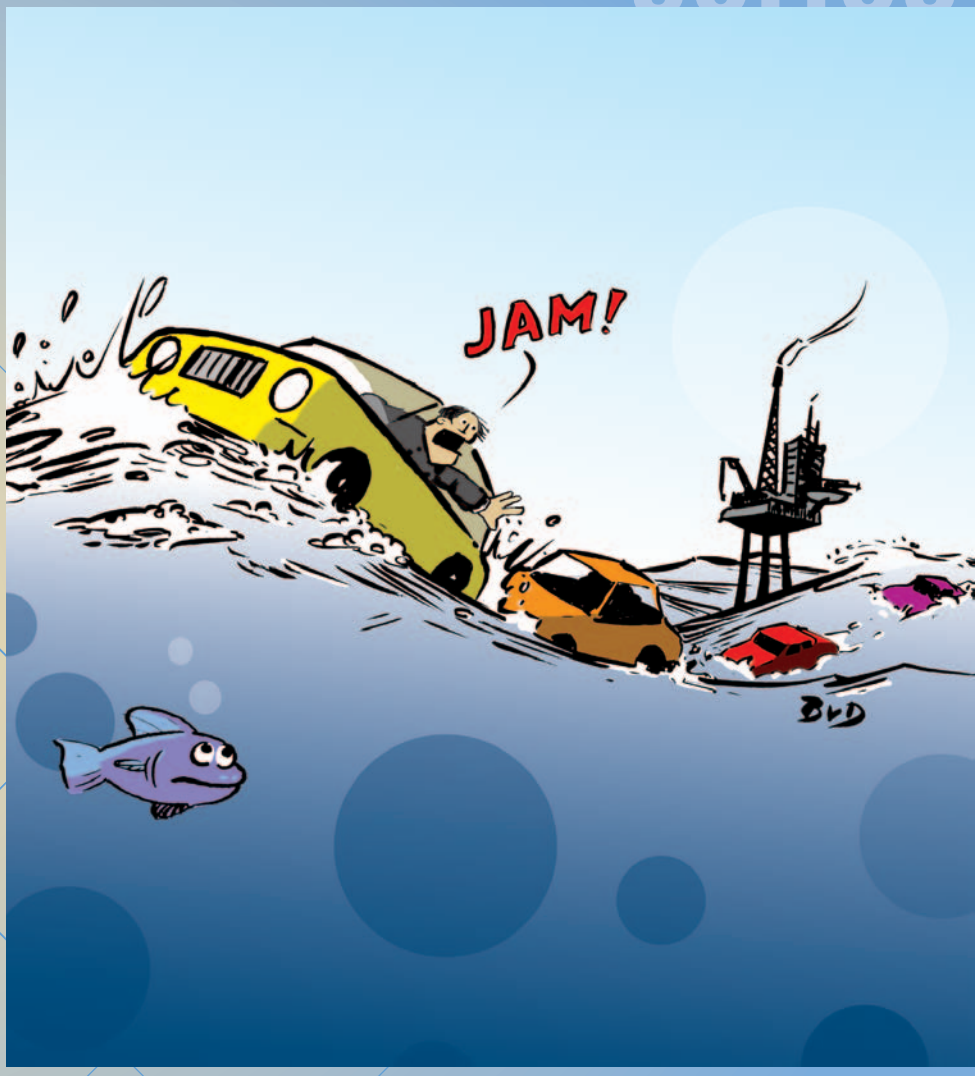
TRAIL

Chris M.J. Tampère

thesis series



Human-Kinetic Multiclass Traffic Flow Theory and Modelling

With Application to Advanced Driver Assistance Systems in Congestion

.

Stellingen

Human-kinetic Verkeersstroom Theorie en Modellering met Meerdere Gebruikersklassen

Toegepast op Rijtaak Ondersteunende Systemen in Congestie

Chris Tampère, 17 december 2004

1. De argumenten waarmee in de gaskinetica de tijdsduur van deeltjesinteracties wordt verwaarloosd, gaan niet op voor verkeer: de duur van zowel acceleratie als deceleratie van voertuigen is immers *niet* verwaarloosbaar ten opzichte van de golfsnelheden in de stroom; noch kan men de kans verwaarlozen dat zich nieuwe voertuiginteracties voordoen vóórdat eerdere interacties beëindigd zijn.
   – Hoofdstuk 4 van dit proefschrift –

2. Datgene wat tot nu toe ‘anticipatie’ genoemd werd in macroscopische verkeersstroommodellen is in feite louter het reageren op stimuli van de voorganger; echte anticipatie staat voor het bij voorbaat reageren op een *voorspelling* van het toekomstige verloop van deze stimuli.
   – Hoofdstuk 4 van dit proefschrift –

