-vehicle information and Variable Message Sign (VMS) information, and on-line (real-time) applications such as making short-term forecasts of the system state that are used by adaptive traffic control and traveler guidance systems.

It may well be clear that DTA models find wide application. However, the theory of DTA is still relatively undeveloped, which necessitates new approaches, developments and improvements that account for challenges from the application domain. In an attempt to fill one of the gaps in current DTA theory (cf. Section 2.3), this thesis presents a new model for Dynamic Network Loading (DNL), being one of the two fundamental components of a typical DTA model.

1.1.1. General framework of the DTA model

DTA models originate from their static counterpart. While traffic conditions in dynamic models are time-dependent, static assignment models assume that traffic conditions do not vary over time. However, congestion occurring in traffic networks is dynamic by nature. Travel times in congested networks highly depend on the accumulation of traffic and generally on the history of the system, which is not taken into account by static models. Ever-increasing congestion necessitated the use of dynamic models. Furthermore, the evaluation of time-dependent ATMS and ATIS strategies obviously requires the use of time-dependent or dynamic models.

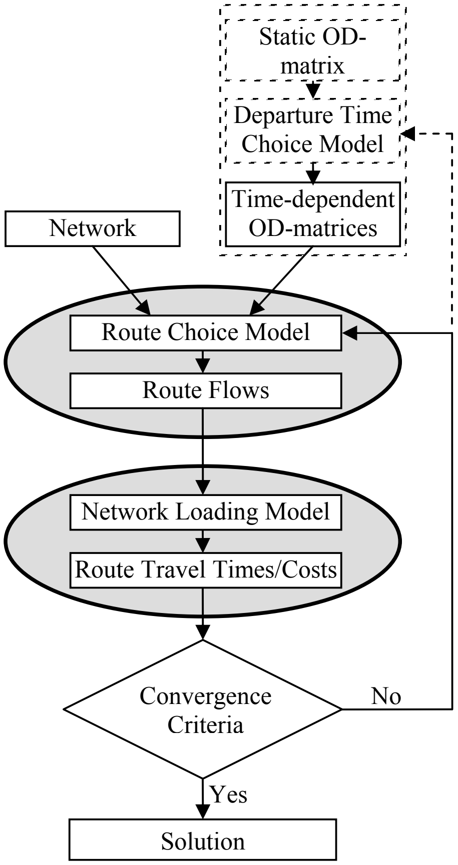

Static OD-matrix
Departure Time Choice Model
Time-dependent OD-matrices
Network
Route Choice Model
Route Flows
Network Loading Model
Route Travel Times/Costs
Convergence Criteria
No
Yes
Solution

Figure 1.1: General framework of the Dynamic Traffic Assignment (DTA) Model

Figure 1.1 illustrates the general framework of a DTA model.

DTA models have 2 fundamental components. The first component of the DTA model is the route choice model (i.e. the assignment mechanism) in which all travelers are assigned to a specific route. The second component of the DTA model is the Dynamic Network Loading (DNL) model in which traffic is propagated through the network along the assigned routes. In this thesis, we focus on developing a model for DNL.

Inputs to the route choice model are a traffic network, a traffic demand and route costs from a previous iteration.

The traffic network consists of links and nodes representing road segments and intersections. The network can be influenced by ATMS, like ramp metering, tolling, etc. Traffic demand is given by time-dependent Origin-Destination (OD) matrices, which for each time instant determine the departure rates from each origin node to each destination node. Time-dependent OD-matrices usually originate from a static OD-matrix, which constitutes the first three steps of the classical four-step planning process, namely trip production, trip distribution and modal split.

Although in this thesis time-dependent OD-matrices are assumed to be given, implying that travelers have chosen their departure times and do not alter these anymore, general DTA models include both route- and departure time choice models which can be run either sequentially or simultaneously (Bliemer (2001)).

Some DTA models even include long-term mobility decisions, such as land-use, residential locations and car ownership, which affect the static OD-matrix. Examples are given in Chapter 2.

The route choice model determines which route travelers take for their journey. Three different types of route-assignment are distinguished in the next section. Resulting route flow rates are input for the DNL model. DNL models propagate traffic through the network along the assigned routes and they determine link and route speeds and densities, and more importantly, link and route travel times and travel costs.

1.1.2. Types of assignment

In a Dynamic User Equilibrium (DUE) assignment, the two major components of the DTA model, being the route choice model and the DNL model, are typically somehow iterated in sequence until they converge into a Dynamic User Equilibrium (DUE). Route choices are based on actual experienced travel times. The DUE can be defined as an extension of Wardrop's first principle (Wardrop, 1952): the DUE requires that, at network equilibrium, no traveler who departed during the same time interval can reduce his or her travel costs by unilaterally changing routes. Each user non-cooperatively seeks to minimize his cost of transportation. An alternative but equivalent statement is that all routes used between an origin-destination pair have the same minimal cost, and no unused route has a lower cost for travelers that departed during the same time interval. DUE assignments are typically used for planning purposes, for estimating 'typical day' traffic conditions, and for testing ATMS (Advanced Traveler Management Systems).

A similar iterative procedure is used for a System Optimal (SO) assignment. The SO requires that total travel costs of all travelers together are minimal. Each user behaves cooperatively in choosing routes to ensure the most efficient use of the whole system. In general, the basic policy for the SO is to assign traffic demand to the network, just equal to the network capacity, but not more than capacity (Kuwahara and Akamatsu (2001)). This type of assignment is typically enforced by control strategies such as tolling. Theoretical specifications and practical applications can be found in Yperman (2005a, 2005b, 2005c) and Logghe and Yperman (2003).

No iterative procedure occurs in a Dynamic User Optimal (DUO) assignment, where route choice is based only on present instantaneous travel times. This type of assignment is also referred to as en-route assignment or reactive assignment. The DUO assignment can be considered as a simplified representation of route choices brought by ATIS (Advanced Traveler Information Systems), where variable message signs, highway radios, in-vehicle navigation equipment, etc. frequently supply current traffic information (Kuwahara and Akamatsu (2001)).

At the heart of all types of above-mentioned assignments is the DNL model. A new DNL model is developed in this thesis. Properties and impacts of the developed DNL model are illustrated within a DTA framework.

1.2. Objectives and scope

1.2.1. Objective

The objective of this thesis is to develop a DNL model that realistically describes traffic propagation on both motorways and urban roads in practical large scale traffic networks. To achieve this main goal, the model attempts to provide:

- A description of traffic propagation on network links that is consistent with first order kinematic wave theory.
- A description of traffic dynamics at signalized and unsignalized urban intersections that is consistent with state-of-the-art queuing theory.
- A high computational efficiency such that DNL in large scale networks can be performed in a reasonable amount of time.

An alternative formulation of this thesis' objective is “the development of a computationally efficient algorithm, where traffic is propagated as in kinematic wave theory and where traffic dynamics at intersections, such as local flow restrictions and intersection delays, are accounted for as in queuing theory”.

1.2.2. Model validation

Both kinematic wave- and queuing theory have extensively been validated in literature (see for example Jin (2002) and Troutbeck and Brilon (2000)). It is not our intention to re-validate these models in this thesis. We rather explore whether the developed algorithm combines both models in a way that they don't affect each other's validity.

1.2.3. Higher order traffic phenomena

Traffic propagates on links as assumed in first order kinematic wave theory where traffic conditions are stable. Acceleration, deceleration, nor anticipation behavior is taken into account.

Therefore, higher order traffic phenomena such as the emergence of stop-and-go waves, oscillatory congested traffic, or capacity drop (which is caused by the so-called hysteresis effect, meaning that the sequence of traffic states follows a different path in the fundamental diagram during breakdown to congestion than during recovery from congestion (Tampère (2004)), are not considered.

1.2.4. Multiple user classes

The kinematic wave traffic flow model described in this thesis is a single user class model. Vehicle characteristics and driver behavior are assumed to be equal for all travelers. Recently, Logghe (2003) came up with a Multiple User Class (MUC) implementation of kinematic wave theory. Though MUC is out of the scope of this thesis, it seems not inconceivable to adopt Logghes approach. This is a topic for future research (cf. Section 8.3.1).

1.2.5. Time-dependent delay at intersections

Traffic dynamics at intersections are described as in queuing theory. Local flow restrictions are taken into account for both signalized and unsignalized intersections. Furthermore, a method to include intersection delays is elaborated for deterministic intersection delays at signalized intersections, which are independent on time.

Detailed solution method descriptions for time-dependent stochastic delays at signalized and unsignalized intersections however, are out of the scope of this thesis.

1.2.6. DTA components

The scope of this thesis is limited to the DNL component of DTA. Other DTA related functionalities, such as route generation, route choice modeling, iteration schemes, convergence criteria etc… may be used in this thesis to demonstrate the properties of the developed DNL model, but they are not the object of evaluation.

1.3. Thesis contributions

The major contribution of this thesis is the development of a computationally efficient algorithm for large scale DNL, called Link Transmission Model (LTM), where both traffic propagation on network links (consistent with kinematic wave theory) and traffic dynamics at intersections (consistent with queuing theory) are described in a realistic way.

The algorithm's computational efficiency is enhanced by using large time steps to walk through simulations. This procedure was enabled due to:

- the development of a numerical solution procedure to integrate Newell's (1993) method for updating kinematic wave traffic conditions on long links

- the development of a method to implicitly include average flow restrictions at intersections and average intersection delays in a flow-based model

Other contributions include

- Development of a Multi-Commodity DNL model, where vehicles are disaggregated by route (each commodity corresponds to a specific route). A multi-commodity model keeps track of the routes of the vehicles at all times, when describing the collective motion of traffic. Due to the disaggregation of traffic flows by route, vehicles can actually be assigned to specific pre-defined routes, as governed by the route choice model.

- Extension of the traditional kinematic wave approach where traffic is characterized by a triangular shaped fundamental diagram, to include characterization by piecewise linear fundamental diagrams.

- Formulation of algorithm specifications, based on which LTM has been implemented as DNL model within state-of-the-art DTA model INDY (Bliemer (2003, 2004, 2005)).

- Development of urban node models for signalized and unsignalized intersections where the average capacity effects of traffic lights and priority streams are taken into account implicitly.

- Demonstration of the practical feasibility of the developed model for large sized transportation networks.

1.4. Thesis outline

Chapter 2 of this thesis gives a concise overview of DTA approaches, with special emphasis on their DNL component. It is shown how the developed LTM fits in and contributes to the field of DTA and DNL modeling. The remainder of the thesis is structured as indicated in Figure 1.2:

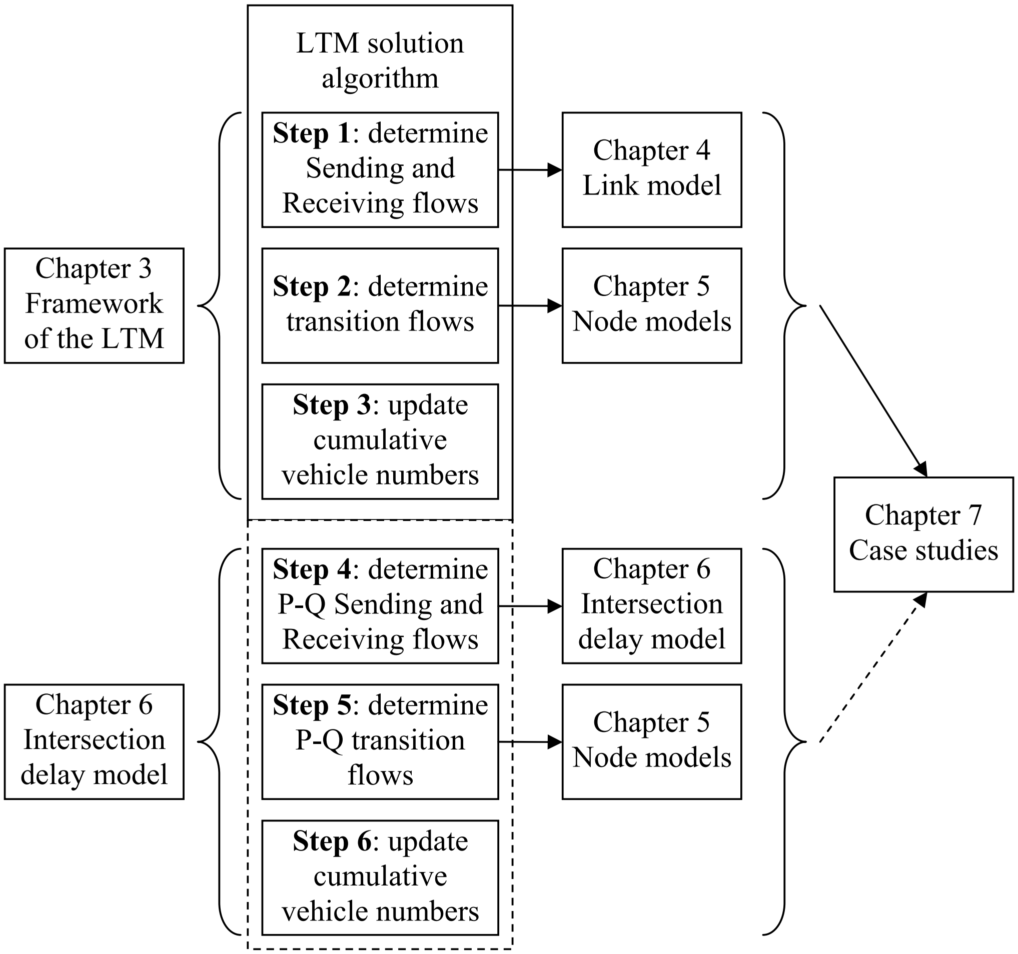

LTM solution algorithm
Step 1: determine Sending and Receiving flows
Step 2: determine transition flows
Step 3: update cumulative vehicle numbers
Step 4: determine P-Q Sending and Receiving flows
Step 5: determine P-Q transition flows
Step 6: update cumulative vehicle numbers
Chapter 3 Framework of the LTM
Chapter 4 Link model
Chapter 5 Node models
Chapter 6 Intersection delay model
Chapter 5 Node models
Chapter 6 Intersection delay model
Chapter 7 Case studies

Figure 1.2: Structure of the thesis chapters

Chapter 3 presents the basic LTM solution algorithm. In three successive steps, Sending and Receiving flows are determined, transition flows are determined and cumulative vehicle numbers are updated. To accomplish these different algorithm steps, link and node models are developed in Chapters 4 and 5.

The link model in Chapter 4 propagates traffic on network links as assumed in kinematic wave theory.

Node models are developed in Chapter 5 to represent traffic dynamics at both motorway junctions and urban intersections.

Chapter 6 extends the basic LTM solution algorithm to include the modeling of intersection delays at urban intersections. The chapter introduces a method to implicitly include intersection delays in flow-based continuum models like LTM. This method is used to determine P-Q Sending flows, which constitutes the fourth step of the extended algorithm.

Two case studies are elaborated in Chapter 7 to illustrate some characteristics of the developed LTM. The first case study considers a small theoretical network, where queue propagation in LTM is compared with queue propagation in a Dynamic Queuing Model (DQM). A second case study demonstrates the applicability of LTM in a large scale network. Queue propagation characteristics in LTM and DQM are compared and a small sub-network is considered to show the impact of taking into account intersection delays. Finally in Chapter 8, the main conclusions of this thesis research are formulated and possible directions for further research are indicated.

2

OVERVIEW OF DTA AND DNL APPROACHES

In this thesis, we propose a DNL model that fits within a DTA framework. To clarify our approach and its properties, a concise state-of-the-art overview of DTA models is given below, with special emphasis on DNL procedures. It is not our intention to give an exhaustive literature review on DTA. Such a review can for example be found in Peeta and Ziliaskopoulos (2001). We rather discuss a few state-of-the-art models that serve as an example of a certain approach.

Two distinct approaches have dominated the methodologies applied to the DTA research: the analytical and the simulation-based approach.

2.1. Analytical approach

Analytical DTA approaches typically focus on dynamic user equilibrium (DUE) or System Optimum (SO) objectives. Achieving (unique) DUE or SO solutions is the ultimate objective; other issues like the realism of traffic propagation are of secondary importance (Szeto, 2003).

The DTA problem (how to find DUE or SO solutions) is formulated as a sound mathematical problem - either a mathematical programming problem (e.g. Merchant and Nemhauser (1978a, 1978b), Carey (1987), Janson (1991a, 1991b) and Ziliaskopoulos (2000)), an optimal control problem (e.g. Friesz et al. (1989), Wie (1991) and Ran et al. (1993)), or a variational inequality problem (e.g. Friesz et al. (1993), Ran and Boyce (1996) and Bliemer and Bovy (2003)) - which is directly solved using well-known mathematical optimization techniques. The DNL model thereby typically takes the form of a set of constraints that describe traffic propagation. Solving the mathematical problem is equivalent to solving the DUE or SO problem.

Because of the limitations of mathematical programming (MP) and optimal control theory (OCT) in DTA context and the advantages offered by variational inequalities (VI), analytical DTA models have in recent years migrated towards the VI approach. An extended overview of MP, OCT and VI approaches can be found in Peeta and Ziliaskopoulos (2001).

Advantages of the analytical approach are the ability to apply existing mathematical solution algorithms to solve the DTA problem, and the ability to determine solution properties such as existence and uniqueness beforehand.
However, a theoretical guarantee of properties such as existence and uniqueness imposes restrictions on the so-called mapping function (mapping route flows on travel times), which prevents realistic representation of traffic propagation, as further discussed in the next paragraph.
Another drawback is the limited applicability of analytical models. The successful use of analytical models is usually limited to small hypothesized networks, as these models use solution procedures that do not take advantage of the specific characteristics of the transportation problems (Bliemer (2006, 2007)).

Dynamic Network Loading (DNL) in analytical DTA models

Peeta and Ziliaskopoulos (2001) distinguish three main approaches on capturing traffic flow propagation (DNL) within an analytical DTA context: exit functions, link performance functions and cell transmission models.

Exit functions have been used extensively to propagate traffic in DTA models, first by Merchant and Nemhauser (1978a, 1978b) and later also by Carey (1987), Friesz et al. (1989) and Wie et al. (1995). Exit functions determine the outflow from a link given the number of vehicles on it, implicitly assuming that changes in density propagate instantaneously across the link. Travel times of vehicles depend on traffic conditions behind these vehicles, which aside form being unrealistic, typically leads to a violation of the first-in-first-out (FIFO) condition on the link, as demonstrated in Carey (1986, 1987). Furthermore, these models fail to capture some fundamental traffic dynamics such as the spillback of queues.

Link performance functions have a.o. been used by Janson (1991), Ran et al. (1996), Chen and Hsueh (1998), and Bliemer and Bovy (2003). They are typically straightforward temporal extensions of static Bureau of Public Roads (BPR) functions, which express travel time as a function of traffic volume. At the time of link entrance, link travel times are determined as a function of the number of vehicles on it, again allowing violations to link FIFO behavior. Capacity constraints are not explicitly taken into account. Similar to exit functions, link performance functions cannot capture queue spillbacks.

Cell transmission models are discrete versions of the continuum kinematic wave model of traffic flow, introduced by Lighthill and Whitham (1955) and Richards (1956). These models provide more realism in the propagation of traffic. They describe dynamic traffic conditions on a road network, including shock waves and the propagation of queues. Daganzo (1994) and Lebacque (1996) came up with the first cell-transmission model (CTM). Every link of the road network is divided into homogeneous cells such that the length of each cell is equal to the distance traveled by the free-flow moving vehicles in one time interval. Traffic conditions are updated in successive time steps. Ziliaskopoulos (2000) suggested that the CTM could be captured by a set of constraints describing traffic propagation in analytical DTA. This approach has also been explored by Waller and Ukkusuri (2003), Lo (1999), Lo and Szeto (2002) and Szeto and Lo (2004). Kuwahara and Akamatsu (2001) and Roels and Perakis (2004) propose a set of constraints that is consistent with Newell's (1993) discrete version of the continuum kinematic wave model of traffic flow.

As stated above, a theoretical guarantee of properties such as existence and uniqueness imposes restrictions on the mapping function, which prevents a realistic representation of traffic propagation.
The existence of solutions requires the mapping function of the problem to be continuous (Nagurney 1993) whereas the uniqueness of solution requires the mapping function to be strictly monotonic (Nagurney 1993).Thus, solution existence (uniqueness) requires route travel times to be continuous (strictly monotone) with respect to route flows. Route travel times resulting from exit functions and link performance functions typically meet these requirements. For route travel times resulting from the CTM however, Waller and Ukkususri (2003) found the assumption of strict monotonicity with respect to route flows 'clearly problematic' in bottleneck networks. Moreover, Szeto and Lo (2006) even found that under congested conditions, route travel times obtained with the CTM may become discontinuous, making it impossible in certain cases to find solutions that satisfy the equilibrium route choice principle.
The finding that capturing detailed traffic dynamics, such as queue spillback, may violate the requirement on solution existence, resulting in the non-existence of DUE-solutions (Szeto (2003)) shakes the very foundation of the analytical DTA approach. In general, Szeto and Lo (2006) clearly observe the trade-off between the existence of solutions and the levels of traffic dynamics captured; point-queue DTA solutions always exist whereas those for physical-queue problems may not.
The CTM does not yield well-behaved mathematical formulations. Moreover, analytical representations of traffic flow that adequately replicate traffic theoretic relationships and yield well-behaved mathematical formulations are currently unavailable (Peeta and Ziliaskopoulos (2001)). Likewise, traffic flows and interactions at urban intersections do not have acceptable analytical formulations. Heuristic approaches provide far more flexibility.

2.2. Simulation-based approach

Simulation-based DTA models focus on enabling practical deployment for realistic networks. They are designed to handle transportation problems in real-life networks. Important objective thereby is to provide realism in traffic propagation and driver behavior. Properties like existence and uniqueness of DUE or SO solutions are of secondary importance.

Advantages of simulation-based models are the applicability in real-life networks and the ability to adequately capture traffic dynamics and driver behavior.

Drawbacks include the fact that (i) theoretical insights cannot be analytically derived, since traffic propagation is modeled using simulation, and that (ii) solution properties like existence and uniqueness are not guaranteed and cannot be determined in advance.

However, claims of existence and uniqueness of DUE or SO solutions may not be essential nor particularly meaningful from a practical standpoint. If the objective is to reflect reality as closely as possible and if there is no equilibrium in reality, why then should the solution be an equilibrium? The question whether equilibrium actually takes place or is a mathematical construct is a very old issue. But even if equilibrium does not actually occur, the concept still provides insight, it gives policy-makers something to hold on to, and it still has a direction-finding role in dynamic networks.

2.2.1. Equilibrium versus en-route DTA models

While analytical approaches focus on equilibrium objectives, simulation-based approaches distinguish between two mechanisms to modeling route choice: equilibrium assignment and en-route assignment.

In the equilibrium approach, user equilibrium typically results from an iterative procedure, where route choices in one iteration step are based on experienced travel times in the previous iteration. It is assumed that travelers have ‘perfect knowledge’ of travel times on all links at all times. The iterative procedure can be thought of as travelers’ day-to-day learning and adaptation to experienced traffic conditions, until equilibrium is reached. Day-to day behavior refers to the response of travelers to changes in the characteristics of the system, given long-term decisions. The equilibrium assignment is typically used for estimating ‘typical day’ traffic conditions, forecasting the impact of proposed changes to the transportation system, and testing and evaluating ATMS (Advanced Traveler Management Systems).

In the en-route approach, there is no iterative procedure and the solution is not necessarily (or probably not) equilibrium. Kuwahara (2001) refers to the solution as a Dynamic User Optimum (DUO). Route choices are based on instantaneous travel times (at the present time) and as opposed to the equilibrium approach, they are independent on future travel times. Travelers already on the network can modify their route during the journey (within-day behavior), as link travel times are updated after each time interval. One can think of within-day behavior as the response of travelers to disturbances in the transportation system. The en-route assignment is used to estimate the impact of incidents and unusual events (in an equilibrium assignment, travelers would know beforehand the incident was going to happen) and to evaluate ATIS (Advanced Traveler Information Systems) strategies. The route choice for a typical day, resulting from the equilibrium assignment, is in both cases used as “do-nothing” alternative.

2.2.2. Micro-, meso- and macroscopic DTA models

Table 2.1 gives an overview of the properties of simulation-based DTA models.

Table 2.1: General properties of simulation-based DTA models

	Microscopic simulation model	Mesoscopic simulation model	Macroscopic simulation model
Description of traffic flow propagation	Microscopic	Macroscopic	Macroscopic
Representation of trip-maker decisions	Microscopic	Microscopic	Macroscopic
Calibration efforts	High	Medium	Low
Computation times	High	Medium	Low
Applicability	Small networks	Medium sized networks	Large networks
Assignment type	mostly En-route	En-route & Equilibrium	mostly Equilibrium

These properties are discussed in the remainder of this section.

2.2.2.1. Microscopic simulation-based DTA models

Microscopic simulation-based DTA models, such as TRANSIMS (Nagel (1998)), DYNAMEQ (Florian et al. (2006)), and the micro-simulation models AIMSUN2 (Barcelo, 2002), PARAMICS (Quadstone Limited, 2000), MITSIM (Yang (1997)) and VISSIM (PTV (2005)), describe traffic flow propagation (DNL) on the level of individual vehicles. Trip-maker decisions such as route choice are represented on the individual level as well.

To propagate traffic, micro-simulation models like AIMSUN2, PARAMICS, MITSIM and VISSIM combine mathematical car-following and gap acceptance models, with heuristic models that represent driver behavior (aggressiveness, lane-changing behavior etc…). Individual vehicles are moved according to rules operating on the individual level. Traffic conditions are sensitive to a significant number of model parameters which are not directly measurable. Therefore, these models are hard to calibrate and they provide some kind of illusive accuracy.

In the microscopic TRANSIMS model, traffic propagation is based on a cellular automata technique for car-following and lane-changing, enhanced by additional rules for elements such as signals, unprotected turns, weaving lanes, etc… (Nagel (1998)). The lanes of all network links are divided into cells of equal size that are either empty or occupied by one single vehicle. Local rules determine the speed and position of each individual vehicle. Since there are less model parameters, cellular automata models are easier to calibrate. Maerivoet (2006) also proposes a cellular automaton to propagate traffic in a simulation-based DTA context.

The DNL model in the microscopic model DYNAMEQ attempts to capture the effects of car-following, lane-changing and gap acceptance with a minimum number of model parameters. This leads to a reduction in calibration effort. The simulation, which is based on a simplified car-following relationship, is a discrete-event procedure. This leads to a sharp reduction in computational effort as well, when compared to microscopic discrete time approaches. Mahut (2000) provides a detailed description of the DNL model. Though categorized as a microscopic simulation-based DTA model, DYNAMEQ properties such as calibration efforts, computation times and applicability, have actually more in common with mesoscopic approaches.

Since data have to be stored for each individual vehicle, computation times in microscopic simulation models are high; they are proportional to the amount of vehicles on the network. Due to high computation and calibration efforts, the successful use of microscopic simulation-based DTA models is mostly limited to small size networks.

The microscopic simulation of traffic propagation has a stochastic nature. One simulation run represents only one sample in a whole spectrum of possible solutions. Since route choices are based on average travel times, one network-loading step should be composed of several simulation runs. Such a procedure is not always done, since it substantially increases computation times of one single network loading.

While most micro-simulation models use an en-route approach, Casas (2004) recently showed that the AIMSUN2 traffic flow model can also be used in conjunction with an iterative assignment method. TRANSIMS and DYNAMEQ deal with equilibrium assignment as well.

Some microscopic simulation models constitute more than the simple DTA pointed out in Figure 1.1. For example, the approach taken in the agent-based micro-simulation model TRANSIMS is to represent each individual agent in a metropolitan region and to simulate all aspects of his/her decision-making as long as it is related to transportation. TRANSIMS explicitly simulates the first three steps of the classical four step planning process.

2.2.2.2. Mesoscopic simulation-based DTA models

Mesoscopic simulation models such as CONTRAM (Taylor (2003)), DYNASMART (Mahmassani et al. (2001)) and DYNAMIT (Ben-Akiva et al. (1998)), move individual (packets of) vehicles according to macroscopic traffic flow relations. Vehicle movements are governed on an aggregate level, while trip-maker decisions such as route choice are made individually; i.e. a microscopic level of representation of individual trip-maker decisions is combined with a macroscopic description of traffic flow propagation.

The macroscopic simulation of traffic propagation can be described with less complex deterministic models, delivering a repeatable average result for a given data set.
Traffic conditions depend on few flow parameters that are directly measurable, which significantly reduces calibration efforts.
Mesoscopic models are generally less precise in the representation of traffic dynamics. However, the representation varies substantially in different DNL models. The DNL model in CONTRAM is based on time-dependent queuing theory. Capacity constraints at the link end yield queues whenever demand exceeds capacity. The model describes an unrealistic Point-Queuing process which does not take into account the physical space occupied by the queue. The travel-time based traffic model yields incorrect densities and FIFO behavior is not obeyed.
DYNASMART and DYNAMIT use flow-based traffic models that propagate individual vehicles on links according to a modified Greenshield type speed-density relationship. These models consider physical queues on links that are artificially split up into a moving part and a queuing part. Traffic conditions in the queuing part are fixed, which results in a less realistic description of traffic propagation, as will be shown in Chapter 7 of this thesis.

Computation times in mesoscopic models, though still in proportion with the number of vehicles on the network, are significantly reduced compared to microscopic models due to an aggregate description of traffic flow. Mesoscopic models can computationally succeed in the analysis of medium-sized networks.
The microscopic simulation of trip-maker decisions such as route choice, departure time choice and en-route behavior and response to information, requires complex behavioral models that can be hard to calibrate. The stochastic nature of such models furthermore necessitates multiple simulation runs to compose a spectrum of possible solutions.
Mesoscopic models typically carry out both equilibrium assignments and en-route assignments. The microscopic representation of trip-maker decisions allows for a simple incorporation of multiple user classes in terms of information availability, and in terms of behavior and response to information. The en-route approach requires a distinction of these user classes.
Some mesoscopic models also constitute more than the simple DTA described above. DynaMIT for example, estimates and predicts OD-demand using Kalman filtering methodology. It considers both historical information and the driver response to information to estimate and predict in real-time current and future traffic conditions.

2.2.2.3. Macroscopic simulation-based DTA models

Macroscopic simulation models, such as INDY (Bliemer et al. (2004), Bliemer (2005)) or METANET (Messmer and Papageorgiou (1990)), describe both traffic flow propagation and trip-maker decisions on the aggregate level. Traffic is considered to be a continuum, both as far as vehicle movements and trip-maker decisions are concerned.

The accuracy of traffic flow representation in macroscopic models highly depends on the DNL mechanism.
Two different DNL models are available within INDY, which is further referred to as INDY-TT or INDY-DQM depending on which of these two DNL models is used. The DNL model in INDY-TT is based on Travel Time (TT) functions.
At the time of link entrance, link travel times are determined as a function of the number of vehicles on the link. Capacity constraints are not explicitly taken into account. Traffic description in this model suffers from drawbacks of realism.
The DNL model in INDY-DQM is based on a Dynamic Queuing Model (DQM), which considers physical queues on links that are split up into a moving part and a queuing part. Instead of predicting link travel times at the time of link entrance, the model first determines link flows, based on traffic conditions. Only afterwards, link travel times can be derived from these link flows. The DNL model in INDY-DQM, though being far more realistic than the one in INDY-TT, still lacks some realism in the representation of traffic dynamics, as discussed in Chapter 7.

Trip-maker decisions are simulated with macroscopic models. These models are usually deterministic, delivering one repeatable average result for a given data set, whereas mesoscopic models rather deliver a spectrum of possible results.
Computation times no longer depend on the amount of vehicles on the network, and DTA is possible for large networks with millions of vehicles (see e.g. Bliemer (2006, 2007)).
To be able to simulate the result of route choice decisions (i.e. to be able to simulate route flow rates that are consistent with the route choice decisions), traffic flows in INDY are disaggregated by route so that traffic can actually be assigned to a specific route, as governed by the route choice model. Flows that are disaggregated by route are often referred to as ‘multi-commodity’ flows.
Traffic flows in METANET are not disaggregated by route. Single commodity flows are routed by splitting proportions at nodes of the network. The routes that are followed in a random network loading are generally not consistent with the route flow rates resulting from the route choice model, since splitting proportions that correspond to a given route flow set are not (easily) determinable.
Existing macroscopic DTA models typically use an equilibrium approach, requiring traffic flows to be disaggregated by route. For an en-route approach, simulating the result of en-route trip-maker decisions like behavior and response to information would require a further disaggregation of traffic flows (disaggregation by behavior etc…). This is an interesting topic for future research.

2.3. Discussion

As pointed out in the above overview, a lot of different approaches on DTA exist. Each approach has its own specific characteristics and each approach addresses specific problems or questions that other approaches cannot address that well. Therefore, it seems interesting to keep developing the different approaches simultaneously. The discussion is not whether one approach is better than the other, but it is about the most appropriate use of each approach.

Analytical models for example, are especially useful to generate theoretical insights, to analyze system properties and to explore new directions to address problems, rather than to perform close-to-reality simulations of real-world networks.

Micro-simulation models for example, should only be used if the application asks for a detailed representation of each individual vehicle.

In the field of simulation-based DTA models, significant progress is made by many researchers in the last decade. However, a model describing realistic traffic dynamics on both motorways and urban regions of large scale networks in a reasonable amount of time is still lacking.

The Link Transmission Model (LTM) presented in this thesis attempts to fill this gap. LTM is a DNL model for a macroscopic simulation-based DTA model; vehicles are moved as a continuum. In LTM, traffic propagation on network links is consistent with kinematic wave theory. This theory provides substantially more realism in the representation of queue-propagation (blocking back) and queue-dissipation, compared to existing approaches in macroscopic DTA models. Furthermore, LTM considers a detailed description of traffic dynamics at signalized and unsignalized intersections. Local flow restrictions and experienced intersection delays are consistent with state-of-the-art queuing theory. Since the LTM solution algorithm is computationally efficient and walks through simulations in large time steps, large scale networks can be dealt with in a small amount of time (Yperman (2005d, 2006)).

3

FRAMEWORK OF THE LINK TRANSMISSION MODEL

This chapter presents the basic LTM solution algorithm. First, some relevant concepts are defined.

3.1. Definitions

LTM

The Link Transmission Model (LTM) is a model for Dynamic Network Loading (DNL): it determines time-dependent link volumes, link travel times τ_a and route travel times τ^p in traffic networks, given the time-dependent route flow rates $f^p(t)$ for a fixed time period.

Network

Traffic networks consist of homogeneous unidirectional links a, which start at place x_a^0 and end at place x_a^L. The links can have any length L_a and they are connected to each other via nodes.

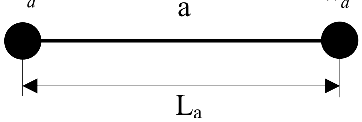


Figure 3.1: Length and boundaries of link *a*

A route *p* is a series of links *a* and nodes *n* between an origin node *r* and a destination node *s*. *P* is the set of all routes *p* on the network. Nodes have no physical length. They act merely as a flow exchange medium. Figure 3.2 shows some possible node configurations: inhomogeneous node, origin node, destination node, diverge node, merge node and cross node.

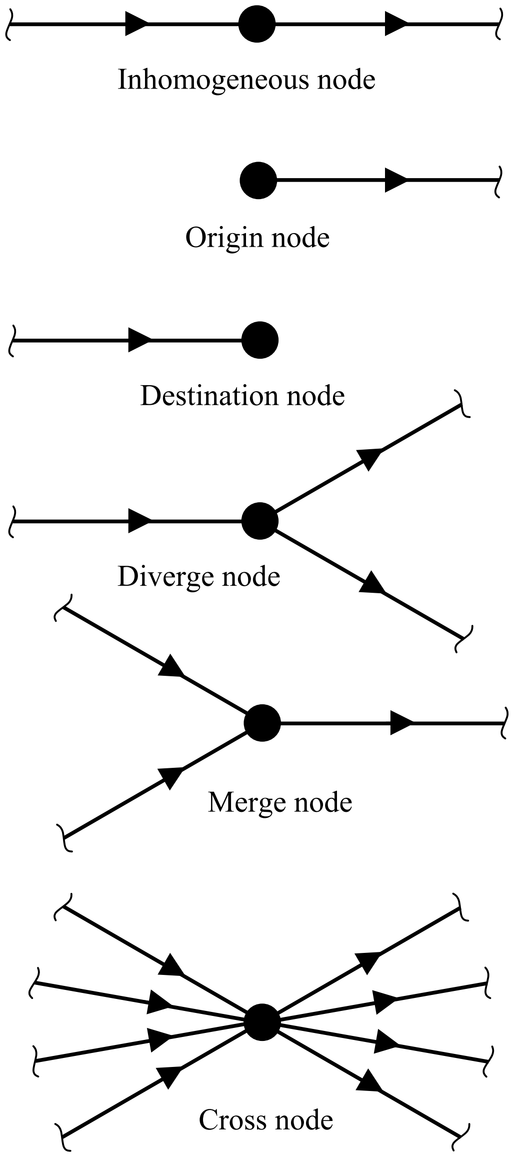

Inhomogeneous node
Origin node
Destination node
Diverge node
Merge node
Cross node

Figure 3.2: Different node configurations

A general traffic network can be represented by a combination of links and these basic nodes. Traffic is loaded on to the network in an origin node and it leaves the network in a destination node. An inhomogeneous node can be used to model a change in capacity or in any other characteristic on a road. Diverge nodes and merge nodes are respectively used to model diverging lanes/off-ramps and merging lanes/on-ramps in motorway networks.

While the maximum number of links entering and/or leaving a merge or a diverge node is 3, cross nodes connect an arbitrary number of incoming links i to an arbitrary number of outgoing links j. Cross nodes are used to represent urban intersections.

The behavioral rules behind these node models are discussed in Chapter 5.

Cumulative vehicle numbers and link travel times

The cumulative number of vehicles that pass location x by time t is indicated as $N(x,t)$. Suppose that an observer at location x numbers the vehicles consecutively as they pass him, and he attaches the numbers to the vehicles, then $N(x,t)$ represents the number of the last vehicle to pass the observer before time t.

LTM primarily determines the cumulative number of vehicles $N(x,t)$ that pass locations x_a^0 and x_a^L of each link a by time t. Only afterwards, when vehicles have left the link, link volumes and link travel times are derived from these cumulative vehicle numbers, as shown in Figure 3.3.

In Figure 3.3, the vertical distance between the curves $\mathrm{N}(\mathrm{x}_\mathrm{a}^0,\mathrm{t}_1)$ and $\mathrm{N}(\mathrm{x}_\mathrm{a}^\mathrm{L},\mathrm{t}_1)$ represents the number of vehicles on link a at time t_1 (traffic volume). The link travel time τ_a of the hth vehicle on link a is represented by the horizontal distance between these curves at height h, if vehicles do not pass each other. The determination of link travel times thus requires first-in-first-out (FIFO) behavior on each network link. LTM ensures this FIFO condition.

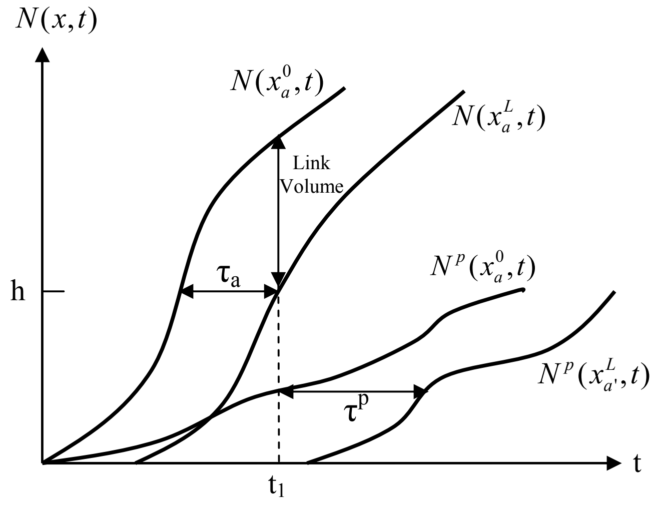


Figure 3.3: Cumulative vehicle numbers as a function of time

Multi-commodity traffic and route travel times

LTM is a multi-commodity (MC) model, where each commodity corresponds to a specific (pre-defined) route. Vehicles are disaggregated by route. We keep track of the routes of the vehicles at all times, when describing the collective motion of the traffic stream. This disaggregation by routes is necessary to use route choice information within the model.

$N^p(x_a^0,t)$ represents the cumulative number of vehicles on route p, that pass location x_a^0 by time t. The representation in terms of disaggregated cumulative vehicle numbers allows for a simple derivation of route travel times.

If origin node r and destination node s of route p are respectively connected to links a and a', i.e. if link boundary x_a^0 ($x_{a'}^L$) borders on node r (s), then the route travel time τ^p of route p is represented by the horizontal distance between the curves $N^p(x_a^0,t)$ and $N^p(x_{a'}^L,t)$.

Figure 3.3 indicates the travel time of route p for a vehicle departing at time t_1.

Since nodes have no physical length, route travel times only consist of link travel times. Times spent on nodes are not taken into account.

For all locations x and times t, the cumulative vehicle number $N(x,t)$ is the sum of the cumulative vehicle numbers disaggregated by route:

$$N(x,t) = \sum_{p \in P} N^p(x,t) \qquad \textit{for all } x \in X,\ t \in T \tag{3.1}$$

Inverse cumulative vehicle function

The inverse function of the cumulative vehicle number $N_x^{-1}(N)$ determines the time $t_x(N)$ at which vehicle number N passed location x.

Since the LTM solution algorithm only calculates cumulative vehicle numbers on discrete time steps $t + m\Delta t$ (where m is an integer), an interpolation procedure might be necessary to calculate $t_x(N)$. As shown in Figure 3.4, we propose a linear interpolation procedure.

$$t_x(N) = t + \frac{N - N(x,t)}{N(x,t+\Delta t) - N(x,t)}\,\Delta t = t + \alpha\Delta t \tag{3.2}$$

where $N(t) \le N \le N(t+\Delta t)$.

The disaggregated cumulative vehicle numbers are calculated as follows:

$$N^p(x,t_x(N)) = N^p(x,t) + \alpha\left(N^p(x,t+\Delta t) - N^p(x,t)\right) \qquad \textit{for all } p \in P \tag{3.3}$$

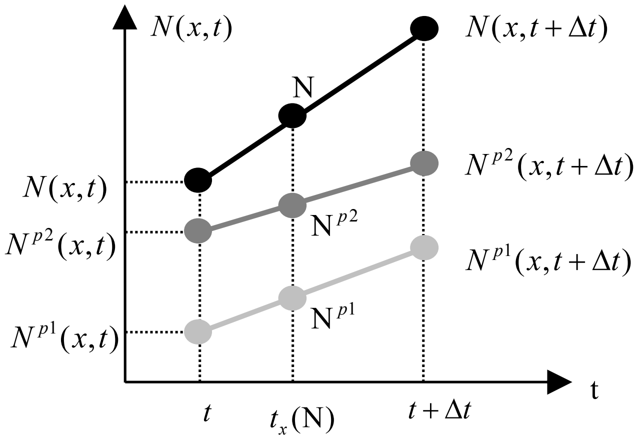


Figure 3.4: Linear interpolation of cumulative vehicle numbers

Sending flows, Receiving flows and transition flows

The Sending flow $S_i(t)$ of link *i* at time *t* is defined as the maximum amount of vehicles that could leave the downstream end of this link during time interval [*t, t+Δt*], if this link end would be connected to a traffic reservoir with an infinite capacity.
The Receiving flow $R_j(t)$ of link *j* at time *t* is defined as the maximum amount of vehicles that could enter the upstream end of this link during time interval [*t, t+Δt*], if a traffic reservoir with an infinite traffic demand would be connected to this link end.
Section 4.6 explains how Sending and Receiving flows are constrained by the traffic flow model.
Transition flow $G_{ij}(t)$ is defined as the amount of vehicles that are actually transferred from link *i* to link *j* during time interval [*t , t+Δt*].
Chapter 5 explains how transition flows are determined by node models.

Note that the Sending, Receiving and transition flows actually represent flow increments (numbers of vehicles). Actual flows are in this thesis referred to as flow rates.