3. Het belang van de pure reactietijd van bestuurders voor stabiliteit van de verkeersstroom wordt sterk overschat, daar waar de dynamische relaxatietijd veel belangrijker is.
   – Hoofdstuk 5 van dit proefschrift –

4. Door het onderbouwen van het human-kinetic model vanuit individueel bestuurdersgedrag lijkt het probleem om congestie adekwaat te beschrijven alleen maar te zijn verlegd naar de gedragswetenschappen. Dit is geen pas op de plaats maar een belangrijke stap in de goede richting: het gedrag van de verkeersstroom is immers slechts een afgeleide van bestuurdersgedrag.
   – Hoofdstuk 6 van dit proefschrift –

5. De aanname van variabel bestuurdersgedrag is niet *noodzakelijk* om een verkeersstroommodel te verkrijgen dat ‘capacity drop’ en ‘capacity funnel’ effecten vertoont. De aanname is echter plausibel én relevant, want ze heeft een significante invloed op de omvang van deze fenomenen en de omstandigheden waarin ze voorkomen. Variabel bestuurdersgedrag is daarom onontbeerlijk in modellen die ambiëren om deze fenomenen *correct* te beschrijven.
   – Hoofdstuk 6 van dit proefschrift –

6. Het is verdacht dat overgangen van vrij stromend naar congestie verkeer gemodelleerd kunnen worden door een model dat peletonvorming negeert, terwijl het uit empirische gegevens en verkeersstroomtheorie duidelijk is dat peletons en vooral de openingen daartussen een cruciale rol moeten spelen.
   – Hoofdstukken 2 en 7 van dit proefschrift –

7. Filestaart waarschuwingen hebben een gunstig effect op de relaties tussen aanhangers van de eerste en hogere orde verkeersstroomtheorie: door het elimineren van instabiliteit gedraagt de verkeersstroom zich als een eerste orde stroom, ongeacht het model waarmee men hem modelleert.
   – Hoofdstuk 8 van dit proefschrift –

8. Het heeft geen zin eerste orde modellen te verdedigen ten koste van hogere orde modellen of omgekeerd: zij houden namelijk geen tegenspraak in. Hogere orde modellen houden naast eerste orde effecten ook rekening met meer verfijnde effecten, met name potentiële instabiliteit. Men kan dus slechts discussiëren over de noodzaak van deze verfijning voor de beoogde toepassing en over de instelling van de parameters.

9. Ook voor verkeersmanagement en intelligente voertuigen geldt: $1+1=3$: de grootste efficiëntie verbetering wordt niet bereikt door de systemen afzonderlijk, maar zodra een samenwerking ontstaat tussen de bestuurder, zijn voertuig, andere verkeersdeelnemers en de verkeersmanager, waarbij iedere partij onderhandelt en rekening houdt met wederzijdse belangen.

10. Ingenieurs houden vaak te weinig rekening met belangrijke aspecten, louter omdat ze moeilijk te meten, te modelleren of te kwantificeren zijn; terwijl toch belangrijke beslissingen worden genomen op basis van hun pragmatische, vereenvoudigde en dus onvolledige modellen.

11. Gedragswetenschappers controleren steeds – en terecht – zorgvuldig de omstandigheden waarin zij hun experimenten uitvoeren. Hierdoor durven zij echter zelden hun theorieën te veralgemenen en in breed inzetbare kwantitatieve modellen te gieten, waardoor hun kennis in de praktijk te weinig benut wordt.

12. Bij het ontwerp van rijtaak ondersteunende systemen zou men vanaf het eerste conceptuele ontwerp rekening moeten houden met de effecten ervan op de verkeersafwikkeling. Dit betekent dat de rol van overheden in de ontwikkeling van deze systemen veel minder vrijblijvend moet worden dan nu het geval is.

13. Onderzoeksresultaten die de a priori verwachting van de onderzoeker bevestigen zijn de gevaarlijkste: het is immers verleidelijk om minder kritisch te zijn en ze zonder meer te geloven.

14. Omwille van het kleine aantal voor de eindscore beslissende momenten in een voetbalmatch, is Vrouwe Fortuna meestal de belangrijkste speler op het veld. Wellicht draagt het gevoel van onrecht dat hierdoor ontstaat in belangrijke mate bij aan het ontstaan van voetbalgeweld.