3.2. LTM solution algorithm

The solution procedure divides the simulation time period *T* into time steps *Δt*. The time step should be smaller than the smallest link travel time to prevent vehicles from traversing a link within one time period (this condition is known as the Courant-Friedrichs-Lewy (CFL) condition):

$$\Delta t \le \frac{L_a}{v_{f,a}} \quad \text{for all } a \in A \tag{3.4}$$

where $v_{f,a}$ is the free-flow speed, i.e. the maximum speed of vehicles in link a.
For each time interval Δt, the algorithm involves three steps:

LTM solution algorithm

For each time interval Δt,
For each node n,

Step 1: For each incoming link $i \in I_n$, determine the Sending flow S_i at the downstream link end (x_i^L), and for each outgoing link $j \in J_n$, determine the Receiving flow R_j at the upstream link end (x_j^0).
I_n and J_n are the sets of incoming respectively outgoing links into node n.

Step 2: Determine the transition flows $G_{ij}(t)$ from incoming links $i \in I_n$ to outgoing links $j \in J_n$, i.e. determine which parts of the Sending and Receiving flows can actually be sent and received.

Step 3: For the downstream link boundary (x_i^L) of each incoming link $i \in I_n$, and for the upstream link boundary (x_j^0) of each outgoing link $j \in J_n$, update the cumulative vehicle numbers $N(x,t)$:

$$N(x_i^L, t+\Delta t) = N(x_i^L, t) + \sum_{j \in J_n} G_{ij}(t) \qquad \text{for all } i \in I_n$$

$$N(x_j^0, t+\Delta t) = N(x_j^0, t) + \sum_{i \in I_n} G_{ij}(t) \qquad \text{for all } j \in J_n$$

The disaggregation of cumulative vehicle numbers at the downstream boundary of an incoming link is adopted from the upstream link boundary:

$$N^p(x_i^L, t+\Delta t) = N^p(x_i^0, t_{x_i^0}(N(x_i^L, t+\Delta t)) \qquad \text{for all } i \in I_n \text{ , } p \in P$$

The determination of disaggregate cumulative vehicle numbers thus requires FIFO behavior on each network link. Note that this procedure also requires satisfaction of the CFL condition ($N(x_i^L, t+\Delta t)$ should be smaller than $N(x_i^0, t)$).

The disaggregated cumulative vehicle numbers at the upstream link boundary of an outgoing link are given by:

$$N^p(x_j^0, t+\Delta t) = N^p(x_j^0, t) + \sum_{i \in I_n} \delta_{jp} G_{ij}(t) \qquad \text{for all } j \in J_n \text{ , } p \in P$$

where δ_{jp} is an element of link-route incidence matrix δ; $\delta_{jp} = 1$ if route p contains link j and $\delta_{jp} = 0$ otherwise.

FIFO conditions

In LTM, First-In-First-Out (FIFO) behavior is required on links for both a correct determination of link travel times and a correct disaggregation of cumulative vehicle numbers. FIFO at link level means that travelers who enter a link earlier, also leave this link earlier. LTM ensures this link-FIFO condition.
However, since link travel times are not necessarily multiples of the time step, and since the node at the upstream link end might operate at a different time step compared to the node at the downstream link end (see next section), it is possible that travelers who entered a link earlier, leave this link at the same time. Therefore, LTM does not guarantee "strict" link-FIFO behavior.

Furthermore, LTM also ensures node-FIFO and route-FIFO conditions to be satisfied. Route-FIFO behavior is required for a correct determination of route travel times and for a proper functioning of departure time choice models (choosing later departure times to arrive earlier makes no sense in DTA-context).

Event-based solution method

For each node *n*, the different steps of the LTM solution algorithm are completed for each time interval Δt. Alternatively formulated, each node *n* is updated with a time step Δt.
In LTM, each node can be updated with a different time step, depending on the particular needs of each individual node. Therefore, the solution method is referred to as an event-based method.
To satisfy the CFL conditions, the largest possible time step to update a node equals the link travel time of the shortest link that is connected to that node. Nodes that are connected to a short link need to be updated with a small time step. The shortest link is also the weakest link; it determines the required length of the time step Δt.
In time-based methods, all nodes would generally be updated with the same fixed time step Δt. Since computation times are proportional to the length of the time step (cf. Section 4.8), event-based methods obviously outperform time-based methods computational-wise.

4

LINK MODEL

LTM propagates traffic on links as assumed in kinematic wave theory. Traffic is characterized by three macroscopic variables: flow q, density k and average speed v. Only two of these three variables are independent (cf. Section 4.1). A fundamental relationship between two of the remaining independent variables (Section 4.2) reduces this amount to only one independent variable. The evolution over time and space of a traffic state, represented by such a variable (e.g. density k), is described by the conservation law (Section 4.3). The basics of kinematic wave theory are illustrated by means of a simple example in Section 4.4. This section further explains how kinematic wave theory can be used to determine cumulative vehicle numbers. Newell substantially simplifies this procedure (Section 4.5) and LTM uses Newell's efficient method (Section 4.6) to determine Sending and Receiving flows, which is the first step in the LTM solution algorithm. Section 4.7 extends this method to include a piecewise linear fundamental diagram. The chapter ends with a comparison of the computational efficiency and accuracy in LTM and CTM (Section 4.8).

4.1. Macroscopic variables

Only two of the three macroscopic variables that characterize a traffic state are independent.

The flow q is defined as the number of vehicles m, observed by a stationary observer (i.e. at one point in space) during a given time interval Δt, divided by the length of this time interval:

$$q = {}^{m}\!/_{\Delta \mathrm{t}} \tag{4.1}$$

By multiplying the numerator and denominator of this expression by a small differential of space, *dx*, the formula for flow becomes:

$$q = {}^{mdx}\!\big/_{\Delta \mathrm{tdx}} \tag{4.2}$$

The denominator of (4.2) represents a certain time-space region, and the numerator represents the total distance traveled by vehicles inside this time-space region (e.g., in vehicle-kilometers). The generalized definition of flow is formulated as follows:

$$q = \frac{total\ \text{distance}\ in\ time-space\ region\ (vehicle-kilometers)}{Area\ of\ \ time-space\ region\ (kilometer-hours)} \tag{4.3}$$

The density *k* is defined as the number of vehicles *m'*, observed on a road section (at one point in time), divided by the length of the road section, L:

$$k = {}^{m'}\!\big/_{L} \tag{4.4}$$

By multiplying the numerator and denominator of this expression by a small differential of time, *dt*, the formula for density becomes:

$$k = {}^{m'dt}\!\big/_{Ldt} \tag{4.5}$$

The denominator of (4.5) represents a certain time-space region, and the numerator represents the total vehicular time that is spent inside this time-space region (e.g., in vehicle-hours). The generalized definition of density is formulated as follows:

$$k = \frac{total\ time\ in\ time-space\ region\ (vehicle-hours)}{Area\ of\ \ time-space\ region\ (kilometer-hours)} \tag{4.6}$$

The (generalized) average speed *v* in a time-space region is defined as the ratio of the total distance traveled in this time-space region to the total time spent in this time-space region:

$$v = \frac{total\ \text{distance}\ in\ time-space\ region\ (vehicle-kilometers)}{total\ time\ in\ time-space\ region\ (vehicle-hours)} \tag{4.7}$$

which using (4.3) and (4.6) yields:

$$v = {}^{q}\!\big/_{k} \tag{4.8}$$

Average speed, flow and density are associated with each other by Equation (4.8). When two of these macroscopic variables are known, Relationship (4.8) immediately yields the third variable.

The following section establishes a relationship between the two remaining independent variables.

4.2. The fundamental diagram

The kinematic wave theory assumes that there exists a functional relation between traffic flow *q* and density *k*, also known as the fundamental diagram of traffic flow. The fundamental diagram represents all possible stationary traffic states. The diagram reflects a local relationship between macroscopic traffic variables; it should not be perceived as a causal relationship.

The diagram is a property of the road (e.g. number of lanes, slope), the environment (e.g. weather conditions), and the population of travelers (e.g. commuters or Sunday drivers). The assumed existence of this fundamental relationship is plausible since one can reasonably expect drivers to behave the same on average under the same average conditions (Daganzo (1997)). In reality however, the population of travelers is never exactly the same, and traffic conditions are seldom stationary – they change over time and space. Therefore, the diagram is only an approximation of reality.

Based on (limited) empirical data, Greenshields (1934) was the first to propose a relationship between the traffic flow *q* and density *k*.

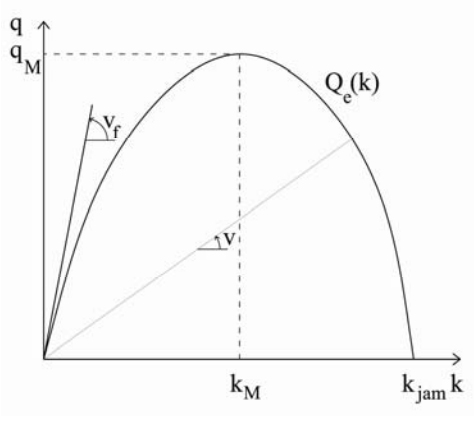


Figure 4.1: Fundamental diagram of Greenshields

The equation corresponding to Figure 4.1 can be written as:

$$Q_e(k) = \frac{v_f}{k_{jam}} k(k_{jam} - k) \tag{4.9}$$

The equilibrium flow $Q_e(k)$ equals *zero* at minimum density ($k = 0$) and at maximum density (k_{jam} = jam density). The maximum flow q_M, also called the capacity, occurs at critical density k_M. As indicated in the figure, the average speed of the vehicles in a traffic state (v) is represented by the slope of the straight line connecting the traffic state with the origin. The rising part of the curve corresponds to the free-flow regime, the descending part to congestion. The maximum possible speed, occurring for k approximating zero, is referred to as free-flow speed v_f.

As in Newell's simplified theory (cf. Section 4.5), LTM uses a triangular shaped fundamental diagram, defined by three parameters: a fixed free-flow speed (v_f), a maximum flow or capacity (q_M) and a jam density (k_{jam}) (see Figure 4.2).

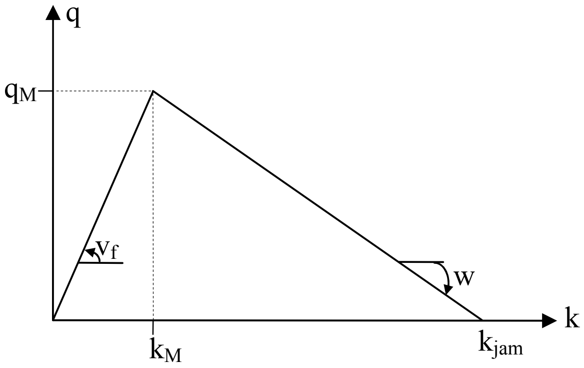


Figure 4.2: Triangular shaped fundamental diagram

Traffic states on the increasing branch of the triangular shaped fundamental diagram ($k < k_M$) hold vehicles traveling with a fixed free-flow speed v_f. Traffic states on the decreasing branch ($k > k_M$) are congested. Vehicles travel with a speed q/k. The maximum flow or capacity (q_M) occurs at critical density k_M, whereas jam density (k_{jam}) corresponds to zero flow (all vehicles stand still).

4.3. The conservation law

The conservation law describes the evolution of a traffic state over time and space. In the most general way, the conservation law is expressed in terms of the cumulative function $N(x,t)$ (Daganzo (1997)). This cumulative function $N(x,t)$ represents the number of the last vehicle to pass location x before time t.

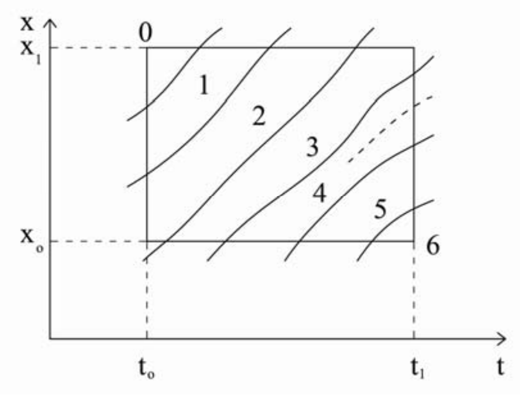


Figure 4.3: Vehicle trajectories and values of the cumulative vehicle function *N(x,t)*

Figure 4.3 depicts a space-time interval $(\Delta x,\Delta t)$ with some vehicle trajectories representing the position of the vehicles as a function of time. The values of the cumulative vehicle function are indicated on the figure as well. These values remain constant between the different vehicle trajectories. The vehicles have been numbered in increasing order in the direction of increasing time.
Note that the difference $N(x_0,t) - N(x_1,t)$ represents the number of vehicles on section $[x_0, x_1]$ at time t, whereas the difference $N(x,t_1) - N(x,t_0)$ represents the number of vehicles observed during time interval $[t_0, t_1]$ at location x.

The function $N(x,t)$ will in general be discontinuous, but one can also define a smooth approximation to make this function differentiable.
The partial derivatives of $N(x,t)$ can be interpreted as the instantaneous flow and (the negative of) local density (prevailing at point (x,t)):

$$q(x,t) = \frac{\partial N(x,t)}{\partial t} \tag{4.10}$$

$$k(x,t) = \frac{-\partial N(x,t)}{\partial x} \tag{4.11}$$

The negative sign in (4.11) arises because N decreases in the direction of increasing x (see Figure 4.3).

Assume that a vehicle enters the road between locations x_0 and x_1 so that a trajectory was created somewhere inside the space-time interval (dotted line). If one would count the cumulative vehicle number at the bottom right corner of the space-time interval starting from the top left corner, then he would count 6 if he counted along the path $(x_1,t_0) \rightarrow (x_1,t_1) \rightarrow (x_0,t_1)$, but he would count 7 if he counted along the path $(x_1,t_0) \rightarrow (x_0,t_0) \rightarrow (x_0,t_1)$.
It would be impossible to assign a vehicle number to the bottom right corner of the space-time interval that would be correct for both paths.
This means that our ability to define a function $N(x,t)$ for all t, x in a region of interest

implies no entering (or exiting) traffic. The mere existence of a function *N(x,t)* is a conservation condition that ensures vehicles are not created nor lost along the road. This is the most general way in which the conservation law can be stated.

Provided that the function *N(x,t)* exists in a certain region *(Δx,Δt)*, as well as its first and second derivatives, the identity

$$\frac{\partial^2 N(x,t)}{\partial x \partial t} = \frac{\partial^2 N(x,t)}{\partial t \partial x} \tag{4.12}$$

together with Equations (4.10) and (4.11) becomes:

$$\frac{\partial k(x,t)}{\partial t} + \frac{\partial q(x,t)}{\partial x} = 0 \tag{4.13}$$

which is the well-known form of the conservation law.

The conservation law can also be reformulated in terms of line integrals of the gradient of *N(x,t)*. Green's theorem (1828) states that a line integral of the projection of the gradient along any curve joining two points is independent of the curve, and that the result equals the difference in *N(x,t)* at the extremes of the curve.

Green's theorem yields:

$$N(x_2,t_2) - N(x_1,t_1) = \int_C dN(x,t) = \int_C \frac{\partial N(x,t)}{\partial t} dt + \frac{\partial N(x,t)}{\partial x} dx = \int_C q dt - k dx \tag{4.14}$$

for all *C*, where *C* is an arbitrary curve from (x_1,t_1) to (x_2,t_2) in a region *(Δx,Δt)* in which vehicles are conserved.
This form of the conservation equation is useful because it applies even if *q* and *k* are discontinuous. It will intensively be used in the remainder of this chapter.

4.4. Kinematic wave theory

The kinematic wave theory is in this section explained by means of a simple example. For a complete overview of the theory, we refer to Daganzo (1997).
The theory of kinematic waves, as described originally by Lighthill, Whitham and Richards (LWR), arises from the assumption that stationary fundamental diagrams also apply when traffic is not stationary:

$$q(x,t) = Q_e(k(x,t)) \tag{4.15}$$

Equation (4.15) holds, independently of the flows, densities and speeds prevailing upstream and downstream of x, and also independently of these conditions at prior times. While this may be inaccurate for a detailed description, it seems like a reasonable assumption for long roads in large scale networks (Daganzo (1997)).

The evolution of traffic conditions on a link is described by a combination of the conservation law and the fundamental diagram.
Replacing $q(x,t)$ by $Q_e(k(x,t))$ in conservation Equation (4.13) yields the LWR form of the conservation condition (containing only one independent variable, k):

$$\frac{\partial k(x,t)}{\partial t} + \frac{dQ_e(k(x,t))}{dk}\frac{\partial k(x,t)}{\partial x} = 0 \tag{4.16}$$

This partial differential equation can be solved given initial and boundary conditions. The solution is a function F that takes the form:

$$k(x,t) = F\left(x - \frac{dQ_e(k(x,t))}{dk} t\right) \tag{4.17}$$

Equation (4.17) expresses that the density $k(x,t)$ is constant on a straight line with slope $\frac{dQ_e(k(x,t))}{dk}$ in the t-x plane.
The solutions of the conservation Equation (4.16) are straight lines in the t-x plane, issued from the boundary, on which k – and correspondingly also q and u – is constant and equal to the value at the boundary. These lines are called characteristics or waves. Characteristics can be constructed given the initial and boundary conditions.

Example 4.1

In Figure 4.4a, we assume the following initial boundary condition:
$k(x_0, t_0) = k_0$

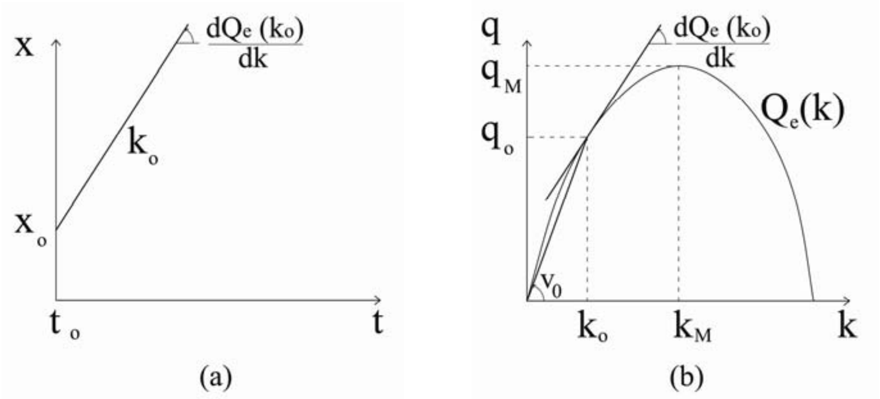


Figure 4.4: *t-x* diagram (a) and *k-q* fundamental diagram (b)

Slope $\frac{dQ_e(k_0)}{dk}$ is determined as the slope of the tangent line of the fundamental diagram in k_0 (Figure 4.4b). In the t-x plane (Figure 4.4a), one can then construct a characteristic with slope $\frac{dQ_e(k_0)}{dk}$ passing through boundary point (t_0,x_0). Traffic conditions (k_0,q_0,u_0) are constant on this straight line.

Characteristics (i.e. solutions of the conservation Equation (4.16) describe the evolution of a traffic state. In the *t-x* plane, traffic states k,q,u travel at wave speed $\frac{dQ_e(k(x,t))}{dk}$. In Figure 4.4a, for example, traffic state k_0 starts at point (x_0,t_0) and travels in the direction of increasing x and t, at (positive) wave speed $\frac{dQ_e(k_0)}{dk}$.
Generally, free flow traffic states travel with a positive wave speed, i.e., they travel in the direction of the traffic, since the slope of the tangent line of the fundamental diagram is positive for free-flow traffic states. Congested traffic conditions travel with negative speed, against the direction of the traffic.

The kinematic wave theory further describes shockwaves. Different boundary conditions (on different locations on the boundary) give rise to different traffic states travelling at different speeds. These different traffic states may intersect one another. However, on one specific point in the *t-x* plane, there can only be one traffic state. Therefore, two traffic states intersect each other in a shock, separating the two different states. Shocks travel in the *t-x* plane with speed

$$v_{shock} = \frac{q_1 - q_2}{k_1 - k_2} \tag{4.18}$$

where k_1 (q_1) and k_2 (q_2) represent the densities (flows) of the different traffic states.

Example 4.2

Assume a link $[x_a^0, x_a^L]$ with length L_a, representing a homogeneous road section (see Figure 4.5). Possible traffic states are in this example given by a triangular shaped fundamental diagram. Vehicles drive from x_a^0 towards x_a^L. Traffic state k_1,q_1,v_1 occurs on boundary x_a^0. At time t_1, a capacity restriction q_2 occurs at the downstream link end (x_a^L). Initial and boundary conditions are as follows:

$$k(x,t_0) = 0 \tag{4.19}$$

$$k(x_a^0,[t_0,t_n]) = k_1 \tag{4.20}$$

$$k(x_a^L,[t_1,t_n]) = k_2 \tag{4.21}$$

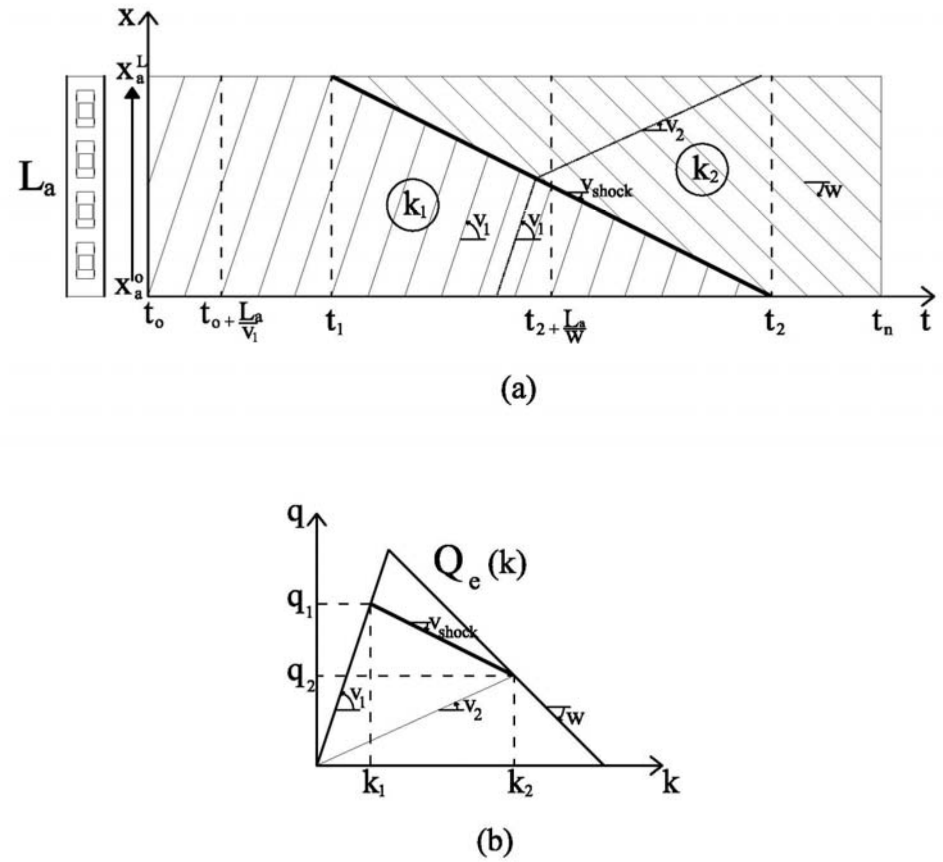


Figure 4.5: shockwave in the *t-x* diagram (a)
and in the *k-q* fundamental diagram (b) in Example 4.2

The dotted line in Figure 4.5 represents a vehicle trajectory, i.e. a vehicle path. Vehicles in free-flow traffic state k_1,q_1,v_1 travel at the same speed v_1 as the traffic state itself. At (x_a^L,t_1), traffic demand (q_1) exceeds capacity (q_2) which leads to congested traffic state k_2,q_2,v_2 (see Figure 4.5b). Vehicles in the congested traffic state travel at speed v_2, while the traffic state itself travels with negative speed w, against the direction of traffic.

The two traffic states k_1,q_1,v_1 and k_2,q_2,v_2 intersect each other in a shock. This shock between the two states travels with speed v_{shock} (cf. Equation (4.18)), which in this case is negative. The shock thus travels against the direction of traffic as well. At t_2, the congested state reaches link boundary x_a^0 and from that time, the whole road section is congested.

The theory of kinematic waves can be used to determine cumulative vehicle numbers on the link boundaries x_a^0 and x_a^L. By following the path of the shock, one can straightforwardly determine cumulative vehicle numbers, as illustrated in the example below.

Example 4.3

Consider Example 4.2. By following the path of the shock, it appears that:

$$q(x_a^0,[t_0,t_2]) = q_1 \tag{4.22}$$

$$q(x_a^0,[t_2,t_n]) = q_2 \tag{4.23}$$

$$q(x_a^L,[t_0,t_1]) = q_1 \tag{4.24}$$

$$q(x_a^L,[t_1,t_n]) = q_2 \tag{4.25}$$

Cumulative vehicle curves are constructed using Formula (4.10):

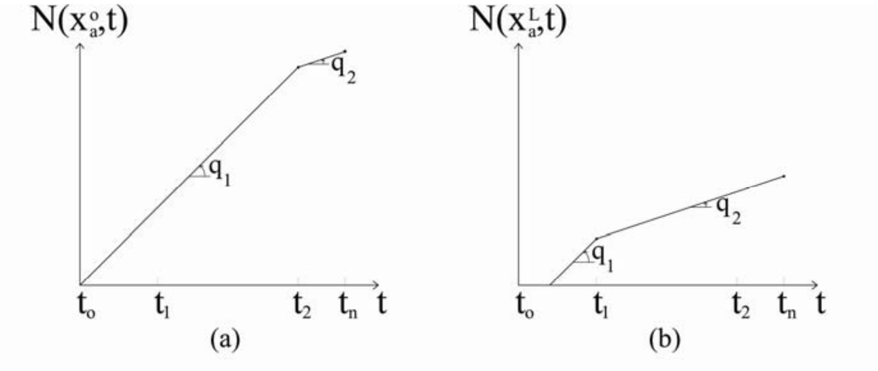


Figure 4.6: Cumulative vehicle numbers at the upstream boundary x_a^0 (a) and at the downstream boundary x_a^L (b) in Example 4.3

4.5. Newell's simplified theory of kinematic waves

Newell (1993) substantially simplifies the procedure to determine cumulative vehicle numbers. Instead of using kinematic wave theory to evaluate flows or densities, Newell uses it to directly evaluate the cumulative flow $N(x,t)$ past any point x by time t.

From conservation Equation (4.14), if one knows $N(x_0,t_0)$ and $k(x_0,t_0)$ at some boundary point (x_0,t_0), one can easily determine $N(x,t)$ at all points along the characteristic curve through (x_0,t_0). If characteristic curves intersect, the cumulative curve $N(x,t)$ is still easily constructed but it gives a multiple-valued function of x and t. Each of the values of $N(x,t)$ at the same point (x,t) is derived from different initial or boundary conditions. However, on one specific point in the t-x plane, there can only be one value of $N(x,t)$. For concave fundamental diagrams, Newell (1993) claims that the unique solution one is looking for (i.e. the unique solution that is physically meaningful) is the lower envelope of the multiple-valued solution derived from properly set initial and boundary conditions. One can think of this as the most restrictive constraint being determinant.

Furthermore, Newell uses a triangular shaped fundamental diagram (cf. Figure 4.2) in his simplified theory of kinematic waves. There are only two possible values of the wave speed, one positive (v_f) for traffic states that hold free-flow traffic and one negative (w) for congested traffic states. This avoids some unpleasant mathematical complications, as explained below.

Example 4.4

Consider Example 4.2 (corresponding to Figure 4.5). The boundary conditions (4.22) to (4.25) can be presented in the form of cumulative vehicle numbers:

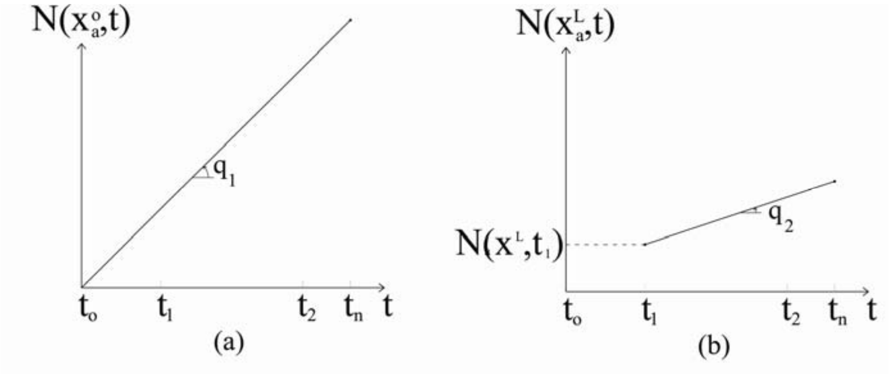


Figure 4.7: upstream (x_a^0) (a) and downstream (x_a^L) (b) boundary conditions on link a in Example 4.4

Free-flow traffic state q_1 travels from the upstream link boundary (x_a^0) in the direction of the downstream link boundary (x_a^L) (see Figure 4.5a). If this state does not intersect with another state, then it will reach the downstream link boundary L_a/v_1 time units later, since this state travels with speed v_1. State q_1 travels for example from (x_a^0,t_0) to (x_a^L, t_0+L_a/v_1). Traffic conditions in (x_a^0,t_0) and ($x_a^L,t_0+L_a/v_1$) are thus identical. Equation (4.14) yields the change in cumulative vehicle number between these points of interest:

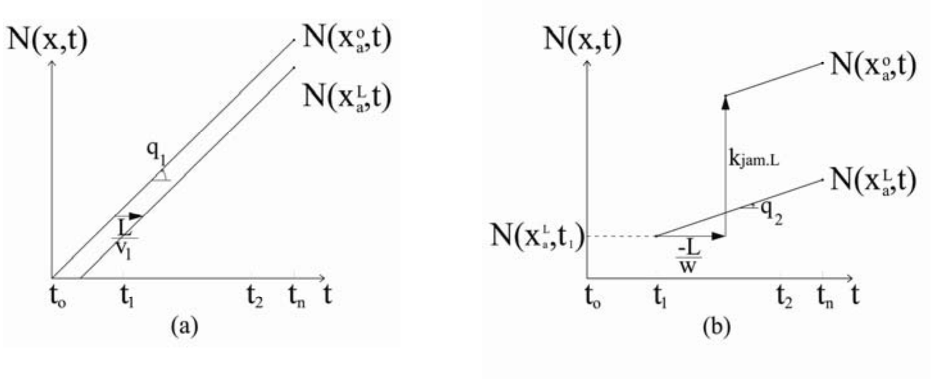
$$N(x_a^0,t_0)-N(x_a^L,t_0+\frac{L_a}{v_1})=q_1(-\frac{L_a}{v_1})-k_1(-L_a)=(L_a)(-\frac{q_1}{v_1}+k_1)=0 \qquad (4.26)$$

Note that curve C in Equation (4.14) is in this case a curve between these two points along which k and q are constant (e.g. a straight line in the t-x plane between these two points). The difference in cumulative vehicle number between these points is *zero*, which is logical since a straight line between (x_a^0, t_0) and $(x_a^L, t_0+L_a/v_1)$ corresponds to a vehicle trajectory. This means that, based solely on the upstream boundary conditions, cumulative vehicle numbers on the downstream boundary are translations of the cumulative vehicle numbers at the upstream boundary over L/v_1 time units (see Figure 4.8a). This is how cumulative vehicle numbers at the downstream boundary are constrained by boundary conditions at the upstream link end.

Congested traffic state q_2 travels from the downstream link boundary (x_a^L) in the direction of the upstream link boundary (x_a^0) (see Figure 4.5a). If this state does not intersect with another state, then it will reach the upstream link boundary -L_a/w time units later, since this state travels with negative speed w. State q_2 travels for example from $(x_a^L,\ t_2+L_a/w)$ to $(x_a^0,\ t_2)$. Traffic conditions in $(x_a^L,\ t_2+L_a/w)$ and $(x_a^0,\ t_2)$ are thus identical. Equation (4.14) yields the change in cumulative vehicle number between these points of interest:

$$N(x_a^0, t_2) - N(x_a^L, t_2 + \frac{L_a}{w}) = q_2(-\frac{L_a}{w}) - k_2(-L_a) = (L_a)(-\frac{q_2}{w} + k_2) = k_{jam}(L_a) \quad (4.27)$$

Note that curve C in Equation (4.14) is in this case a curve between these two points along which k and q are constant (e.g. a straight line in the t-x plane between these two points). The difference in cumulative vehicle number between these points is $k_{jam}L_a$. Based solely on downstream boundary conditions, cumulative vehicle numbers on the upstream boundary are translations of the cumulative vehicle numbers at the downstream boundary over -L/w time units and $k_{jam}L$ vehicle units (see Figure 4.8b). This is how cumulative vehicle numbers at the upstream boundary are constrained by boundary conditions at the downstream link end.



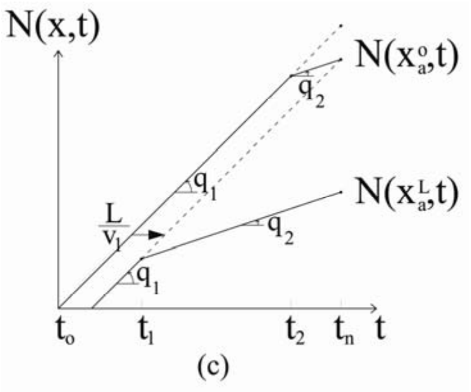


Figure 4.8: Translations of upstream (a) and downstream (b) boundary conditions and the lower envelope of the double-valued solution (c)

> Now we have double-valued solutions for $N(x_a^0,t)$ and $N(x_a^L,t)$, represented in Figure 4.8a and 4.8b. The unique solutions are the lower envelopes of the double-valued solutions: see Figure 4.8c. These curves indeed correspond to the curves determined earlier in Figure 4.6. Discontinuities in slope describe the passage of a shock.

The main advantage of Newell's method is that the cumulative vehicle functions on the link boundaries can be determined without following the actual path of the shock. The solution for $N(x,t)$ can be evaluated directly from initial and boundary conditions, without evaluation at intermediate times and positions.

Furthermore, when using a triangular shaped fundamental diagram, with only two possible wave speeds, free-flow (congested) boundary conditions are translated over a fixed number of L/v_f ($-L/w$) time units and a fixed number of *zero* ($k_{jam}L$) vehicle units, independent on the values of the boundary conditions.
This allows for an efficient procedure, especially when compared to using a concave fundamental diagram like in Figure 4.1, where boundary conditions are translated over a number of time and vehicle units, that has to be re-determined for every new value of the boundary condition.

Based on Newell's simplified theory of kinematic waves, a very efficient yet accurate procedure to determine Sending and Receiving flows can and will be elaborated (next section). Determining Sending and Receiving flows corresponds to the first step of the LTM solution procedure (cf. Section 3.2).

4.6. Sending and Receiving flows in LTM

In this section, Newell's simplified method is used to determine Sending and Receiving flows, which is the first step in the LTM solution algorithm. By determining Sending and Receiving flows that are consistent with Newell's simplified theory, LTM ensures kinematic wave behavior on its links.

Constraints on Sending flows

The Sending flow $S_i(t)$ of link *i* at time *t* is defined as the maximum amount of vehicles that could leave the downstream end of this link during $[t, t+\Delta t]$, if this link end were connected to a traffic reservoir with an infinite capacity.

The Sending flow is constrained by the boundary conditions at the upstream end of the link (x_i^0). If a free-flow traffic state occurs at the downstream link boundary (x_i^L) at time $t+\Delta t$, then Newell's simplified theory of kinematic waves states that this state must have been emitted from the upstream boundary $L_i/v_{f,i}$ time units earlier, since a free-flow traffic state travels with speed $v_{f,i}$ (see Figure 4.9a).

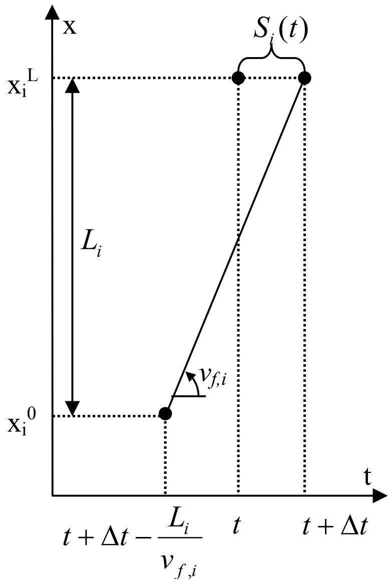


Figure 4.9a: Propagation of a free-flow traffic state in the *t-x* plane

This means that traffic conditions at (x_i^L, $t+\Delta t$) and at (x_i^0, $t+\Delta t-L_i/v_{f,i}$) are identical. Conservation Equation (4.14) yields the change in cumulative vehicle number between these points of interest:

$$N(x_i^0, t+\Delta t-\frac{L_i}{v_{f,i}})-N(x_i^L, t+\Delta t)=q(-\frac{L_i}{v_{f,i}})-k(-L_i)=(L_i)(-\frac{q}{v_{f,i}}+k)=0 \qquad (4.28)$$

where C is a curve between these two points along which k and q are constant (e.g. a straight line in the t-x plane between these two points).

The constraints on the Sending flow can be formulated as:

$$S_i(t)\leq N(x_i^L, t+\Delta t)-N(x_i^L, t)=N(x_i^0, t+\Delta t-\frac{L_i}{v_{f,i}})-N(x_i^L, t) \qquad (4.29)$$

One can recognize the above-mentioned translation (Figure 4.8a) in this Equation (4.29), which expresses that no vehicle can leave the link sooner than $L_i/v_{f,i}$ time units after having entered the link, i.e. the minimal link travel time needs to be respected.

The Sending flow is also constrained by the link's capacity:

$$S_i(t)\leq q_{M,i}\Delta t \qquad (4.30)$$

The Sending flow is the maximum flow taking into account these constraints. This is formulated as:

$$S_i(t)=\min((N(x_i^0, t+\Delta t-\frac{L_i}{v_{f,i}})-N(x_i^L, t)), q_{M,i}\Delta t) \qquad (4.31)$$

Constraints on Receiving flows

The Receiving flow $R_j(t)$ of link j at time t is defined as the maximum amount of vehicles that could enter the upstream end of this link during $[t, t+\Delta t]$, if a traffic reservoir with an infinite traffic demand were connected to this link end. The Receiving flow is constrained by the boundary conditions at the downstream end of the link (x_j^L). If a congested traffic state occurs at the upstream boundary (x_j^0) at time $t+\Delta t$, then the simplified kinematic wave theory states that this state must have been emitted from the downstream boundary ($-L_j/w_j$) time units earlier, since a congested traffic state travels with a (negative) speed w_j (see Figure 4.9b).

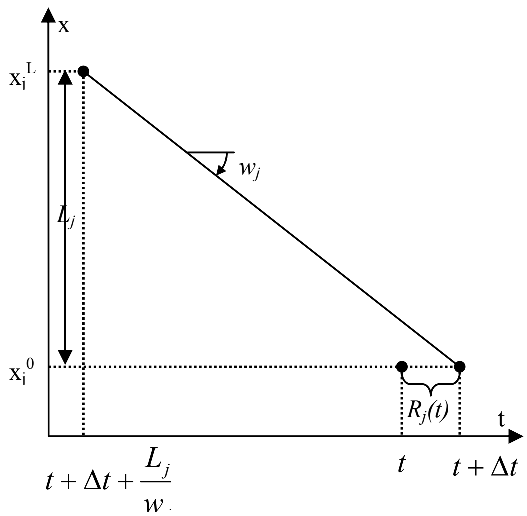


Figure 4.9b: Propagation of a congested traffic state in the *t-x* plane

This means that traffic conditions at (x_j^0, $t+\Delta t$) and at (x_j^L, $t+\Delta t+L_j/w_j$) are identical. Conservation Equation (4.14) yields the change in cumulative vehicle number between these points of interest:

$$N(x_j^0,t+\Delta t)-N(x_j^L,t+\Delta t+\frac{L_j}{w_j})=q(-\frac{L_j}{w_j})-k(-L_j)=(L_j)(-\frac{q}{w_j}+k)=k_{jam}(L_j) \quad (4.32)$$

where C is a curve between these two points along which k and q are constant (e.g. a straight line in the *t-x* plane between these two points).

The constraints on the Receiving flow can be formulated as:

$$R_j(t)\le N(x_j^0,t+\Delta t)-N(x_j^0,t)=N(x_j^L,t+\Delta t+\frac{L_j}{w_j})+k_{jam}L_j-N(x_j^0,t) \quad (4.33)$$

One can recognize the above-mentioned translation (Figure 4.8b) in this Equation (4.33), which expresses that no vehicle can enter the link sooner than ($-L_j/w_j$) time units after the ($k_{jam}L$)th vehicle before him has left the link. An inflow limitation is imposed due to the spilling back of queues.

The Receiving flow is also constrained by the link's capacity:

$$R_j(t)\le q_{M,j}\Delta t \quad (4.34)$$

The Receiving flow is the maximum flow taking into account these constraints. This is formulated as:

$$R_j(t)=\min((N(x_j^L,t+\Delta t+\frac{L_j}{w_j})+k_{jam}L_j-N(x_j^0,t)),q_{M,j}\Delta t) \quad (4.35)$$

4.7. Extension to a piecewise linear fundamental diagram

In the previous section, Sending and Receiving flows are determined based on Newell's simplified theory of kinematic waves, which uses a triangular shaped fundamental diagram. However, road properties, driving behavior and environmental conditions may in some cases be more precisely characterized by a piecewise linear fundamental diagram (Henn (2003)). This section explains how Sending and Receiving flows can be determined when using a concave piecewise linear fundamental diagram.
The procedure is explained by means of an example with a fundamental diagram with four linear pieces, and is later on extended for fundamental diagrams with any number of linear pieces. Figure 4.10 depicts the considered piecewise linear fundamental diagram.

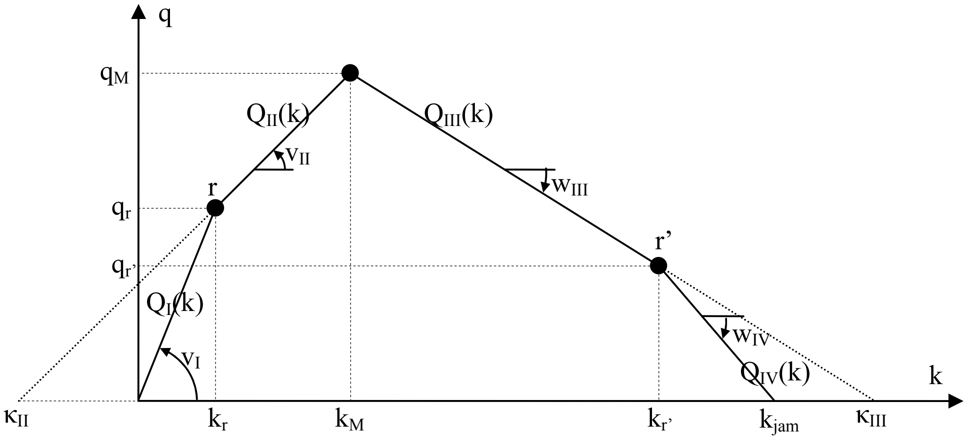


Figure 4.10: Piecewise linear fundamental diagram

The diagram includes two free-flow (*I* and *II*) and two congested (*III* and *IV*) branches, respectively separated by rotation points *r* and *r'*. All four branches have different wave speeds: v_I, v_{II}, w_{III} and w_{IV}. They are given by the following equations:

- Branch *I*: $q = Q_{e,I}(k) = kv_I$ *for all* $k \in [0, k_r]$ (4.36)
- Branch *II*: $q = Q_{e,II}(k) = (k - \kappa_{II})v_{II}$ *for all* $k \in [k_r, k_M]$ (4.37)
- Branch *III*: $q = Q_{e,III}(k) = (k - \kappa_{III})w_{III}$ *for all* $k \in [k_M, k_{r'}]$ (4.38)
- Branch *IV*: $q = Q_{e,IV}(k) = (k - \kappa_{IV})w_{IV}$ *for all* $k \in [k_{r'}, k_{jam}]$ (4.39)

Note that constants κ_{II} and κ_{III} do not represent physically feasible densities.
Determination of Sending and Receiving flows is somewhat more complicated when using a piecewise linear fundamental diagram.

Constraints on Sending flows

The Sending flow $S_i(t)$ of link i at time t is constrained by the boundary conditions at the upstream end of the link (x_i^0).

If a free-flow traffic state I occurs at the downstream link boundary (x_i^L) at time $t+\Delta t$, then the kinematic wave theory states that this state must have been emitted from the upstream boundary $L_i/v_{I,i}$ time units earlier, since a free-flow traffic state I travels with speed $v_{I,i}$.

This means that traffic conditions at (x_i^L, $t+\Delta t$) and at (x_i^0, $t+\Delta t-L_i/v_{I,i}$) are identical. Conservation Equation (4.14) yields the change in cumulative vehicle number between these points of interest:

$$N(x_i^0, t+\Delta t-\frac{L_i}{v_{I,i}})-N(x_i^L, t+\Delta t)=q(-\frac{L_i}{v_{I,i}})-k(-L_i)=(L_i)(-\frac{q}{v_{I,i}}+k)=0 \qquad (4.40)$$

where C is a curve in $(\Delta x, \Delta t)$ along which $k\in[0, k_r]$ and $q\in[0, q_r]$ are constant.

The constraints on the Sending flow can in this case be formulated as:

$$S_i(t)\le N(x_i^L, t+\Delta t)-N(x_i^L, t)=N(x_i^0, t+\Delta t-\frac{L_i}{v_{I,i}})-N(x_i^L, t) \qquad (4.41)$$

If a free-flow traffic state II occurs at the downstream link boundary (x_i^L) at time $t+\Delta t$, then the kinematic wave theory states that this state must have been emitted from the upstream boundary $L_i/v_{II,i}$ time units earlier, since a free-flow traffic state II travels with speed $v_{II,i}$.

This means that traffic conditions at (x_i^L, $t+\Delta t$) and at (x_i^0, $t+\Delta t-L_i/v_{II,i}$) are identical. Conservation Equation (4.14) yields the change in cumulative vehicle number between these points of interest:

$$N(x_i^0, t+\Delta t-\frac{L_i}{v_{II,i}})-N(x_i^L, t+\Delta t)=q(-\frac{L_i}{v_{II,i}})-k(-L_i)=(L_i)(-\frac{q}{v_{II,i}}+k)=\kappa_{II}L_i \qquad (4.42)$$

where C is a curve in $(\Delta x, \Delta t)$ along which $k\in[k_r, k_M]$ and $q\in[q_r, q_M]$ are constant.

The constraints on the Sending flow can in this case be formulated as:

$$S_i(t)\le N(x_i^L, t+\Delta t)-N(x_i^L, t)=N(x_i^0, t+\Delta t-\frac{L_i}{v_{II,i}})-\kappa_{II}L_i-N(x_i^L, t) \qquad (4.43)$$

Sending flow $S_i(t)$ is restricted either by condition (4.41), if a free-flow traffic state I occurred on the upstream link boundary at time $t+\Delta t-L_i/v_{I,i}$ *(i)*, or by Condition (4.43), if a free-flow traffic state II occurred on the upstream link boundary at time $t+\Delta t-L_i/v_{II,i}$ *(ii)*.

Note that if *(i)* and *(ii)* occur at the same time, these two traffic states coincide at the downstream link boundary and Conditions (4.41) and (4.43) equal each other.

For concave fundamental diagrams, Newell (1993) shows that (4.41) will be the more restrictive condition if a free-flow traffic state *I* occurred on the upstream link boundary at time $t+\Delta t-L_i/v_{I,i}$, and that Condition (4.43) will be the more restrictive one if a free-flow traffic state *II* occurred on the upstream link boundary at time $t+\Delta t-L_i/v_{II,i}$. Therefore, constraints on Sending flow $S_i(t)$ from upstream boundary conditions can be formulated by the following expression, which is valid for all possible upstream boundary conditions, no matter to which free-flow branch they precisely belong to:

$$S_i(t)=\min(N(x_i^0,t+\Delta t-\frac{L_i}{v_{I,i}})-N(x_i^L,t)\ ,\ N(x_i^0,t+\Delta t-\frac{L_i}{v_{II,i}})-\kappa_{II}.L_i-N(x_i^L,t))\tag{4.44}$$

Furthermore, the Sending flow is also constrained by the link's capacity:

$$\begin{aligned}S_i(t)=\min(&N(x_i^0,t+\Delta t-\frac{L_i}{v_{I,i}})-N(x_i^L,t)\ ,\\ &N(x_i^0,t+\Delta t-\frac{L_i}{v_{II,i}})-\kappa_{II}.L_i-N(x_i^L,t)\ ,\ q_{M,i}\Delta t)\end{aligned}\tag{4.45}$$

The different linear free-flow branches *z* in the free-flow part of the fundamental diagram are given by the following equation:

$$q=Q_{e,z}(k)=(k-\kappa_z)v_z \qquad \text{for all } z\in Z\tag{4.46}$$

where *Z* is the set of piecewise linear fundamental diagram branches in the free-flow part of the fundamental diagram and where $\kappa_I = 0$ and $v_z < v_{z-1}$ and $\kappa_z < \kappa_{z-1}$ for all $z\in Z$, since only concave fundamental diagrams are considered.

The Sending flow can be formulated as follows:

$$S_i(t)=\min(q_{M,i}\Delta t\ ,\ \min_{z\in Z}(N(x_i^0,t+\Delta t-\frac{L_i}{v_{z,i}})-\kappa_z L_i-N(x_i^L,t)))\tag{4.47}$$

Constraints on Receiving flows

The Receiving flow $R_j(t)$ of link *j* at time *t* is constrained by the boundary conditions at the downstream end of the link (x_j^L).

The procedure is completely analogous to the above described procedure to determine Sending flows. The Receiving flow can be formulated as follows:

$$R_j(t) = \min(q_{M,j}\Delta t \ , \ \min_{z'\in Z'}(N(x_j^L, t + \Delta t + \frac{L_j}{w_{z',j}}) + \kappa_{z'} L_j - N(x_j^0, t))) \qquad (4.48)$$

where Z' is the set of piecewise linear fundamental diagram branches in the congested part of the fundamental diagram. The different linear branches $z' \in Z'$ are given by the following equation:

$$q = Q_{e,z'}(k) = (k - \kappa_{z'})w_{z'} \qquad \textit{for all } z' \in Z' \qquad (4.49)$$

where $\kappa_{final} = k_{jam}$ and where $w_{z'} < w_{z'-1}$ and $\kappa_{z'} < \kappa_{z'-1}$ for all $z' \in Z'$, since only concave fundamental diagrams are considered.

4.8. Computational efficiency and accuracy in comparison with CTM

In this section, the computational complexity and the accuracy of LTM is compared to that of the Cell Transmission Model (CTM), which to the present day is a frequently used model to propagate traffic in DTA models (cf. Section 2.1). LTM and CTM are different numerical solution methods for the same kinematic wave traffic flow model.

In his CTM, Daganzo (1994) and (1995a) proposes a system of finite difference equations to approximate the partial differential equations of the kinematic wave model. This method assumes that the road has been divided into homogeneous sections (cells), whose lengths equal the distance traveled by free-flowing traffic in one time interval. For each time step, CTM calculates Sending and Receiving flows for each cell and it determines the flows through the cell boundaries by a minimum rule. Afterwards, the number of vehicles contained in each cell is recalculated, allowing for an update of the cumulative vehicle numbers at the cell or link boundaries.

In LTM, only one Sending flow and one Receiving flow is determined for a whole link. Flows and cumulative vehicle numbers are determined solely at the link boundaries.

To determine these cumulative vehicle numbers in a homogeneous link with *m* cells, CTM requires about *m* times more computational effort compared to LTM for a given time step Δt. In other words, the computational complexity is about *m* times higher in CTM compared to LTM.

On the other hand, LTM requires more memory storage. Where in CTM, cumulative vehicle numbers only need to be stored for one time step, LTM requires them to be stored for $L/-w$ time units (cf. Equation (4.35)). Note that using a piecewise linear fundamental diagram further increases the required memory storage, as cumulative vehicle numbers need to be stored for $max(L/v_z \ , \ L/-w_z)$ time units (cf. Equations (4.47) and (4.48)). Memory storage can be limited by choosing wave speed values that are not too small.

To reduce computational efforts in CTM, one could enlarge the length of the time step Δt. Such operation would reduce the number of cells in a link (*m*), since the length of a

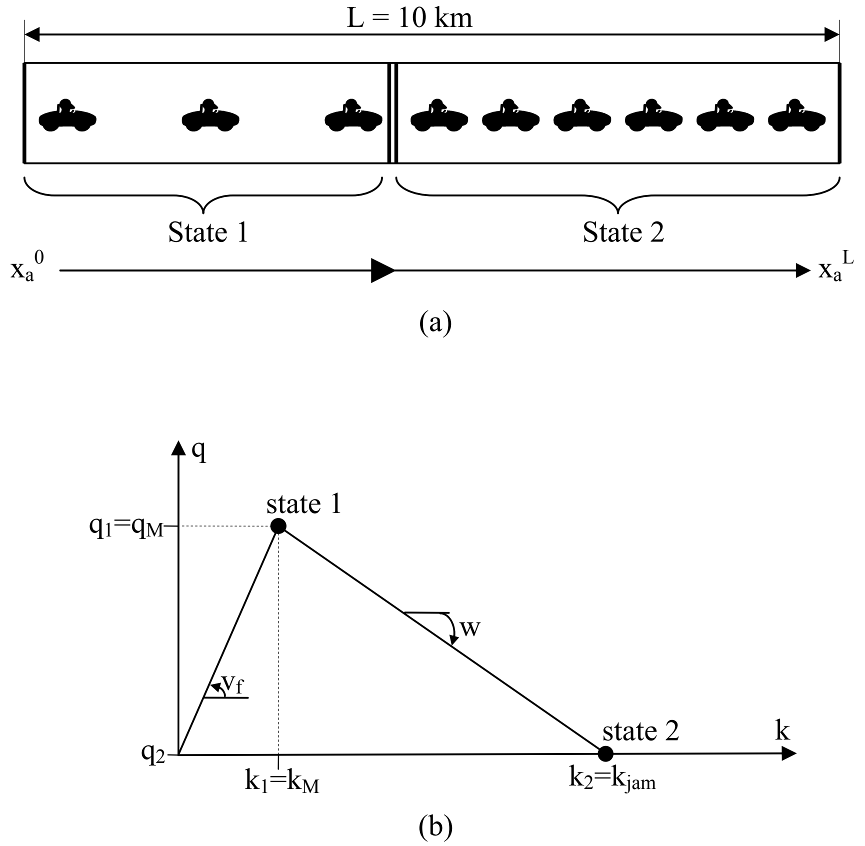
cell equals the distance traveled by free-flowing traffic in one time step. However, such operation would go at the cost of some accuracy as well.
CTM introduces some error into the calculation by means of numerical diffusion. Vehicles are assumed to be spread homogeneously over the length of a cell. This assumption violates kinematic wave behavior if shockwaves do not coincide with cell boundaries, which generally is the case. In kinematic wave theory, shockwaves that are somewhere in the middle of a cell divide this cell in two parts that are characterized by different traffic states. CTM however spreads this cell's contents homogeneously over the length of the cell, thereby replacing the original shockwave by two artificial shockwaves that coincide with the cell's boundaries. In a next time step, these two artificial shockwaves will have traveled a bit further and CTM will again spread the contents of the cells in which these shockwaves are located etc… This process of numerical diffusion introduces errors into the calculation. Shockwaves do not travel with their intended wave speed and cumulative vehicle numbers deviate from the analytical solution. As shown in Example 4.5, errors are higher for a smaller amount of cells in a link, i.e., for a higher time step. Daganzo (1995b) shows that errors introduced by CTM are approximately proportional to the length of the time step Δt.

The accuracy of CTM can be enhanced if the downstream density that is used to calculate the Receiving flow is read with a time lag l. The so-called Lagged CTM (Daganzo (1999)) reduces numerical diffusion, but it requires more memory storage since cell densities must be stored for $l + 1$ time slices.

In LTM, shockwaves do travel through links with their intended wave speed. During their journey through a link, numerical diffusion around a shockwave only occurs in the one time step where this shockwave reaches the link boundary. Vehicles are spread homogeneously over this time step. Introduced errors are relatively small, they are proportional to the length of the time step Δt, and they can be referred to as interpolation errors. Example 4.5 compares the errors that are introduced by the different numerical solution methods CTM, LCTM and LTM.

Example 4.5

In a simple example, we compare the accuracy of CTM, Lagged CTM (LCTM) and LTM. The example considers a single homogeneous link with length $L = 10\ km$, capacity $q_M = 3600\ veh/h$, free flow-speed $v_f = 120\ km/h$, and jam density $k_{jam} = 225\ veh/km$. Traffic is simulated during two hours. At time $t = 0h$, a free-flow traffic regime ($q_1 = q_M = 3600\ veh/h;\ k_1 = q_1/v_f = 30\ veh/km$) starts flowing from the upstream link boundary on an empty road. During the whole simulation period, traffic is prevented from passing the downstream link boundary. A congested regime ($q_2 = 0\ veh/h;\ k_2 = k_{jam} = 225\ veh/km$) will occur at this downstream link end. Figure 4.11 pictures the considered link with its different traffic states:



(a)

(b)

Figure 4.11: Homogeneous link (a) with its different traffic states (b) in Example 4.5

Traffic demand is constant such that $\Delta q = q_1 - q_2 = 3600$ *veh/h* and $\Delta k = k_1 - k_2 = -195$ *veh/km*. At time $t = 5$ *min*, a shock separating the free-flow and the congestion regime travels from the downstream link boundary to the upstream link boundary with a (negative) wave speed w. It reaches the upstream link boundary $L/-w = 32.5$ *minutes* later, i.e. at $t = 37{,}5$ *min*).

For all three models, the cumulative vehicle numbers at the upstream link boundary deviate from the analytical solution due to errors introduced by the numerical scheme.

Parameters in this example are chosen to end up with worst case scenarios, i.e. to end up with the highest possible errors, both for CTM and LTM. In the analytical solution, the shockwave reaches the upstream link boundary precisely halfway in an LTM time step ($\Delta t = [35$ *min* , 40 *min*$]$) to maximize the LTM interpolation error. In CTM, numerical diffusion is maximized by choosing the highest possible shockwave speed (w).

Figure 4.12 depicts the cumulative vehicle numbers at the upstream link boundary for the following scenarios:

- Analytical solution
- CTM solution with a time step *Δt = 30 s*, which means that the link is divided into 10 homogeneous cells
- CTM solution with a time step *Δt = 5 min*, which means that the link corresponds to only 1 cell
- LCTM solution with a time step *Δt = 5 min* and an optimal time lag of 2 time intervals
- LTM solution with a time step *Δt = 5min*

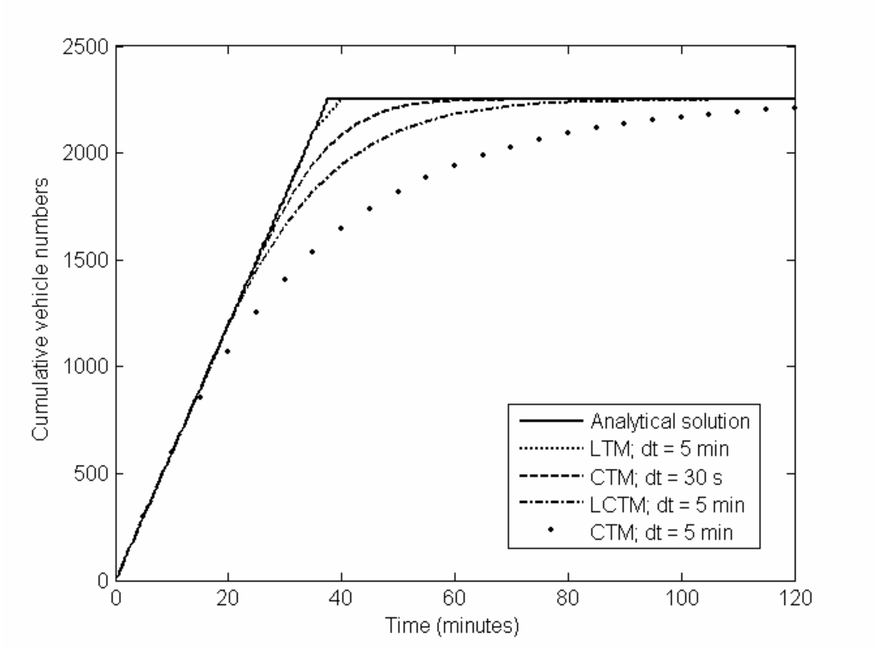


Figure 4.12: Cumulative vehicle numbers at the upstream link boundary in Example 4.5

It can be seen in this example, that enlarging the length of time step *Δt* seriously affects the accuracy of CTM. Errors (measured by the surface between the simulated and the analytical solution) are 9.5 times higher for *Δt=5 min* compared to *Δt = 30 s.* The Lagged CTM (LCTM) enhances the accuracy, but the error for *Δt = 5 min* is in this case still 2.7 times higher compared to the CTM simulation with *Δt = 30 s.* The interpolation error in the LTM solution with *Δt = 5 min* is 38 times smaller compared to the LCTM solution with *Δt = 5 min*.
Note that in LCTM, a flow restriction on the upstream boundary occurs *17,5 min* early (at *t = 20 min* instead of *t = 37.5 min*), i.e., traffic is influenced *17,5 minutes* early by traffic conditions *10 km* downstream, compared to only 2,5 minutes of earliness in LTM. Improved accuracy is an important quality of LTM.

4.9. Conclusions

LTM propagates traffic on its links as assumed in kinematic wave theory. This theory is based on a combination of the conservation of vehicles principle and the assumption that there exists a functional relation between traffic flow q and density k, also known as the fundamental diagram of traffic flow. Cumulative vehicle numbers are generally determined by following the path of the shock.

In his simplified theory of kinematic waves, Newell substantially simplifies this procedure by evaluating cumulative vehicle numbers directly from initial and boundary conditions, without following the actual path of the shock.

Based on Newell's method, an efficient numerical procedure to determine Sending and Receiving flows of network links is developed (first step of the LTM solution algorithm, cf. Section 3.2). Cumulative vehicle numbers only have to be determined at link boundaries. The basic procedure which uses triangular shaped fundamental diagrams, can straightforwardly be extended to include piecewise linear fundamental diagrams.

By determining Sending and Receiving flows that are consistent with Newell's simplified theory, LTM ensures that traffic is propagated as in kinematic wave theory.

Compared to CTM, higher efficiency and better accuracy are the main attractions of LTM. For a homogeneous link with m cells, the computational complexity of CTM is about m times higher for a given time step Δt.

Computational efforts of CTM can be reduced by enlarging the length of the time step Δt. Such operations affect the accuracy, which at the cost of some memory storage can partially be recovered by using the LCTM. However, LTM is far more accurate. Small interpolation errors can occur only once over the length of a link.

In general, LTM requires considerably less computational effort for the same level of accuracy, or, LTM is considerably more accurate for the same level of computational effort.

5

NODE MODELS

Node models that are used to build up motorway networks are presented in Section 5.1 of this chapter. Section 5.2 extends this set of node models with cross nodes for urban intersections. Both signalized and unsignalized intersections are represented by node models that take into account local flow restrictions.

Node models determine transition flows $G_{ij}(t)$ from the node's incoming links i to its outgoing links j. For node n, transition flow $G_{ij}(t)$ is the maximum amount of vehicles that can be transferred from incoming link $i \in I_n$ to outgoing link $j \in J_n$.
Node models always obey to the conservation of vehicles concept. Besides that, each node model holds some particular priority- and behavioral rules that impose constraints on the transition flows and determine which part of the Sending and Receiving flows can actually be sent and received.
Sending flows $S_i(t)$ are disaggregated by their Receiving links:

$$S_i(t) = \sum_{j \in J_n} S_{ij}(t) \qquad \textit{for all } i \in I_n \tag{5.1}$$

where $S_{ij}(t)$ is the Sending flow from incoming link $i \in I_n$ to outgoing link $j \in J_n$, i.e., $S_{ij}(t)$ is the fraction of Sending flow S_i that wants to go to link j.
Sending flow $S_i(t)$ is disaggregated into parts $S_{ij}(t)$, according to the disaggregation of Sending flow $S_i(t)$ at the time of passing the upstream link boundary:

$$S_{ij}(t) = \sum_{p \in P} \delta_{jp} \left\{ N^p \left[x_i^0, t_{x_i^0}\left(N(x_i^L, t) + S_i(t)\right) \right] - N^p(x_i^L, t) \right\} \qquad \textit{for all } i \in I_n,\ j \in J_n \tag{5.2}$$

where δ_{jp} is an element of link-route incidence matrix δ; $\delta_{jp} = 1$ if route p contains link j and $\delta_{jp} = 0$ otherwise, and where $t_{x_i^0}(N(x_i^L, t) + S_i(t))$ is the time at which vehicle number $(N(x_i^L, t) + S_i(t))$ passed the upstream link boundary x_i^0.
Note that the determination of disaggregate Sending flows requires both the link-FIFO- and CFL conditions to be satisfied (cf. Section 3.2).

5.1. Node models for motorway intersections

5.1.1. Inhomogeneous node

Inhomogeneous nodes are used to model a change in capacity or a change in free-flow speed on a road. They connect one incoming link i to one outgoing link j (see Figure 5.1).

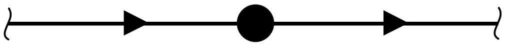


Figure 5.1: Inhomogeneous node

The transition flow $G_{ij}(t)$ through an inhomogeneous node is the maximum that can be sent by the incoming link i unless prevented to do so by the outgoing link j (the time variable t is dropped for simplicity of notation):

$$G_{ij} = \min(S_i, R_j) \tag{5.3}$$

5.1.2. Origin node

Traffic enters the network in an origin node, which is connected to one outgoing link j (see Figure 5.2).

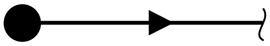


Figure 5.2: Origin node

The time-dependent route flow rates $f^p(t)$ are given in each origin node r $(r \in R)$. These flow rates are easily transformed to cumulative vehicle numbers as a function of time, representing traffic demand in each origin node. Let $N_r(t)$ denote the cumulative vehicle

curve representing total traffic demand over time in origin node r; $N_r^p(t)$ then represents the traffic demand for route p.
The number of vehicles that would like to enter the network during $[t, t+\Delta t]$ is calculated by $(N_r(t+\Delta t) - N(x_j^0, t))$, where j is the first network link connected to node r. $N(x_j^0,t)$ denotes the cumulative number of vehicles that already entered the network at time t. Vehicles that are unable to enter the network are implicitly waiting in a Point-Queue (P-Q) in the origin node. The number of vehicles in the P-Q at time t is given by $N_r(t) - N(x_j^0,t)$.

The (transition) flow G_j out of an origin node is the number of vehicles that would like to enter the network, unless this number exceeds the Receiving flow of the outgoing link j:

$$G_j(t) = \min((N_r(t+\Delta t) - N(x_j^0, t)),\ R_j) \tag{5.4}$$

An origin node can be considered as an inhomogeneous node where the Sending flow of the incoming link equals traffic demand.

5.1.3. Destination node

Traffic leaves the network in a destination node, which is connected to one incoming link i (see Figure 5.3).

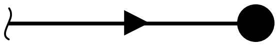


Figure 5.3: Destination node

The (transition) flow G_i towards a destination node is the Sending flow of the incoming link.

$$G_i = S_i \tag{5.5}$$

The destination node can be considered as an inhomogeneous node where the Receiving flow of the outgoing link is infinite.

5.1.4. Diverge node

Diverge nodes are used to model diverging lanes/off-ramps in motorway networks. They connect one incoming link to two outgoing links (see Figure 5.4).

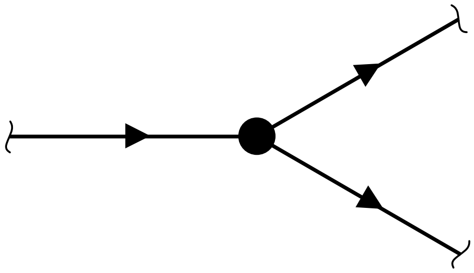


Figure 5.4: Diverge node

The flow through a diverge node is the maximum that can be sent by the incoming link, unless one of the outgoing links is unable to receive its allocated part of the Sending flow. In that case, it will be assumed as in Newell (1993) and Daganzo (1995a) that all the flow is restricted, which implies that vehicles at the intersection are served in a first-in-first-out (FIFO) sequence. The transition flows to each of the outgoing links are given by:

$$G_{ij} = \min_{j'\in J_n}\left(\frac{R_{j'}S_{ij}}{S_{ij'}}, S_{ij}\right) \qquad for\ all\ \ j \in J_n \tag{5.6}$$

However, when the Sending flow consists of vehicles having entered the link in different time intervals, application of Formula (5.6) does not guarantee perfect compliance to FIFO. To ensure perfect FIFO behavior on the incoming link, it is necessary to keep track of the time intervals vehicles entered the link. Vehicles having entered in an earlier time interval must advance sooner. The following iterative procedure assures the FIFO condition to hold on the incoming link (see also Figure 5.5).

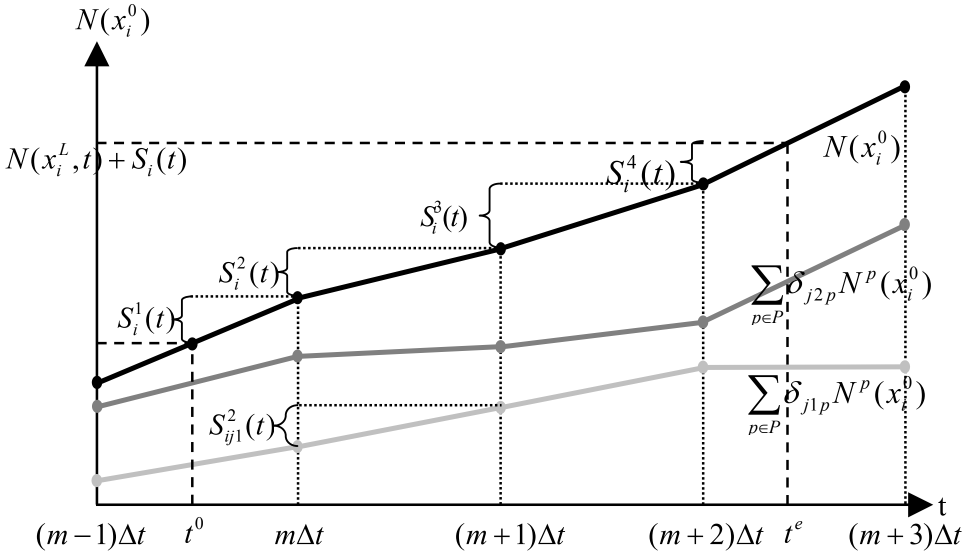


Figure 5.5: Sending flow and its sub-flows at a diverge node

LTM solution algorithm step 2 (cf. Section 3.2): Procedure to ensure perfect FIFO at a diverge node

```
Initialization

$t^0 = t_{x_i^0}(N(x_i^L, t))$
$t^e = t_{x_i^0}(N(x_i^L, t) + S_i(t))$
Determine m such that $(m-1)(\Delta t) \le t^0 < m(\Delta t)$, where m is an integer
For all $j \in J_n$: $R_j^1 = R_j(t)$
$k = 1$
$G_{ij} = 0$


Step

$t^k = \min((m+k-1)\Delta t, t^e)$
$S_i^k = N(x_i^0, t^k) - N(x_i^0, t^{k-1})$
For all $j \in J_n : S_{ij}^k = \sum_p \delta_{jp}(N^p(x_i^0, t^k) - N^p(x_i^0, t^{k-1}))$
        $G_{ij}^k = \min_{j' \in J_n}(\frac{R_{j'}^k}{S_{ij'}^k} S_{ij}^k, S_{ij}^k)$
        $R_j^{k+1} = R_j^k - G_{ij}^k$
        $G_{ij} = G_{ij} + G_{ij}^k$

Stop criterion

if $t^k = t^e$ then Stop
if $R_j^{k+1} = 0$ then Stop
else $k = k+1$; repeat Step.
```

This procedure to ensure perfect FIFO at a diverge node is time-consuming. It may be used for (theoretical) applications that require perfect compliance to FIFO. For most practical applications however, this procedure may be skipped. Flows are then averaged over each time interval without keeping track of the time interval vehicles entered. However, it is believed that the effect on overall results is small and can be minimized by using shorter time intervals. The procedures for perfect FIFO at merge nodes (Section 5.1.5) and urban cross nodes (Section 5.2) are analogous; they are not further elaborated in this thesis.

5.1.5. Merge node

Merge nodes are used to model merging lanes/on-ramps in motorway networks. They connect two incoming links to one outgoing link (see Figure 5.6).

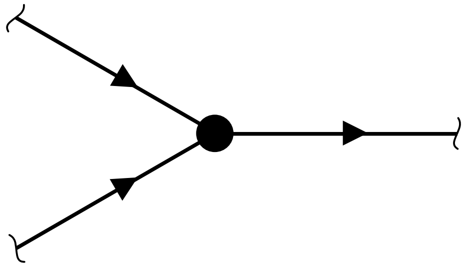


Figure 5.6: Merge node

The flow through a merge node is the sum of the Sending flows of the incoming links $i \in I_n$, unless this total flow exceeds the Receiving flow of the outgoing link j. In that case, it will be assumed that the maximum amount that can be received by link j will be transferred according to the priority parameters p_{ij} associated with the incoming links. Parameter p_{ij} denotes the fraction of the total amount of vehicles, coming from link i $(\sum_i p_{ij} = 1)$. These parameters provide flexibility in modeling prioritized merge junctions or the action of a ramp metering signal. Daganzo (1995a) provided a detailed treatment of this aspect. The transition flows are written as:

$$G_{ij} = median(S_{ij}\ ,\ R_j - ((\sum_i S_{ij}) - S_{ij})\ ,\ p_{ij}R_j) \qquad for\ all\ \ i \in I_n \tag{5.7}$$

Simplified merge models were suggested by Lebacque (1996) and Jin & Zhang (2003). Lebacque assumes that a fixed fraction p_{ij} of the outgoing link j is reserved exclusively to vehicles from incoming link i. The outflow from an incoming link can never exceed the reserved fraction of the capacity of the outgoing link. The transition flows are calculated as follows:

$$G_{ij} = \min(S_{ij}\ ,\ p_{ij}R_j) \qquad for\ all\ \ i \in I_n \tag{5.8}$$

Priority proportions in Jin & Zhang's model are not fixed, but they are equal to the proportions of the upstream Sending flows. The so called 'demand proportional' transition flows are given by:

$$G_{ij} = \min(\frac{R_j S_{ij}}{\sum_{i' \in I_n} S_{i'j}}, S_{ij}) \qquad for\ all\ \ i \in I_n \tag{5.9}$$

5.2. Node models for urban intersections

The previous section described node models for motorway networks where merge and diverge node models represent basic motorway junctions. To describe traffic conditions on urban networks, we need to include urban intersection models, both for signalized and unsignalized intersections.
Signalized and unsignalized intersections have two main effects:

(i) they impose capacity constraints $q_{n,ij}$ on the transition flows G_{ij} from incoming links $i \in I_n$ to outgoing links $j \in J_n$ and

(ii) vehicles on the incoming links experience intersection delays $D_{int,ij}$.

Even if the constrained capacities $q_{n,ij}$ (further called node capacities) are on average high enough to handle traffic demand, some vehicles still experience delays waiting before red lights or waiting before conflicting priority streams. These delays are further referred to as intersection delays.

To incorporate capacity constraints and intersection delays, there are two possible solution methods:

(i) explicitly simulating green and red stages (for signalized intersections) and gaps in conflicting priority streams (for unsignalized intersections), or

(ii) considering the average effects of traffic lights and priority streams, without simulating them directly.

The first solution method has three main disadvantages in the context of DTA:

(i) Explicit simulations of traffic light stages and gaps in traffic streams impose constraints on the simulation time step. Frequently changing boundary conditions require small simulation time steps. However, small time steps substantially increase computational complexity (cf. Section 4.8). For signalized intersections, the highest possible time step would be the greatest common divisor of the green times of all turning movements.
For unsignalized intersections, constraints can even be more severe: the highest possible time step would be the greatest common divisor of critical gap times and follow-up times between successive vehicles. Time steps would typically correspond to those of micro-simulation models.

(ii) Explicit simulations generate high-frequent fluctuations in route travel times over time due to the occurrence of stage alternations etc… However, drivers do not alter their route choice (nor their departure time choice) based on these small accidental travel time fluctuations. Expected average travel times (and distributions around these averages) are more relevant.

(iii) Due to their stochastic nature, explicit simulations yield one single solution in a whole spectrum of possible solutions. This single solution depends on an accidental combination of traffic light offsets, accidental gap distributions in traffic streams etc… Since route choices are made based on expected average travel times, an average solution would be more relevant.

In an attempt to overcome these disadvantages, we explore the second solution method, where the average effects of traffic lights and priority streams are considered indirectly. A similar approach is followed by Durlin and Henn (2005). A basic urban cross node model that takes into account average capacity constraints is presented below. Chapter 6 proposes a method to include average intersection delays $D_{int,ij}$.

Basic urban cross node

Cross nodes are used to represent urban intersections. They are multi-legged junctions connecting two or more incoming links i to two or more outgoing links j (see Figure 5.7). Urban cross nodes are typically composed out of 4 incoming and 4 outgoing links.

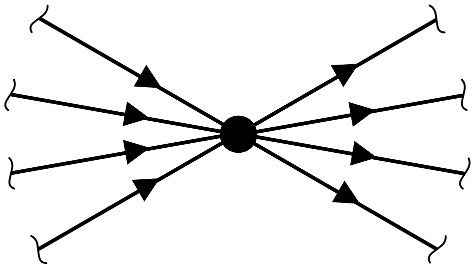


Figure 5.7: Cross node

Transition flows G_{ij} through cross node n are constrained by the Sending flows S_{ij} of the incoming links $i \in I_n$ and the Receiving flows R_j of the outgoing links $j \in J_n$.
Receiving flows R_j are distributed over their different incoming links $i \in I_n$ according to priority parameters p_{ij}:

$$R_{ij} = p_{ij} R_j \tag{5.10}$$

The Receiving flow R_{ij} of outgoing link j is reserved exclusively to vehicles from incoming link i. Priority parameters p_{ij} denote the fractions of the total amount of link j vehicles that are coming from link i:

$$\sum_{i \in I_n} p_{ij} = 1 \tag{5.11}$$

However, if Receiving flow R_{ij} is not completely used by transition flow G_{ij}, then the remaining part of R_{ij} may be used by other flows $G_{i'j}$, should these flows $G_{i'j}$ be restricted by their allocated Receiving flows $R_{i'j}$.
Parameters p_{ij} are characteristics of the intersection. They provide flexibility in modeling traffic lights and priorities for unsignalized intersections.

Next to the constraints imposed by Sending and Receiving flows S_{ij} and R_{ij}, transition flows G_{ij} are also subject to average capacity constraints $q_{n,ij}$ that are imposed by traffic lights or by conflicting priority streams on node n, and they are also subject to FIFO constraints. FIFO constraints imply that all the Sending flow from link i is restricted if one of its sub-flows S_{ij} is restricted. The transition flow G_{ij} is the maximum flow taking into account these constraints:

$$G_{ij} = \min_{j' \in J_n}\left(\frac{q_{n,ij'}}{S_{ij'}}, \frac{R_{ij'}}{S_{ij'}}, 1\right). S_{ij} \tag{5.12}$$

Hereafter, a distinction is made between signalized and unsignalized intersections.

5.2.1. Basic signalized urban cross node

Signalized intersections occur in all shapes and sizes. It is not our intention to give an exhaustive solution method description for all types of signalized intersections. We rather describe how one typical signalized intersection could be dealt with and then discuss possible adjustments to deal with other types of intersections.

5.2.1.1. Prototype signalized intersection

The prototype intersection pictured below (Figure 5.8) has 4 incoming and 4 outgoing links and it has the following characteristics:

- The FIFO principle applies to all links

- All turning movements have the disposal of at least one turning lane and they have enough storage capacity on their turning lanes and on the intersection area itself, so that they don't block other sub-flows on the same link, as long as the turning flows can on average be handled. If one of the turning flows is oversaturated however, then the other sub-flows are restricted due to FIFO constraints.

- Traffic lights are controlled without conflicts, i.e. there are no conflicting traffic streams, so that the time during which vehicles can actually proceed equals the effective green time for these vehicles.

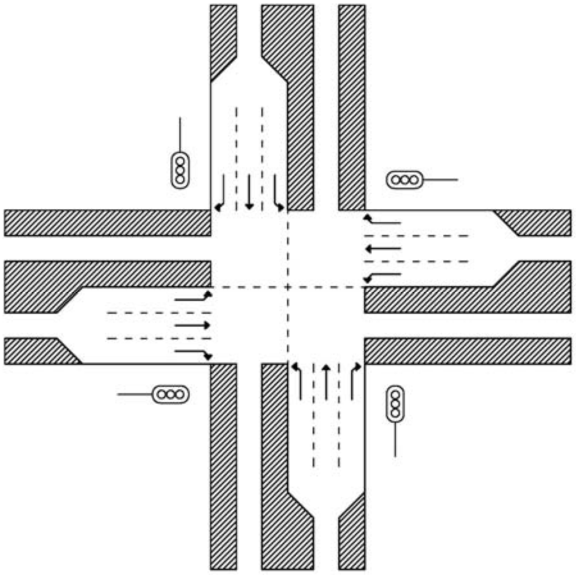


Figure 5.8: Prototype Signalized intersection

Node capacity $q_{n,ij}$

Signalized intersections on average impose the following trivial capacity constraints for each turning movement:

$$q_{n,ij} = \frac{g_{ij}}{c} q_{M,ij} \tag{5.13}$$

where $q_{n,ij}$ = average node capacity for turning movement ij, imposed by node n (veh/s)
g_{ij} = effective green time for turning movement ij (s)
c = cycle length (s)
$q_{M,ij}$ = saturation flow for turning movement ij (veh/s).

Saturation flow $q_{M,ij}$ depends on the amount of available lanes for this particular turn, and on the length of the turning lanes.

Receiving flow R_{ij}

The Receiving flow R_j is distributed over its Sending links according to priority parameters p_{ij}. For signalized intersections, these priority parameters correspond to the ratios of the products of green time and saturation flow:

$$R_{ij} = p_{ij} R_j = \frac{g_{ij} q_{M,ij}}{\sum_{i \in I_n} g_{ij} q_{M,ij}} R_j = \frac{q_{n,ij}}{\sum_{i \in I_n} q_{n,ij}} R_j \tag{5.14}$$

Solution method for the prototype signalized intersection

In a first step, transition flows G_{ij} are determined using Formula (5.12), where $q_{n,ij}$ and R_{ij} are calculated with Formulae (5.13) and (5.14) respectively. For each outgoing link j, all transition flows G_{ij} that are not restricted by their corresponding Receiving flows R_{ij} are sent definitively. If there is some spare Receiving flow R_j resulting from transition flows G_{ij} being smaller than their allocated Receiving flows R_{ij}, then transition flows $G_{i'j}$ that are restricted by their corresponding Receiving flows $R_{i'j}$ are re-evaluated in a second step. The unused Receiving flow $\sum_{i \in I_n} \left(R_j - G_{ij} \right)$ is then re-assigned to the turning movements $i'j$ that were restricted in the first step. This re-distribution typically depends on the prevailing sequence of signal phases (Viti (2006)). In absence of any information on the signal plan, an equal distribution over the different restricted turning movements is assumed, as indicated in the algorithm below. If there are transition flows $G_{i'j}$ that are still restricted by their Receiving flow $R_{i'j}$ after the second step, and if there is still some spare Receiving flow R_j left, then Receiving flows $R_{i'j}$ are re-evaluated in a third step, etc…

LTM solution algorithm step 2 (cf. Section 3.2): Solution method for the prototype signalized intersection

```
for all $j \in J_n$:

Initialization

$k = 0$
for all $i \in I_n$:
    $G_{ij}^0 = \min_{j' \in J_n}(\frac{q_{n,ij'}}{S_{ij'}}, \frac{R_{ij'}^0}{S_{ij'}}, 1)\ S_{ij}$
    where $R_{ij}^0 = \frac{g_{ij} q_{M,ij}}{\sum_{i \in I_n} g_{ij} q_{M,ij}} R_j$
Let $I'^0 = \left\{ i \in I_n \middle| \left( G_{ij}^0 = R_{ij}^0 \ \& \ R_{ij}^0 < S_{ij}^0 \ \& \ R_{ij}^0 < q_{n,ij} \right) \right\}$
Let $I^0 = \left\{ i \in I_n \middle| i \notin I'^0 \right\}$

Stop criterion

If $\#I^0 = 0$ or $\#I'^0 = 0$, then Stop
Else $k = 1$; go to Step

Step

for all $i \in I^{k-1}$: $G_{ij}^k = G_{ij}^{k-1}$

for all $i \in I'^{k-1}$:
        $G_{ij}^k = \min_{j' \in J_n}(\frac{q_{n,ij'}}{S_{ij'}}, \frac{R_{ij'}^k}{S_{ij'}}, 1) . S_{ij}$
    where $R_{ij}^k = R_{ij}^{k-1} + \frac{1}{\# I'^{k-1}} \sum_{i \in I^{k-1}} (R_{ij}^{k-1} - G_{ij}^{k-1})$
Let $I'^k = \left\{ i \in I'^{k-1} \middle| \left( G_{ij}^k = R_{ij}^k \ \& \ R_{ij}^k < S_{ij}^k \ \& \ R_{ij}^k < q_{n,ij} \right) \right\}$
Let $I^k = \left\{ i \in I'^{k-1} \middle| i \notin I'^{k-1} \right\}$

Stop criterion

If $\#I^k = 0$ or $\#I'^k = 0$, then Stop
Else $k = k+1$; go to Step

for all $i \in I_n$:    $G_{ij} = G_{ij}^k$
```

5.2.1.2. Adjustments for other types of signalized intersections

Other signalized intersections are treated analogously. Some adjustments might be desirable:

- If one of the turning movements doesn't have a separate lane, neither a (long enough) turning lane to prevent the potential blocking of other sub-flows on the same link, then the saturation flow $q_{M,ij}$ for this turning movement should be determined by the shared-lane formula, where different movements operate on the same lane.
  The shared-lane formula arises by assuming homogeneously spread sub-flows S_{ij} and by applying the FIFO principle on this one lane:

  $$q_{M,ij} = \frac{S_{ij}}{S_i} q_{M,i} \tag{5.15}$$

  Saturation flow $q_{M,i}$ should be partitioned according to the ratios between sub-flows S_{ij} and total Sending flow S_i. This partition is the only one that is consistent with the FIFO principle.

- If traffic lights are not controlled without conflicts, i.e. if vehicles, for example left-turners, have to yield to conflicting traffic streams during their green time, then the average node capacity $q_{n,ij}$ for these left-turners depends on conflicting flows. Saturation flow $q_{M,ij}$ should in this case be determined by Equation (5.16) for unsignalized intersections, which is dependent on conflicting priority streams. A situation of mutual interdependency occurs, since the capacity of left-turning flows is constrained by conflicting non-left-turning flows, and non-left-turning flows are restricted by left-turning flows due to FIFO constraints. This topic is dealt with in the next Section 5.2.1.

- For spacious intersections with long separate lanes for each turning movement, the FIFO principle might considerably be violated in reality. In such cases, the different turning lanes should be modeled as separate links. This goes, however, at the cost of higher computational efforts, especially since turning lanes are typically short compared to average link lengths, which requires the up- and downstream nodes of these short links to be updated with small time steps (cf. Section 3.2).

5.2.2. Basic unsignalized urban cross node

At unsignalized intersections, there is a hierarchy of traffic streams. Some streams have absolute priority, some have to yield to higher order streams, and others have to yield to some streams which in turn have to yield to others. Different levels of priority ranking

can be distinguished:
Rank 1 streams have absolute priority and do not need to yield to other streams.
Rank 2 streams have to yield to Rank 1 streams.
Rank 3 streams have to yield to Rank 2 streams and in turn to Rank 1 streams.
Rank 4 streams have to yield to Rank 3 streams and in turn to Rank 2 and Rank 1 streams (e.g. left turners from the minor street).
To express which traffic movement has the right of way over another one, Miltner (2003) proposed a so-called conflict matrix. This matrix (Table 5.1) represents all existing priority rules. If one turning movement *ij* interferes with another one ($\overline{ij}$), then the corresponding cell of the matrix contains a value $A_{\overline{ij},ij}$. If $A_{\overline{ij},ij} = 1$, then movement $\overline{ij}$ (conflicting movement) has priority over movement *ij* (subject movement). If $A_{\overline{ij},ij} = 0$, then $\overline{ij}$ has to give way to *ij*.