Deze stellingen worden verdedigbaar geacht en zijn als dusdanig goedgekeurd door de promoter Prof. Dr. H.J. van Zuylen

Propositions

Human-Kinetic Multiclass Traffic Flow Theory and Modelling

With Application to Advanced Driver Assistance Systems in Congestion

Chris Tampère, 17 december 2004

1. *Neglecting the duration of vehicular interactions cannot be justified by the same arguments as in gas-kinetics: the duration of vehicular acceleration or deceleration is* not *negligible with respect to the macroscopic time scale (wave speeds); neither can one neglect the probability of new interactions occurring between vehicles before earlier interactions have been completed.*
   – Chapter 4 of this thesis –

2. *The term traditionally indicated as 'anticipation term' in macroscopic traffic flow models actually represents* non-local *responses to stimuli of the predecessor; real anticipation means that drivers respond to a* prediction *of these stimuli.*
   – Chapter 4 of this thesis –

3. *The impact of pure reaction time delay of drivers on traffic flow stability is overestimated, whereas the importance of the dynamic relaxation time is much more important.*
   – Chapter 5 of this thesis –

4. *The foundation of the human-kinetic model on individual driver behaviour seems to pass on the problem of adequately describing congestion to the domain of behavioural science. This is not a status quo, but rather an important step forward. After all, traffic flow behaviour is a resultant of individual drivers' actions.*
   – Chapter 6 of this thesis –

5. *Assuming variable driver behaviour is not* necessary *to obtain a traffic flow model that exhibits capacity drop and capacity funnel. However, it is a plausible and relevant assumption with significant influence on the magnitude of these phenomena and the conditions in which they occur. Variable driver behaviour is therefore indispensable in traffic flow models that aim at describing these phenomena* correctly.
   – Chapter 6 of this thesis –

6. *The fact that phase transitions from free flowing traffic to all kinds of congested traffic can be reproduced by a model that neglects platoon formation should raise suspicion. After all, both empirical observations and theoretical considerations indicate that platoons and inter-platoon gaps play a crucial role.*
   – Chapters 2 and 7 of this thesis –

7. *Queue tail warning systems have a beneficial effect on the relation between propagators of first and higher order traffic flow theory. Because instability is suppressed, traffic flow behaves as a first order flow, no matter which model is used to describe it.*
   – Chapter 8 of this thesis –

8. *It is useless to propagate first order traffic flow theory and avert higher order theory or vice versa, since both theories are not contradictory. Higher order models are* extensions *of first order models that account for the more refined behaviour of traffic flow, especially potential instability. At most, one can question the necessity to model such refinements, given the application of interest, and argue about the parameters choice.*

9. *Traffic management and intelligent vehicles comply with the rule* $1+1=3$*: the largest efficiency improvement is not obtained by the separate systems, but when the driver, his vehicle, other road users, and the traffic operator co-operate, negotiate, and take mutual objectives into account.*

10. *Engineers often disregard important aspects, simply because it is difficult to measure, to model, or to quantify them. Nevertheless, their pragmatic, simplified, and therefore incomplete models are used as a motivation for many important decisions.*

11. *Behavioural scientists always carefully control the experimental conditions in their research – and with good reason. However, this attitude is also an obstacle that impedes generalisation of their theories into widely applicable, quantitative models; therefore their knowledge is often underexploited in practice.*

12. *The impact on traffic flow should be one of the major concerns from the earliest stages of development of Advanced Driver Assistance Systems. This can only be achieved if authorities are more actively engaged in the development of ADAS than they are now.*

13. *Research results that confirm the a priori expectations of the researcher are dangerous: it is very tempting to consider them less critically and believe them right away.*

14. *Due to the small number of decisive phases in a football game, Lady Fortune is usually the most important player on the field. This invokes a feeling of injustice that probably contributes significantly to the emergence of aggression in and around the stadium.*

These propositions are considered defendable and thus have been approved by the supervisor Prof. Dr. H.J. van Zuylen

Human-Kinetic Multiclass Traffic Flow Theory and Modelling

With Application to Advanced Driver Assistance Systems in Congestion

Chris Magda Jules Tampère

This thesis is a result from a project
funded by the co-operative research programme T3 of the
Netherlands Organisation of Applied Scientific Research TNO
and the Netherlands Research School for Transport, Infrastructure
and Logistics TRAIL.