5.2.2.1. Prototype unsignalized intersection

Again, we describe here one typical unsignalized intersection and then discuss possible adjustments to deal with other types of unsignalized intersections.
Figure 5.9 illustrates the prototype intersection which has the following characteristics:

- The FIFO principle applies to all links.

- There are no turning lanes. All lanes are shared lanes, where all movements operate on the same lane.

- The level of priority ranking of the different traffic streams is as follows:
  Priority rank 1: turning movements *ab*, *ac*, *ba*, *bd*
  Priority rank 2 - major street: turning movements *ad*, *bc*
  Priority rank 2 - minor street: turning movements *cb, da*
  Priority rank 3: turning movements *cd*, *dc*
  Priority rank 4: turning movements *ca*, *db*
  Higher rank streams *ij* have absolute priority over lower rank streams *i'j*, implying that $p_{ij} = 1$ and $p_{i'j} = 0$.
  The conflict matrix is given in Table 5.1. Conflicting movement $\overline{ad}$ for example, has the right of way over turning movements *ca*, *cd*, *db* and *dc*, but it has to give priority to turning movements *ba* and *bd*.

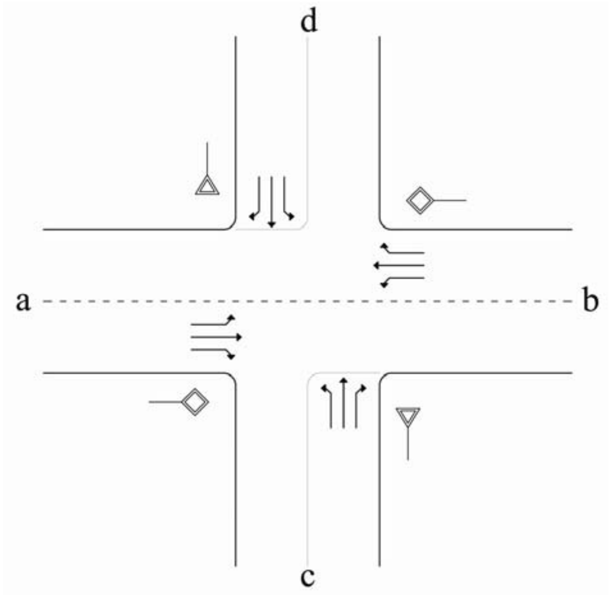


Figure 5.9: Prototype unsignalized intersection

Table 5.1: Conflict matrix representing priority rules for the prototype unsignalized intersection

$A_{\overline{ij},ij}$	turning movement $\overline{ij}$											
		$\overline{ad}$	$\overline{ab}$	$\overline{ac}$	$\overline{ca}$	$\overline{cd}$	$\overline{cb}$	$\overline{bc}$	$\overline{ba}$	$\overline{bd}$	$\overline{db}$	$\overline{dc}$	$\overline{da}$
turning movement ij	ad	▨	▨	▨	1	1	▨	▨	0	0	1	1	▨
	ab	▨	▨	▨	1	1	1	1	▨	▨	1	1	▨
	ac	▨	▨	▨	▨	▨	▨	1	▨	▨		1	▨
	ca	0	0	▨	▨	▨	▨	0	0			0	0
	cd	0	0	▨	▨	▨	▨	0	0	0	1		▨
	cb	▨	0	▨	▨	▨	▨	▨	▨	▨	1		▨
	bc	▨	0	0	1	1	▨	▨	▨	▨	1	1	▨
	ba	1	▨	▨	1	1	▨	▨	▨	▨	1	1	1
	bd	1	▨	▨	▨	1	▨	▨	▨	▨	▨	▨	▨
	db	0	0	▨	▨	0	0	0	0	▨	▨	▨	▨
	dc	0	0	0	1	▨	▨	0	0	▨	▨	▨	▨
	da	▨	▨	▨	1	▨	▨	▨	0	▨	▨	▨	▨

Node capacity $q_{n,ij}$

For unsignalized intersections, Brilon and Miltner (2005) recently proposed a capacity formula based on so-called conflict technique. This method allows calculating the capacity for one minor movement out of one single equation:

$$q_{n,ij} = \frac{3600}{t_{s,ij}} \left[1 - \frac{1}{3600} \sum_{\overline{ij} \in D_{\overline{ij}}} A_{\overline{ij},ij}\, q_{\overline{ij}}\, t_{s,\overline{ij}} \right] e^{\frac{1}{3600} \sum_{\overline{ij} \in D_{\overline{ij}}} A_{\overline{ij},ij}\, q_{\overline{ij}}\, t_{a,\overline{ij}}} \tag{5.16}$$

where $q_{n,ij}$ = average node capacity for turning movement ij, imposed by node n (veh/h)
$\overline{ij}$ = index for a turning movement that is conflicting with turning movement ij (-)
$D_{\overline{ij}}$ = set of turning movements that are conflicting with turning movement ij (-)
$q_{\overline{ij}}$ = traffic flow of conflicting movement $\overline{ij}$ (veh/h)
$t_{s,ij}$ = occupation time caused by a vehicle from movement ij (s)
$t_{a,ij}$ = time margin during which an approaching major stream vehicle from movement ij is blocking the intersection in advance of its arrival (s)
$A_{\overline{ij},ij}$ = conflict factor to consider the priority rules (-).

A minor stream vehicle can only enter the intersection if the intersection is not used by a major vehicle and if no major vehicle is approaching the intersection. To express the latter effect, a time margin $t_{a,ij}$ (in advance of major vehicle ij's arrival) is defined, during which no minor vehicle is able to enter the intersection. Major stream gaps are assumed to be exponential distributed.

Receiving flow R_{ij}

Receiving flow R_j is distributed over its Sending links according to priority parameters p_{ij} (cf. Equation (5.10)). It is assumed that a high rank (major) stream has absolute priority over a lower rank (minor) stream. Absolute priority to incoming link i implies that $p_{ij} = 1$ and $p_{i'j} = 0$ for all $i' \in \{I_n \setminus i\}$. If transition flow G_{ij} on an incoming link with absolute priority is smaller than its allocated Receiving flow R_{ij}, then the remaining part of R_{ij} is redistributed over turning movements $i'j$ that have a lower priority rank.

Solution method for the prototype unsignalized intersection

Transition flows G_{ij} depend on each other: node capacities $q_{n,ij}$ and Receiving flows R_{ij} depend on conflicting transition flows, and conversely transition flows depend on node capacities $q_{n,ij}$ and Receiving flows R_{ij}. Therefore, transition flows G_{ij} cannot be determined for all turning flows ij together. The solution method consists in first determining higher rank transition flows and then proceeding to lower rank transition flows. In a second phase, higher rank streams are re-evaluated based on the results of the lower rank flows. Re-evaluation of lower rank streams might be considered as well, depending on the application's desired accuracy level.

- Priority rank 1 streams

First, we determine the transition flows for rank 1 streams, provisionally ignoring all lower rank streams. There are no conflicting flows $q_{\overline{ij}}$ ($D_{\overline{ij}}$ is empty) and the complete Receiving flows are at the disposal of these highest rank streams:

For G_{ab} and G_{ac}: $G_{ij} = \min\limits_{j' \in J'}(\frac{q_{n,ij'}}{S_{ij'}}, \frac{R_{ij'}}{S_{ij'}}, 1) . S_{ij}$

where $J' = \{b, c\}$

$R_{ij'} = R_{j'}$ (i.e. $R_{ab} = R_b$ and $R_{ac} = R_c$)

$q_{n,ij}$ is determined by Formula (5.16), where $D_{\overline{ij}} = \{/\}$

For G_{ba} and G_{bd}: $G_{ij} = \min\limits_{j' \in J'}(\frac{q_{n,ij'}}{S_{ij'}}, \frac{R_{ij'}}{S_{ij'}}, 1) . S_{ij}$

where $J' = \{a, d\}$

$R_{ij'} = R_{j'}$

$q_{n,ij}$ = (5.16), where $D_{\overline{ij}} = \{/\}$

- Priority rank 2 – major street – streams

Secondly, we determine the transition flows for rank 2 streams, given the higher rank transition flows. The remainder of the Receiving flows of links *d* and *c* are now attributed to the left-turning flows *ad* and *bc*. Note that rank 1 streams might restrict the left-turning rank 2 streams due to FIFO constraints.

For G_{ad}: $G_{ij} = \min\limits_{j' \in J'}(\frac{q_{n,ij'}}{S_{ij'}}, \frac{R_{ij'}}{S_{ij'}}, 1) . S_{ij}$

where $J' = \{b, c, d\}$

$R_{ad} = R_d - G_{bd}$

$q_{n,ij}$ = (5.16), where $D_{\overline{ad}} = \{ba, bd\}$; $q_{ba} = G_{ba}$; $q_{bd} = G_{bd}$

For G_{bc}: $G_{ij} = \min_{j' \in J'}(\frac{q_{n,ij'}}{S_{ij'}}, \frac{R_{ij'}}{S_{ij'}}, 1) . S_{ij}$

where $J' = \{a, c, d\}$

$R_{bc} = R_c - G_{ac}$

$q_{n,ij}$ = (5.16), where $D_{\overline{bc}} = \{ab, ac\}$; $q_{ab} = G_{ab}$; $q_{ac} = G_{ac}$

- Re-evaluation of rank 1 streams

The FIFO principle on links *a* and *b* implies that all the Sending flow from link *i* is restricted if one of its sub-flows S_{ij} is restricted:

For G_{ab} and G_{ac}: $G_{ij} = \min_{j' \in J'}(\frac{G_{ij'}}{S_{ij'}}) . S_{ij}$

where $J' = \{b, c, d\}$

For G_{ba} and G_{bd}: $G_{ij} = \min_{j' \in J'}(\frac{G_{ij'}}{S_{ij'}}) . S_{ij}$

where $J' = \{a, c, d\}$

If left-turning flows G_{ad} and G_{bc} both impose the most stringent restrictions on the transition flows, then one might consider to recalculate rank 2 transition flows, since R_{ij} and $q_{n,ij}$ were calculated based on rank 1 transition flows that now appear to be too high. Rank 1 and rank 2 streams can be harmonized with each other in an iterative procedure. However, such an iterative procedure is time-consuming; it should only be used for (theoretical) applications that require the highest accuracy.

- Priority rank 2 – minor street – streams

Transition flows for rank 2 – minor street – streams are determined, given the higher rank transition flows. The remainder of the Receiving flows of links *b* and *a* are now attributed to the flows *cb* and *da*.

For G_{cb}: $G_{ij} = \min_{j' \in J'}(\frac{q_{n,ij'}}{S_{ij'}}, \frac{R_{ij'}}{S_{ij'}}, 1) . S_{ij}$

where $J' = \{b\}$

$R_{cb} = R_b - G_{ab}$

$q_{n,ij}$ = (5.16), where $D_{\overline{cb}} = \{ab\}$; $q_{ab} = G_{ab}$

For G_{da}: $G_{ij} = \min_{j' \in J'}\left(\frac{q_{n,ij'}}{S_{ij'}}, \frac{R_{ij'}}{S_{ij'}}, 1\right).S_{ij}$

where $j' = \{a\}$

$R_{da} = R_a - G_{ba}$

$q_{n,ij}$ = (5.16), where $D_{\overline{da}} = \{ba\}$; $q_{ba} = G_{ba}$

- Priority rank 3 streams

Transition flows for rank 3 streams are determined analogously, given the higher rank transition flows. The remainders of the receiving flows (parts that are not used by higher rank streams) are now attributed to rank 3 streams. Note that rank 2 – minor street – streams might restrict rank 3 streams due to FIFO constraints.

For G_{cd}: $G_{ij} = \min_{j' \in J'}\left(\frac{q_{n,ij'}}{S_{ij'}}, \frac{R_{ij'}}{S_{ij'}}, 1\right).S_{ij}$

where $J' = \{d, b\}$

$R_{cd} = R_d - G_{bd} - G_{ad}$

$q_{n,ij}$ = (5.16), where $D_{\overline{cd}} = \{ad, ab, bc, ba, bd\}$; $q_{ad} = G_{ad}$ etc...

For G_{dc}: $G_{ij} = \min_{j' \in J'}\left(\frac{q_{n,ij'}}{S_{ij'}}, \frac{R_{ij'}}{S_{ij'}}, 1\right).S_{ij}$

where $J' = \{c, a\}$

$R_{da} = R_a - G_{ba}$

$q_{n,ij}$ = (5.16), where $D_{\overline{dc}} = \{ad, ab, ac, bc, ba\}$; $q_{ad} = G_{ad}$ etc...

- Priority rank 4 streams

Finally, we determine the transition flows for rank 4 streams, given the higher rank transition flows. The procedure is analogous to that of rank 2 – major street – streams:

For G_{ca}: $G_{ij} = \min_{j' \in J'}\left(\frac{q_{n,ij'}}{S_{ij'}}, \frac{R_{ij'}}{S_{ij'}}, 1\right).S_{ij}$

where $J' = \{a, b, d\}$

$R_{ca} = R_a - G_{ba} - G_{da}$

$q_{n,ij}$ = (5.16), where $D_{\overline{ca}} = \{ad, ab, bc, ba, dc, da\}$; $q_{ad} = G_{ad}$ etc...

For G_{db}: $G_{ij} = \min_{j' \in J'}(\frac{q_{n,ij'}}{S_{ij'}}, \frac{R_{ij'}}{S_{ij'}}, 1) . S_{ij}$

where $J' = \{a, b, c\}$

$R_{db} = R_b - G_{ab} - G_{cb}$

$q_{n,ij}$ = (5.16), where $D_{\overline{db}} = \{ad, ab, cd, cb, bc, ba\}$; $q_{ad} = G_{ad}$ etc...

5.2.2.2. Adjustments for other types of unsignalized intersections

Other unsignalized intersections are treated analogously. Some adjustments might be desirable:

- If one of the turning movements *ij* has a separate lane or a long enough presort-lane in order not to block other sub-flows as long as the turning flows can on average be handled, then the saturation flow $q_{M,ij}$ for this turning movement may be determined as the capacity of the corresponding separate or presort-lane, instead of using the shared-lane formula.

- The partition $p_{ij} = 1$ and $p_{i'j} = 0$ for stream *ij* having absolute priority over stream *i'j* might be on the sharp side. Based on observations of 'priority reversal', typically due to either gap forcing or politeness, Brilon and Miltner (2005) propose a less distinct partition of the priority parameters.

- Some types of unsignalized intersections, like roundabouts, suffer from a high degree of interdependency of transition flows, which would require the use of complex iterative procedures to determine transition flows.
  To speed up the process in such cases, one might consider – at the cost of some accuracy – to determine average node capacities $q_{n,ij}$ and Receiving flows R_{ij}, as a function of the transition flows in the previous time interval $G_{ij}(t-1)$.
  This approach, which might as well be applied for the prototype unsignalized intersection, is illustrated in Example 5.1, where the transition flows for a classical roundabout (see Figure 5.10) are determined.

Example 5.1

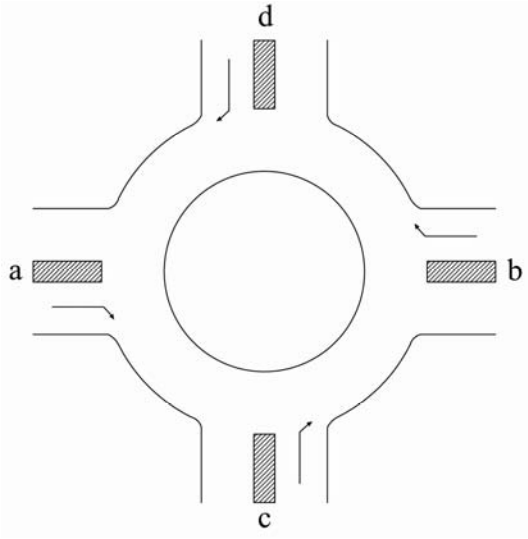


Figure 5.10: Classical Roundabout

Solution method for a classical roundabout:

$$\textit{for all } G_{ij}: \; G_{ij} = \min_{j' \in J_n}\left(\frac{q_{n,ij'}}{S_{ij'}}, \frac{R_{ij'}}{S_{ij'}}, 1\right) . S_{ij}$$

where

$$q_{p,aj} = G_{bc}(t-1) + G_{dc}(t-1) + G_{db}(t-1)$$

$$q_{p,bj} = G_{ad}(t-1) + G_{cd}(t-1) + G_{ca}(t-1)$$

$$q_{p,cj} = G_{db}(t-1) + G_{ab}(t-1) + G_{ad}(t-1)$$

$$q_{p,dj} = G_{ca}(t-1) + G_{ba}(t-1) + G_{bc}(t-1)$$

$$R_{ij} = \frac{G_{ij}(t-1)}{\sum_{i' \in I_n} G_{i'j}(t-1)} R_j$$

5.3. Conclusions

Node models determine transition flows from the node’s incoming links to its outgoing links, as prescribed in the second step of the LTM solution algorithm. Node models that represent basic motorway junctions, such as merge and diverge nodes, are easily integrated within LTM.

To describe traffic conditions at signalized and unsignalized intersection, urban cross node models are introduced. The approach here is to consider the average effects of traffic lights and priority streams, without explicitly simulating green and red stages (for signalized intersections) or gaps in conflicting priority streams (for unsignalized intersections).

This methodology has two main advantages in the context of DTA: it allows walking through simulations in large time steps and it directly yields average expected travel times that travelers use for their route choice. After describing the procedures for prototypes of a signalized and an unsignalized intersection, possible adjustments are discussed to deal with other types of intersections.

6

INTERSECTION DELAY MODEL

Even if signalized or unsignalized intersections can on average handle traffic demand, vehicles still experience delays waiting before red lights or waiting before conflicting priority streams. These delays are further referred to as intersection delays.

In general, currently used intersection delay models are based on queuing theory. They are quite easily integrated within travel-time based traffic models. However, integration within continuum traffic flow models is not self-evident because travel time is not a basic variable in these models. This chapter introduces a method to implicitly include intersection delays in flow-based continuum models like LTM.

Section 6.1 explains how vehicles are hold up in Point-Queues at the downstream link end, in which they experience their intended intersection delay.

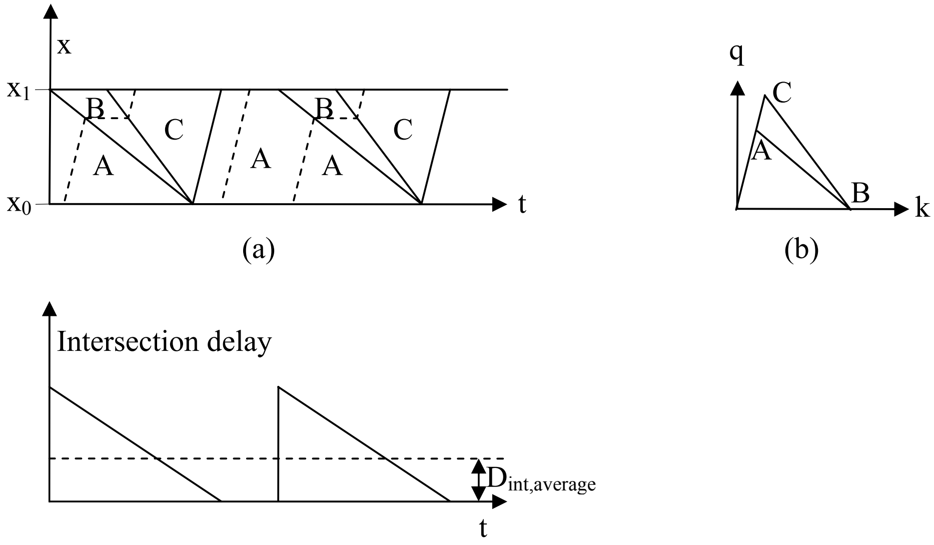
Section 6.2 presents the extended LTM solution algorithm which includes the modeling of intersection delays. Sections 6.2.1, 6.2.2, 6.2.3 and 6.2.4 describe the different algorithm steps in detail. Furthermore, Section 6.2.4 presents an example where the proposed solution method to implicitly include intersection delays is compared with an explicit simulation of traffic flows. Conclusions are formulated in Section 6.3.

6.1. Introduction of Point-Queues to implicitly realize average intersection delays

6.1.1 Objective

LTM realizes intersection delays in an indirect way. The objective is to realize average delays, where high-frequent fluctuations due to a repeated process of queue formation and queue dissipation are filtered out.

Figure 6.1 shows an example where traffic lights are simulated explicitly. Figure 6.1a plots densities in a *t-x* plane, where vehicles drive from location x_0 to traffic light x_1. The dotted lines represent vehicle trajectories. Densities *A*, *B* and *C* are indicated in the fundamental diagram of Figure 6.1b. Figure 6.1c pictures corresponding intersection delays as a function of time.



(a) (b)

(c)

Figure 6.1: *t-x* diagram (a), *k-q* fundamental diagram (b) and average intersection delay (c) at a signalized intersection

The objective of the intersection delay model is to realize average intersection delay $D_{int,average}$.

6.1.2. Implicit versus explicit consideration of intersection delays

LTM is a flow-based model which primarily determines cumulative vehicle numbers (or link flows), based on traffic conditions. Travel times can only be derived afterwards, when vehicles have completed their journey.

When explicitly simulating traffic light stages and gaps in conflicting priority streams, LTM also primarily determines cumulative vehicle numbers. Realized travel times include intersection delays, which are derived from these cumulative vehicle numbers.

To implicitly consider intersection delays however, a reverse procedure needs to be introduced: it is first determined what the expected average intersection delay would be, based on prevailing traffic conditions. Existing average intersection delay formulae are generally based on queuing theory (cf. Section 6.2.1).

Subsequently, original cumulative vehicle numbers need to be modified such that travel times derived from these modified curves include the intended intersection delays.

6.1.3. Implicit consideration of intersection delays: flow based versus travel-time based models

Average intersection delay formulae that are based on queuing theory are typically used in travel-time based models such as CONTRAM (Taylor (2003)), INDY-TT (Bliemer (2003)), MARPLE (Taale and Westerman (2004)), etc., to include average intersection delays. In these models, link travel times are determined at the time of link entrance, as a function of the number of vehicles on the link. Average intersection delays are included by simply adding them to the link travel times.

As opposed to travel-time based models, it is not self-evident to include average intersection delays in a flow-based model such as LTM. The reason is that travel time is not a basic variable in flow-based models. LTM first determines cumulative vehicle numbers (or link flows), based on traffic conditions. Travel times can only be derived afterwards, when vehicles have completed their journey.

Figure 6.2 compares the sequence in which cumulative vehicle numbers are determined in the flow-based LTM to the one in which cumulative vehicle numbers would be determined in travel-time based models.

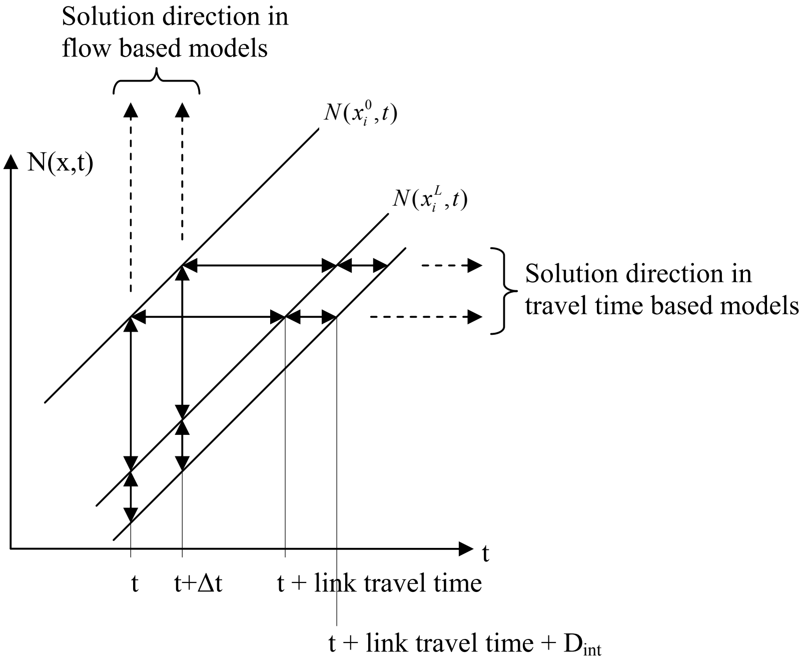

Solution direction in flow based models
$N(x_i^0,t)$
N(x,t)
$N(x_i^L,t)$
Solution direction in travel time based models
t
t
t+Δt
t + link travel time
t + link travel time + D_{int}

Figure 6.2: Solution direction for flow-based versus travel-time based models

Travel-time based solution schemes determine consecutive points in horizontal direction (cf. Figure 6.2). They do not guarantee link FIFO behavior nor do they take into account explicit capacity constraints or queue spillbacks.

Contrarily, the algorithm of the flow-based LTM does not determine points in the future. The algorithm determines cumulative vehicle numbers, time step by step (vertical solution direction, cf. Figure 6.2). Therefore, it provides the ability to deal with changing traffic conditions, to guarantee FIFO behavior, to impose explicit capacity constraints and to capture the spillback of queues. However, due to this vertical solution direction, travel times can only be derived from cumulative vehicle numbers after vehicles have completed their journey.

Therefore, to include average intersection delays in flow-based models, one should modify the original cumulative vehicle numbers (or traffic flows). Travel times derived from these modified curves should correspond to the sum of the original link travel time and the intended intersection delay.

6.1.4. Point-Queues versus kinematic wave queues

In a first attempt to indirectly include intersection delays in LTM, one could consider the artificial construction of physical queues at the downstream link ends. These queues could have an average density and an average length, such that vehicles would experience the intended average delay on top of their original travel time, by passing through this queue.

However, a physical queue at the downstream link end would induce a backward regime in which traffic information travels backwards. Information from a forward regime located upstream of this queue, would not be able to reach the downstream link end, which is however necessary for the determination of transition flows and intersection delays. Artificially transferring this information would not only be inconvenient, it would simply be intractable once such a queue would cross the upstream link boundary.

For these reasons, LTM uses Point-Queues (P-Q's) to indirectly include intersection delays. P-Q's are concentrated in one point of space, in this case the downstream link end. That way, information from a forward regime reaches the downstream link boundary as would happen in the original LTM. A disadvantage of using P-Q's is the neglect of the spatial extent of the temporary queues that would occur in reality. However, this effect is considered to be only of secondary importance.

LTM features two types of queues: kinematic wave queues are queues due to over-saturation, which are typically characterized by a low-frequent process of queue-formation and –dissipation. They are represented by physical queues with a certain spatial extent.

Temporary queues due to stage alternations or priority rules, typically characterized by a high-frequent process of queue-formation and –dissipation, are modeled as Point-Queues (P-Q's), thereby ignoring their spatial extent.

LTM thus realizes intersection delays in an indirect way by holding up vehicles in P-Q's at the downstream link end. The next sections describe a way to do this, so that times spent in P-Q's correspond to intersection delays. The length of the P-Q should be such that vehicles experience the intended average delay.

Since different intersection delays are incurred for different turning movements, P-Q's are associated with turning movements rather than with entire links. Each turning movement has its own P-Q.

The entrance and the exit of a P-Q coincide in horizontal space, but for notational convenience, they are respectively indicated by space-coordinates $x^0_{P-Q,ij}$ and $x^L_{P-Q,ij}$. Figure 6.3 indicates the space-coordinates of the different links and P-Q's.

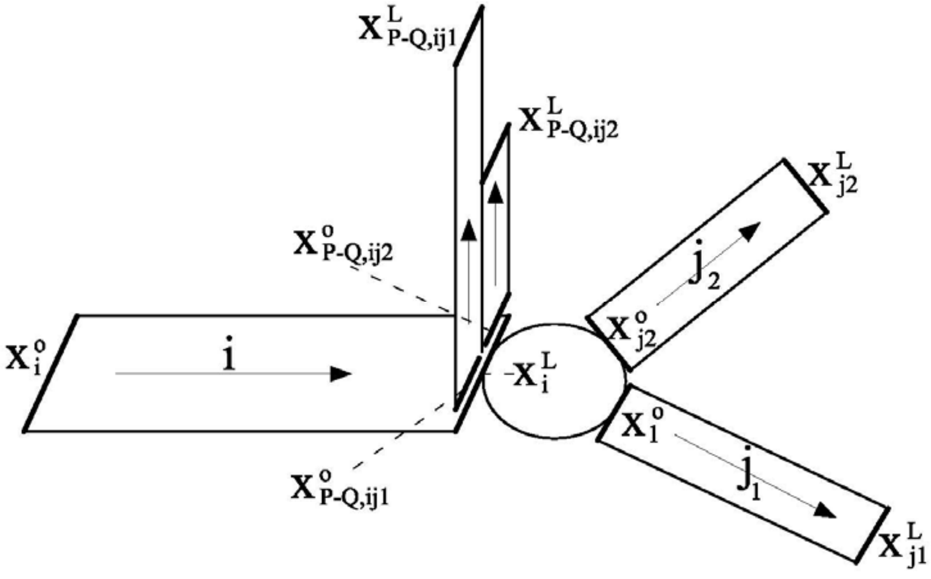


Figure 6.3: Space-coordinates of links and P-Q's

6.1.5. Point-Queues: definitions and characteristics

P-Q's are defined by the cumulative vehicle functions at the P-Q entrance ($N(x^0_{P-Q,ij},t) = \delta_{jp} \sum_{p\in P} N^p(x_i^L,t)$) and P-Q exit ($N(x^L_{P-Q,ij},t)$) (cf. Figure 6.4). These cumulative functions unambiguously determine all P-Q properties.

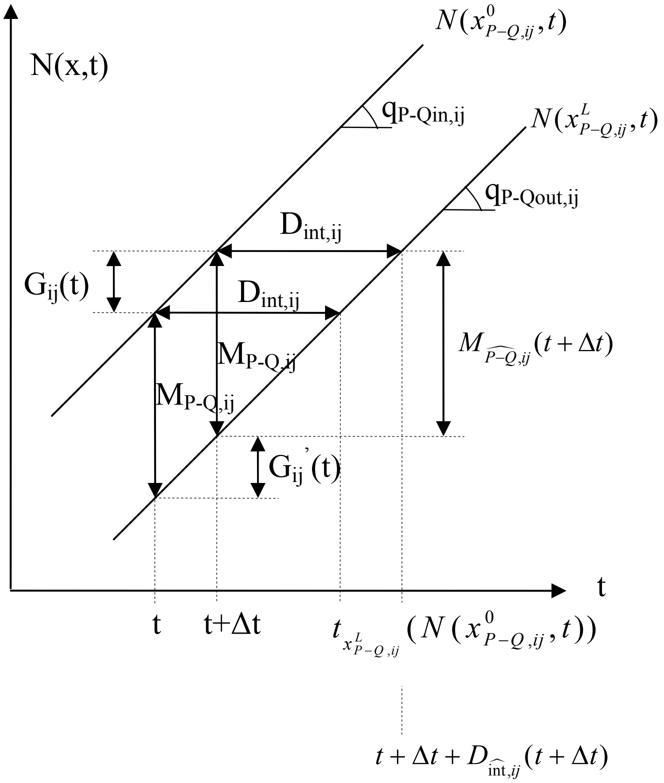

N(x,t)
$N(x^0_{P-Q,ij},t)$
$q_{P\text{-}Qin,ij}$
$N(x^L_{P-Q,ij},t)$
$q_{P\text{-}Qout,ij}$
$D_{int,ij}$
$G_{ij}(t)$
$D_{int,ij}$
$M_{P\text{-}Q,ij}$
$M_{P\text{-}Q,ij}$
$M_{\widehat{P-Q},ij}(t+\Delta t)$
$G_{ij}'(t)$
t
t
t+Δt
$t_{x^L_{P-Q,ij}}(N(x^0_{P-Q,ij},t))$
$t+\Delta t+D_{\widehat{\text{int}},ij}(t+\Delta t)$

Figure 6.4: Cumulative vehicle numbers in a P-Q

P-Q in- and outflow rates ($q_{P\text{-}Qin}$ and $q_{P\text{-}Qout}$) are defined as the derivatives of the cumulative vehicle function *N(x,t)* with respect to time:

$$q_{P-Qin,ij}(t) = q(x^0_{P-Q,ij},t) = \frac{\partial N(x^0_{P-Q,ij},t)}{\partial t} = \frac{N(x^0_{P-Q,ij},t+\Delta t) - N(x^0_{P-Q,ij},t)}{\Delta t} \quad (6.1)$$

$$q_{P-Qout,ij}(t) = q(x^L_{P-Q,ij},t) = \frac{\partial N(x^L_{P-Q,ij},t)}{\partial t} = \frac{N(x^L_{P-Q,ij},t+\Delta t) - N(x^L_{P-Q,ij},t)}{\Delta t} \quad (6.2)$$

Note that flow rates *q(t)* represent flows during time interval *[t , t+Δt]* and that flow rates are constant (homogeneous) during *[t , t+Δt]*.

The number of vehicles in the queue at time *t* (further referred to as $M_{P-Q,ij}(t)$) is the vertical distance between the curves $N(x^0_{P-Q,ij},t)$ and $N(x^L_{P-Q,ij},t)$:

$$M_{P-Q,ij}(t) = N(x^0_{P-Q,ij},t) - N(x^L_{P-Q,ij},t) \quad (6.3)$$

This number of vehicles in the queue evolves according to

$$\frac{M_{P-Q,ij}(t+\Delta t)-M_{P-Q,ij}(t)}{\Delta t}=q(x^0_{P-Q,ij},t)-q(x^L_{P-Q,ij},t) \tag{6.4}$$

implying that the conservation law is satisfied in P-Q's.

Times spent in P-Q's are represented by the horizontal distance between the curves $N(x^0_{P-Q,ij},t)$ and $N(x^L_{P-Q,ij},t)$. These times in the P-Q's correspond to intersection delays:

$$D_{\text{int},ij}(t)=t_{x^L_{P-Q,ij}}(N(x^0_{P-Q,ij},t))-t \tag{6.5}$$

where $t_{x^L_{P-Q,ij}}(N(x^0_{P-Q,ij},t))$ represents the point in time where the vehicle with serial number $N(x^0_{P-Q,ij},t)$ reaches P-Q exit $x^L_{P-Q,ij}$ (see Figure 6.4).

Transition flow $G_{ij}(t)$ is defined as the amount of vehicles entering the P-Q of turning movement *ij* during time interval $[t\,,\, t+\Delta t]$:

$$G_{ij}(t)=q_{P-Qin,ij}(t)\Delta t=N(x^0_{P-Q,ij},t+\Delta t)-N(x^0_{P-Q,ij},t) \tag{6.6}$$

The P-Q Sending flow $S_{ij}^{'}(t)$ from incoming link $i \in I_n$ to outgoing link $j \in J_n$ at time t is defined as the maximum number of vehicles that could leave the P-Q of this turning movement *ij* during time interval $[t\,,\, t+\Delta t]$, if the end of this P-Q would be connected to a traffic reservoir with an infinite capacity.

The P-Q transition flow $G_{ij}^{'}(t)$ is defined as the number of vehicles that are actually transferred from the P-Q of turning movement *ij* to link *j* during time interval $[t\,,\, t+\Delta t]$:

$$G^{'}_{ij}(t)=q_{P-Qout,ij}(t)\Delta t=N(x^L_{P-Q,ij},t+\Delta t)-N(x^L_{P-Q,ij},t) \tag{6.7}$$

Note that (P-Q) Sending flows and (P-Q) transition flows actually represent flow increments. Actual flows are in this thesis referred to as flow rates.

6.2. Extended LTM solution algorithm including intersection delays

To realize intersection delays, the extended LTM solution algorithm includes the construction of P-Q's. The extended algorithm builds on the original LTM solution algorithm (cf. Section 3.2), which is used to determine cumulative vehicle numbers at the downstream link boundary (x_i^L), i.e. at the P-Q entrances ($x^0_{P-Q,ij}$) of each incoming link *i* (first three steps of the extended algorithm).

The last three steps of the extended algorithm determine cumulative vehicle numbers at the P-Q exits ($x^L_{P-Q,ij}$) of the turning movements of each incoming link *i* and at the upstream link boundary (x_j^0) of each outgoing link *j*.
The extended LTM solution algorithm determines cumulative vehicle numbers, time step by step (vertical solution direction, cf. Figure 6.2). Only afterwards, when vehicles have completed their journey, link travel times and intersection delays can be derived from these cumulative numbers.
For each time step, the last three steps of the extended LTM solution algorithm actually determine P-Q lengths. P-Q lengths are determined such that vehicles would experience their intended average intersection delays.

Extended LTM solution algorithm

For each time interval Δt,
For each node *n*,

Step 1: For each incoming link $i \in I_n$, determine the Sending flow S_i at the downstream link end (x_i^L), and for each outgoing link $j \in J_n$, determine the Receiving flow R_j at the upstream link end (x_j^0).

Step 2: Determine the transition flows $G_{ij}(t)$ from incoming links $i \in I_n$ to outgoing links $j \in J_n$ by applying the appropriate node model (cf. Chapter 5).

Step 3: For the downstream link boundary (x_i^L) of each incoming link $i \in I_n$, and for the P-Q entrances ($x^0_{P-Q,ij}$) of the turning movements of each incoming link *i*, update the cumulative vehicle numbers $N(x,t)$:

$$N(x_i^L, t+\Delta t) = N(x_i^L, t) + \sum_{j \in J_n} G_{ij}(t) \qquad \text{for all } i \in I_n$$

$$N(x^0_{P-Q,ij}, t+\Delta t) = N(x^0_{P-Q,ij}, t) + G_{ij}(t) \qquad \text{for all } i \in I_n,\ j \in J_n$$

The disaggregation of cumulative vehicle numbers is adopted from the upstream link boundary:

$$N^p(x_i^L, t+\Delta t) = N^p(x_i^0,\ t_{x_i^0}(N(x_i^L, t+\Delta t)) \qquad \text{for all } i \in I_n\ ,\ p \in P$$

$$N^p(x^0_{P-Q,ij}, t+\Delta t) = \sum_{p \in P} \delta_{jp} N^p(x_i^L, t+\Delta t) \qquad \text{for all } i \in I_n\ ,\ j \in J_n\,,\ p \in P$$

Note that the cumulative vehicle numbers at the P-Q exits and at the upstream link boundaries are not yet updated at this point.

Step 4: For each turning movement *ij* of each incoming link $i \in I_n$, determine the P-Q Sending flow $S_{ij}'(t)$. This step is composed out of three sub-steps:

Step 4.1: Determine the expected average intersection delay at time $t + \Delta t$, which is further referred to as the target value of intersection delay, $D_{\widehat{\text{int},ij}}(t+\Delta t)$ (cf. Section 6.2.1).

Step 4.2: Determine the target number of vehicles in the P-Q at time $t + \Delta t$, $M_{\widehat{P-Q},ij}(t+\Delta t)$ (cf. Section 6.2.2).

Step 4.3: Determine the P-Q Sending flow $S_{ij}'(t)$ (cf. Section 6.2.3).

Step 5: Determine the P-Q transition flows $G_{ij}'(t)$ from the P-Q's of turning movements ij ($i \in I_n$, $j \in J_n$) to the outgoing links $j \in J_n$ (cf. Section 6.2.4).

Step 6: Update the cumulative vehicle numbers at the P-Q exits ($x^L_{P-Q,ij}$) of the turning movements ij of each incoming link $i \in I_n$ and at the upstream link boundary (x_j^0) of each outgoing link $j \in J_n$:

$$N(x^L_{P-Q,ij}, t+\Delta t) = N(x^L_{P-Q,ij}, t) + G'_{ij}(t) \qquad \textit{for all } i \in I_n, j \in J_n$$

$$N(x^0_j, t+\Delta t) = N(x^0_j, t) + \sum_{i \in I_n} G'_{ij}(t) \qquad \textit{for all } j \in J_n$$

The disaggregation of cumulative vehicle numbers at the P-Q exits and at the upstream link boundaries is adopted from the P-Q entrances:

$$N^p(x^L_{P-Q,ij}, t+\Delta t) = N^p(x^0_{P-Q,ij}, t_{x^0_{P-Q,ij}}(N(x^L_{P-Q,ij}, t+\Delta t))) \quad \textit{for all } i \in I_n, j \in J_n, p \in P$$

$$N^p(x^0_j, t+\Delta t) = \delta_{jp} N^p(x^L_{P-Q,ij}, t+\Delta t) \qquad \textit{for all } i \in I_n, j \in J_n, p \in P$$

Steps 4 and 5 are now elaborated in detail.

6.2.1. Target value of intersection delay (algorithm step 4.1)

In the first step of the extended part of the LTM solution algorithm (i.e. in step 4.1), expected average intersection delays are determined based on prevailing traffic conditions.

To estimate average intersection delays, many formulae based on queuing theory have been proposed in the last 5 decades (Troutbeck and Brilon (2000); Rouphail et al. (2000)). Some of the most widely accepted ones are formulated below.

The target value of intersection delay at time $t + \Delta t$ for turning movement ij, $D_{\widehat{\text{int},ij}}(t+\Delta t)$ is determined by one of these formulas, where the degree of saturation $x_{ij} = S_{ij}(t)/q_{n,ij}(t)$.

6.2.1.1. Deterministic intersection delay for signalized intersections

The first, widely used deterministic intersection delay formula for signalized intersections was developed by Webster (1958) from a combination of theoretical and numerical simulation approaches:

$$D_{\text{int},\text{det},ij} = \frac{c(1 - {}^{g_{ij}}\!/_{c})^2}{2[1 - ({}^{g_{ij}}\!/_{c})\min(1, x_{ij})]} \tag{6.8}$$

where $D_{int,det,ij}$ = average deterministic intersection delay per vehicle for turning movement ij (s)

x_{ij} = degree of saturation for turning movement ij (flow to capacity ratio).

Formula (6.8) represents delay when traffic can be considered arriving at a uniform rate. It is referred to as the deterministic component of intersection delay.
Figure 6.5 depicts this intersection delay function, which is a monotonically increasing function of traffic flow:

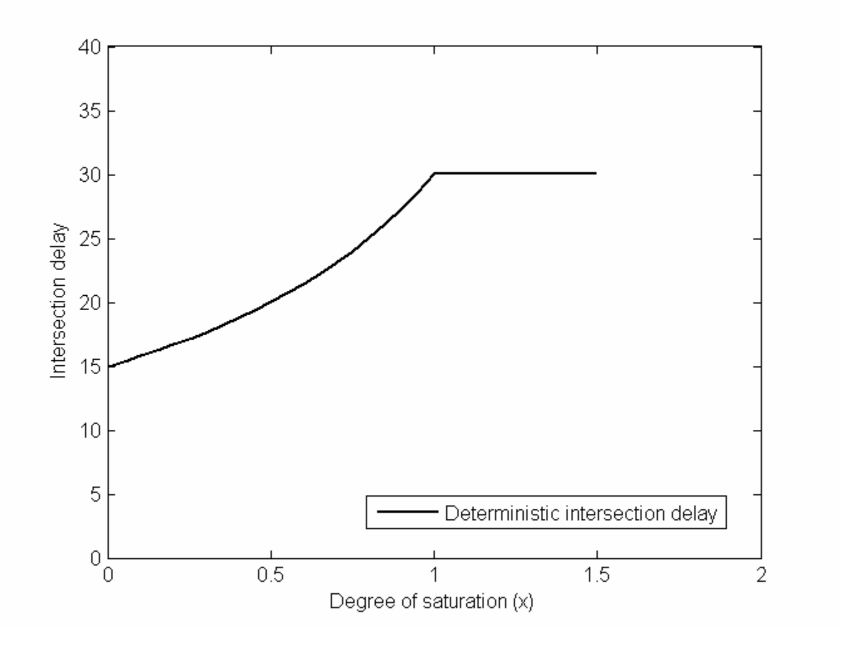


Figure 6.5: Deterministic component of intersection delay

Deterministic intersection delay Formula (6.8) is independent on the time of operation and on the history of the system. The formula is valid both for (over-) saturated flows and for non-saturated flows that are sustained over a period of at least one cycle length. Formula (6.8) is used for applications where only the deterministic component of intersection delay is taken into account.