Cover illustration: Bram van Driel

Human-Kinetic Multiclass Traffic Flow Theory and Modelling

With Application to Advanced Driver Assistance Systems in Congestion

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof.dr.ir. J.T. Fokkema,
voorzitter van het College van Promoties,
in het openbaar te verdedigen op vrijdag 17 december 2004 om 10:30 uur

door Chris Magda Jules TAMPÈRE

burgerlijk bouwkundig ingenieur
geboren te Antwerpen, België

Dit proefschrift is goedgekeurd door de promotor:

Prof. dr. H.J. van Zuylen

Samenstelling Promotiecommissie:

Rector Magnificus,	Voorzitter
Prof. dr. H.J. van Zuylen,	Technische Universiteit Delft, promotor
Dr. Ir. S.P. Hoogendoorn,	Technische Universiteit Delft, toegevoegd promotor
Prof. dr. ir. B. van Arem,	Universiteit Twente
Prof. dr. ir. P.H.L. Bovy,	Technische Universiteit Delft
Prof. dr. K.A. Brookhuis,	Technische Universiteit Delft & Rijksuniversiteit Groningen
Prof. ir. L.H. Immers,	Katholieke Universiteit Leuven, België
Prof. Dr.-Ing. M. Papageorgiou	Technical University of Crete, Greece

TRAIL Thesis Series nr. T2004/11, The Netherlands TRAIL Research School

This thesis is the result of a Ph.D. study carried out from 2000 to 2004 at TNO Inro, Delft University of Technology, Faculty of Civil Engineering and Geosciences, Transportation and Planning Section and the Katholieke Universiteit Leuven, Faculty of Engineering, Department of Civil Engineering, Transportation Planning and Highway Engineering Section.

Published and distributed by:
TRAIL Research School

P.O. Box 5017
2600 GA Delft, The Netherlands
T +31 15 278 60 46
F +31 15 278 43 33
E info@rsTRAIL.nl
I www.rsTRAIL.nl

ISBN 90-5584-060-2

Keywords: Traffic Flow Theory, Congestion, Advanced Driver Assistance Systems

Printed in The Netherlands

“I have not failed.
I’ve just found 10,000 ways that don’t work”
- Thomas Alva Edison

in memory of my beloved mother

PREFACE

When I graduated as a civil engineer in 1997, I lost the student's privilege of being occupied with all kinds of subjects as widely varying as thermodynamics, electronics, hydrology, construction, traffic engineering etcetera. I was confronted with the necessity of choosing a job in only one of these areas, a though dilemma for someone with interest in almost any scientific or technical domain. Soon, the final choice was between hydraulics and traffic, and it was the enthusiasm of my professor of traffic engineering that convinced me to accept a job at the Traffic and Transportation department of TNO Inro in Delft, The Netherlands. Who could have predicted then that seven years later, I would finalise a Ph.D. research in traffic flow modelling, a discipline that was once founded as a cross-over between traffic engineering and hydraulic modelling?

Is it a coincidence that my supervisor professor Henk van Zuylen worked, among others, at the Delft Hydraulics laboratories for years before joining the Transportation and Traffic Engineering Section of the Delft University of Technology? Henk, I appreciate the way in which we overcame our initial conflict of roles (with me in the unfortunate combination of project leader and Ph.D. student), built up mutual respect, and finally produced a joint piece of research and a dissertation of which I am very proud. I thank you for bringing focus in my work, and for raising the scientific level of my texts.

If I have been able to successfully conduct the research reported in this dissertation, this is not in the least because I was privileged with two extraordinary daily supervisors: Bart van Arem and Serge Hoogendoorn. Bart, I owe a lot to you. From my first application talk at TNO Inro, you have given me an incredible amount of trust and opportunities. You were the instigator of this project and have supervised the progress critically but positively ever since. And apart from being my mentor, and during some period even my chief, you remained in the first place a reliable and sympathising

colleague. Serge, your role was different: you were my ‘wizard’. Any time I had an idea, but did not know how to work it out, or any time I ran out of ideas and did not know how to proceed, you came up with one of your brilliant solutions. Usually it took us one or two hours to exchange thoughts, after which it took me another month to elaborate them. I admire your intelligence and professionalism, especially since at the same time you are spontaneous, gentle and encouraging.

It is no use to have such a professional team and great ideas, if there were no people giving trust and financial support. I am therefore very grateful to Ben Janssen and Piet Bovy of the co-operative research programme T3 of the Netherlands Organisation of Applied Scientific Research TNO and the Netherlands Research School for Transport, Infrastructure and Logistics TRAIL. We were one of the first projects being sponsored by you, so things were as new for you as they were for us. I think you found a great mixture of conciseness with respect to organisation, progress and financial status on the one hand, and flexibility and honest scientific interest on the other hand. Additional financing came from TNO Inro, who gave me the chance to combine this research with numerous other interesting projects in a very rich professional and friendly environment, and from the Katholieke Universiteit Leuven, where I got the opportunity to finalise the research and dissertation text and where new challenges are awaiting.