6.2.1.2. Stochastic intersection delay for signalized intersections

Webster's (1958) stochastic intersection delay formula makes some allowance for the random nature of the arrivals. It is known as the "random delay", assuming a Poisson arrival process (indices *ij* are dropped for simplicity of notation):

$$D_{\text{int},stoch,steady-state} = \frac{x^2}{2q_{arr}(1-x)} \qquad for\ x < 1 \tag{6.9}$$

where $D_{int,stoch,steady\text{-}state}$ = average stochastic steady-state intersection delay per vehicle (s)

q_{arr} = arrival rate (veh/s)

x = degree of saturation for the prevailing turning movement (flow to capacity ratio).

This stochastic component of intersection delay is only valid for non-saturated flows ($x < 1$) in steady-state conditions, assuming an infinite time period of stable traffic conditions to be achieved. This steady-state formula is only an acceptable approximation of the real-world process for low flow to capacity ratios, where equilibrium is reached in a reasonable period of time. When traffic flow approaches signal capacity, however, the time to reach equilibrium usually exceeds the period over which demand is sustained (Rouphail et al. (2000)).

To circumvent the limiting assumptions of steady-state conditions, many authors proposed time-dependent delay formulae, where the predicted average delay is dependent on the time of operation. Time-dependent delay formulae are valid for all flow to capacity ratios, including over-saturated flows ($x \geq 1$) in congested traffic states.

A widely accepted time-dependent delay formula is given by the Highway Capacity Manual (TRB (2000)):

$$D_{time-dep} = 900T\left[(x-1) + \sqrt{(x-1)^2 + \frac{4x}{q_n T}}\right] \tag{6.10}$$

where $D_{time\text{-}dep}$ = average time-dependent delay per vehicle (s)

T = time of operation over which demand and capacity are sustained (h)

q_n = capacity of the prevailing turning movement (veh/h)

x = degree of saturation for the prevailing turning movement (flow to capacity ratio).

Formula (6.10) contains two components: the stochastic intersection delay and the deterministic delay due to over-saturation. The latter is not an intersection delay; it is already taken into account by LTM. The average deterministic delay per vehicle due to over-saturation equals half of the time that would be needed to clear the queue that has built up during period *T* due to over-saturation:

$$D_{\det,over-sat} = \frac{1}{2}\max(0,\frac{q_{arr}-q_n}{q_n})3600T = 1800T(\max(1,x)-1) \quad (6.11)$$

where $D_{det,over\text{-}sat}$ = average deterministic delay due to over-saturation per vehicle (s).

It needs to be subtracted from total delay (6.10) to be left with the pure stochastic time-dependent intersection delay:

$$D_{\mathrm{int},stoch,time-dep} = 900T\left[(x-1)+\sqrt{(x-1)^2+\frac{4x}{q_nT}}\right]-1800T(\max(1,x)-1) \quad (6.12)$$

where $D_{int,stoch,time\text{-}dep}$ = average stochastic time-dependent intersection delay per vehicle (s).

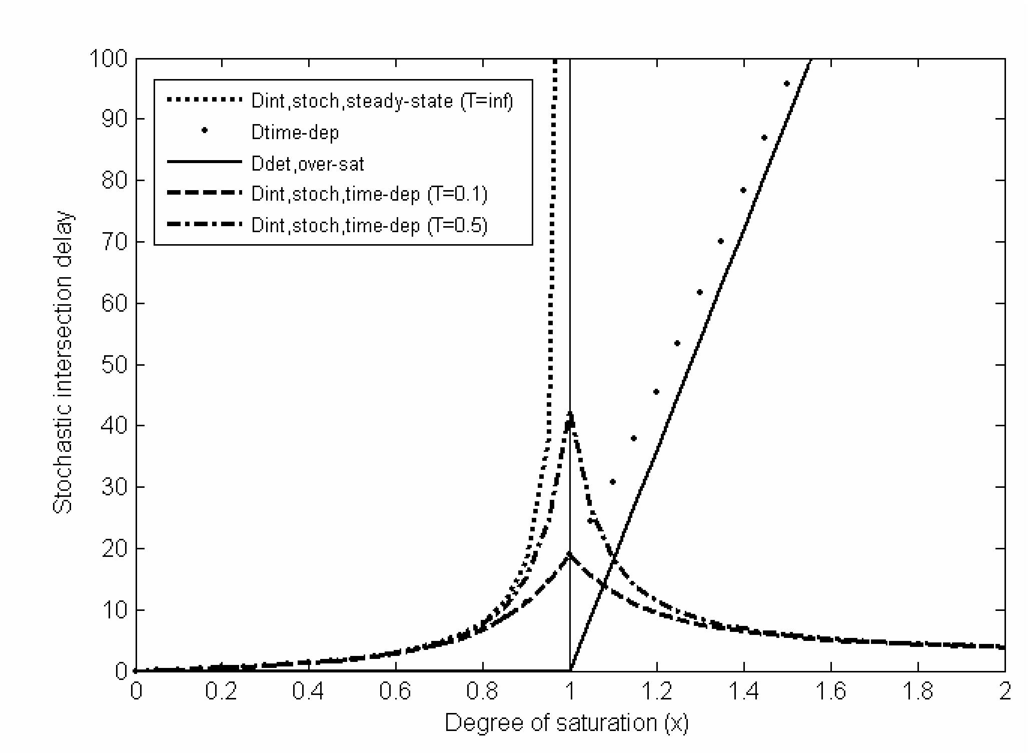


Figure 6.6: Steady-state versus time-dependent stochastic intersection delay

Figure 6.6 compares steady-state Formula (6.9) with time-dependent Formula (6.12). Intersection delay $D_{int,stoch}$ is plotted as a function of the degree of saturation x. Subtraction of deterministic delay due to over-saturation, $D_{det,over\text{-}sat}$, from time-dependent delay $D_{time\text{-}dep}$ yields the pure stochastic time-dependent intersection delay $D_{int,stoch,time\text{-}dep}$. For higher degrees of over-saturation, the stochastic time-dependent intersection delay decreases since arrivals at the intersection become more uniform (queue discharge rate). Higher times of operation yield higher stochastic time-dependent intersection delays (see Figure 6.6: $T = 0.1$ *(h)* vs. $T = 0.5$ *(h)*).

6.2.1.3. Total time-dependent intersection delay for signalized intersections

Total intersection delay for signalized intersections consists of a deterministic and a stochastic component to reflect both the fluid and random properties of traffic flow. Furthermore, initial intersection delays should be considered in combination with time-dependent Formula (6.12). When a residual queue (due to intersection delay) from a previous time period causes an initial queue to occur at the start of the analysis period (*T*), additional delay is experienced by vehicles arriving in this period since the initial queue must first clear the intersection. Formula (6.13) represents the initial intersection delay:

$$D_{\mathrm{int},initial} = \frac{1800 Q_b (1+u) t}{q_n T} \tag{6.13}$$

where $D_{int,initial}$ = average initial intersection delay per vehicle (s)
Q_b = initial queue (due to intersection delay) at the start of period T (veh)
t = duration of unmet demand in T (h)
u = delay parameter (-).

For a precise determination of this delay component, we refer to the HCM (TRB (2000) pp. 142-151). The total intersection delay for signalized intersections is the sum of deterministic intersection delay, stochastic intersection delay and initial intersection delay. It is dependent on the history of the system and on the time of operation:

$$\begin{aligned} D_{\mathrm{int}} = {} & \frac{c(1-{}^{g}\!/_{c})^2}{2[1-({}^{g}\!/_{c})\min(1,x)]} + 900T\left[(x-1)+\sqrt{(x-1)^2+\frac{4x}{q_n T}}\right] \\ & -1800T(\max(1,x)-1) + \frac{1800 Q_b (1+u) t}{q_n T} \end{aligned} \tag{6.14}$$

where D_{int} = average total time-dependent intersection delay per vehicle (s).

6.2.1.4. Total time-dependent intersection delay for unsignalized intersections

For unsignalized intersections, existing delay formulae do not distinguish between deterministic and stochastic delay components. A widely used average delay formula that accounts for the history of the system and the time of operation, has been given by Akcelik (1991):

$$D_{time-dep} = \frac{1}{\mathrm{q_n}} + \frac{1800 Q_b}{\mathrm{q_n}} + 900T\left[\left((x-1)+\frac{7200 Q_b}{\mathrm{q_n} T}\right)^2 + \frac{8x}{\mathrm{q_n} T}\right] \tag{6.15}$$

Subtraction of deterministic delay due to over-saturation (6.11) yields the unsignalized counterpart of Formula (6.14):

$$D_{\text{int}} = \frac{1}{\text{q}_{\text{n}}} + \frac{1800Q_b}{\text{q}_{\text{n}}} + 900T\left[\left((x-1) + \frac{7200Q_b}{\text{q}_{\text{n}}T}\right)^2 + \frac{8x}{\text{q}_{\text{n}}T}\right] - 1800T(\max(1,x)-1) \quad (6.16)$$

6.2.2. Target number of vehicles in the P-Q (algorithm step 4.2)

In step 4.2 of the extended LTM solution algorithm, the target value of intersection delay is converted into a target number of vehicles in the P-Q. $M_{\widehat{P-Q},ij}(t+\Delta t)$, the target number of vehicles in the P-Q of turn *ij* at time *t+Δt*, is determined such that the target value of intersection delay $D_{\widehat{\text{int}},ij}(t+\Delta t)$ would precisely be experienced passing through this P-Q at constant flow rate $q_{P-Qin,ij}(t)$:

$$M_{\widehat{P-Q},ij}(t+\Delta t) = D_{\widehat{\text{int}},ij}(t+\Delta t)\, q_{P-Qin,ij}(t) \quad (6.17)$$

6.2.3. P-Q Sending flow $S'_{ij}(t)$ (algorithm step 4.3)

P-Q Sending flow $S'_{ij}(t)$ is determined in step 4.3 of the extended LTM solution algorithm.
We first determine P-Q Sending flows $S_{ij}'(t)$ when only considering deterministic intersection delays at signalized intersections.

Deterministic intersection delays at signalized intersections have the following characteristics:

- They are monotonically increasing as a function of traffic flow (cf. Figure 6.5)
- They are independent on the time of operation and on the history of the system

At the end of this section, we shortly discuss possible adjustments to deal with time-dependent intersection delays at signalized and unsignalized intersections. These delays are dependent on the time of operation and on the history of the system, and they are not necessarily monotonically increasing as a function of traffic flow. A detailed solution method description for these types of intersection delays is however out of the scope of this thesis.

6.2.3.1. Solution method for deterministic intersection delays at signalized intersections

The P-Q Sending flow $S_{ij}{}'(t)$ from incoming link $i \in I_n$ to outgoing link $j \in J_n$ at time t is defined as the maximum number of vehicles that could leave the P-Q of this turning movement *ij* during time interval $[t, t+\Delta t]$, if the end of this P-Q would be connected to a traffic reservoir with an infinite capacity.

Given the P-Q length at time t, $M_{\widehat{P-Q},ij}(t)$, and given the P-Q inflow rate during time interval $[t, t+\Delta t]$, $q_{P\text{-}Qin,ij}(t)$, the P-Q Sending flow $S_{ij}{}'(t)$ to realize this target number of vehicles in the P-Q at time $t+\Delta t$ equals:

$$S_{ij}^{'}(t) = N(x_{P-Q,ij}^{0}, t+\Delta t) - M_{\widehat{P-Q},ij}(t+\Delta t) - N(x_{P-Q,ij}^{L}, t) \tag{6.18}$$

The P-Q Sending flow is restricted by non-negativity and capacity constraints:

$$0 \le S_{ij}^{'}(t) \le q_{M,ij}\Delta t \tag{6.19}$$

Furthermore, we impose some additional constraints on the P-Q Sending flow to make sure that the intended intersection delay will actually be reached. Note that in order for $D_{\widehat{\text{int}},ij}(t+\Delta t)$ to actually be reached, the P-Q outflow $G_{ij}{}'$ during time interval $[t+\Delta t, t+\Delta t + D_{\widehat{\text{int}},ij}(t+\Delta t)]$ should equal target number $M_{\widehat{P-Q},ij}(t+\Delta t)$ (cf. Figure 6.4), i.e., $N(x_{P-Q,ij}^{L}, t+\Delta t + D_{\widehat{\text{int}},ij}(t+\Delta t)) - N(x_{P-Q,ij}^{L}, t+\Delta t) = M_{\widehat{P-Q},ij}(t+\Delta t)$ or, using (6.17):

$$\begin{aligned} q_{P-Qout,ij}(t+\Delta t) &= q_{P-Qout,ij}(t+2\Delta t) = \ldots \\ \ldots &= q_{P-Qout,ij}(t+\Delta t + D_{\widehat{\text{int}},ij}(t+\Delta t)) = q_{P-Qin,ij}(t) \end{aligned} \tag{6.20}$$

Condition (6.20) expresses that the P-Q outflow rate during time interval $[t+\Delta t, t+\Delta t + D_{\widehat{\text{int}},ij}(t+\Delta t)]$ should equal P-Q inflow rate $q_{P\text{-}Qin,ij}(t)$ of the previous time interval $[t, t+\Delta t]$, that was used to determine the target value of intersection delay. If Condition (6.20) is fulfilled, the intended target value of intersection delay $D_{\widehat{\text{int}},ij}(t+\Delta t)$ will exactly be realized.

Next to constraints (6.19), we impose some additional constraints on the P-Q Sending flow in an attempt to satisfy Condition (6.20) such that the target value of intersection delay $D_{\widehat{\text{int}},ij}(t+\Delta t)$ would actually be reached. Mathematically, the solution method for deterministic intersection delays at signalized intersections may be summarized as follows:

Extended LTM solution algorithm step 4.3: Solution method for deterministic intersection delays at signalized intersections

If $M_{\widehat{P-Q},ij}(t+\Delta t) < M_{P-Q,ij}(t)$ and if $G_{ij}(t) < G'_{ij}(t-\Delta t)$, then

$$S'_{ij}(t) = \max\left\{0\ ,\ \min\begin{bmatrix} q_{M,ij}\Delta t\ ,\ G'_{ij}(t-\Delta t)\ , \\ (N(x^0_{P-Q,ij}, t+\Delta t) - M_{\widehat{P-Q},ij}(t+\Delta t) - N(x^L_{P-Q,ij}, t)) \end{bmatrix}\right\}$$

If $M_{\widehat{P-Q},ij}(t+\Delta t) > M_{P-Q,ij}(t)$ and if $G_{ij}(t) > G'_{ij}(t-\Delta t)$, then

$$S'_{ij}(t) = \min\left\{q_{M,ij}\Delta t\ ,\ \max\begin{bmatrix} G'_{ij}(t-\Delta t)\ , \\ (N(x^0_{P-Q,ij}, t+\Delta t) - M_{\widehat{P-Q},ij}(t+\Delta t) - N(x^L_{P-Q,ij}, t)) \end{bmatrix}\right\}$$

Else $S'_{ij}(t) = \max\left\{0\ ,\min\begin{bmatrix} q_{M,ij}\Delta t\ , \\ (N(x^0_{P-Q,ij}, t+\Delta t) - M_{\widehat{P-Q},ij}(t+\Delta t) - N(x^L_{P-Q,ij}, t)) \end{bmatrix}\right\}$

The remainder of this section discusses this solution method and it explains how this solution method originates from constraints (6.19) and from the additional constraints that are imposed on the P-Q Sending flow in an attempt to satisfy Condition (6.20). Furthermore, we claim that with this solution method, the target value of intersection delay $D_{\widehat{\text{int}},ij}(t+\Delta t)$ will exactly be reached (or Condition (6.20) will be satisfied) if all of the four following conditions (further referred to as Conditions (6.21)) are fulfilled:

1) before time *t*, P-Q outflow rate $q_{P\text{-}Qout,ij}(t\text{-}\Delta t)$ equaled P-Q inflow rate $q_{P\text{-}Qin,ij}(t\text{-}\Delta t)$
2) at time *t*, the realized number of vehicles in the P-Q, $M_{P-Q,ij}(t)$, equals target number $M_{\widehat{P-Q},ij}(t)$
3) Assuming that the P-Q inflow rate changes at time *t*, the P-Q inflow rate does not change again after time *t*, until the new target number of vehicles in the P-Q is reached
4) P-Q outflow rate $q_{P\text{-}Qout,ij}$ is not influenced by downstream boundary conditions during $[t\ ,\ t + D_{\widehat{\text{int}},ij}(t+\Delta t)\ ]$

These four conditions (6.21) typically occur for low-frequently changing traffic flows that are not restricted by downstream boundary conditions. Intended intersection delays are exactly realized and a linear transition occurs between different states with different values of intersection delay. Illustrations are given in Examples 6.1 and 6.4 (see below). In other circumstances, which would be typical for high-frequently changing traffic flows or for traffic flows that are restricted by downstream boundary conditions, realized intersection delays might for some vehicles slightly differ from the intended delays, as respectively shown in Examples 6.2 and 6.3 (see below).

The following three cases are distinguished:

> a) The target number of vehicles in the P-Q at time $t+\Delta t$ is smaller than the realized number of vehicles in the P-Q at time t:

$$M_{\widehat{P-Q},ij}(t+\Delta t) < M_{P-Q,ij}(t) \tag{6.22}$$

This case further breaks up into two sub-cases:

> a.1) The transition flow during time interval $[t\,,\ t+\Delta t]$ (as determined in step 2 of the extended LTM solution algorithm) is smaller than the P-Q transition flow from the previous time update, i.e. during time interval $[t-\Delta t\,,\ t]$ (as determined in step 5 of the extended LTM solution algorithm) (cf. Figure 6.7):

$$G_{ij}(t) < G_{ij}^{'}(t-\Delta t) \tag{6.23}$$

In this case, the following additional constraint is imposed on the P-Q Sending flow:

$$S_{ij}^{'}(t) \leq G_{ij}^{'}(t-\Delta t) \tag{6.24}$$

P-Q Sending flow $S_{ij}'(t)$ is restricted from being higher than the P-Q transition flow $G_{ij}'(t-\Delta t)$ in the previous time interval. The P-Q Sending flow is formulated as:

$$S_{ij}^{'}(t) = \max\left\{0\ ,\min\left[\begin{array}{l} q_{M,ij}\Delta t\ ,\ G_{ij}^{'}(t-\Delta t)\ , \\ (N(x_{P-Q,ij}^{0},t+\Delta t) - M_{\widehat{P-Q},ij}(t+\Delta t) - N(x_{P-Q,ij}^{L},t)) \end{array}\right]\right\} \tag{6.25}$$

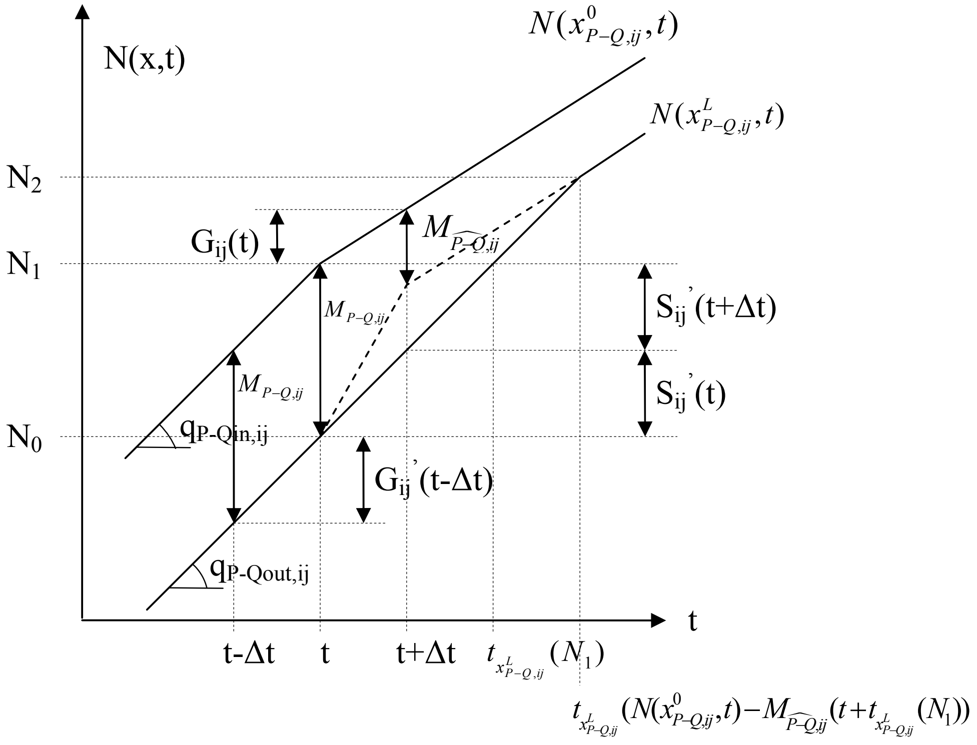
It is now explored how well this P-Q Sending flow satisfies the intended Condition (6.20). Therefore, we distinguish two different scenarios in which situation (6.22) - (6.23) can occur:

> a.1.1) if before time t, P-Q outflow rate $q_{P-Qout,ij}(t-\Delta t)$ equaled P-Q inflow rate $q_{P-Qin,ij}(t-\Delta t)$, and if at time t, the realized number of vehicles in the P-Q $M_{P-Q,ij}(t)$ equals target number $M_{\widehat{P-Q},ij}(t)$, then (6.22) implies (6.23) and vice versa.

Since intersection delays are monotonically increasing functions of traffic flow, situation (6.22) must in this case have been caused by a decrease in P-Q inflow rate after time t

$(q_{P-Qin,ij}(t) \le q_{P-Qin,ij}(t-\Delta t))$: a smaller target value of vehicles in the P-Q must have been induced by a smaller P-Q inflow rate and correspondingly a smaller target value of intersection delay (cf. Formula (6.17)).

Figure 6.7 depicts an example of such a situation:



$N(x_{P-Q,ij}^{0},t)$
N(x,t)
$N(x_{P-Q,ij}^{L},t)$
N_2
$G_{ij}(t)$
$M_{\widehat{P-Q},ij}$
N_1
$M_{P-Q,ij}$
$S_{ij}'(t+\Delta t)$
$M_{P-Q,ij}$
$S_{ij}'(t)$
N_0
$q_{P-Qin,ij}$
$G_{ij}'(t-\Delta t)$
$q_{P-Qout,ij}$
t
t-Δt
t
t+Δt
$t_{x_{P-Q,ij}^{L}}(N_1)$
$t_{x_{P-Q,ij}^{L}}(N(x_{P-Q,ij}^{0},t)-M_{\widehat{P-Q},ij}(t+t_{x_{P-Q,ij}^{L}}(N_1))$

Figure 6.7: P-Q Sending flows in Example 6.1

Example 6.1

At time *t*, steady state [*t-Δt* , *t*] breaks off due to a decreased P-Q inflow rate $q_{P-Qin,ij}(t)$. A transition state appears, where the P-Q outflow needs to be higher than the P-Q inflow to decrease the number of vehicles $M_{P-Q,ij}(t)$ in the P-Q. Unconstrained P-Q Sending flow (6.18) is indicated by the dotted line in Figure 6.7.

The P-Q Sending flow is however restricted by (6.24). By imposing this constraint, the causality principle is satisfied.

Causality principle

> The causality principle states that vehicles should not be influenced by traffic conditions behind them, but only by stimuli ahead of them. P-Q travel times at a given time should depend on traffic entering the P-Q at that time and earlier, but not on traffic entering later. In other words, the P-Q outflow rate should not be influenced by a changing P-Q inflow rate before the last vehicle N_1 (see Figure 6.7) that entered the P-Q at the original inflow rate has left the P-Q, i.e. before time $t_{x^L_{P-Q,ij}}(N_1)$.

If P-Q outflow rate $q_{P\text{-}Qout,ij}(t)$ would follow the dotted line in Figure 6.7, vehicles N_0 to N_1 would first experience a decreasing delay (note that the delay is represented by the horizontal distance between the two cumulative curves in Figure 6.7) followed by an increasing delay while the P-Q inflow rate stays constant. This would be an undesirable effect. Due to constraining the P-Q Sending flow, the intersection delay stays constant until vehicle N_1 has left the P-Q, i.e., until time $t_{x^L_{P-Q,ij}}(N_1)$. The causality principle is satisfied.

Since P-Q in- and outflow rates equaled each other during [t-Δt , t] ($q_{P-Qout,ij}(t-\Delta t) = q_{P-Qin,ij}(t-\Delta t)$), and since the P-Q outflow rate is not influenced by downstream boundary conditions during [t , t+$t_{x^L_{P-Q,ij}}(N_1)$] = [t , t+$D_{\widehat{\text{int}},ij}(t)$], the causality principle ensures that the target value of intersection delay $D_{\widehat{\text{int}},ij}(t)$ is actually reached. Condition (6.20) is in this case fulfilled.

This point in time t+$t_{x^L_{P-Q,ij}}(N_1)$ completes the first part [t , t+$t_{x^L_{P-Q,ij}}(N_1)$] of the transition state. The number of vehicles in the P-Q has decreased towards number $M_{P-Q,ij}(t+t_{x^L_{P-Q,ij}}(N_1))$. By passing through this P-Q at new flow rate $q_{P\text{-}Qin,ij}$ (t+2$t_{x^L_{P-Q,ij}}(N_1)$), vehicles would experience original intersection delay $D_{\text{int},ij}(t)$:

$$M_{P-Q,ij}(t+t_{x^L_{P-Q,ij}}(N_1)) = D_{\text{int},ij}(t)\, q_{P-Qin,ij}(t+t_{x^L_{P-Q,ij}}(N_1))$$

In other words, the intersection delay does not yet change in the first part of the transition state.

Therefore, the process of decreasing the number of vehicles in the P-Q needs to continue in a second part of the transition state, in order to realize target number $M_{\widehat{P-Q},ij}(t+t_{x^L_{P-Q,ij}}(N_1)) < M_{P-Q,ij}(t+t_{x^L_{P-Q,ij}}(N_1))$.

In the second part of the transition state, a linear transition from $M_{P-Q,ij}(t+t_{x^L_{P-Q,ij}}(N_1))$ to $M_{\widehat{P-Q},ij}(t+t_{x^L_{P-Q,ij}}(N_1))$ takes place by keeping P-Q outflow rate $q_{P\text{-}Qout,ij}$(t+$t_{x^L_{P-Q,ij}}(N_1)$) constant and equal to $q_{P\text{-}Qout,ij}(t)$. Target number $M_{\widehat{P-Q},ij}(t+t_{x^L_{P-Q,ij}}(N_1))$ is in this example reached at time

$t+t_{x^L_{P-Q,ij}}(N(x^0_{P-Q,ij},t)-M_{\widehat{P-Q},ij}(t+t_{x^L_{P-Q,ij}}(N_1))$ (see Figure 4.18).

For case a.1.1, the point in time $t_{x^L_{P-Q,ij}}(N(x^0_{P-Q,ij},t)-M_{\widehat{P-Q},ij}(t+t_{x^L_{P-Q,ij}}(N_1))$ always follows after the point in time $t_{x^L_{P-Q,ij}}(N_1)$ where the first part of the transition state is completed, due to the monotonically increasing character of $D_{\widehat{\text{int}},ij}$ (smaller P-Q inflow rates always yield smaller target values of intersection delay). Therefore, the causality principle is always obeyed in this case a.1.1. It can be concluded that the target value of intersection delay $D_{\widehat{\text{int}},ij}(t+\Delta t)$ will actually be reached (or Condition (6.20) will actually be satisfied) if all of the four conditions (6.21) are fulfilled.

The evolution of intersection delay between two different states with different values of intersection delay, is modeled here as a linear process with constant P-Q outflow rate. It might also be interesting to explore the possibility of modeling an exponential relaxation in the second part of the transition state, since recent research results (Viti, 2004) indicate an exponential nature of this evolution.

a.1.2) Situation (6.22) - (6.23) can also occur when before time *t*, P-Q outflow rate $q_{P\text{-}Qout,ij}(t\text{-}\Delta t)$ differed from P-Q inflow rate $q_{P\text{-}Qin,ij}(t\text{-}\Delta t)$, or when at time *t*, the realized number of vehicles in the P-Q $M_{P-Q,ij}(t)$ differs from target number $M_{\widehat{P-Q},ij}(t)$.

Figure 6.8 shows an example where at time *t*, though P-Q outflow rate $q_{P\text{-}Qout,ij}(t\text{-}\Delta t)$ did not yet equal P-Q inflow rate $q_{P\text{-}Qin,ij}(t\text{-}\Delta t)$, the P-Q inflow rate changes again.

Example 6.2

Although P-Q Sending flow $S_{ij}'(t)$ is restricted by constraint (6.24), the figure shows that the causality principle might be violated in this case.

Time $t_{x^L_{P-Q,ij}}(N(x^0_{P-Q,ij},t)-M_{\widehat{P-Q},ij}(t))$ on which target number $M_{\widehat{P-Q},ij}(t)$ is reached, now precedes the point in time $t_{x^L_{P-Q,ij}}(N_2)$ where the first part of the transition state would normally be completed (see Figure 6.8).

At time $t_{x^L_{P-Q,ij}}(N(x^0_{P-Q,ij},t)-M_{\widehat{P-Q},ij}(t))$, the P-Q outflow rate changes due to a changing P-Q inflow rate $q_{P\text{-}Qin,ij}(t)$, though the last vehicle N_2 that entered the P-Q at inflow rate $q_{P\text{-}Qin,ij}(t\text{-}\Delta t)$ has not yet left the P-Q. The causality principle is violated for vehicles N_1 to N_2, which do not experience their intended intersection delay.

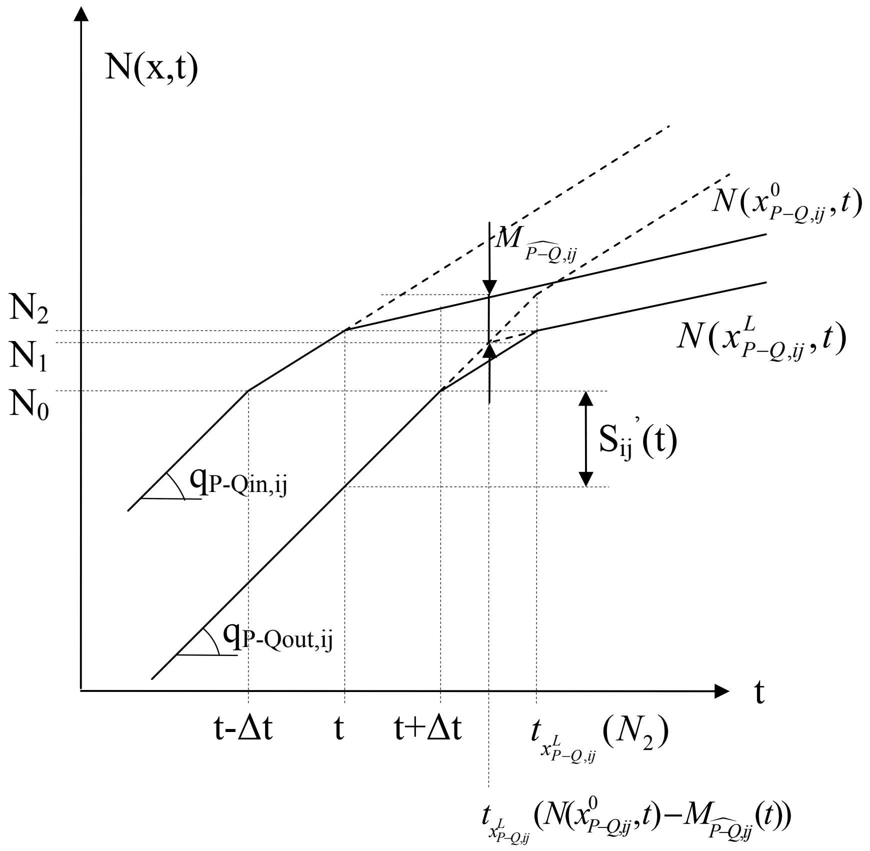

N(x,t)
N_2
N_1
N_0
$M_{\widehat{P-Q},ij}$
$N(x^0_{P-Q,ij},t)$
$N(x^L_{P-Q,ij},t)$
$S_{ij}^{'}(t)$
$q_{P\text{-}Qin,ij}$
$q_{P\text{-}Qout,ij}$
t
t-Δt
t
t+Δt
$t_{x^L_{P-Q,ij}}(N_2)$
$t_{x^L_{P-Q,ij}}(N(x^0_{P-Q,ij},t)-M_{\widehat{P-Q},ij}(t))$

Figure 6.8 P-Q Sending flows in Example 6.2

a.2) The transition flow during time interval [*t , t+Δt*] equals or is higher than the P-Q transition flow during time interval [*t-Δt , t*]:

$$G_{ij}(t) \geq G_{ij}^{'}(t-\Delta t) \tag{6.26}$$

In this case, no additional constraints are imposed on the P-Q Sending flow:

$$S_{ij}^{'}(t) = \max\left\{0\ ,\ \min\begin{bmatrix} q_{M,ij}\Delta t\ , \\ (N(x^0_{P-Q,ij},t+\Delta t) - M_{\widehat{P-Q},ij}(t+\Delta t) - N(x^L_{P-Q,ij},t)) \end{bmatrix}\right\} \tag{6.27}$$

Situation (6.22) - (6.26) exclusively occurs when before time *t*, P-Q outflow rate $q_{P\text{-}Qout,ij}$ *(t-Δt)* differed from P-Q inflow rate $q_{P\text{-}Qin,ij}$*(t-Δt)*, or when at time *t*, the realized number of vehicles in the P-Q $M_{P-Q,ij}(t)$ differs from the target number $M_{\widehat{P-Q},ij}(t)$.

Figure 6.9 shows an example where before time *t*, a congested traffic regime occurred, where target numbers $M_{\widehat{P-Q},ij}$ could not be reached during $[t_0, t]$ since P-Q Sending flows S_{ij}' could not completely be sent during $[t_0, t]$ due to restricting downstream boundary conditions.

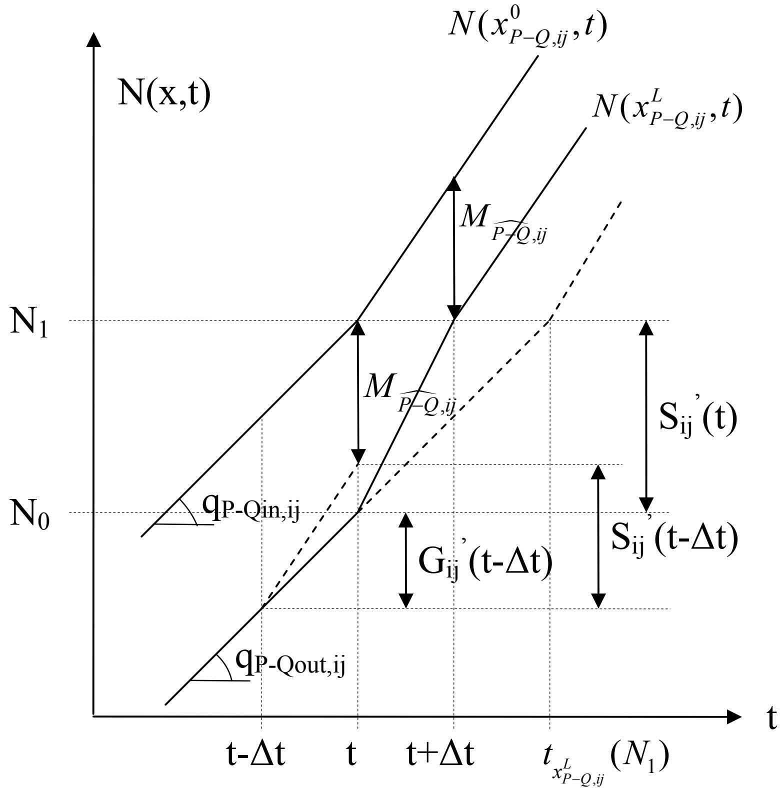

$N(x^0_{P-Q,ij},t)$
N(x,t)
$N(x^L_{P-Q,ij},t)$
$M_{\widehat{P-Q},ij}$
N_1
$M_{\widehat{P-Q},ij}$
$S_{ij}'(t)$
N_0
$q_{P\text{-}Qin,ij}$
$S_{ij}'(t\text{-}\Delta t)$
$G_{ij}'(t\text{-}\Delta t)$
$q_{P\text{-}Qout,ij}$
t
t-Δt
t
t+Δt
$t_{x^L_{P-Q,ij}}(N_1)$

Figure 6.9: P-Q Sending flows in Example 6.3

Example 6.3

At time *t*, the restriction ends, the P-Q inflow rate increases, but the realized number of vehicles in the P-Q still exceeds the target number.
By keeping P-Q outflow rate $q_{P\text{-}Qout,ij}(t\text{-}\Delta t)$ constant (cf. lower dotted line in Figure 6.9), the realized number of vehicles in the P-Q would evolve in the wrong direction: it would increase whereas it should actually decrease following (6.22). Therefore, in this case no additional constraints are imposed on the P-Q Sending flow. Figure 6.9 shows that intersection delays might be overestimated due to restrictive downstream boundary conditions.

b) The target number of vehicles in the P-Q at time *t+Δt* is higher than the realized number of vehicles in the P-Q at time *t*:

$$M_{\widehat{P-Q},ij}(t+\Delta t) > M_{P-Q,ij}(t) \tag{6.28}$$

The methodology is completely analogous to case *a*. Mathematically, it can be summarized as follows:

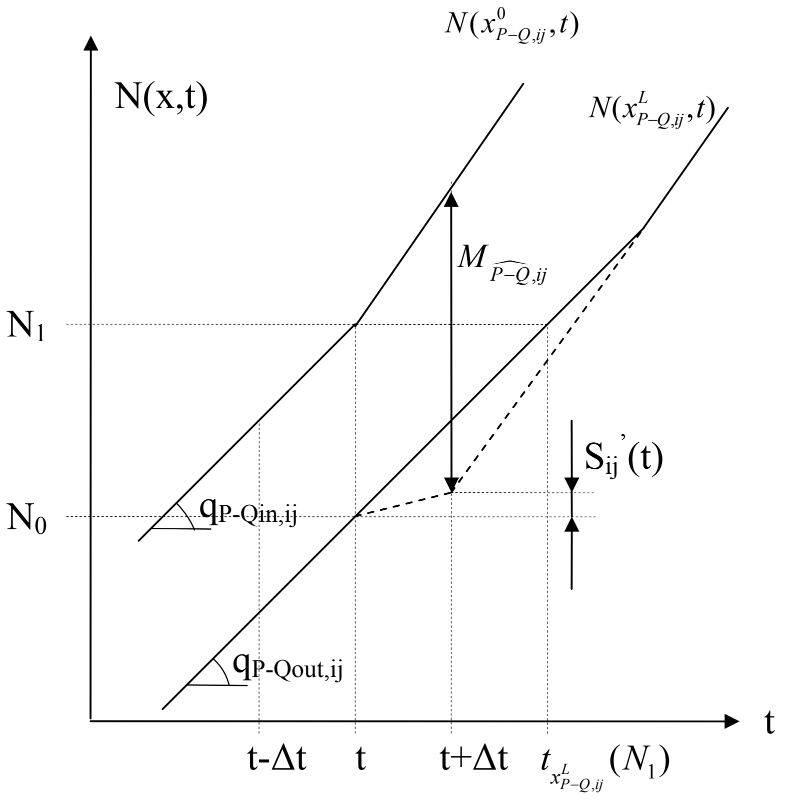
b.1) If $G_{ij}(t) > G_{ij}^{'}(t-\Delta t)$, then

$$S_{ij}^{'}(t) = \min\left\{ q_{M,ij}\Delta t \text{ , } \max\left[\begin{matrix} G_{ij}^{'}(t-\Delta t) \text{ ,} \\ (N(x_{P-Q,ij}^{0}, t+\Delta t) - M_{\widehat{P-Q},ij}(t+\Delta t) - N(x_{P-Q,ij}^{L}, t)) \end{matrix} \right] \right\}$$

b.2) If $G_{ij}(t) \le G_{ij}^{'}(t-\Delta t)$, then

$$S_{ij}^{'}(t) = \max\left\{ 0 \text{ , } \min\left[\begin{matrix} q_{M,ij}\Delta t \text{ ,} \\ (N(x_{P-Q,ij}^{0}, t+\Delta t) - M_{\widehat{P-Q},ij}(t+\Delta t) - N(x_{P-Q,ij}^{L}, t)) \end{matrix} \right] \right\}$$

Figure 6.10 depicts an example of a situation for case *b.1*:



$N(x_{P-Q,ij}^{0}, t)$
N(x,t)
$N(x_{P-Q,ij}^{L}, t)$
$M_{\widehat{P-Q},ij}$
N1
$S_{ij}^{'}(t)$
N0
qP-Qin,ij
qP-Qout,ij
t
t-Δt
t
t+Δt
$t_{x_{P-Q,ij}^{L}}(N_1)$

Figure 6.10: P-Q Sending flows in Example 6.4

Example 6.4

At time *t*, a transition state appears, where the P-Q outflow rate needs to be smaller than the P-Q inflow rate to increase the number of vehicles in the P-Q. To satisfy Condition (6.20), the following additional constraint is imposed on the P-Q Sending flow:

$$S'_{ij}(t) \geq G'_{ij}(t - \Delta t) \tag{6.29}$$

As for case a, the target value of intersection delay $D_{\widehat{\text{int}},ij}(t+\Delta t)$ will exactly be reached in this case b, (or condition (6.20) will be fulfilled) if all of the four conditions (6.21) are satisfied.

c) The target number of vehicles in the P-Q at time $t+\Delta t$ equals the realized number of vehicles in the P-Q at time *t*:

$$M_{\widehat{P-Q},ij}(t+\Delta t) = M_{P-Q,ij}(t) \tag{6.30}$$

In this case, no additional constraints are imposed on the P-Q Sending flow:

$$S'_{ij}(t) = \max\left\{0\ ,\ \min\begin{bmatrix} q_{M,ij}\Delta t\ , \\ (N(x^0_{P-Q,ij}, t+\Delta t) - M_{\widehat{P-Q},ij}(t+\Delta t) - N(x^L_{P-Q,ij}, t)) \end{bmatrix}\right\} \tag{6.31}$$

If before time *t*, P-Q outflow rate $q_{P\text{-}Qout,ij}(t\text{-}\Delta t)$ equaled P-Q inflow rate $q_{P\text{-}Qin,ij}(t\text{-}\Delta t)$, and if at time *t*, the realized number of vehicles in the P-Q $M_{P-Q,ij}(t)$ equals the target number $M_{\widehat{P-Q},ij}(t)$, then (6.30) implies that the P-Q inflow rate does not change at time *t*. Formula (6.31) in this case guarantees that the target value of intersection delay $D_{\widehat{\text{int}},ij}(t)$ will exactly be reached, if the P-Q outflow rate is not influenced by downstream boundary conditions during time interval $[t\ ,\ t + D_{\widehat{\text{int}},ij}(t)\ ]$.

6.2.3.2. Adjustments for time-dependent intersection delays at signalized and unsignalized intersections

As shown in Figure 6.6, time-dependent stochastic intersection delays at signalized intersections are not monotonically increasing as a function of traffic flow. Time-dependent intersection delays at unsignalized intersections are not even monotonically increasing in their non-saturated part ($x < 1$), since they depend on conflicting traffic flows. Therefore, imposing constraints on the P-Q Sending flow as proposed in the solution method for deterministic intersection delays not necessarily implies that the causality principle would be satisfied, nor that the intended intersection delay would actually be reached.

More importantly, time-dependent intersection delays are (obviously) dependent on the time of operation. The target number of vehicles in the P-Q generally changes every time step, even if the P-Q inflow rate stays constant. Figure 6.11 depicts an example with a constant non-saturated P-Q inflow rate.