Then, there are numerous people who contributed in one way or another to the project and to whom I owe a big ‘thank you’. Paul Wewerinke and Jeroen Hogema of TNO Human Factors Institute, your behavioural research has guided my traffic flow theoretical work and vice versa. Although the concrete results of our joint effort are only modestly present in this dissertation, I hope there will be opportunity soon to proceed along the unique path we were following. Arthur van Dam, although you were ‘only’ a student, your Traflow Pack software for traffic flow models looked rather professional, and still proofs its flexibility and applicability every day. Dong Ngoduy, your numerical scheme has never let me down, so that I could concentrate on the development of the modelling theory rather than on numerical issues. I expect spectacular improvements to traffic flow theory from you too! Marieke Collins, you protected me from getting lost in the algebra of the stability analysis. Luisa Palomba, Klaas van der Staal, Marjolein Baart, Conchita van der Stelt and Mirjam Zuil, your support in secretary, administrative and organisational issues has substantially relieved my work. And to all my friends, family, colleagues and sympathisers: the interest you showed and your support and sympathy served as a continuous encouragement.

Finally, saying only ‘thank you’ is by far an understatement for what I owe to you, my beloved wife Katrien. In a hectic period where I was writing this dissertation, commuted daily from Belgium to Delft, renovated the house, and started a new job, you still found the energy and time to give me your love and care and to combine your job with the housekeeping, the administration, and so much more. And as if that were not enough, you gave me the most precious gift, our daughter. How can I ever reward you? Lien, I hope someday you will understand why your daddy was typing so often on that silly computer all evenings and weekends, instead of cuddling and hugging you still more. Maybe someday when you will be driving your car and wonder where that traffic jam suddenly came from?

CONTENTS

NOTATION

Kinetic and macroscopic variables and parameters

A	Average activation level of drivers in the flow
B $b_0,..,b_3$	Shape factor function for the speed variance and parameters of this function
$N(t,L)$	Total number of vehicles at time t on a stretch of road with length L
P	Traffic pressure
Q	(Continuous) flow rate
V	Average speed of vehicles in the flow
$V^{ant}_{c_{j-1},c_j}$	Subjectively anticipated average speed of predecessors (of class c_{j-1} as perceived by class c_j)
W	Average desired speed of the driver population in the flow
$f(k,V)$	Macroscopic acceleration function; i.e. The rhs or source term of the speed equation in the primitive formulation
$f(x)$	Function defining the spatial spreading of merging along an on-ramp
f_{acc}	Acceleration opportunity of vehicles in a multiclass flow
k k_{jam} k_{crit}	(Continuous) traffic density; jam density; critical density
n_{sent}, $n_{received}$	Number of warning messages sent/received by the ADA system
p_{noise}	Noise threshold below which the probability of sending warning messages by the ADA system is neglected

$p_{overtake}$	Overtaking probability, i.e. Probability that a faster vehicle can change lanes immediately when interacting with a slower vehicle
q_{in}	Merging intensity
x_0	Location of the beginning of the merging zone of the on-ramp
Δx_{merge}	Length of the merging zone of the on-ramp
β	Preferred merging point, i.e. Fraction indicating the position along the merging zone of an on-ramp where merging intensity is maximal
π_a, π_v	Transition rate of the attention/speed in case of an interaction
$\rho(t,x,S)$	Generalised density of vehicles with state S at time t and location x
$\tilde{\rho}$	Reduced generalised density, i.e. The generalised density that is integrated over only a subset of its state variables
Θ	Speed variance
Φ	Constant transition rate for the activation level
Π	Interaction rate, i.e. Number of interactions per unit of time

Microscopic / individual state variables and parameters

$S(s_1,..,s_n)$	State vector of a vehicle/driver combination, consisting of n independent state variables $s_1,..,s_n$
T	Reaction time of a driver
a	Activation level of a driver
acc_{min}, acc_{max}	Minimal/maximal acceleration capability of the vehicle
e	Error function, determining the amplitude of the acceleration response of a driver with respect to the gap error
f^{ant}	Anticipation strength of a driver
h	Headway of a vehicle, i.e. Distance between the own rear-bumper and that of the predecessor
l	Length of the vehicle
p_{indiff}	Indifference band with respect to shorter-than-desired gaps
s_d, s_1^d, s_2^d	Desired gap (function) of a driver with respect to a predecessor and parameters of this function

s_j	Gap with respect to the predecessor
s_j^0	Gross vehicle length, i.e. Sum of vehicle length and a safety margin at standstill
s_j^{pp}	Gap with respect to the pre-predecessor
v	Speed of the vehicle
w	Desired speed of the driver
τ	Relaxation constant
Δx^{ant}	Anticipation distance

Probabilities, distribution functions and parameters

$LgN(\mu,\sigma)$	Lognormal probability density function
$N(\mu,\sigma)$	Normal probability density function
b	Exponent determining the variance of the gap distribution
$p_y(y_0\|cond)$	Conditional probability of variable y having value y_0, given condition *cond*
μ	Average or expected value of a probability density function
σ	Standard deviation of a probability density function