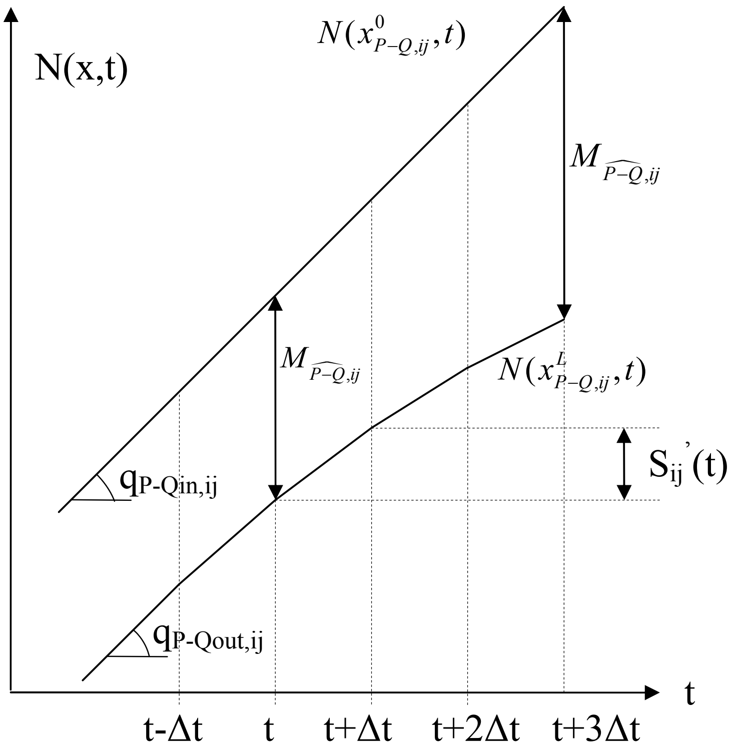


Figure 6.11: Target numbers of vehicles in the P-Q for time-dependent intersection delays

Situations where both *i)* before time *t*, P-Q outflow rate $q_{P\text{-}Qout,ij}(t\text{-}\Delta t)$ equaled P-Q inflow rate $q_{P\text{-}Qin,ij}(t\text{-}\Delta t)$, and *ii)* at time *t*, the realized number of vehicles in the P-Q $M_{P-Q,ij}(t)$ equals the target number $M_{\widehat{P-Q},ij}(t)$, are rather exceptional when working with time-dependent delays. The causality principle is not (necessarily) satisfied. Moreover, imposing constraints (6.24) or (6.29) on P-Q Sending flows does not guarantee that the realized number of vehicles in the P-Q would evolve in the right direction.

For these reasons, above-mentioned constraints are in this case not to be imposed on P-Q Sending flows. We propose to use P-Q Sending flow (6.31) for all scenarios. Target values of intersection delay are calculated by Formula (6.14) for signalized intersections, or Formula (6.16) for unsignalized intersections.

Since these time-dependent delay Formulae (6.14) and (6.16) also depend on the history of the system by taking into account initial queues, gradual transitions between different states are realized automatically; they do not need to be constructed. A solution method for time-dependent intersection delays would however require keeping track of the times of operation of traffic flows. Detailed solution method descriptions for these types of intersection delay are out of the scope of this thesis.

6.2.4. P-Q transition flows $G'_{ij}(t)$ (algorithm step 5)

P-Q transition flow $G'_{ij}(t)$ is determined in step 5 of the extended LTM solution algorithm.
The P-Q transition flow $G_{ij}'(t)$ is defined as the number of vehicles that are actually transferred from the P-Q of turning movement *ij* to link *j* during time interval [t, $t+\Delta t$].
The P-Q transition flow G_{ij}' is determined by applying the appropriate node model (cf. Chapter 5), where Sending flow S_{ij} is now replaced by P-Q Sending flow S_{ij}'.

For intersections where all turning movements have the disposal of at least one turning lane which has enough storage capacity so that it doesn't block other sub-flows on the same link, as long as the turning flows can on average be handled, each turning movement has its own P-Q and FIFO constraints implying that total Sending flow S_i' is restricted if one of its sub-flows S_{ij}' is restricted do not have to be included, since P-Q's contain only one sub-flow S_{ij}':

$$G_{ij}' = \min(S_{ij}', R_{ij}, q_{n,ij}) \tag{6.32}$$

For intersections where multiple turning movements operate on the same shared-lane, a P-Q's is associated with this shared-lane, rather than with its individual turning movements. For such cases, FIFO constraints should be imposed on the P-Q transition flows of these turning movements that operate on the same shared-lane:

$$G_{ij}' = \min_{j' \in J}\left(\frac{q_{n,ij'}}{S_{ij'}'}, \frac{R_{ij'}}{S_{ij'}'}, 1\right) S_{ij}' \tag{6.33}$$

where J is the set of all Receiving links j that are served by the considered shared-lane.

Transition flows G_{ij}' that are restricted by their corresponding Receiving flows R_{ij} might be re-evaluated in a second step, as pointed out in the solution method for the basic urban cross node (cf. Section 5.2.1).

The same downstream boundary conditions are imposed on transition flow G_{ij} and P-Q transition flow G_{ij}' (cf. Formula (5.12) vs. Formula (6.32)). Therefore, a backward propagating congested traffic state enters a link at the P-Q exit and P-Q entrance at the same time. Shockwaves travel through P-Q's with an infinite backward wave speed.

In Section 6.2.3.1, it was stated that intended intersection delays are exactly realized for low-frequently changing traffic flows that are not restricted by downstream boundary conditions.

For flows that are restricted by downstream boundary conditions, Counterexample 6.3 illustrated how realized delays differ from intended delays. A similar Example 6.5 is presented below, only now the proposed solution method to implicitly include deterministic intersection delays at signalized intersections is compared with an explicit simulation of traffic flows. The microscopic simulation model *AIMSUN 2* (Barcelo (2002)) is used to run this explicit simulation. Traffic flows are uniform to exclude stochastic intersection delays.

Example 6.5

If downstream boundary conditions restrict both P-Q transition flow G_{ij}' and transition flow G_{ij}, then the number of vehicles in the P-Q cannot decrease. Therefore, if the restricting P-Q transition flow G_{ij}' is smaller than P-Q Sending flow S_{ij}', it is possible that the realized number of vehicles in the P-Q is higher than the target number, though this target number would have been realized by P-Q Sending flow S_{ij}'. Figure 6.12 pictures an example of such a situation.

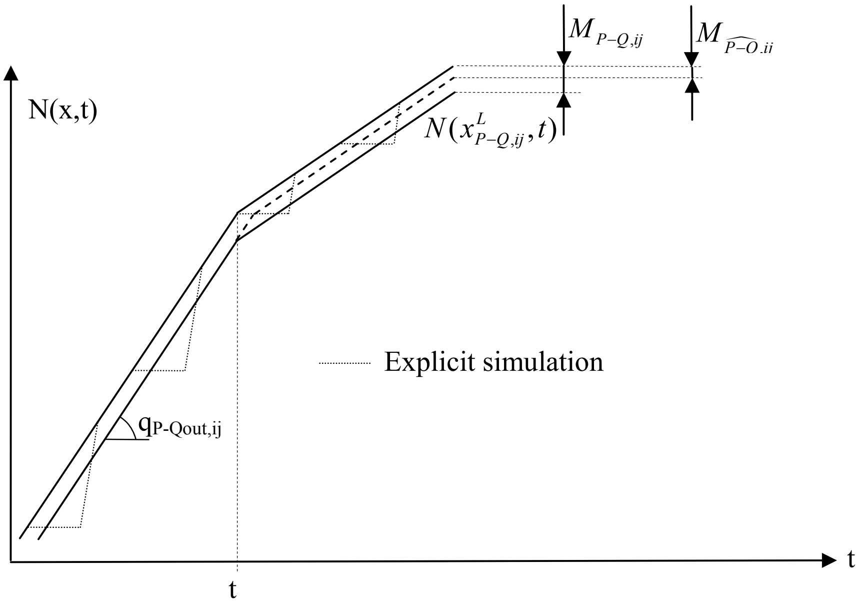


Figure 6.12: P-Q cumulative vehicle numbers in Example 6.5

> Assume that at time *t*, a restriction is imposed by the Receiving flow of the downstream link. Both P-Q in- and outflow rates are restricted at the same time *t*. Before time *t*, the P-Q outflow rate is such that the implicitly experienced intersection delay exactly equals the average intersection delay of the explicit simulation. This implies that using Websters Formula (6.8) yields exact results. After time *t*, the P-Q outflow rate that would have occurred in absence of a restriction imposed by a downstream boundary condition (indicated in Figure 6.12 by the dotted line), is again such that the implicitly experienced intersection delay exactly equals the average intersection delay of the explicit simulation. However, due to a restriction from downstream, the proposed implicit solution method in this case overestimates intersection delays during congestion.

Example 6.5 illustrates that the intended intersection delay, though not always perfectly realized in practice, exactly equals the average of the delays that are experienced in an explicit simulation of traffic flows.

6.3. Conclusions

This chapter introduces a method to implicitly include intersection delays in LTM. Objective of the intersection delay model is to realize average delays, where high-frequent fluctuations due to a repeated process of queue-formation and queue- dissipation are filtered out.
To implicitly account for average intersection delays, it is first determined what the expected average intersection delay would be. Existing intersection delay formulae are generally based on queuing theory. Though these formulae are quite easily integrated in travel-time based models, it is not self-evident to incorporate them within flow-based models such as LTM, where travel time is not a basic variable. Point-Queues are introduced to realize intersection delays in an indirect way.

The expected (or target) values of intersection delay are subsequently converted into a target number of vehicles in the P-Q. The length of the P-Q should be such that vehicles experience their intended average delay by passing through this P-Q. Point-Queues only represent temporary queues due to stage alternations or priority rules. They occur next to kinematic wave queues that are due to over-saturation.

Subsequently, P-Q Sending flows are determined to realize the target numbers of vehicles in the P-Q. However, constraints are imposed on these P-Q Sending flows to make sure that the intended intersection delay will actually be reached.
For low-frequently changing traffic flows that are not restricted by downstream boundary conditions, intended intersection delays are exactly realized and a linear transition occurs between different states with different values of intersection delays.

Finally, P-Q transition flows are determined by applying the appropriate node model, where Sending flow S_{ij} is now replaced by P-Q Sending flow S_{ij}'.

For traffic flows that are restricted by downstream boundary conditions or for high-frequently changing traffic flows, realized intersection delays might slightly differ from the intended delays, which exactly equal the average of the delays that are experienced in explicit traffic flow simulations.

Though further research on the theoretical validity is necessary, the proposed intersection delay model appears to be a promising method to implicitly include intersection delays in flow-based models like LTM.

7

CASE STUDIES

In this chapter, two case studies are elaborated to illustrate some characteristics of the developed LTM. The first case study considers a small theoretical diverge network. Queue propagation in LTM is compared with queue propagation in a Dynamic Queuing Model (DQM) that is used for Dynamic Network Loading in state-of-the-art DTA models such as DynaMIT (Ben-Akiva et al. (1998)), DYNASMART (Mahmassani et al. (2001)) and INDY-DQM (Bliemer (2005, 2006, 2007)).
A second case study demonstrates the applicability of LTM in practical large scale networks. The considered Ghent-Brussels network includes about 12000 possible routes between the different OD-pairs. Queue propagation characteristics in LTM and DQM are compared in this large scale network. Furthermore, a small sub-network is considered to show the impact of taking into account intersection delays.

7.1. Case 1: Simple diverge network

7.1.1. Aims

This case aims to provide insight into LTM queue propagation characteristics by comparing them with DQM queue propagation characteristics.

7.1.2. Model properties: LTM versus DQM

LTM is described extensively in the previous chapters. Dynamic Queuing Models (DQM) are used to propagate traffic in state-of-the-art DTA models such as DynaMIT, DYNASMART and INDY-DQM. In DQM, links are artificially split up into a moving part and a queuing part. Vehicles traverse the moving part of the link with a certain speed v, which is assumed to be a function of traffic density. Different speed-functions can be applied. Here, it will be assumed that speed v equals free-flow speed v_f and does not change while traversing the moving part of the link.

The queuing part of the link is characterized by a fixed queue density k_{queue}. Capacity constraints are imposed on the downstream boundary of each link and queues are always located at this downstream link end.

In LTM, the storage capacity of an outgoing link j depends on the flow on this link. The considered DQM however has a fixed storage capacity of $k_{queue}L_j$ vehicles (where L_j is the length of link j). Spillback to an upstream link occurs if (and only if) this outgoing link j is completely filled with vehicles at queue density k_{queue}. For a detailed description of traffic propagation in DQM, we refer to Chabini (2001).

The fundamental diagram corresponding to the above-described DQM is pointed out in Figure 7.1:

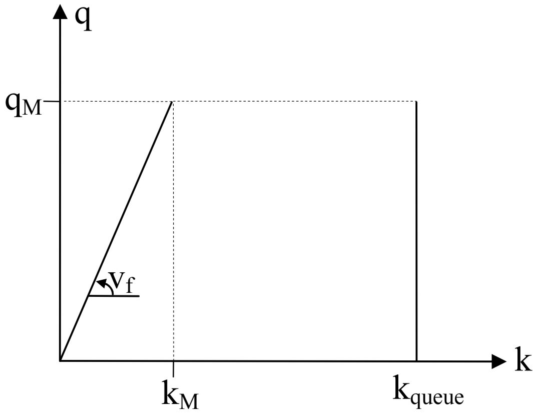


Figure 7.1: Fundamental Diagram in DQM

Traffic states on the free-flow branch ($k < k_M$) of the fundamental diagram occur in the moving part of a link. Traffic states on the vertical branch ($k = k_{queue}$) are congested. They occur in the queuing part of a link. Once critical link density k_M is exceeded, there is only one possible congestion density, no matter which flow is realized in the congested part of the link: k_{queue}. Figure 7.2 compares this fundamental diagram with the triangular one that is used in LTM, which includes a continuous spectrum of possible congestion densities:

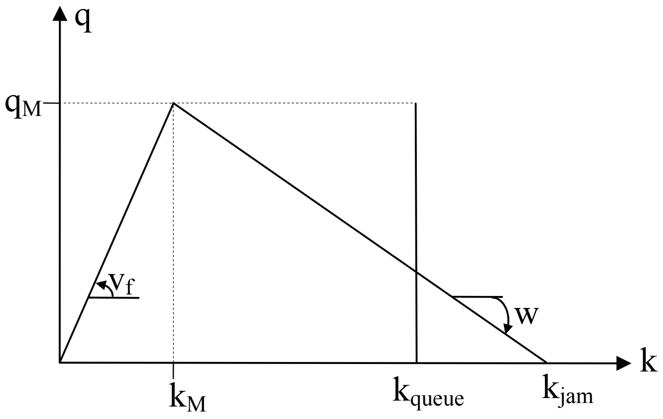


Figure 7.2: Fundamental diagram in LTM versus DQM

To compare queue propagation characteristics in LTM and DQM, we use the Dynamic Traffic Assignment (DTA) model INDY (Bliemer (2005)), where both LTM and DQM are available as DNL model. Depending on which of these DNL models is used, INDY is referred to as INDY-LTM or INDY-DQM.

7.1.3. Inputs

Case 1 considers a simple diverge network with two origin nodes r_1 and r_2, two destination nodes s_1 and s_2, one diverge node *n* and one merge node *m* (see Figure 7.3).

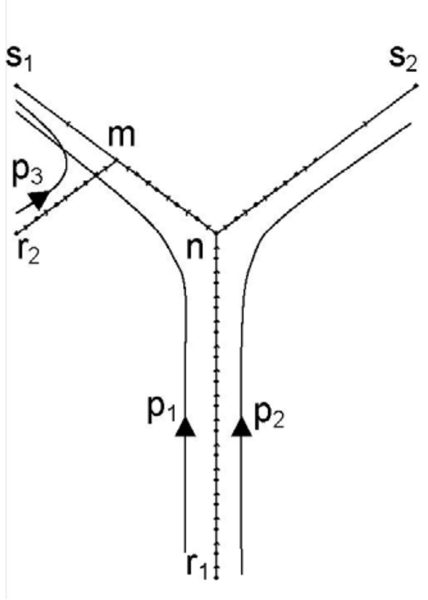


Figure 7.3: Simple diverge network

Network characteristics are given in Table 7.1:

Table 7.1: Network characteristics of the simple diverge network

Link	r_1n	nm	r_2m	ms_1	ns_2
L (km)	7	2,5	2,5	2,5	5
# lanes	4	2	2	2	2
v_f *(km/h)*	120	120	120	120	120
q_M *(veh/h)*	8000	4000	4000	4000	4000
k_{jam} *(veh/h)*	600	300	300	300	300
k_{queue} *(veh/h)*	400	200	200	200	200

Each of the five links in Table 7.1 is subdivided in short links of *500 m*. This subdivision is not necessary for the LTM solution algorithm nor does it provide more accurate results; however it allows us to visualize queue propagation in more detail. The simulation time step $\Delta t = 15\ s$ (which is the minimum link travel time) and total simulation time $T = 90$ *minutes*. Route flow rates are as follows:

$f_{r1s1}^{p1}(t)$ = *3000 veh/h* for all $t \in T$

$f_{r1s2}^{p2}(t)$ = *3000 veh/h* for all $t \in T$

$f_{r2s1}^{p3}(t)$ = *3000 veh/h* for all $t \in T' \subset T \,|\, t < 30 \text{ min}$

$f_{r2s1}^{p3}(t)$ = *0 veh/h* for all $t \in T' \subset T \,|\, t \geq 30 \text{ min}$

The two major flows along routes p_1 and p_2 originate in r_1 and leave the network in s_1 and s_2 respectively. The flow along route p_3 only functions as a temporarily bottleneck-creator (bottleneck in merge node *m*, see later). A demand-proportional node model is used for merge node *m*.

There is no route choice in this case. One single iteration of the DNL algorithm is sufficient to obtain final results.

7.1.4. Outputs

Figures 7.4a to 7.4p compare queue propagation throughout time in LTM and DQM. In-between each screenshot is a time gap of six minutes. The blocks on the different links represent link densities. Thicker and darker blocks represent higher densities.

Initially, link ms_1 has insufficient capacity to accommodate both flows $f_{r1s1}^{p1}(t)$ and $f_{r2s1}^{p3}(t)$. Therefore, merge node m will act as a bottleneck. Queues develop on both links nm and r_2m (cf. Figures 7.4b to 7.4g). The queues in LTM are less dense and they propagate faster through the network compared to DQM.

Figure 7.5 explains the origin of this observation. The situation on link nm (which at this point in time equals the situation on link r_2m) is represented in both LTM and DQM.

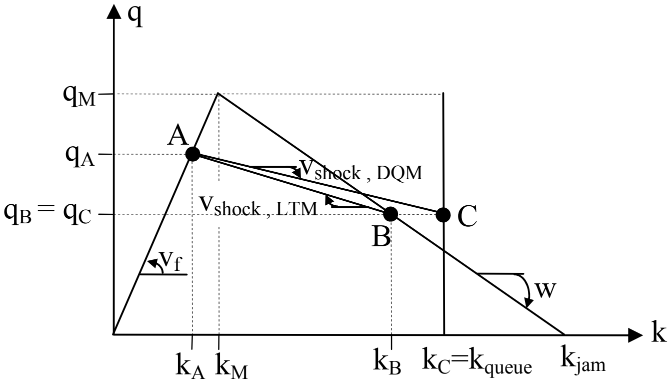
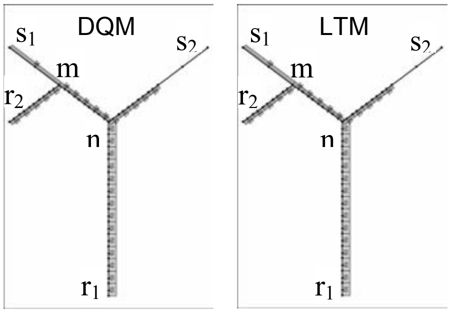
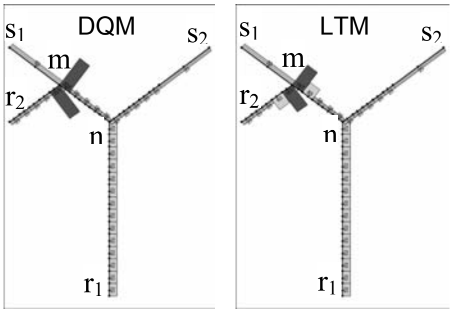


Figure 7.5: Traffic states on link *nm* in LTM and DQM

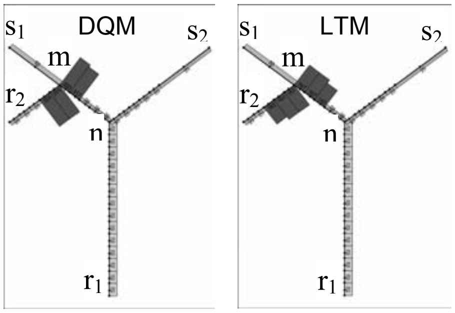
In both models, link nm has a free-flow zone on its upstream side, and a congested zone on its downstream side. Free-flow traffic states in both models correspond to traffic state A ($q_A = 3000$ *veh/h*; $k_A = q_A/v_f = 25$ *veh/km*). The congested traffic zone is characterized by traffic state B for LTM ($q_B = 2000$ *veh/h*; $k_B = k_M + \frac{q_M - q_B}{q_M}(k_{jam,LTM} - k_M) = 167$ *veh/km*) and by traffic state C for DQM ($q_C = 2000$ *veh/h*; $k_C = k_{queue} = 200$ *veh/km*). In LTM, the queue on link nm is less dense compared to DQM ($k_B < k_C$). Since queue inflow rate q_A exceeds queue outflow rates q_B and q_C, the congested zone on link nm will expand, i.e., the queue on link nm will grow. The boundary between the free-flow and the congested state travels backward through link nm with (negative) speed $v_{shock} = \frac{q_{in} - q_{out}}{k_{in} - k_{out}}$ (cf. Formula (4.18)). This shockwave speed is higher in LTM ($v_{shock,LTM} = \frac{q_A - q_B}{k_A - k_B} > \frac{q_A - q_C}{k_A - k_C} = v_{shock,DQM}$), so the queue spills back faster through the network in LTM.


s_1
DQM
s_2
m
r_2
n
r_1
s_1
LTM
s_2
m
r_2
n
r_1

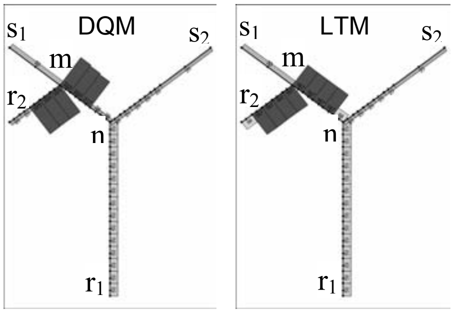
(a) *t = 4 min*


s_1
DQM
s_2
m
r_2
n
r_1
s_1
LTM
s_2
m
r_2
n
r_1

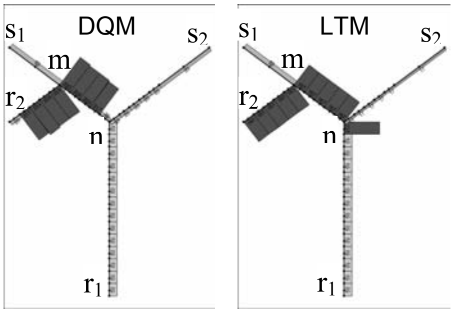
(b) *t = 10 min*


s_1
DQM
s_2
m
r_2
n
r_1
s_1
LTM
s_2
m
r_2
n
r_1

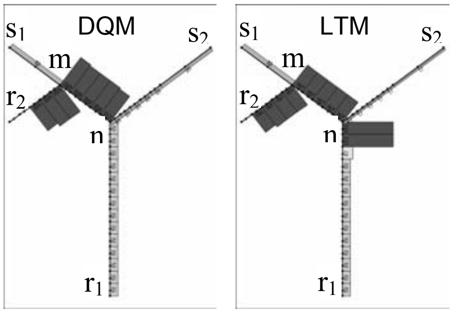
(c) *t = 16 min*


s_1
DQM
s_2
m
r_2
n
r_1
s_1
LTM
s_2
m
r_2
n
r_1

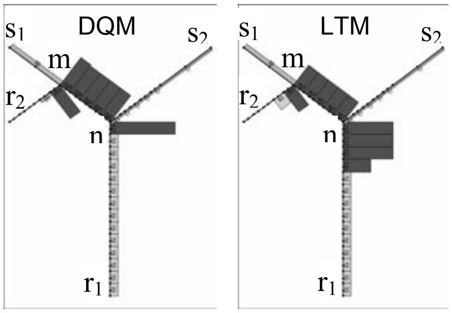
(d) *t = 22 min*


s_1
DQM
s_2
m
r_2
n
r_1
s_1
LTM
s_2
m
r_2
n
r_1

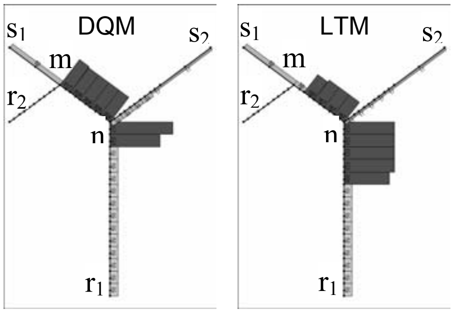
(e) *t = 28 min*


s_1
DQM
s_2
m
r_2
n
r_1
s_1
LTM
s_2
m
r_2
n
r_1

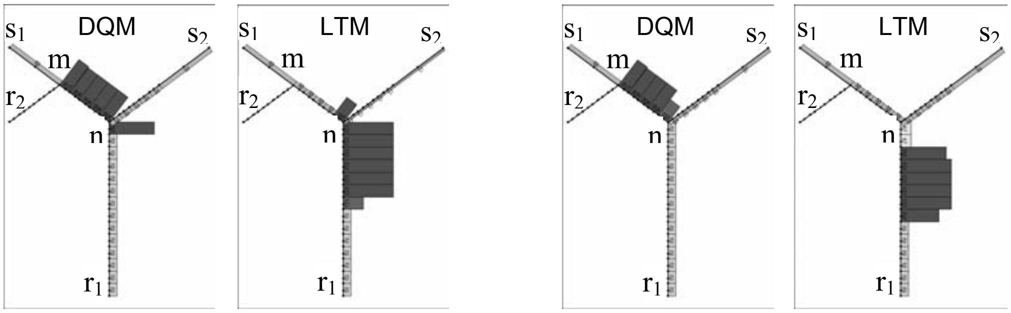
(f) *t = 34 min*


s_1
DQM
s_2
m
r_2
n
r_1
s_1
LTM
s_2
m
r_2
n
r_1

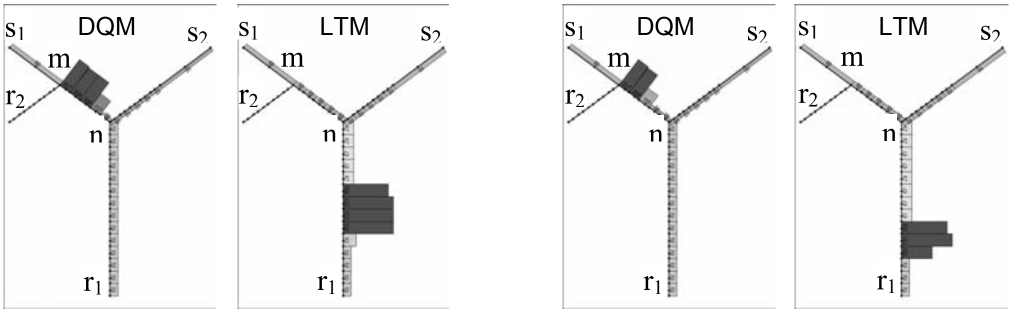
(g) *t = 40 min*


s_1
DQM
s_2
m
r_2
n
r_1
s_1
LTM
s_2
m
r_2
n
r_1

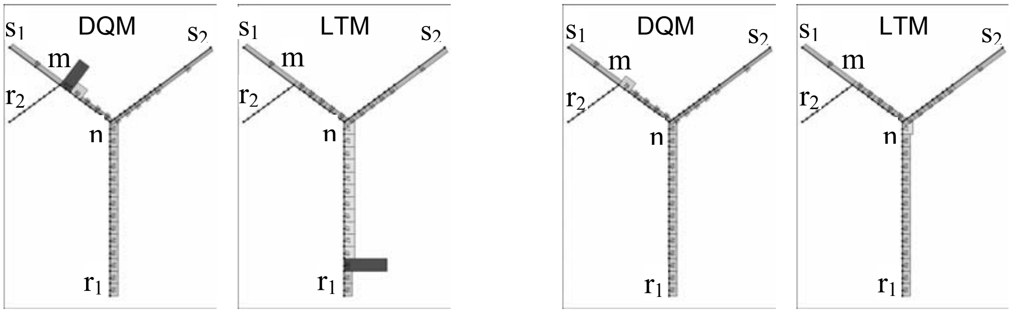
(h) *t = 46 min*



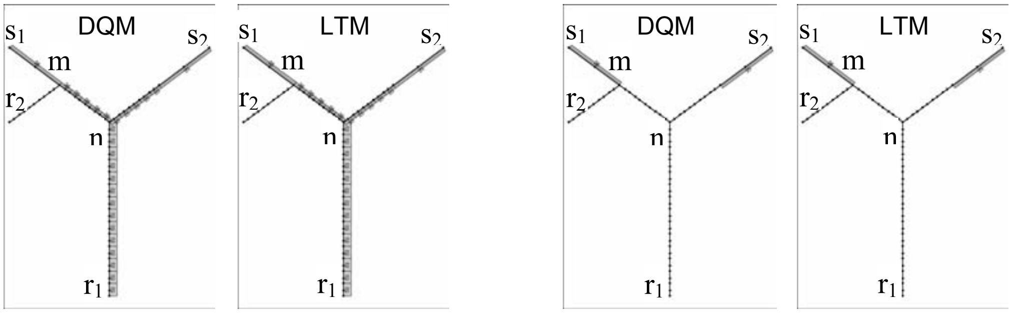
(i) *t = 52 min* (j) *t = 58 min*



(k) *t = 64 min* (l) *t = 70 min*



(m) *t = 76 min* (n) *t = 82 min*



(o) *t = 88 min* (p) *t = 94 min*

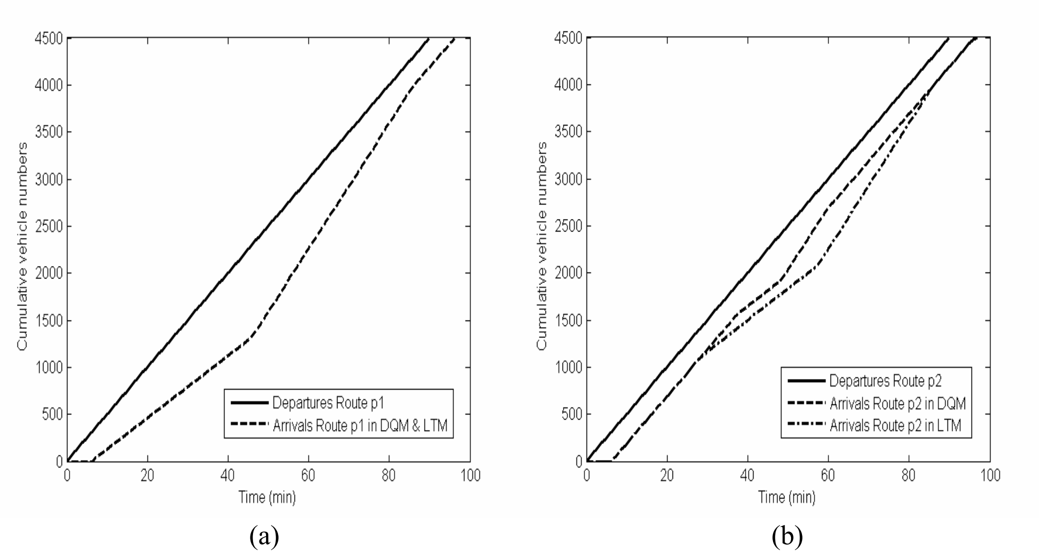
Figure 7.4: Queue propagation throughout time in DQM (left) and LTM (right)

Considering the results of Figure 7.4 again, observe that at time t = *34 min*, flow $f_{r2s1}^{p3}(t)$ no longer exists and the queue on link r_2m gradually disappears. However, the queue on link *nm* keeps growing and even spills back to link r_1n (cf. Figures 7.4f to 7.4h). At time t = *46 min*, link r_2m is cleared and merge node *m* no longer acts as a bottleneck. As from that time, the outflow rate of the queue on links r_1-*n*-*m* ($q_{out} = q_{M,ms1} = q_{M,nm}$ = *4000 veh/h*) exceeds the queue inflow rate ($q_{in} = f_{r1s1}^{p1}(t)$ = *3000 veh/h*) and the queue starts to shrink.

In DQM, the queue dissipates starting from the tail in the direction of the head (cf. Figures 7.4h to 7.4n). In LTM, the direction of queue dissipation is reverse: the queue dissipates starting from the head of the queue in the direction of the tail. The queue grows further at the back side while it shrinks at the front side (cf. Figures 7.4h to 7.4m). This opposite direction of queue dissipation only occurs if queue dissipation is due to a bottleneck becoming inactive. If the bottleneck stays active during queue dissipation (as happens for example on link r_2m from t = *28 min* to t = *46 min*, cf. Figures 7.4e to 7.4h), then the direction of queue dissipation equals the one in DQM, i.e. in that case queues dissipate from the back to the tail of the queue.

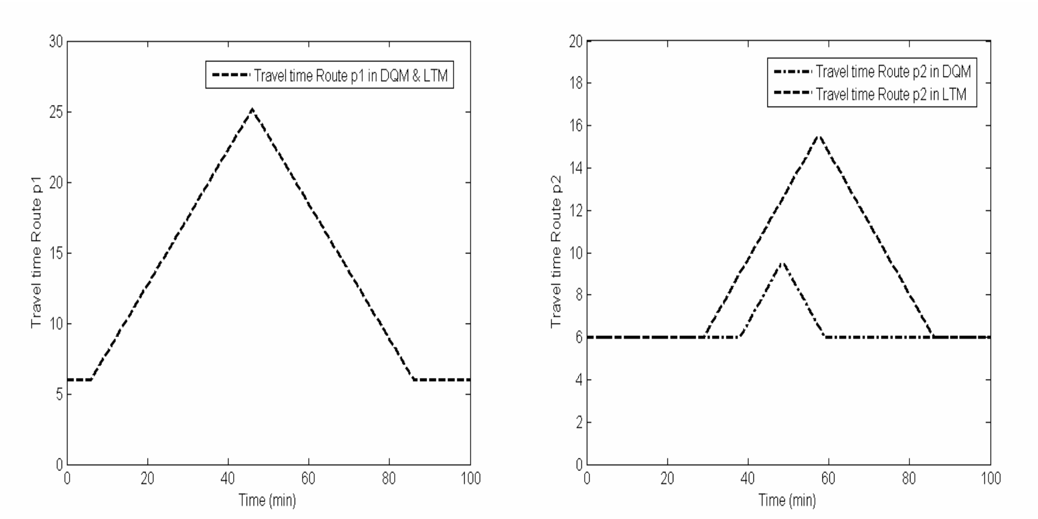
The different spillback speed of the queues and the different direction of queue dissipation significantly influence route travel times.

Figure 7.6 depicts the cumulative departures and arrivals for routes p_1 and p_2 in both LTM and DQM. The horizontal distances between the departure and the arrival curves represent route travel times. Travel times as a function of time are plotted in Figure 7.7.



(a) (b)

Figure 7.6: Cumulative departures and arrivals for routes p_1 (a) and p_2 (b)



(a) (b)

Figure 7.7: Travel times for routes p_1 (a) and p_2 (b)

For route p_1, though the queues behind bottleneck *m* have different spatial extents, route travel times are equal in LTM compared to DQM. For route p_2 however, travel times significantly differ in both models. Figure 7.4 shows that in DQM, the queue on link r_1n turns up on $t = 40\ min$ and disappears on $t = 58\ min$. For this time period, increased travel times on route p_2 are observed (Figure 7.7b). In LTM, the queue on link r_1n initiates sooner ($t = 28\ min$) and disappears later ($t = 82\ min$) (cf. Figure 7.4). As shown in Figure 7.7b, travel times on route p_2 are substantially higher in LTM.

7.1.5. Conclusions of the diverge case study

Queue propagation characteristics of two DNL models, LTM and DQM, are compared in a small theoretical network. Differences in both models include:

- different queue densities, different queue lengths, and different spillback speeds of queues
- different direction of queue dissipation in case queue dissipation is due to a bottleneck becoming inactive.

The case study shows that these different queue propagation properties can significantly influence route travel times.

7.2. Case 2: Ghent-Brussels network

7.2.1. Aims

This case aims to:

- demonstrate the applicability of LTM in practical large scale networks
- compare LTM and DQM queue propagation characteristics in large scale networks
- show the impact of taking into account intersection delays on route travel times and on route choice

7.2.2. Model description

The Dynamic Traffic Assignment (DTA) model INDY is used to run this case. The general principle is illustrated in Figure 7.8.

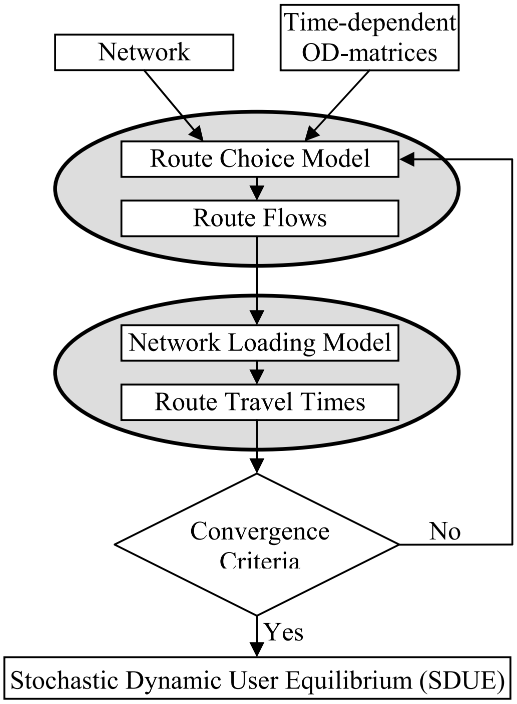

Network
Time-dependent OD-matrices
Route Choice Model
Route Flows
Network Loading Model
Route Travel Times
Convergence Criteria
No
Yes
Stochastic Dynamic User Equilibrium (SDUE)

Figure 7.8: General principle of the DTA model

First, the route choice model assigns all travelers to a specific route. Subsequently, the Dynamic Network Loading (DNL) model propagates traffic through the network along the assigned routes and it computes route travel times. These two major components of the DTA model are iterated in sequence until they converge into a Stochastic Dynamic User Equilibrium (SDUE), where all travelers choose their perceived cheapest routes.

In INDY, both LTM and DQM are available as DNL model. The route choice model is based on discrete choice theory. A simple Multinomial Logit (MNL) model is used to determine the choice probability of route $p \in P_{rs}$ as a function of the difference between the utility of that route and all other alternative routes between the same origin-destination pair *rs*:

$$\Pr(p(t)) = \frac{\exp(-\mu C_{rs}^{p}(t))}{\sum_{p' \in P_{rs}} \exp(-\mu C_{rs}^{p'}(t))} \qquad \textit{for all } p \in P_{rs} \tag{7.1}$$

where *Pr(p(t))* = probability of choosing route $p \in P_{rs}$ at time *t* (-)

$C_{rs}^{p}(t)$ = travel cost of route *p* between origin *r* and destination *s* at time *t*, representing a negative utility (€)

μ = user-defined scale parameter of the MNL model, describing the accuracy of route cost perception (-).

Travelers choose routes from a set of available routes, which is determined a priori, based on a static traffic assignment.

The route choice model and the DNL model are performed iteratively until a Stochastic Dynamic User Equilibrium (SDUE) is reached. The iterative procedure is based on the Method of Successive Averages (MSA) and works as follows:

- Based on currently experienced route costs C_{rs}^{p}, the route choice model determines route flows.
- MSA averages these route flows with route flows from the previous iteration into new route flow rates.
- Based on these new route flow rates, the DNL model then determines new route travel costs.
- These route travel costs are input for the route choice model, which closes the iteration loop (see Figure 7.8).

Since the Method of Successive Averages is a heuristic method, convergence into equilibrium is not guaranteed. Within INDY however, it usually provides sufficient convergence, since INDY starts with a predefined route choice set and a good initial solution based on a static traffic assignment (Bliemer (2006, 2007)). In the SDUE, all travelers choose their perceived cheapest route, i.e., no traveler can reduce his perceived travel cost by unilaterally changing routes. All routes used between an origin-destination pair have the same minimal perceived costs, and no unused route has a lower perceived cost. For a more detailed description of the route choice model and the iterative procedure in INDY, we refer to Bliemer (2005).

7.2.3. Ghent-Brussels network: Inputs

This case considers a practical large scale network with 3540 links, 1146 nodes and 11936 possible routes between the different origin-destination pairs. The network includes highway E40 between Ghent and Brussels, and all important secondary roads in the surrounding area (see Figure 7.9).

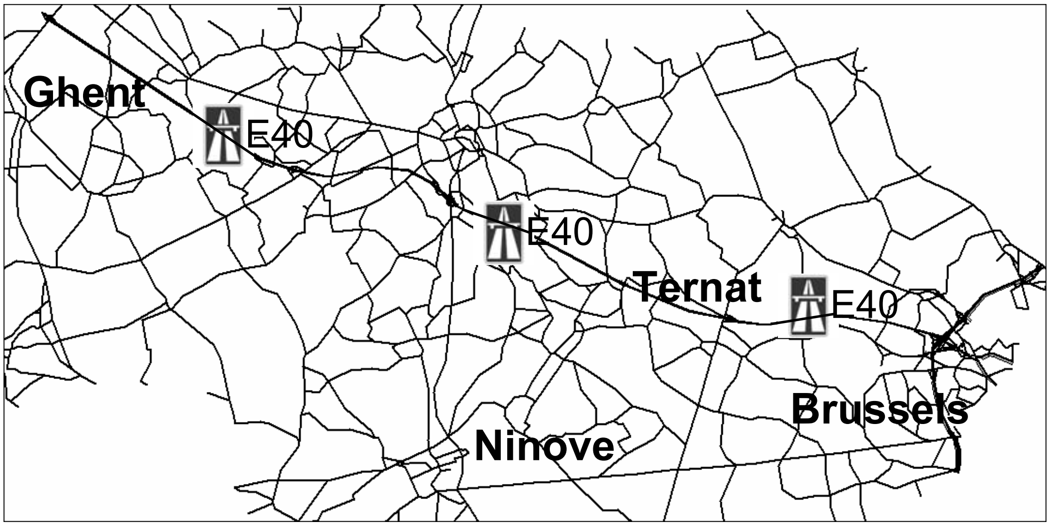


Figure 7.9: Ghent-Brussels network

The simulation time step $\Delta t = 10\ s$ (which is the minimum link travel time) and total simulation time $T = 5\ h$. Time-dependent OD-matrices were imported every 5 minutes to represent a typical morning rush between *5u00* and *10u00*. Dynamic OD-matrices and network characteristics were estimated and calibrated in a previous study (Tampère et al. (2007)). For DQM, the fixed queue density $k_{queue} = 100\ veh/km/lane$, and for LTM, jam density $k_{jam} = 150\ veh/km/lane$.

Routes are chosen from a set of available routes, which are predefined based on a static traffic assignment. Route travel costs C_p^{rs} equal route travel times τ_p^{rs}. The scale factor for the Logit model equals $\mu = 1/0.07$, and 10 iterations are performed in order to approach SDUE.