Numerical variables and parameters

$\mathbf{U}$	Vector containing independent model variables
$\mathbf{F}(\mathbf{U})$	Numerical flux function
$\mathbf{G}(\mathbf{U})$	Numerical source term function
c_0	Threshold for numerical information flow inversion
p_{prior}	Fraction of numerical source term accounted for before convective part of the traffic model
λ	Eigenvalue

Indices, subscripts, superscripts

a	Subscript referring to activation level
c	Subscript referring to user class c
e	Superscript referring to equilibrium conditions
i	Index for discrete time (in Markov approximate model)
j	Index for individual vehicles (counting from downstream to upstream)
j	Index for cells and cell interfaces (counting from upstream to downstream)
max	Subscript referring to maximum value
min	Subscript referring to minimum value
$normal$	Subscript referring to the normal value
s	Subscript referring to the gap
tot	Subscript referring to the total value
t, Δt	Time (continuous variable); time increment
v	Subscript referring to speed
w	Subscript referring to desired spee
x, Δx	Longitudinal position (continuous variable); space increment
κ	Index for the order of moments

Mathematical functions and operators

RHS/LHS	Right/left hand side of an equation
Δ	Increment
∇	Nabla operator; in combination with the inner product it forms the divergence of a vector
$\cdot$	Inner product
$\langle y_1 \rangle_{y_2}$	Expected value of y_1 when averaged over all values of y_2
δ	Kronecker delta function
Ξ	Heavyside function
$\overline{y}$	Average value of y
$\dot{y}$	Total time derivative of y

1 INTRODUCTION

Since the 1950's, traffic jams have been the subject of scientific study. Now, 50 years later, they still are. Researchers with backgrounds ranging from engineering over mathematics and physics to chemistry and thermodynamics have continuously strived for a better description of traffic flow, and the jams occurring therein.

Nowadays, traffic flow theorists are faced with a new challenge: to describe or even predict the dynamics of traffic flows that do not yet exist. Intelligent Transportation Systems, either controlled from traffic management centres, at the road-side, or in the vehicles themselves, change the behaviour of drivers and vehicles and thereby possibly invalidate the existing descriptions of traffic flow. Especially the widely applied class of macroscopic traffic flow models that take as inputs empirical traffic characteristics by means of a fundamental diagram need to be reconsidered critically.

The work described in this dissertation thesis builds macroscopic models of traffic flow in a most flexible way. The approach is based on a behavioural specification at the microscopic level of the individual vehicle and driver. Its description of current traffic flows and congestion is comparable to modern state-of-the-art models. However, thanks to the micro-macro link and the introduction of the *activation level* of drivers as an additional variable that governs driver behaviour, it paves the way to all kinds of further refinements to the basic model behaviour. The multiclass version of this *human-kinetic* model enables modelling of future traffic flows that will consist of a mixture of traditional vehicles and vehicles equipped with new Advanced Driver Assistance systems.

The remainder of this chapter explains the context and background of this research (section 1.1), its objectives and scope (section 1.2), the approach of the research (section

1.3), a summary of the theoretical contributions to traffic flow theory (section 1.4), and the practical relevance (section 1.5).

1.1 Research background and context

1.1.1 Traffic flow and congestion in the age of Intelligent Transportation Systems

Today, a congestion free traffic network has become almost unthinkable. For example in the Netherlands, over 200 km of congestion is reported on an average working day in a motorway network of approximately 2200 km length (Bovy, 2001). In Belgium, estimated delay due to congestion on motorways totals 9 million hours yearly on a total of 26 billion km driven. This corresponds to an average daily economic loss of € 460 000 each day (Vanhove & De Ceuster, 2003). At the same time, unsafety in the traffic system is high: each year, traffic kills 40 000 persons in the European Union (15 countries), while another 1.7 million are injured. The total cost of traffic accidents in the EU is estimated at € 160 billion per year, approximately 2% of the gross national product of the EU (European Commission, 2003).

To relieve these negative impacts of increased mobility, public and private sectors are co-operating in an effort to make the traffic system safer and more efficient. An integrated approach is needed, covering town and country planning, infrastructure planning and design, (multimodal) travel demand management, logistics, traffic management, traffic behaviour, and active and passive safety and driver support systems in the vehicle. Modifications to infrastructure or spatial planning are important long-term solutions that are however not always desirable from an economic or environmental point of view (noise, emissions). Especially in densely populated areas, the societal and political support for (large) infrastructural extensions is nowadays difficult to obtain. Much attention is therefore paid to short- and medium-term solutions aimed at controlling or improving the behaviour of travellers and vehicles. The idea is to fight congestion in traffic networks by better utilisation of existing infrastructure in such a way that safety is not affected or even improved. An important contribution to this goal is expected from Intelligent Transportation Systems (ITS).