7.2.4. Ghent-Brussels network: Outputs

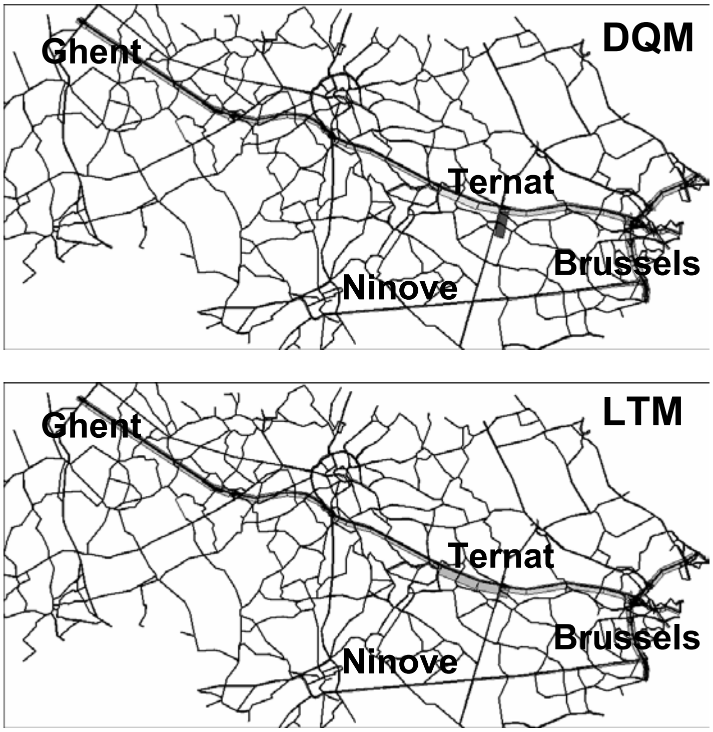
Applicability of LTM in large scale networks

First, the practical feasibility for large sized transportation networks is illustrated. Computational tests show that the computation time is independent on the number of considered vehicles, but that the computation time increases proportional to the number of considered routes and inversely proportional to the time step Δt.
For the Ghent-Brussels network (3540 links, 1146 nodes, 11936 routes, $\Delta t = 10\ s$, $T = 5\ h$, *# iterations* $= 10$), each of the 10 Dynamic Network Loadings took *5 minutes* of computing time on a Pentium IV-3 GHz (dual core) computer with 2 GB RAM, thereby showing the applicability of LTM in a large sized network.

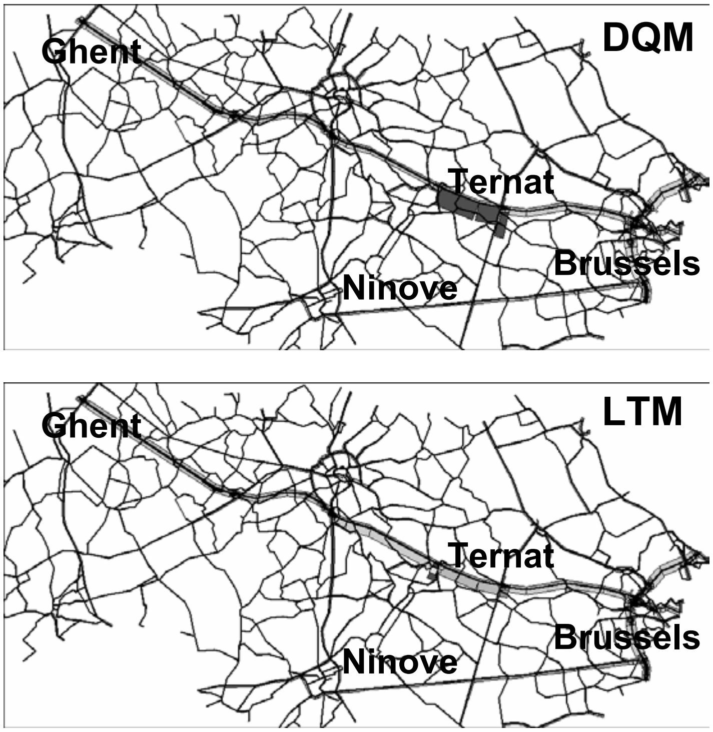
Comparison of queue propagation throughout time in LTM vs. DQM

Figures 7.10a to 7.10f compare queue propagation throughout time in LTM and DQM. The first screenshot depicts the situation at *6u00*, and in-between each screenshot is a time gap of *45 minutes*. The blocks on the different links represent link densities. Thicker and darker blocks represent higher densities.

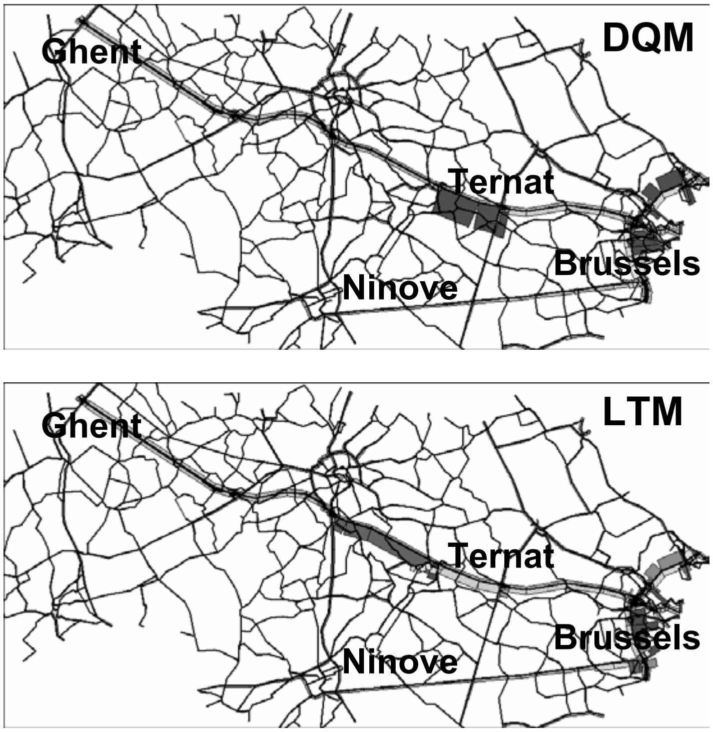
At *6u00*, the bottleneck at the merge of highway E40 and the on-ramp in Ternat becomes active and a queue develops on highway E40 behind this bottleneck. As shown in Figures 7.10b to 7.10d, queues in LTM are less dense, longer, and they propagate faster through the network compared to DQM.
At *7u30*, another queue occurs on the ring road around Brussels. In DQM, this queue does not spill back further than the ring road itself. In LTM however, this queue spills back on the road towards Ninove (cf. Figures 7.10d and 7.10e). Queue locations significantly differ in both models.


Ghent
DQM
Ternat
Brussels
Ninove
Ghent
LTM
Ternat
Brussels
Ninove

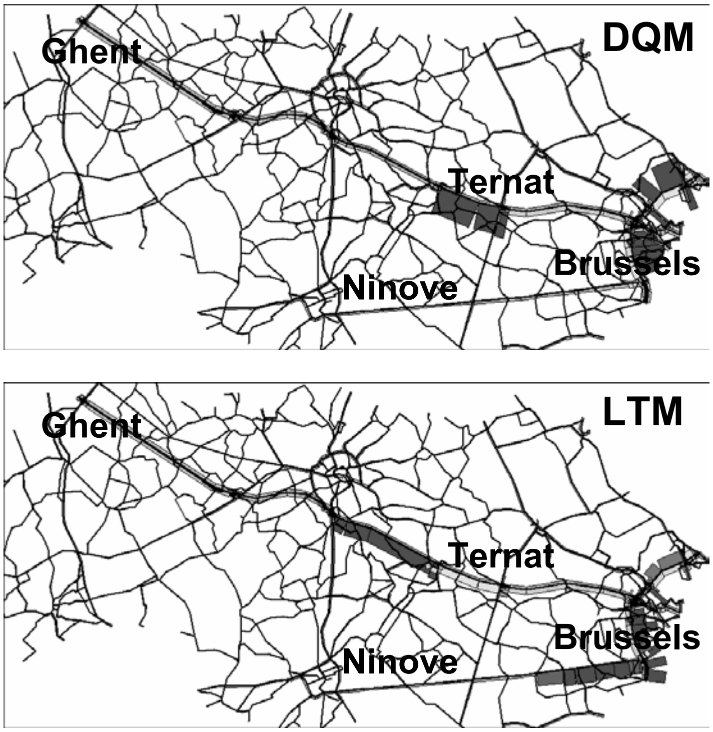
(a) 6 u 00


Ghent
DQM
Ternat
Brussels
Ninove
Ghent
LTM
Ternat
Brussels
Ninove

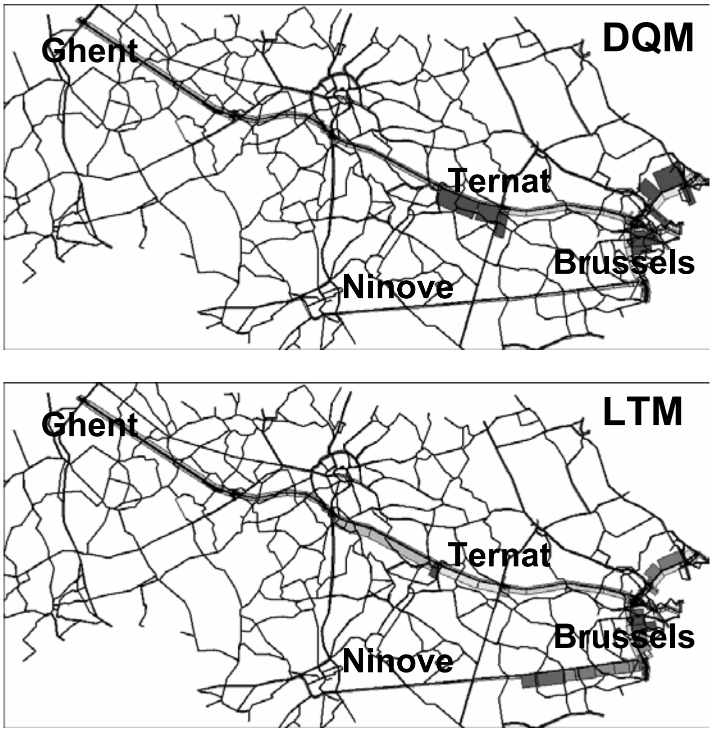
(b) 6 u 45


Ghent
DQM
Ternat
Brussels
Ninove
Ghent
LTM
Ternat
Brussels
Ninove

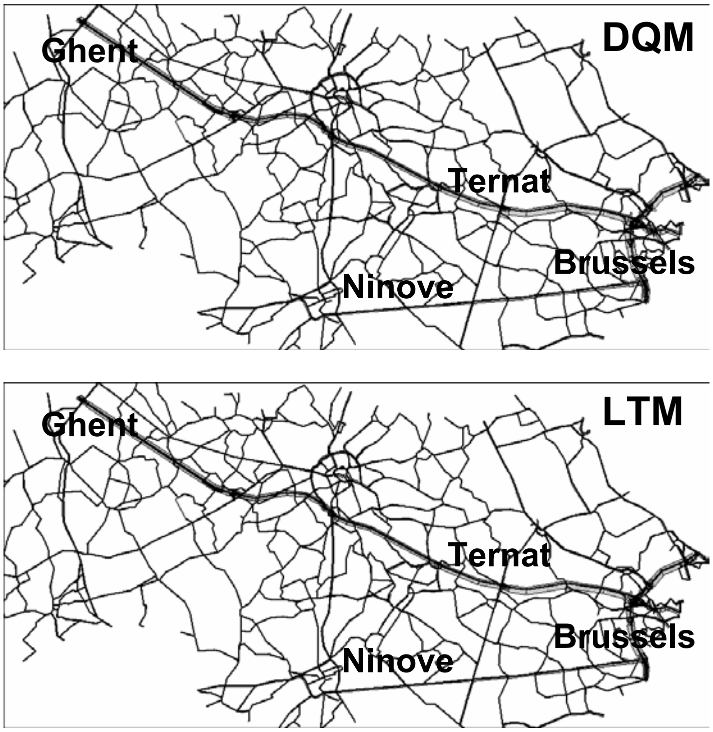
(c) 7 u 30


Ghent
DQM
Ternat
Brussels
Ninove
Ghent
LTM
Ternat
Brussels
Ninove

(d) 8 u 15


Ghent
DQM
Ternat
Brussels
Ninove
Ghent
LTM
Ternat
Brussels
Ninove

(e) 9 u 00

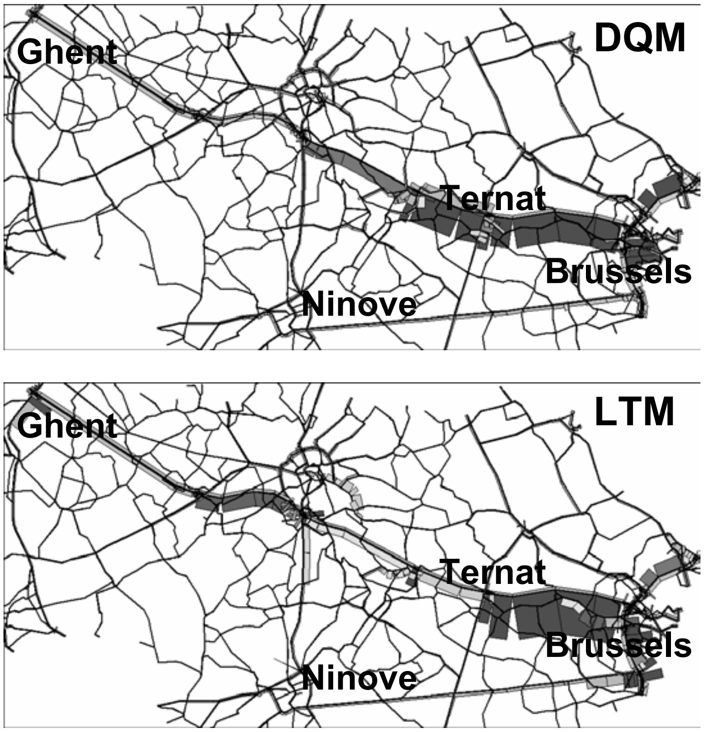

Ghent
DQM
Ternat
Brussels
Ninove
Ghent
LTM
Ternat
Brussels
Ninove

(f) 9 u 45

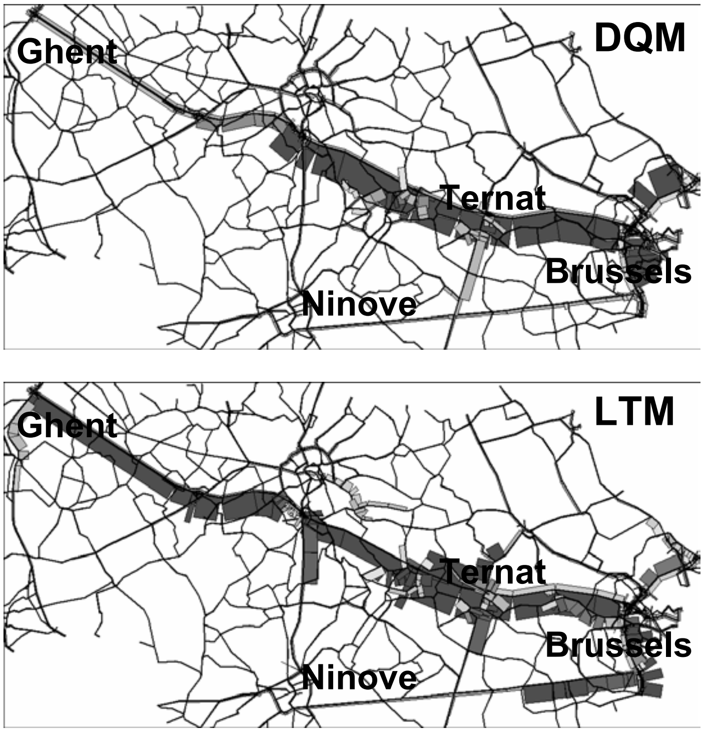
Figure 7.10: Queue propagation throughout time in DQM (on top) and LTM (below) in a typical morning rush

Figures 7.11a to 7.11e compare queue propagation throughout time in LTM and DQM, when at *7u00*, an incident occurs on highway E40 Ghent-Brussels between Ternat and Brussels (direction of Brussels). At the incident location, the capacity of highway E40 is first (from *7u00* to *7u30*) reduced to *0%* , and later on (from *7u30* to *8u30*) to *50%* of its original capacity. At *8u30*, the original capacity is re-established.

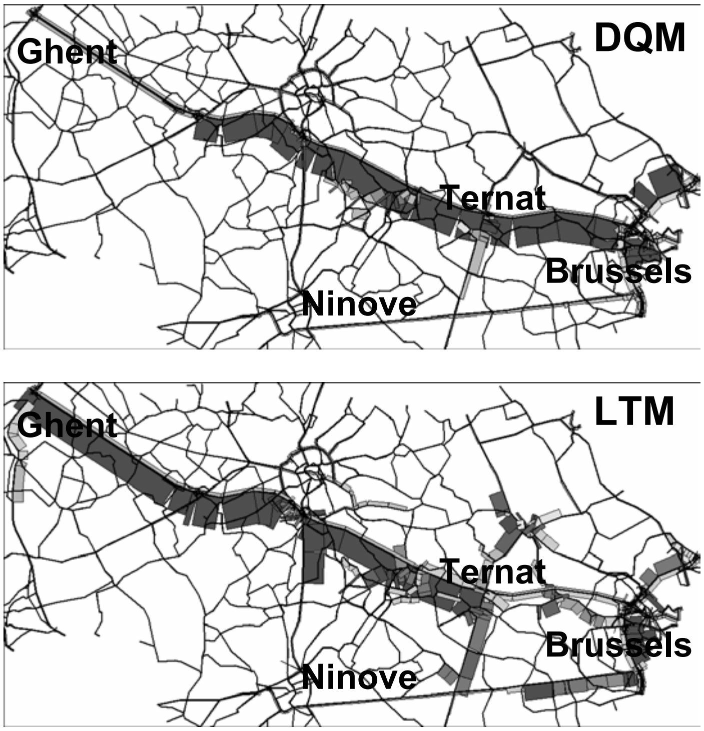
Queues in LTM are less dense, longer, and they propagate faster through the network. Furthermore, queues in LTM dissipate in the opposite direction compared to DQM, since queue dissipation occurs due to a bottleneck becoming inactive. The bottleneck at the incident location disappears at *8u30*. Queues in LTM dissipate from the head to the tail of the queue, i.e. from Brussels to Ghent, while queues in DQM dissipate from Ghent to Brussels (cf. Figures 7.11c to 7.11e). Queue locations significantly differ in both models (cf. Figure 7.11a to 7.11e).



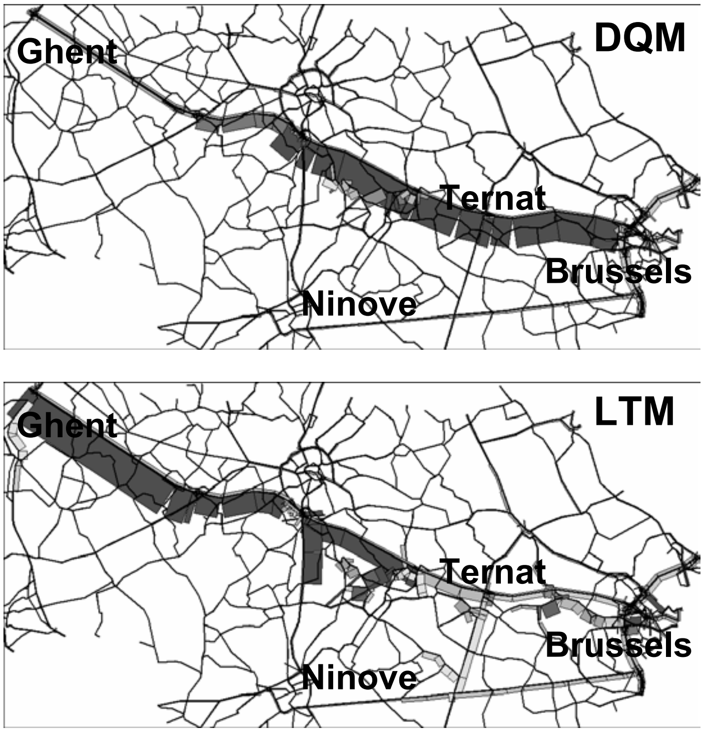
(a) 7 u 30


Ghent
DQM
Ternat
Brussels
Ninove
Ghent
LTM
Ternat
Brussels
Ninove

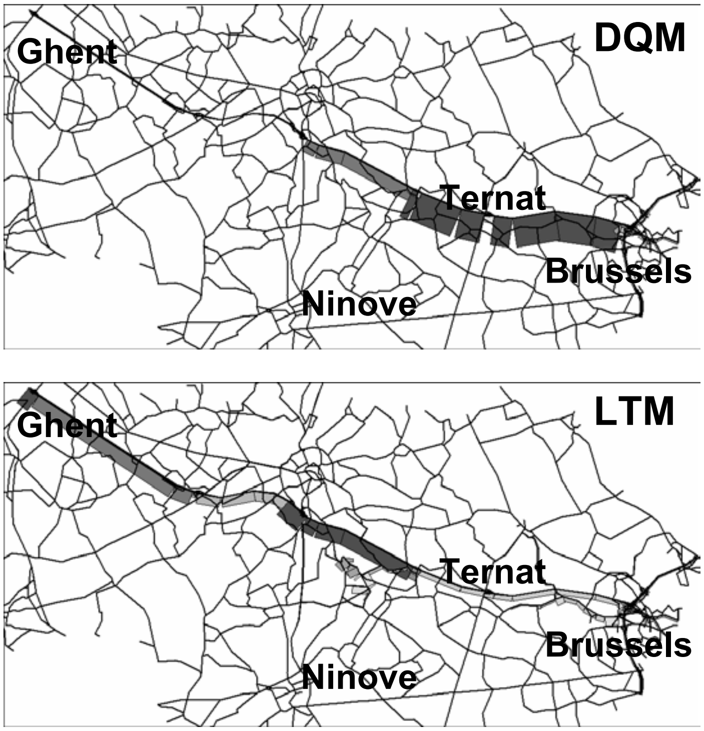
(b) 8 u 15


Ghent
DQM
Ternat
Brussels
Ninove
Ghent
LTM
Ternat
Brussels
Ninove

(c) 9 u 00


Ghent
DQM
Ternat
Brussels
Ninove
Ghent
LTM
Ternat
Brussels
Ninove

(d) 9 u 45


Ghent
DQM
Ternat
Brussels
Ninove
Ghent
LTM
Ternat
Brussels
Ninove

(e) 10 u 30

Figure 7.11: Queue propagation throughout time in DQM (on top) and LTM (below) in a morning rush with incident

7.2.5. Affligem-Brussels network: Inputs

The sub-network Affligem-Brussels cuts a part out of the Ghent-Brussels network to mainly focus on two different routes between Affligem and Brussels: Route p_2 takes travelers along the secondary road between Affligem and Brussels. Figure 7.12 depicts three signalized intersections on this route. Route p_1 takes travelers along highway E40.

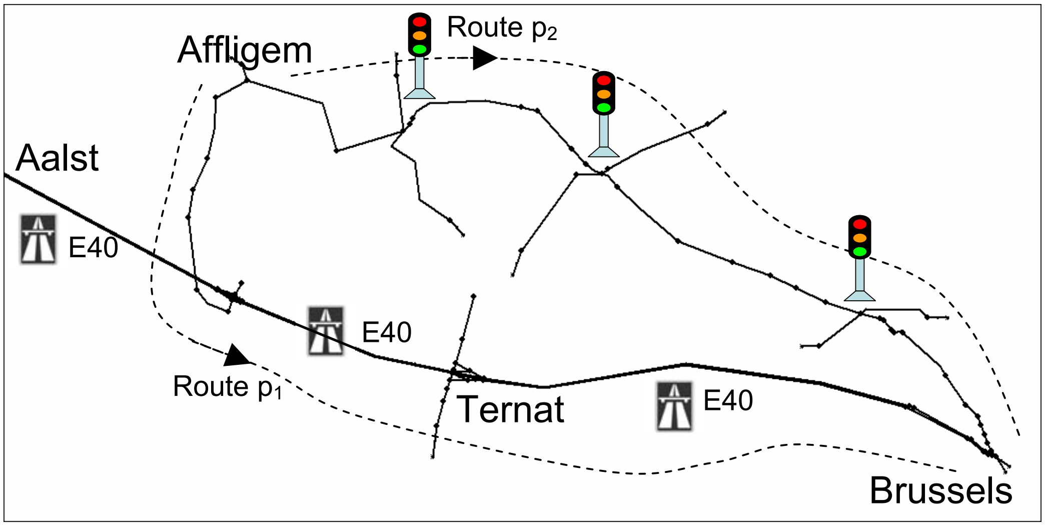


Figure 7.12: Affligem-Brussels network

As indicated in Table 7.2, only one (static) Origin-Destination matrix is used, with three Origin-Destination pairs. For the first OD-pair, Affligem – Brussels, there are two possible routes: p_1 and p_2 (cf. Figure 7.12).

Table 7.2: OD-matrix for Affligem-Brussels network

From \ To	Brussels
Affligem	1500 *veh/h*
Aalst	5000 *veh/h*
Ternat	1000 *veh/h*

Simulation parameters for this case are as follows:

- $\Delta t = 10\ s$
- $T = 2\ h$
- $k_{queue\ ,\ DQM} = 100\ veh/km/lane$
- $k_{jam\ ,\ LTM} = 150\ veh/km/lane$
- $C_{rs}^{p} = \tau_{rs}^{p}$
- $\mu = 1/0.07$
- *# iterations = 50*

7.2.6. Affligem-Brussels network: Outputs

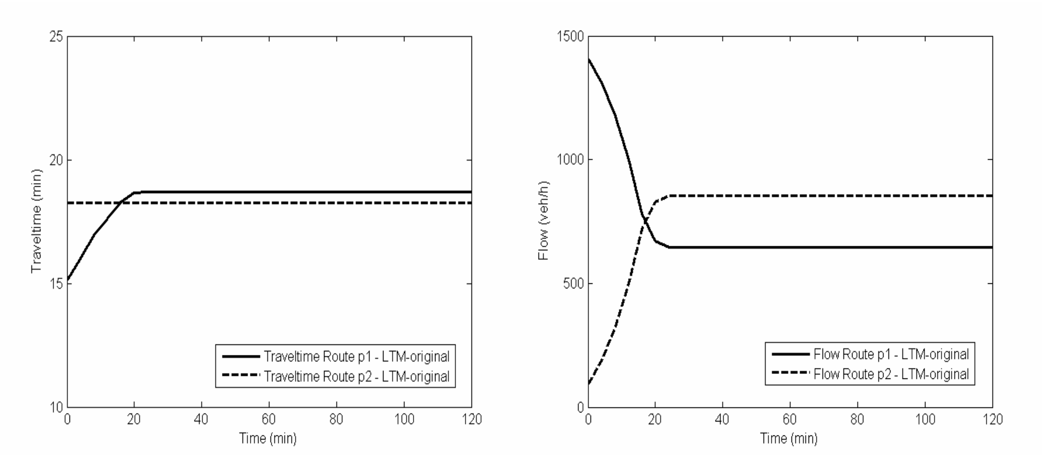
To demonstrate the impact of taking into account intersection delays on route travel times and on route choice, the distribution of the Affligem – Brussels flows over routes p_1 and p_2 is compared in LTM-original (i.e. without taking into account intersection delays) and in LTM-including-intersection-delays.

The version of LTM-including-intersection-delays where capacity constraints and intersection delays are taken into account implicitly, is not yet available as DNL model within INDY. In the version that is used here, capacity constraints and intersection delays are simulated explicitly. The drawbacks of this approach in DTA context (highest possible time step equals greatest common divisor of all green times; high-frequently fluctuating travel times due to stage alternations) can in this simple theoretical case be avoided by choosing fixed and equal cycle times for all intersections, by choosing fixed green times that are a multiple of the time step, and by averaging travel times over this fixed cycle length. Note that this operation is not applicable in large scale practical networks.

Figures 7.13a and 7.14a compare travel times throughout time in LTM-original and LTM-including-intersection-delays. Figures 7.13b and 7.14b compare the corresponding flow distribution over routes p_1 and p_2 throughout time.

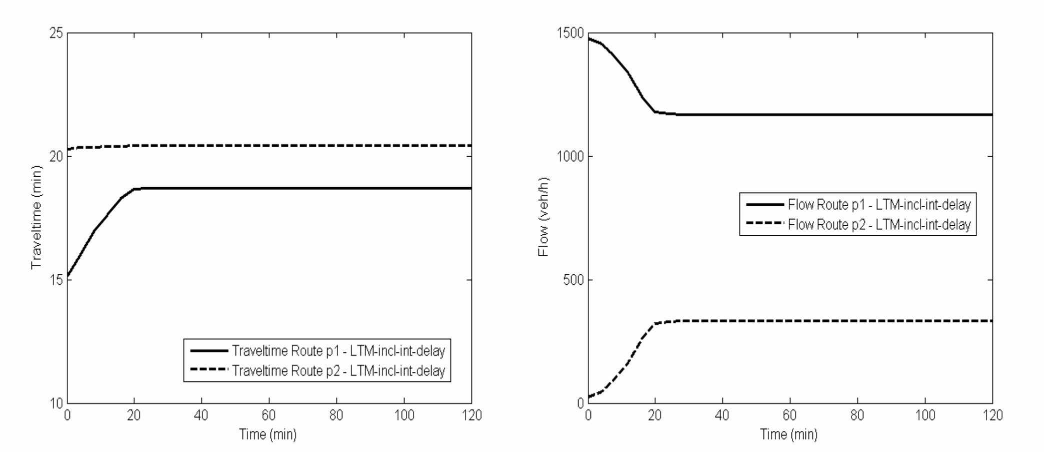
Initially, at time $t = 0$, travel times on route p_1 are in both models substantially shorter compared to travel times on route p_2 (cf. Figures 7.13a and 7.14a). Route p_1 is far more popular (cf. Figures 7.13b and 7.14b).

However, travel times on route p_1 immediately start to increase. Highway E40 between Ternat and Brussels has insufficient capacity to accommodate both flows Aalst - Brussels and Ternat – Brussels (cf. Figure 7.12). Therefore, the merge of highway E40 and the on-ramp in Ternat will act as a bottleneck. A queue develops behind this bottleneck, from Ternat towards Aalst. Travel times on route p_1 increase and a growing number of travelers switches to route p_2. At time $t = 20\ min$, the queue on highway E40 spilled back from the on-ramp in Ternat to the previous on-ramp below Affligem. As from that time, travel times on route p_1 stay constant.



(a) (b)

Figure 7.13: Travel times (a) and route flows (b) in LTM-original



(a) (b)

Figure 7.14: Travel times (a) and route flows (b) in LTM-including-intersection-delays

For route p_2, travel times are invariant in LTM-original, since flows are smaller than capacity. In LTM-including-intersection-delays, travel times on route p_2 slightly increase as the flow on this route increases (cf. Figure 7.14a). Higher flows yield higher intersection delays, however the effect is in this case small, since flows are far away from saturation.

Figure 7.13a shows that after 20 minutes of simulation in LTM-original ($t > 20\ min$), travel times on route p_1 have become higher than those on route p_2. The majority of travelers chooses route p_2. In LTM-including-intersection-delays however, travel times on route p_1 always stay smaller than those on route p_2. Route p_1 always stays the more popular route. Taking into account intersection delays appears to significantly influence route travel times and flow distribution over the different routes.

One can verify that the results in Figures 7.13 and 7.14 correspond to a Stochastic Dynamic User Equilibrium (SDUE). Figure 7.15 depicts the used multinomial Logit function with scale factor $\mu = 1/0.07$. The probability of choosing route p_1 is set as a function of the difference between route costs, i.e. the difference between route travel times ($C^{p2} - C^{p1} = \tau^{p2} - \tau^{p1}$).

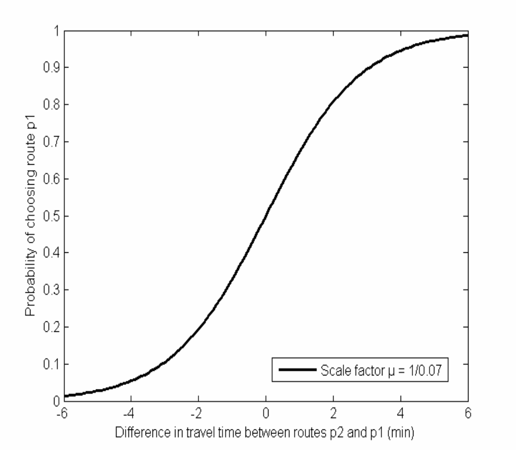


Figure 7.15: Multinomial Logit function

For all time steps, the proportion of vehicles choosing route p_1 corresponds to what the Logit function would predict based on the difference in route costs, using scale factor $\mu = 1/0.07$. Therefore, the solution corresponds to a SDUE for all time steps.

7.2.7. Conclusions of the Ghent-Brussels case study

This case shows the practical feasibility of LTM in large scale networks. For the Ghent-Brussels network with about 12000 routes, one DNL run of *1 hour* takes about *1 minute* of computing time on a Pentium IV-3 GHz (dual core) computer with 2 GB RAM.
Secondly, different queue propagation characteristics in LTM and DQM are shown to substantially influence queue locations in this large scale network. Compared to DQM, queues in LTM are less dense, they spill back faster through the network, and they dissipate in the opposite direction (from the head to the tail of the queue) in case the queue dissipates due to a bottleneck becoming inactive.
And finally, this case also shows that taking into account intersection delays significantly affects route travel times and route choice in a Stochastic Dynamic User Equilibrium.

8

CONCLUSIONS

In this final chapter, general conclusions of this thesis research are presented and some recommendations for further research are formulated. First, an overview of the properties of the developed model is given.

8.1. Properties of the developed model

In this thesis, we developed the Link Transmission Model (LTM) for DNL which has the following characteristics:

- It is suited for use in macroscopic simulation-based DTA models
- It provides a realistic representation of traffic- and queue propagation in traffic networks
- It is compatible with existing node models
- It provides a realistic representation of flow restrictions and delays at intersections
- It considers multi-commodity traffic flows, where vehicles are disaggregated by route
- It has a high computational efficiency and numerical accuracy
- It is suited for DNL in practical large scale networks with both motorways and urban regions

8.2. General conclusions

In the overview of DTA and DNL approaches (Chapter 2), it is shown how LTM with its above-mentioned characteristics, fits in and contributes to the field of DTA and DNL modeling.
The ultimate objective of this thesis was the development of a computationally efficient algorithm for DNL in large scale networks, where both traffic propagation on network links (*consistent with kinematic wave theory*) and traffic dynamics at intersections (*consistent with queuing theory*) are described in a realistic way.
The link model, node models and the intersection delay model were developed to attain this objective.

8.2.1. Link model

LTM propagates traffic on its links as assumed in kinematic wave theory, which provides substantially more realism in the representation of queue-propagation (blocking back) and queue-dissipation, compared to existing approaches in macroscopic DTA models (cf. Section 2.2.2.3).
In two case studies, it is shown that LTM's different queue-propagation characteristics (different queue densities, different queue lengths, different queue spillback speeds, different queue dissipation type) significantly influence queue locations and route travel times.

Kinematic wave theory is based on a combination of the conservation of vehicles principle and the assumption that there exists a functional relation between traffic flow q and density k, also known as the fundamental diagram of traffic flow.
Based on Newell's simplified method to determine cumulative vehicle numbers directly from initial and boundary conditions, an efficient numerical procedure to determine Sending and Receiving flows of network links is developed, which constitutes the first step of the LTM solution algorithm.
Time steps with which the algorithm is updated can be as large as (free-flow) link travel times, which substantially contributes to the algorithm's high computational efficiency.

8.2.2. Node models

At signalized and unsignalized intersections, LTM considers flow restrictions for different turning movements as assumed in queuing theory. Urban cross node models are developed to determine transition flows at network nodes, which constitutes the second step of the LTM solution algorithm.
LTM indirectly considers average flow restrictions, without explicitly simulating green and red stages (for signalized intersections) or gaps in conflicting priority streams (for unsignalized intersections).

This approach enables the use of large time steps in the solution algorithm, and it directly yields average expected travel times that travelers use for their route choice.

Furthermore, node models that represent basic motorway junctions, such as merge and diverge nodes, are easily integrated within LTM.

8.2.3. Intersection delay model

LTM implicitly accounts for average intersection delays, where high-frequent fluctuations due to a repeated process of queue-formation and queue-dissipation are filtered out. These average intersection delays are estimated by extensively validated formulae that are generally based on queuing theory.

This approach (implicitly considering average intersection delays) allows for the use of large time steps in the LTM solution algorithm, which significantly enhances its computational efficiency, and it directly yields average expected delays that travelers take into account for their route choice.

The Ghent-Brussels case study shows that taking into account intersection delays significantly affects route travel times and route choice in practical networks.

Point-Queues are introduced to realize intersection delays in an indirect way. The length of the P-Q should be such that vehicles experience their intended average delay by passing through this P-Q. Point-Queues only represent temporary queues due to stage alternations or priority rules. They occur next to kinematic wave queues that are due to over-saturation.

The introduction of P-Q’s does not affect the validity of kinematic wave traffic flow propagation on network links. Reversely, compliance with kinematic wave theory generally does not influence P-Q characteristics either. Intended intersection delays are generally realized and a linear transition occurs between different states with different values of intersection delays.

Only for traffic flows that are restricted by downstream boundary conditions or for high-frequently changing traffic flows, compliance with kinematic wave theory might cause P-Q lengths to slightly differ from their intended values that are derived from queuing theory.

8.3. Further research

8.3.1. Disaggregation of traffic flows to include Multiple User Classes (MUC)

In LTM, traffic flows are disaggregated by route (multi-commodity flows) to be able to assign vehicles to specific routes.
Traffic flows may further be disaggregated by vehicle type, to consider vehicles with different vehicle characteristics and different types of driver behavior, i.e., to deal with Multiple User Classes (MUC) in terms of vehicle type (e.g. trucks, buses, passenger cars). It would be interesting to explore the possibilities for implementation of Logghes (2003) MUC approach of kinematic wave theory.

Furthermore, traffic flows may also be disaggregated by choice behavior or by response to information, to be able to simulate the result of en-route trip-maker decisions, i.e. to make the DNL model suitable for use in en-route DTA models.

8.3.2. Development of a solution method to consider stochastic intersection delays

In this thesis, a method to include intersection delays is elaborated for deterministic intersection delays at signalized intersections. Deterministic intersection delays arise from the assumption of uniform arrival rates at intersections. These delays are independent on the time of operation and on the history of the system.
Stochastic intersection delays account for the random nature of arrivals. These delays are dependent on the time of operation and on the history of the system.
Solution methods for stochastic intersection delays therefore require keeping track of the times of operation of traffic flows. Such methods are not elaborated in this thesis; they are subject of further research.

8.3.3. Model validation

Both kinematic wave- and queuing theory have extensively been validated in literature and are not re-validated in this thesis. The issue is rather whether or not the developed LTM solution algorithm combines both theories in a way that they don't affect each other's validity.
While the validity of kinematic wave traffic flow propagation is always ensured, specific circumstances have been pointed out, for which compliance with kinematic wave theory may cause intersection delays to slightly differ from the intended values that are derived from queuing theory. Though these circumstances are rather exceptional, quantification of possible deviations would be desirable.

8.3.4. Evaluation of computational efforts

In this thesis, we claim the LTM solution algorithm to be computationally efficient. Different indications that support this claim include:

- computation times are proportional to the length of the time step Δt, and LTM can generally be run with large times steps (though this depends on network characteristics)

- for the same level of accuracy, LTM requires considerably less computational effort compared to CTM (cf. Section 4.8)

- computation times no longer depend on the amount of vehicles in the network, as was the case in microscopic and mesoscopic simulation models. Therefore, it can reasonably be assumed that the computational efficiency is substantially higher compared to those models

- the practical feasibility of LTM in large scale networks is demonstrated and quantified in the Ghent-Brussels case study (cf. Section 7.2.4)

However, it would be interesting to more generally quantify LTM's computational complexity and to compare it with other state-of-the-art models for DNL.

REFERENCES

Akcelik, R. and R.J. Troutbeck (1991) *Implementation of the Australian Roundabout Analysis Method in SIDRA*, In: U. Brannolte (ed.) Highway Capacity and Level of Service, Proceedings of the International Symposium on Highway Capacity, Karlsruhe, A.A. Balkema, Rotterdam, 17-34.

Barcelo, J. (2002) *Microscopic traffic simulation: A tool for the analysis and assessment of its systems*. In: Proceedings of the Half Year Meeting of the Highway Capacity Committee. Lake Tahoe.

Ben-Akiva, M., Koutsopoulos H.N., Mishalani R. (1998) *DynaMIT: A simulation-Based System for Traffic Prediction*. Paper presented at the DACCORD Short Term Forecasting Workshop, Delft. (See also its.mit.edu).

Bliemer, M.C.J. (2001) *Analytical Dynamic Traffic Assignment with Interacting User-Classes: Theoretical Advances and Applications using a Variational Inequality Approach*, T2001/1, TRAIL Thesis Series, Delft University Press, The Netherlands.

Bliemer M.C.J and P.H.L. Bovy (2003) Quasi-variational inequality formulation of the multiclass dynamic traffic assignment problem, *Transportation Research B, 37, 501-519.*

Bliemer, M.C.J., H.H. Versteegt and R.J. Castenmiller (2004) *INDY: A New Analytical Multiclass Dynamic Traffic Assignment Model*, Proceedings of the TRISTAN V conference, Guadeloupe.

Bliemer, M.C.J. (2005) *INDY 2.0 Model Specifications*, Delft University of Technology working report.

Bliemer, M.C.J. (2006) *Dynamic queuing and spillback in an analytical multiclass dynamic traffic assignment model*. Proceedings of the First International Symposium on Dynamic Traffic Assignment, Leeds, UK.

Bliemer, M.C.J. (2007) Dynamic Queuing and Spillback in an Analytical Multiclass Dynamic Network Loading Model. Accepted for publication in *Transportation Research Record*, Natl. Acad. Sciences, Washington, DC.

Brilon, W. and T. Miltner (2005) *Capacity and delays at intersections without traffic signals*, Proc. 84th Annual Meeting of the TRB, Washington, DC.

Carey, M. (1986) A constraint Qualification for a Dynamic Traffic Assignment Model. *Transportation Science 20, 55-58.*

Carey, M. (1987) Optimal Time Varying Flows On Congested Networks. *Operations Research 35(1), 58-69.*

Casas, J. (2004) *Estimation of the transport demand for real-time applications*. Ph.D. Thesis, Universitat Politecnica de Catalunya, Barcelona, Spain.

Chabini, I. (2001) Analytical Dynamic Network Loading Problem: Formulation, Solution Algorithms, and Computer Implementations. *Transportation Research Record, 1771*, TRB, National Research Council, Washington, D.C., 191-200.

Chen, H.-K. and Hsueh C.F. (1998) A Model and an Algorithm for the Dynamic User Optimal Route Choice Problem. *Transportation Research 32B(3), 219-234.*

Daganzo, C.F. (1994) The cell-transmission model. A dynamic representation of highway traffic consistent with the hydrodynamic theory. *Transportation Research 28B, 269-288.*

Daganzo, C.F. (1995a) The cell transmission model, part II: network traffic, *Transportation Research 29B, 79-94.*

Daganzo, C.F. (1995b) A finite difference approximation of the kinematic wave model of traffic flow, *Transportation Research 29B, 261-276.*

Daganzo, C.F. (1997) *Fundamentals of transportation and traffic operations,* Pergamon-Elsevier, Oxford, UK.

Daganzo, C.F. (1999) *The lagged cell-transimission model*, Proceedings of the 14th ISTTT, Ed. A. Ceder.

Durlin, T. and V. Henn (2005) *A Delayed Flow Intersection Model for Dynamic Traffic Assignment*. Advanced OR and AI Methods in Transportation, Proc. 10th EWGT Meeting and 16th Mini-EURO Conference, Poznan, Poland, 441-449, Publishing House of Poznan University of Technology.

Florian, M., M. Mahut and N. Tremblay (2006) *A simulation-based Dynamic Traffic Assignment Model: Dynameq*. In Proceedings of the First International Symposium on Dynamic Traffic Assignment, Leeds, UK.

Friesz, T.L., Bernstein D., Mehta N.J., Tobin R.L. and Ganjalizadeh S. (1989) Dynamic Network Assignment Considered as a Continuous Time Optimal Control Problem. *Operations Research 37(6) 893-901.*

Friesz, T.L., D. Bernstein and T.E. Smith, et al. (1993) A variational inequality formulation of the dynamic network user equilibrium problem. *Operations Research 41(1), 179-191.*

Green, G. (1828) *An essay on the application of mathematical analysis to the theories of electricity and magnetism*, http://en.wikipedia.org/wiki/George_Green, accessed April 10, 2007.

Greenshields, B.D. (1934) *A study of traffic capacity*, Proceedings of the highway research board 14, 448-477.

Henn, V. (2003) *A wave-based resolution scheme for the hydrodynamic LWR traffic flow model*. In "Traffic and granular flows", Delft Technical University, The Netherlands.