ITS is a collective noun for technologically supported services for participants in the traffic system: pedestrians, drivers, fleet managers, traffic managers, etcetera. On the one hand, some ITS systems offer commercial (e.g. telemetry, anti-theft tracking of the vehicle) and public services (e.g. reservation and payment for parking, emergency call) and systems aimed at increasing driver comfort (e.g. internet, positioning). On the other hand, ITS also incorporate systems for traffic control, driver support, and even vehicle control, with the purpose of maximising traffic efficiency and/or safety on existing infrastructure. For instance, the ITS industry provides traffic managers with systems to inform drivers of current network conditions (e.g. incidents or location and severity of congestion), to manage capacity (e.g. dynamic lane allocation), to regulate traffic demand (e.g. co-ordinated traffic control through traffic lights, ramp metering systems) and traffic behaviour (e.g. intelligent speed adaptation), and to guide traffic over different routes. Also the drivers and the vehicles themselves are becoming important

players thanks to precise and relatively cheap positioning technology, sensors, and mobile communication. The vehicle will autonomously gather information (e.g. enhanced vision, pedestrian detection) and will communicate with the nearby infrastructure and surrounding vehicles. Although the driver will retain main responsibility for controlling his vehicle, he will be increasingly supported by all kinds of Advanced Driver Assistance (ADA) systems to make the trip safer, more comfortable, and more efficient. Such systems not only provide the driver with extra information, they also help the driver in managing the booming loads of information and partially take over some driving tasks (ADASE, 2003).

Traffic flow operations in the next decades will therefore be driven by the control actions of the driver, as well as his vehicle. These actions will be based on autonomously gathered information, enriched by information that reaches the car or the driver through inter-vehicle communication or communication with the road-side infrastructure. In order to predict the impacts of these systems on future traffic flows, models are needed that account for the role of:

- the driver,
- the vehicle (ADA systems),
- additional information through communication,
- road-side traffic control systems.

It is unclear how traffic flow will develop in the next decades. For instance the wider deployment of Adaptive Cruise Control, a system designed to maintain a preset speed or preset following distance depending on the presence of a slower predecessor, might slightly increase capacity and render traffic flow more smooth (Van Arem et al., 1997a), or might increase capacity significantly at the cost of more critical braking interactions (Minderhoud, 1999), depending on the parameter settings. Such effects will become even more difficult to estimate when more complex systems with co-operative vehicles or vehicle-roadside interaction are considered (e.g. Morsink et al., 2003). This is still more the case for systems aimed at influencing congested traffic or the transition between free flowing traffic and congested traffic patterns, since these phenomena are not yet completely understood in current traffic flows (Hoogendoorn & Alkim, 1999).

1.1.2 Primary influence factors

Let us elaborate briefly on the driver and his intelligent vehicle. These are considered primary influence factors in the sense that the other two aspects, information exchange and traffic control, only have influence on traffic flow if the driver, the vehicle, or both respond to the information or control signals.

1.1.2.1 The driver

Driver behaviour in motorised vehicles has become a subject of extensive scientific research since the 1950's. Initially the research was driven by car manufacturers desiring to build safer and more comfortable cars. The majority of car-following models

still in use today have their origin in the period 1950 – 1970, for example in the seminal experiments at the General Motors Research Laboratories. For an excellent overview of the history of car-following model development, as well as car-following experiments from that period, we refer to Rothery (1999) or Brackstone & McDonald (1999). Apart from the car-following control task, the study of which has been at the cradle of traffic flow theory, other human-factors aspects have received less attention in the traffic flow theorists' community. Nevertheless engineers and psychologists have conducted numerous experiments in driving simulators and instrumented vehicles on test tracks and in real traffic flow (for an introduction: see Koppa, 1999). Such studies examine among others: perception, decision making, workload, vigilance, driving performance under the influence of fatigue, drugs, or alcohol, and driving performance of specific user groups such as disabled and elderly drivers. Behavioural researches have quite rightly noticed that this field of knowledge has been underexploited by traffic flow theorists (e.g. Boer, 1999; Ranney, 1999; Van Winsum, 1999). This thesis therefore aims for a fundamental improvement of traffic flow theory and traffic flow models by bringing these fields closer together.