Janson, B.N. (1991a) Convergent Algorithm for Dynamic Traffic Assignment. *Transportation Research Record 1328, 69-80.*

Janson, B.N. (1991b) Dynamic traffic assignment for urban road networks. *Transportation Reserach 25B(2/3), 143-161.*

Jin, W.-L. (2002) *Development and validation of kinematic wave traffic flow models for road networks*, University of California Transportation Center Dissertation Grant, University of California, Davis, CA.

Jin, W.L. and Zhang H.M. (2003) On the distribution schemes for determining flows through a merge. *Transportation Research 37B, 521-540.*

Kuwahara, M. and Akamatsu T. (2001) Dynamic user optimal assignment with physical queues for a many-to-many OD pattern. *Transportation Research* 35B, 461-479.

Lebacque, J.P. (1996) *The Godunov scheme and what it means for first order traffic flow models*. Proceedings of the 13th ISTTT. Ed. J.B. Lesort.

Lighthill, M.J. and Whitham G.B. (1955) *On kinematic waves (II): A theory of traffic flow on long crowded roads*. Proc. Roy. Soc., A. 229, 281-345.

Lo, H.K. (1999) A dynamic traffic assignment formulation that encapsulates the cell transmission model. In A. Ceder (ed.) *Transportation and Traffic Theory, Pergamon, Oxford, 327-350.*

Lo, H.K. and W.Y. Szeto (2002) A cell-based variational inequality formulation of the dynamic user optimal assignment problem. *Transportation Research B(36), 421-443.*

Logghe, S. (2003) *Dynamic modeling of heterogeneous vehicular traffic*. Ph.D. Thesis, Katholieke Universiteit Leuven, Leuven, Belgium.

Logghe, S. and I. Yperman (2003) *De dynamica van congestie en prijsbeleid*. Proc. 30th Coll. Vervoersplanologisch Speurwerk, Antwerp, Belgium.

Maerivoet, S. (2006) *Modelling traffic on motorways: state-of-the-art, numerical data analysis, and dynamic traffic assignment*. Ph.D. Thesis, Katholieke Universiteit Leuven, Leuven, Belgium.

Mahmassani, H.S., Abdelghany A.F., Huynh N., Zhou X., Chiu Y-C. and Abdelghany K.F. (2001) *DYNASMART-P User's Guide*. Technical Report STO67-85-PIII, Centre for Transportation Research, University of Texas at Austin.

Mahut, M. (2000) *Discrete flow model for dynamic network loading*. Ph.D. Thesis, Université de Montreal, Montreal, Canada.

Merchant, D.K. and Nemhauser G.L. (1978a) A model and an algorithm for the Dynamic Traffic Assignment Problems." *Transportation Science 12(3), 1978, 183-199.*

Merchant, D.K. and Nemhauser (1978b) G.L. Optimality Conditions for a Dynamic Traffic Assignment Model. *Transportation Science 12(3), 1978, 200-207.*

Messmer, A. and M. Papageorgiou (1990) METANET: A macroscopic simulation program for motorway networks. *Traffic Engineering and Control, volume 31, 466-470.*

Miltner, T. (2003) Verkehrsqualitaet an vorfahrtgeregelten Innerortsknotenpunkten (Level of service at unsignalized urban intersections) *Institute for Transportation and Traffic Engineering, Ruhr-University Bochum, no. 27.*

Nagel, K., et al. (1998) *TRANSIMS traffic flow characteristics*. Los Alamos National Laboratory Unclassified Report (LA-UR) 97-3530.

Nagurney, A. (1993) *Network Economics: A variational inequality approach*. Kluwer Academic Publishers, Norwell, Massachusetts, USA.

Newell, G.F. (1993) A simplified theory of kinematic waves in highway traffic, Part I: General theory, Part II: Queuing at freeway bottlenecks, Part III: Multi-destination flows, *Transportation Research 27B, 281-313.*

Peeta, S. and Ziliaskopoulos A.K. (2001) Foundations of Dynamic Traffic Assignment: The Past, the Present and the Future. *Networks and Spatial Economics, 2001 (1), 233-265*.

PTV (2005) *PTV Vision software suite (VISUM/VISEM/VISEVA/VISSIM)*. PTV AG – Traffic Mobility Logistics.

Quadstone Limited (2000) *Paramics-online v3 – System overview*. Technical report, Quadstone Limited, 16 Chester Street, Edinburgh EH3 7RA, Scotland (http://www.paramics-online.com).

Ran, B. and Boyce D.E. (1996) A link-based Variational Inequality Formulation of Ideal Dynamic User Optimal Route Choice Problem. *Transportation Research 4C(1), 1-12*.

Ran, B. and D.E. Boyce (1996) *Modelling dynamic transportation networks: an intelligent transportation systems oriented approach*, 2nd revised edn. Springer-Verlag, New York.

Ran, B., D.E. Boyce, and L.J. LeBlanc (1993) A new class of instantaneous dynamic user-optimal traffic assignment models. *Operations Research 41(1)*.

Richards, P.I. (1956) Shockwaves on the highway. *Operations Research 4, 42-51*.

Roels, G. and G. Perakis (2004) *An analytical Model for Traffic Delays*. Proceedings of the TRISTAN V conference, Guadeloupe, France.

Rouphail, N., Tarko, A. and Li, J. (2000) *Traffic flow at signalized intersections*, Lieu H. Revised Monograph of Traffic Flow Theory, update and expansion of the Transportation Research Board (TRB) special report 165, "Traffic Flow Theory", published in 1975.

Szeto, W.Y. (2003) *Dynamic traffic assignment: formulations, properties, and extensions*. Ph.D. Thesis, Hong Kong University of Science and Technology, Hong Kong.

Szeto, W.Y. and H.K. Lo (2004) A cell-based simultaneous route and departure time choice model with elastic demand. *Transportation Research B(38), 593-612*.

Szeto, W.Y. and H.K. Lo (2006) Dynamic traffic assignment: properties and extensions. *Transportmetrica, Vol.2, No.1, 31-52*.

Taale, H. and M. Westerman (2004) *The application of sustainable Traffic Management in the Netherlands,* Proceedings of the European Transport Conference 2005, October 3-5, 2005, Strasbourg, France, Association for European Transport.

Tampère, C.M.J. (2004) *Human-Kinetic Multiclass Traffic Flow Theory and Modelling: With Application to Advanced Driver Assistance Systems in Congestion,* T2004/11, TRAIL Thesis Series, The Netherlands.

Tampère, C.M.J., J.E. Stada, L.H. Immers, E. Peetermans and K. Organe (2007) *Methodology for Identifying Vulnerable Sections in a National Road Network*. Proc. 86th Annual Meeting of the TRB, Washington, DC.

Taylor, N.B. (2003) The CONTRAM Dynamic Traffic Assignment Model, *Networks and Spatial Economics, Vol. 3, Issue 3, 297-322.*

TRB (2000) *Highway Capacity Manual. Special report 209.* Washington D.C., National Research Council.

Troutbeck, R. J., Brilon W. (2000). *Unsignalized intersection theory*, Lieu H. Revised Monograph of Traffic Flow Theory, update and expansion of the Transportation Research Board (TRB) special report 165, "Traffic Flow Theory", published in 1975.

Viti, F. and H.J. Van Zuylen (2004) Modeling Queues At Signalized Intersections. *Transportation Research Record* No. 1883, 68-77.

Viti, F. (2006) *The Dynamics and the Uncertainty of Delays at Signals,* T2006/7, TRAIL Thesis Series, Delft University Press, The Netherlands.

Waller, S.T., and S.V. Ukkusuri (2003) A linear programming formulation for the user optimal dynamic traffic assignment problem. *Transportation Science*. (In review. Submitted February 2003).

Wardrop, J. G. (1952) *Some theoretical aspects of road traffic research*, Proceedings, Institution of Civil Engineers, Part II, Vol.1, 325-378.

Webster, F.V. (1958) *Traffic Signal Settings*. Road Research Laboratory Technical Paper No. 39, HMSO, London, UK.

Wie, B.W. (1991) *Dynamic analysis of user optimized network flows with elastic travel demand*. Proceedings of the 70th TRB Annual Meeting, Washington, D.C.

Wie, B.W., Tobin R.L., Friesz T.L. and Bernstein D. (1995) A Discrete Time, Nested Cost Operator Approach to the Dynamic Network User Equilibrium Problem. *Transportation Science 29(1), 79-92.*

Yang, Q. (1997) *A simulation laboratory for evaluation of dynamic traffic management systems*. Ph.D. Thesis, Massachusetts Institute of Technology, Boston, USA.

Yperman, I., Logghe S. and Immers L.H. (2005a) *Strictly Pareto-improving congestion pricing in a multi-destination network*. Proc. 84th Annual Meeting of the TRB, Washington, DC.

Yperman, I., S. Logghe and L. Immers (2005b) *Dynamic congestion pricing in a network with queue spillover.* Proc. 12th World Congress on Intelligent Transportation Systems, San Francisco.

Yperman, I., S. Logghe and L. Immers (2005c) *How congestion pricing can increase traffic volumes*. Proc. 8th Conference of the Network on Eur. Communications and Transportation Act. Research, Las Palmas de GC, Gran Canaria, Spain.

Yperman, I., S. Logghe and L. Immers (2005d) *The Link Transmission Model: An Efficient Implementation of the Kinematic Wave Theory in Traffic Networks*. Advanced OR and AI Methods in Transportation, Proc. 10th EWGT Meeting and 16th Mini-EURO Conference, Poznan, Poland, 122-127, Publishing House of Poznan University of Technology.

Yperman, I., S. Logghe, C. Tampère and L. Immers (2006) *The multi-commodity Link Transmission Model for Dynamic Network Loading*. Proc. 85th Annual Meeting of the Transportation Research Board, Washington, DC.

Ziliaskopoulos, A.K. (2000) A Linear Programming Model for the Single Destination System Optimum Dynamic Traffic Assignment Problem. *Transportation Science 34(1), 37-49*.

NEDERLANDSTALIGE SAMENVATTING

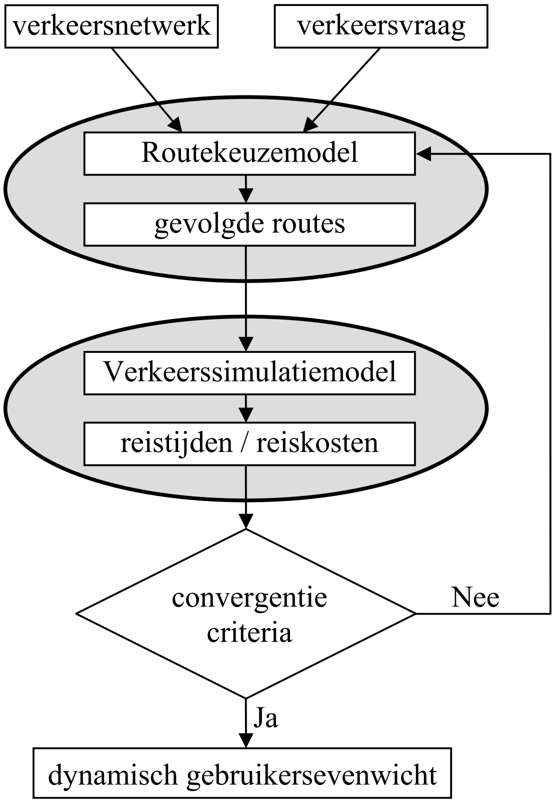
Het Link Transmissie Model voor Verkeerssimulaties in Netwerken

1. Inleiding

Dynamische verkeerstoedelingsmodellen worden gebruikt om de impact van infrastructuuraanpassingen in verkeersnetwerken te voorspellen en om de effecten van informatieverstrekking en verkeersbeheersingsmaatregelen in te schatten.
Deze modellen hebben twee fundamentele compenten: een routekeuzemodel (i.e. het toedelingsmechanisme) waarin reizigers worden toegedeeld aan bepaalde routes, en een dynamisch verkeerssimulatiemodel, waarin het verkeer gesimuleerd wordt langs die bepaalde routes.
In deze thesis wordt een dynamisch verkeerssimulatiemodel ontwikkeld.
Verkeerssimulatiemodellen bepalen reistijden en reiskosten in een verkeersnetwerk, gegeven de gevolgde routes van alle reizigers. Deze routes worden bepaald in het routekeuzemodel op basis van een gegeven verkeersvraag en op basis van de ervaren reiskosten.

In een evenwichtstoedeling worden deze twee fundamentele componenten, het routekeuzemodel en het verkeerssimulatiemodel achtereenvolgens doorlopen in een iteratieproces, totdat convergentie optreedt naar een dynamisch gebruikersevenwicht. Figuur S.1. schetst dit proces.

In een dynamisch gebruikersevenwicht hebben alle reizigers die tijdens een bepaald tijdsinterval vertrekken vanuit een bepaald herkomstpunt naar een bepaald bestemmingspunt, gelijke en minimale reiskosten.
Evenwichtstoedelingen worden vaak gebruikt voor planningsdoeleinden, om de verkeersomstandigheden op een ‘typische dag’ te beschrijven en om de effecten van verkeersbeheersingsmaatregelen (zoals toeritdosering, coördinatie van verkeerslichtenregelingen, tolheffing,...) in te schatten vooraleer deze geïmplementeerd worden in de praktijk.


verkeersnetwerk
verkeersvraag
Routekeuzemodel
gevolgde routes
Verkeerssimulatiemodel
reistijden / reiskosten
convergentie criteria
Nee
Ja
dynamisch gebruikersevenwicht

Figuur S.1: Structuur van een evenwichtstoedelingsmodel

2. Eigenschappen van het Link Transmissie Model

Het Link Transmissie Model (LTM) is een macroscopisch verkeerssimulatiemodel, waarin voertuigen zich verplaatsen als een continuum. Verkeersstromen worden realistisch gemodelleerd in grote praktische netwerken die zowel snelwegen als stedelijke regio’s omvatten. Het LTM heeft volgende kenmerken:

- Verkeersstromen planten zich voort op schakels zoals dat gebeurt in de eerste orde kinematische golftheorie. Het file-opbouw en file-afbouw proces in LTM sluit nauwer aan bij de realiteit dan in state-of-the-art macroscopische verkeerstoedelingsmodellen.
- Verkeersstromen op stedelijke knooppunten gedragen zich zoals in de wachtrijtheorie. Lokale capaciteitsrestricties en knoopvertragingen worden gedetailleerd weergegeven.
- Het LTM oplossingsalgoritme is rekenefficiënt, waardoor verkeersstromen in grote netwerken gesimuleerd kunnen worden in een beperkte rekentijd.
- Voertuigen worden gedisaggregeerd naar hun route. Tijdens hun reis door het netwerk worden de routes van de voertuigen voortdurend bijgehouden, zodanig dat de informatie uit het routekeuzemodel ook echt gebruikt kan worden.

3. LTM Oplossingsalgoritme

3.1. Definities

Verkeersnetwerken worden opgebouwd uit homogene schakels *a*, die afgebakend worden door een startpunt x_a^0 en een eindpunt x_a^L. Deze schakels hebben een willekeurige lengte L_a en zijn met elkaar verbonden via knooppunten *n*. Knooppunten hebben geen fysieke lengte en fungeren als uitwisselingslocatie van verkeersentiteiten. Elk knooppunt heeft een aantal toekomende schakels *i* en een aantal vertrekkende schakels *j*.

Het cumulatief aantal voertuigen $N(x,t)$ is het totaal aantal voertuigen dat locatie x voorbijgekomen is op tijdstip *t*. Voor elke tijdsstap bepaalt LTM het cumulatief aantal voertuigen dat het start- en eindpunt (x_a^0 en x_a^L) van elke schakel *a* voorbijgekomen is. Pas wanneer voertuigen de schakel verlaten hebben, wordt hun reistijd bepaald.

De zendstroom $S_i(t)$ van schakel *i* op tijdstip *t* is de maximum hoeveelheid voertuigen die de schakel zouden kunnen verlaten tijdens tijdsinterval $[t\ ,\ t+\Delta t]$, als deze schakel zou uitmonden in een reservoir met een oneindige capaciteit.
De ontvangstroom $R_j(t)$ van schakel *j* op tijdstip *t* is de maximum hoeveelheid voertuigen die de schakel zouden kunnen binnenstromen tijdens tijdsinterval $[t\ ,\ t+\Delta t]$, als een oneindige verkeersvraag zou losgelaten worden op deze schakel.
De overgangsstroom $G_{ij}(t)$ is de hoeveelheid voertuigen die daadwerkelijk overgaan van link *i* naar link *j* tijdens tijdsinterval $[t\ ,\ t+\Delta t]$. Het LTM oplossingsalgoritme verdeelt de totale simulatieperiode in tijdsintervallen Δt. Voor elk tijdsinterval Δt omvat het algoritme drie stappen:

3.2. LTM oplossingsalgoritme

Voor elk tijdsinterval Δt,
Voor elk knooppunt n,

Stap 1: Bepaal de zendstroom S_i aan het afwaartse schakeleinde (x_i^L) voor elke toekomende schakel $i \in I_n$ en bepaal de ontvangstroom R_j aan het opwaartse schakeleinde (x_j^0) voor elke vertrekkende schakel $j \in J_n$.
I_n (J_n) is de verzameling van toekomende (vertrekkende) schakels van knooppunt n.

Stap 2: Bepaal de overgangsstromen $G_{ij}(t)$ van de toekomende schakels $i \in I_n$ naar de vertrekkende schakels $j \in J_n$, i.e. bepaal welk aandeel van de zend- en ontvangstromen daadwerkelijk verzonden en ontvangen wordt.

Stap 3: Pas het cumulatief aantal voertuigen $N(x,t)$ aan voor het afwaartse schakeleinde (x_i^L) van elke toekomende schakel $i \in I_n$ en voor het opwaartse schakeleinde (x_j^0) van elke vertrekkende schakel $j \in J_n$:

$$N(x_i^L, t+\Delta t) = N(x_i^L, t) + \sum_{j \in J_n} G_{ij}(t) \qquad \textit{voor alle } i \in I_n$$

$$N(x_j^0, t+\Delta t) = N(x_j^0, t) + \sum_{i \in I_n} G_{ij}(t) \qquad \textit{voor alle } j \in J_n$$

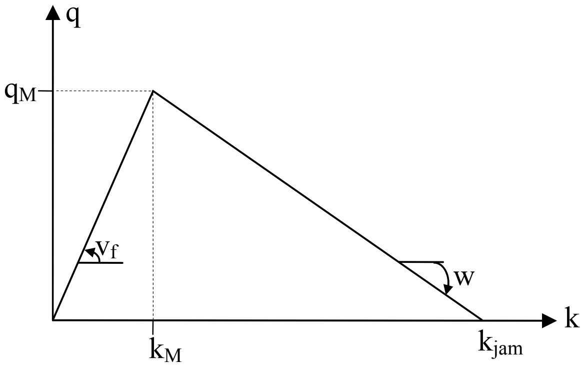
4. Schakelmodel

In LTM planten verkeersstromen zich voort op schakels zoals dat gebeurt in de kinematische golftheorie. Verkeer wordt gekenmerkt door drie macroscopische variabelen: de instensiteit q, de dichtheid k en de gemiddelde snelheid v. Tussen deze drie variabelen geldt de volgende relatie:

$$v = q/k$$

De intensiteit q is het aantal voertuigen dat een bepaalde plaats voorbijkomt tijdens een bepaald tijdsinterval dt.
De dichtheid k is het aantal voertuigen dat op een bepaald tijdsstip waargenomen wordt over een bepaald plaatsinterval dx.

Verder wordt in de kinematische golftheorie aangenomen dat er een functionele relatie bestaat tussen de intensiteit q en de dichtheid k, ook wel het fundamenteel diagram genoemd. Naar analogie met Newell (1993) wordt in LTM een driehoekig fundamenteel diagram gebruikt, zoals aangegeven in Figuur S.2.



Figuur S.2: Driehoekig fundamenteel diagram

De intensiteit q is nul voor een nuldichtheid en voor de maximale dichtheid k_{jam}. Tussenin bereikt de functie een maximale intensiteit q_M, die ook de capaciteit wordt genoemd. De stijgende tak van het fundamenteel diagram geeft verkeerstoestanden weer die zich in het 'vrij-verkeer' regime bevinden, de dalende tak stelt verkeerstoestanden in 'congestie' voor.

De kinematische golftheorie stoelt op de combinatie van dit verondersteld fundamenteel diagram en de wet van behoud van voertuigen. Deze verkeersbehoudswet beschrijft de evolutie van een verkeerstoestand over plaats en tijd:

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

Verkeerstoestanden in het 'vrij-verkeer' regime planten zich voort met snelheid v_f en verkeerstoestanden in 'congestie' verplaatsen zich met negatieve snelheid w (cf. Figuur S.2).

Uit de kinematische golftheorie worden cumulatieve voertuigaantallen afgeleid. Newell (1993) ontwikkelde een methode om het cumulatief aantal voertuigen in schakels rechtstreeks te bepalen uit begin- en randvoorwaarden. Op basis van Newell's theorie werd een efficiënte methode ontwikkeld om zend- en ontvangstromen te bepalen, de eerste stap in het LTM oplossingsalgoritme:

$$S_i(t) = \min((N(x_i^0, t + \Delta t - \frac{L_i}{v_{f,i}}) - N(x_i^L, t)), q_{M,i}\Delta t)$$

$$R_j(t) = \min((N(x_j^L, t + \Delta t + \frac{L_j}{w_j}) + k_{jam}L_j - N(x_j^0, t)), q_{M,j}\Delta t)$$

Door zend- en ontvangstromen op deze manier te bepalen zorgen we ervoor dat het verkeer op schakels zich voortplant zoals in de kinematische golftheorie.

In vergelijking met het Cel Transmissie Model (CTM), heeft LTM een hogere rekenefficiëntie. CTM is een veelgebruikt verkeersmodel met een ander numeriek oplossingsschema voor hetzelfde kinematisch golfmodel. Om dezelfde nauwkeurigheid te bereiken heeft het CTM behoorlijk meer rekentijd nodig.

5. Knoopmodellen

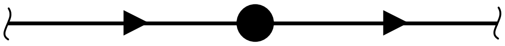
Knoopmodellen bepalen de overgangsstromen $G_{ij}(t)$ van de toekomende schakels i naar de vertrekkende schakels j van een knooppunt n, de tweede stap in het LTM oplossingsalgoritme. Knoopmodellen voldoen aan de verkeersbehoudswet. Elk knoopmodel heeft bepaalde voorrangs- en gedragsregels die beperkingen opleggen aan de overgangsstromen en die uiteindelijk bepalen welk aandeel van de zend- en ontvangstromen daadwerkelijk verzonden en ontvangen wordt.

5.1. Knoopmodellen voor kruispunten op snelwegen

Volgende knoopmodellen worden gebruikt om snelwegnetwerken te modelleren:

- Inhomogeen knooppunt

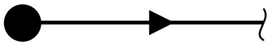
Inhomogene knooppunten worden gebruikt om een verandering in capaciteit of in snelheidslimiet te modelleren. Deze knooppunten verbinden 1 toekomende schakel met 1 vertrekkende schakel (cf. Figuur S.3).



Figuur S.3: Inhomogeen knooppunt

- Herkomstknooppunt

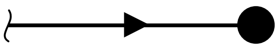
In een herkomstknooppunt wordt verkeer op het netwerk losgelaten. Deze knooppunten fungeren als voedingspunten van het verkeersnetwerk. In dergelijk knooppunt vertrekt 1 schakel (cf. Figuur S.4).



Figuur S.4: Herkomstknooppunt

- Bestemmingsknooppunt

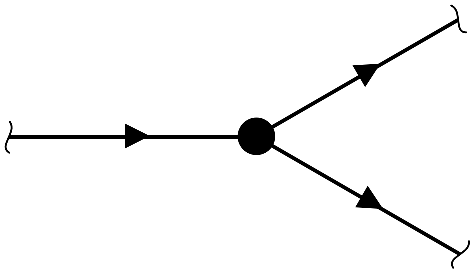
Verkeer verlaat het netwerk in een bestemmingsknooppunt. In dergelijk knooppunt komt 1 schakel toe (cf. Figuur S.5).



Figuur S.5: Bestemmingsknooppunt

- Splitsingsknooppunt

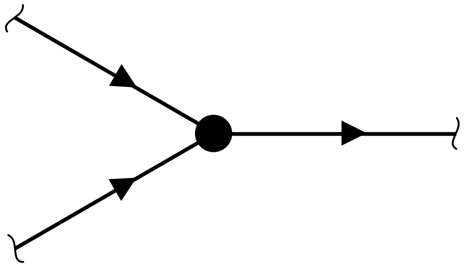
Splitsingsknooppunten worden gebruikt om afritten van een snelweg te modelleren. In een splitsingsknooppunt komt 1 schakel toe en vertrekken precies 2 schakels (cf. Figuur S.6).



Figuur S.6: Splitsingsknooppunt

- Samenvoegingsknooppunt

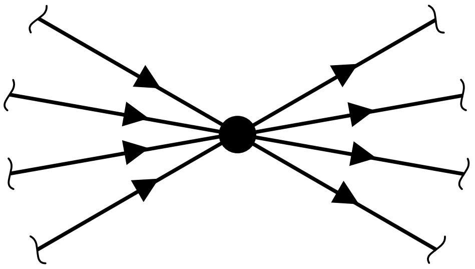
Samenvoegingsknooppunten worden gebruikt om opritten op een snelweg te modelleren. In een samenvoegingsknooppunt komen 2 schakels toe en vertrekt 1 schakel (cf. Figuur S.7).



Figuur S.7: Samenvoegingsknooppunt

5.2. Knoopmodellen voor stedelijke kruispunten

Om verkeerscondities in stedelijke netwerken te beschrijven worden in dit eindwerk knoopmodellen voor stedelijke kruispunten ontwikkeld. Kruisknooppunten verbinden een willekeurig aantal toekomende schakels met een willekeurig aantal vertrekkende schakels (cf. Figuur S.8).



Figuur S.8: Kruisknooppunt

Deze knoopmodellen voor stedelijke kruispunten leggen capaciteitsrestricties op aan de overgangsstromen van de toekomende naar de vertrekkende schakels.

Om deze capaciteitsrestricties in rekening te brengen zijn er twee mogelijke oplossingsrichtingen:

- expliciete simulatie van verkeerslichtenregelingen en hiaten in voorrangsstromen, ofwel
- impliciete beschouwing van de gemiddelde effecten van verkeerslichtenregelingen en voorrangsstromen, zonder deze expliciet te simuleren

De eerste oplossingsmethode heeft volgende nadelen in de context van verkeerstoedelingsmodellen:

- expliciete simulaties van verkeerslichtenregelingen en hiaten in voorrangsstromen vereisen een kleine simulatietijdsstap. Hierdoor neemt de rekentijd gevoelig toe.
- expliciete simulaties genereren frequent fluctuerende reistijden. Reizigers houden echter geen rekening met deze hoogfrequente fluctuaties bij het bepalen van hun route. Gemiddelde of 'te verwachten' reistijden zouden meer relevant zijn.
- expliciete simulaties resulteren in 'mogelijke' reistijden, die het resultaat zijn van een toevallige samenloop van omstandigheden. Gemiddelde of 'te verwachten' reistijden zouden meer relevant zijn.

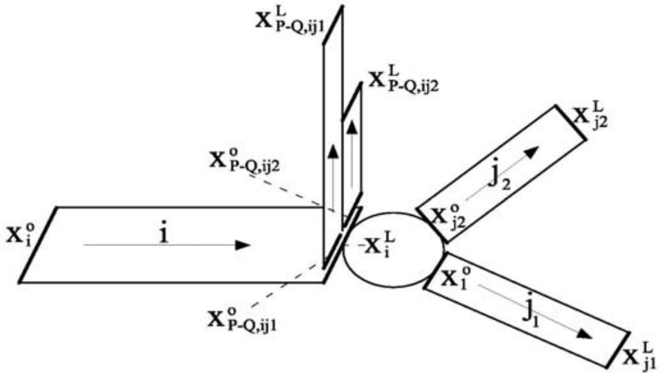
Omwille van deze nadelen wordt geopteerd voor de tweede oplossingsmethode. De ontwikkelde knoopmodellen leggen capaciteitsrestricties op zonder dat verkeerslichtenregelingen of hiaten in voorrangsstromen expliciet gesimuleerd worden. Deze capaciteitsrestricties worden geschat zoals dat gebeurt in de wachtrijtheorie.

6. Knoopvertragingsmodel

Knoopvertragingen zijn verliestijden die voertuigen oplopen voor een kruispunt, omdat ze bijvoorbeeld voor een rood licht staan of voorrang moeten verlenen aan een kruisend voertuig. Deze knoopvertragingen komen voor op stedelijke kruispunten, ook al is de capaciteit gemiddeld genomen voldoende groot om de verkeersvraag te verwerken.

Omwille van hogervermelde redenen wordt ervoor geopteerd om gemiddelde knoopvertragingen te modelleren op een impliciete manier.
Gemiddelde knoopvertragingen zijn knoopvertragingen waarbij de hoogfrequente fluctuaties ten gevolge van het voorturend op- en afbouwen van wachtrijen voor een kruispunt werden uitgemiddeld.
Gemiddelde knoopvertragingen worden geschat zoals dat gebeurt in de wachtrijtheorie. Om deze gemiddelde knoopvertragingen impliciet te modelleren in LTM worden verticale wachtrijen geïntroduceerd aan het einde van een schakel. In deze wachtrijen lopen voertuigen hun beoogde gemiddelde verliestijd op.

Verticale wachtrijen worden geassocieerd met een bepaalde draaibeweging. Figuur S.9 illustreert de twee verticale wachtrijen aan het einde van de toekomende schakel op een splitsingsknooppunt.



Figuur S.9: Verticale wachtrijen voor een splitsingsknooppunt

Deze verticale wachtrijen geven het gemiddelde weer van de tijdelijke files die ontstaan ten gevolge van verkeerslichtenregelingen of voorrangsregels. De (horizontale) plaats op de weg die deze files innemen wordt verwaarloosd. Verticale wachtrijen bestaan naast de horizontale ‘kinematische golf wachtrijen’, die het gevolg zijn van een gebrek aan capaciteit. Deze files nemen wel een (horizontale) plaats in op de weg.

Om knoopvertragingen te modelleren wordt de constructie van verticale wachtrijen toegevoegd aan het LTM oplossingsalgoritme. De ‘te realiseren’ gemiddelde knoopvertragingen worden omgezet in ‘te realiseren’ wachtrijlengtes. Vervolgens worden de zendstromen uit de wachtrij bepaald om deze beoogde wachtrijlengtes te verwezenlijken. Bepaalde restricties worden opgelegd aan de zendstromen om ervoor te zorgen dat de beoogde verliestijden daadwerkelijk opgelopen worden. Vervolgens worden ook de overgangsstromen uit de wachtrij bepaald door het geschikte knoopmodel toe te passen:

Uitgebreid LTM opslossingsalgoritme

Voor elk tijdsinterval Δt,
Voor elk knooppunt n,

Stap 1: Bepaal de zendstroom S_i aan het afwaartse schakeleinde (x_i^L) voor elke toekomende schakel $i \in I_n$ en bepaal de ontvangstroom R_j aan het opwaartse schakeleinde (x_j^0) voor elke vertrekkende schakel $j \in J_n$.
I_n (J_n) is de verzameling van toekomende (vertrekkende) schakels van knooppunt n.

Stap 2: Bepaal de overgangsstromen $G_{ij}(t)$ van de toekomende schakels $i \in I_n$ naar de vertrekkende schakels $j \in J_n$, i.e. bepaal welk aandeel van de zend- en ontvangstromen daadwerkelijk verzonden en ontvangen wordt.

Stap 3: Pas het cumulatief aantal voertuigen $N(x,t)$ aan voor het afwaartse schakeleinde (x_i^L) van elke toekomende schakel $i \in I_n$ en voor de opwaartse wachtrij-uiteindes ($x_{P-Q,ij}^0$) van de draaibewegingen van elke toekomende schakel i:

$$N(x_i^L, t+\Delta t) = N(x_i^L, t) + \sum_{j \in J_n} G_{ij}(t) \qquad \textit{voor alle } i \in I_n$$

$$N(x_{P-Q,ij}^0, t+\Delta t) = N(x_{P-Q,ij}^0, t) + G_{ij}(t) \qquad \textit{voor alle } i \in I_n\,,\ j \in J_n$$

Stap 4: Bepaal de zendstroom uit de wachtrij $S_{ij}^{'}(t)$ voor elke draaibeweging ij van elke toekomende schakel $i \in I_n$. Deze stap omvat 3 sub-stappen:

Stap 4.1: Bepaal de 'te realiseren' gemiddelde knoopvertraging $D_{\widehat{\text{int}},ij}(t+\Delta t)$

Stap 4.2: Bepaal de 'te realiseren' wachtrijlengte $M_{\widehat{P-Q},ij}(t+\Delta t)$

Stap 4.3: Bepaal de zendstroom uit de wachtrij $S_{ij}^{'}(t)$

Stap 5: Bepaal de overgangsstromen uit de wachtrij $G_{ij}^{'}(t)$ van de wachtrijen van de draaibewegingen ij ($i \in I_n\,, j \in J_n$) naar de vertrekkende schakels $j \in J_n$.

Stap 6: Pas het cumulatief aantal voertuigen $N(x,t)$ aan voor de afwaartse wachtrij-uiteindes ($x_{P-Q,ij}^L$) van de draaibewegingen ij van elke toekomende schakel $i \in I_n$, en voor het opwaartse schakeleinde (x_j^0) van elke vertrekkende schakel $j \in J_n$:

$$N(x_{P-Q,ij}^L, t+\Delta t) = N(x_{P-Q,ij}^L, t) + G_{ij}^{'}(t) \qquad \textit{voor alle } i \in I_n\,,\ j \in J_n$$

$$N(x_j^0, t+\Delta t) = N(x_j^0, t) + \sum_{i \in I_n} G_{ij}^{'}(t) \qquad \textit{voor alle } j \in J_n$$

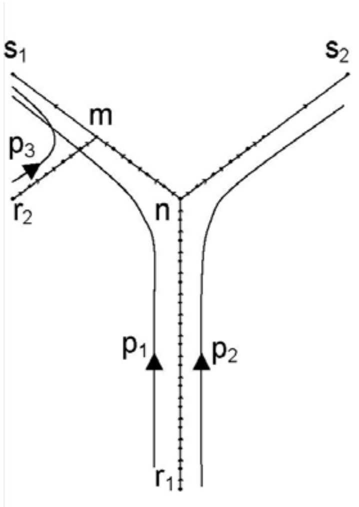
Voor verkeersstromen die beperkt worden door afwaartse randvoorwaarden en voor frequent fluctuerende verkeersstromen kunnen de gerealiseerde verliestijden in beperkte mate verschillen van de beoogde verliestijden. Desalniettemin lijkt het voorgestelde model een veelbelovende methode om knoopvertragingen impliciet te integreren in verkeerssimulatiemodellen zoals LTM.

7. Gevalstudies

Enkele eigenschappen van het ontwikkelde LTM worden geïllustreerd aan de hand van twee gevalstudies.

7.1. Eenvoudig Y-vormig netwerk

In een eerste gevalstudie wordt een eenvoudig Y-vormig netwerk beschouwd met drie mogelijke routes (p_1, p_2, p_3), zoals aangegeven in Figuur S.10.



Figuur S.10: Eenvoudig Y-vormig netwerk

In dit netwerk wordt de voortplanting van files in LTM vergeleken met deze in DQM (Dynamic Queuing Model), een verkeerssimulatiemodel dat gebruikt wordt in de toedelingsmodellen DynaMIT, DYNASMART en INDY-DQM.
Hiertoe gebruiken we het dynamisch verkeerstoedelingsmodel INDY, waar zowel LTM als DQM beschikbaar zijn als verkeerssimulatiemodel.
Volgende verschillen tussen LTM en DQM worden waargenomen:

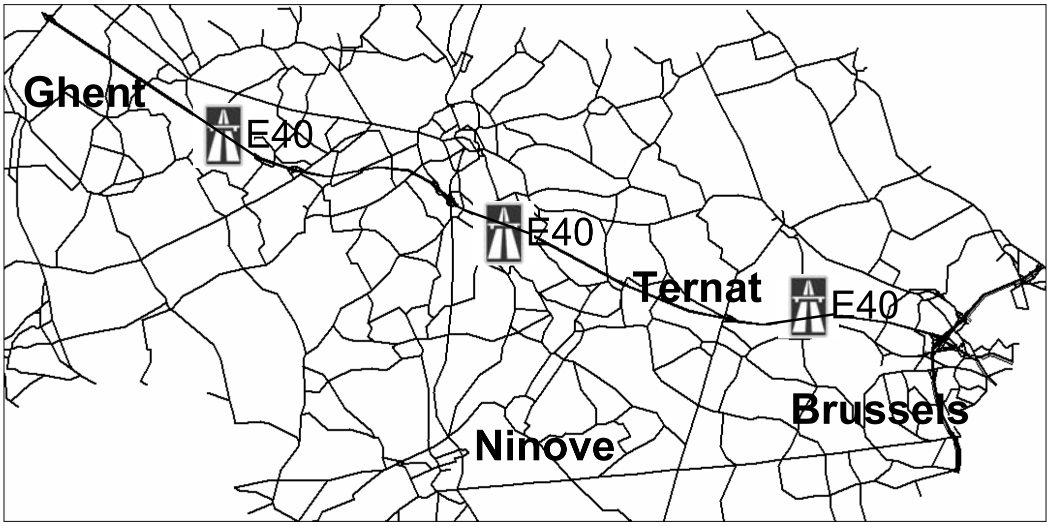
- Files hebben verschillende dichtheden, verschillende lengtes en verschillende terugslagsnelheden

- Files bouwen af in tegenovergestelde richting. In DQM lossen files op vanaf hun staart in de richting van de kop van de file. In LTM is de oplossingsrichting omgekeerd. De file kan zelfs nog aangroeien aan de staart, terwijl hij oplost vanaf de kop. Deze tegenovergestelde richting van file-afbouw komt voor wanneer de file oplost ten gevolge van een flessenhals die inactief wordt.

Deze gevalstudie toont aan dat de verschillende filevoortplantings-eigenschappen in LTM en DQM aanleiding geven tot significant verschillende reistijden.

7.2. Gent-Brussel netwerk

Een tweede gevalstudie illustreert de toepasbaarheid van LTM in een grootschalig feitelijk netwerk. Het beschouwde Gent-Brussel netwerk (cf. Figuur S.11) telt ongeveer 12000 mogelijke routes tussen de verschillende herkomsten en bestemmingen. Een verkeerssimulatie van 1 uur neemt ongeveer 1 minuut rekentijd in beslag op een Pentium IV – 3 GHz (dual core) computer met 2 GB RAM.



Figuur S.11: Gent-Brussel netwerk

De filevoortplantings-eigenschappen in LTM en DQM worden vergeleken in dit grote netwerk. Files in LTM hebben een lagere dichtheid, ze slaan sneller terug in het netwerk, en ze bouwen af in de tegenovergestelde richting (van kop naar staart) in vergelijking met DQM. De gevalstudie toont aan dat de verschillende filevoortplantings-eigenschappen significant verschillende file-locaties tot gevolg hebben. Er wordt ook aangetoond dat het in rekening brengen van knoopvertragingen een belangrijke invloed heeft op de reistijden en op de routekeuze die reizigers maken.

8. Conclusie

In dit eindwerk wordt een rekenefficiënt algoritme ontwikkeld voor verkeerssimulaties in grote feitelijke netwerken die zowel snelwegen als stedelijke regio's omvatten. Het ontwikkelde Link Transmissie Model combineert een realistische beschrijving van de verkeersvoortplanting op schakels (consistent met de kinematische golftheorie) met een realistische beschrijving van de verkeersdynamica op kruispunten (consistent met de wachtrijtheorie).

LIST OF PUBLICATIONS

Yperman, I., C. Tampère and L. Immers, "A Kinematic Wave Dynamic Network Loading Model including Intersection Delays", *Proc. 86th Annual Meeting of the Transportation Research Board*, Washington, DC, January 2007

Yperman, I., S. Logghe, C. Tampère and L. Immers, "The multi-commodity Link Transmission Model for Dynamic Network Loading," *Proc. 85th Annual Meeting of the Transportation Research Board*, Washington, DC, January 2006

Yperman, I., S. Logghe and L. Immers, "Dynamic congestion pricing in a network with queue spillover," *Proc. 12th World Congress on Intelligent Transportation Systems*, San Francisco, November 2005

Yperman, I., S. Logghe and L. Immers, "The Link Transmission Model: An Efficient Implementation of the Kinematic Wave Theory in Traffic Networks," *Advanced OR and AI Methods in Transportation, Proc. 10th EWGT Meeting and 16th Mini-EURO Conference, Poznan, Poland*, 122-127, Publishing House of Poznan University of Technology, September 2005

Yperman, I., S. Logghe and L. Immers, "How congestion pricing can increase traffic volumes," *Proc. 8th Conference of the Network on Eur. Communications and Transportation Act. Research*, Las Palmas de GC, Gran Canaria, Spain, June 2005

De Coensel B., T. De Muer, I. Yperman and D. Botteldooren, "The influence of traffic flow dynamics on urban soundscapes," *Applied Acoustics*, Vol. 66, Issue 2, 175-194, Elsevier Sci. Ltd., February 2005

Yperman, I., S. Logghe and L. Immers, “Strictly Pareto-improving congestion pricing in a multi-destination network,” *Proc. 84th Annual Meeting of the Transportation Research Board*, Washington, DC, January 2005

Yperman, I., S. Logghe and L. Immers, “Robustness of transportation networks,” *Proc. 10th World Conference on Transport Research*, Istanbul, Turkey, July 2004

Immers, L., J. Stada, I. Yperman and A. Bleukx, “Robustness and resilience of transportation networks” *Proc. 9th International Scientific Conference MOBILITA*, Bratislava, Slovakia, May 2004

Immers, L., I. Yperman, J. Stada and A. Bleukx, “Reliability and robustness of transportation networks,” *Proc. 5th Cluster 1 Meeting of the Network on European Communication and Transportation Activities Research*, Amsterdam, The Netherlands, March 2004

Yperman, I., “Zonder files of ongelukken... via een turborotonde,” *Het Ingenieursblad*, Antwerpen, Belgium, December 2003

Yperman, I. and L. Immers, “Capacity of a turbo-roundabout determined by micro-simulation,” *Proc. 10th World Congress on ITS*, Madrid, Spain, November 2003

Yperman, I., J. Stada and L. Immers, “Punctuele logistiek dankzij een robuust netwerk,” *Proc. 10e Vervoerslogistieke Werkdagen*, Corsendonck, Belgium, November 2003

Logghe, S. and I. Yperman, “De dynamica van congestie en prijsbeleid,” *Proc. 30th Coll. Vervoersplanologisch Speurwerk*, Antwerp, Belgium, November 2003

Immers, L., S. Logghe, J. Stada en I. Yperman, “In de rij voor de vracht of de vracht in de rij?,” *Proc. 30th Coll. Vervoersplanologisch Speurwerk*, Antwerp, Belgium, November 2003

Int Panis, L., et al. and I. Yperman, “Mobilee: A study of sustainable mobility at the local level,” *Proc. 1st Symposium Environnement et transport*, Avignon, France, June 2003

Immers, L., H. Meeuwissen, J. Stada, and I. Yperman, “DVM en beter benutten vereisen een robuust netwerk,” *Proc. 5e DVM Symposium*, Rotterdam, Nederland, April 2003

ABOUT THE AUTHOR

Isaak Yperman was born in Leuven on April 29, 1977. He received a Bachelor of Bioscience Engineering degree (Kandidaat Bio-ingenieur), a Bachelor of Applied Sciences and Engineering degree (Kandidaat Burgerlijk Ingenieur) and a Master of Applied Sciences and Engineering: Civil Engineering degree (Burgerlijk Bouwkundig Ingenieur) from the Katholieke Universiteit Leuven. He graduated cum laude in 2002. His undergraduate thesis '*Analyse en Optimalisering van de Verkeersafwikkeling op de Koning Boudewijnlaan te Leuven'*, supervised by Prof. Ir. L.H. Immers, provided him with the title 'laureaat van de KVIV-Ingenieursprijzen'. In September 2002, he started working as a research assistant and PhD student at the Traffic and Infrastructure research group of the Department of Civil Engineering at the Katholieke Universiteit Leuven.