1.1.2.2 Advanced Driver Assistance (ADA) systems

The development of Advanced Driver Assistance (ADA) systems has been one of the major interests in the automotive industry for many years. ADA systems intend to provide assistance to the driver by warning him, informing him, or by partially taking over driving tasks. An overview of ADA functions is given for instance by Van Arem et al. (2002), who distinguish two classes of systems:

- Intervening systems must provide active support to the driver on the control level of the driving task. These systems must be able to take over longitudinal and/or lateral control of the vehicle and perform parts of the driving task automatically.
- Driver information and warning systems are aimed to support the driver on the strategic and manoeuvring level of the driving task. While the driver remains in full control of his vehicle, additional information and warnings will increase his situation awareness and his ability to act appropriately to the actual scenario.

ADA systems are in continuous development. A common future vision on ADA development by car manufacturers, research institutes, and road authorities in Europe is provided by ADASE (2003). This thesis anticipates on future needs for modelling traffic flows with ADA systems (see next section).

1.1.3 Need for model development and model requirements

There is a need to model the traffic flow process and the interaction between the driver, the (intelligent) vehicle, and the environment (infrastructure, traffic control,…) for the following three reasons:

a. *theoretical improvements and insight*: The aim of any modelling effort is to describe the system as good as possible. Because many aspects of driver

behaviour have not yet been accounted for in existing traffic flow theory (see the arguments in 1.1.2.1), theoretical improvements in traffic flow theory are still necessary and possible. A second theoretical argument is that model building as such improves the insight into the system: it requires a thorough understanding of the process under consideration, not only qualitatively, but also quantitatively. During model development and application, the insight into the processes governing the system grows significantly, as well as the insight into the main influence factors, and into the system's sensitivity to changes of these factors. In the context of traffic flow theory, building models increases the insight into congestion formation, the emergence of different congestion patterns and the role of driver behaviour therein, as is indicated by several examples and analyses throughout the main text of this thesis.

b. *assessment of traffic control and ADA systems*: the impacts of a given system design can be simulated and assessed using the model. In the context of traffic flow: given the behaviour of an ADA system (directly from the controller specifications) and the behaviour of the driver assisted by the system (e.g. from driving simulator experiments or small scale field trials), the impact on traffic flow can be extrapolated using a suitable traffic flow model.

c. *design of traffic control and ADA systems*: models and the insight gained during the process of model development are indispensable for designing traffic control systems. Insight into the workings of the system reveal how a desired change of the system's behaviour can be achieved, after which the model is used to provide the proof-of-concept, to refine, and to optimise the system design and its parameter settings. In the context of traffic flow: requirements can be set to the combined behaviour of the driver and the ADA system, so that a desired change in traffic flow dynamics (e.g. stability or capacity improvement) will be achieved.

Three basic requirements to models for traffic flows with ADA systems are:

- A realistic representation of traffic flow dynamics should be provided, in particular with respect to the congested regime and the transitions between flow regimes;
- A behavioural specification at the level of the individual vehicle – driver units should be used, so that the model builder can differentiate between behaviour of units with and without the ADA system;
- The level of detail of the model should be suitable for the desired application or analysis.

The first requirement is self-evident: one cannot analyse impacts of ADA systems if the reference situation is not accurately described. With respect to traffic flow dynamics, the most complex and challenging problem is to properly describe congestion phenomena. Moreover, in ADA system design, an increasing interest in congested traffic and ways to alleviate its undesirable effects can be observed (ADASE, 2003). The second requirement follows from the fact that ADA systems are vehicle-bound systems, so that potential changes to traffic dynamics enter traffic flow through

behavioural changes of the individual driver and/or his vehicle. Note that not only changes from normal (unsupported) behaviour to ADA-supported behaviour of the vehicle-driver unit can be expected, but also modified driver behaviour of unequipped units in the flow as a result of being surrounded by ADA traffic. For that reason, a refined driver and vehicle behavioural model is required, as well as a model that is capable of handling multiclass traffic (mixed equipped and non-equipped traffic). Finally, the third requirement states that the level of detail of the model should also be compatible with the desired application. For instance, designers of sensors and actuators of the ADA system require a level of detail that primarily focuses on the individual vehicle in relation to the system components on the one hand, and the direct traffic flow context (surrounding vehicles) on the other hand. Traffic flow models for ADA assessment at this level of detail are readily available and in continuous development (see section 3.2.4). The interest of traffic managers developing co-operative control strategies for future traffic flows lies at the level of flow patterns and their relation to individual vehicle behaviour (ADA system) on the one hand, and the road-side system on the other hand. Traffic flow models for ADA assessment at this level of detail are rare (see sections 3.3.4 and 3.4.3).

1.2 Objectives and scope

1.2.1 Research objectives

The research reported in this thesis is part of the project “Modelling of traffic flow during congestion for ADA applications” within the T3 TNO-TRAIL co-operation program of the Dutch Organisation of Applied Science TNO and the TRAIL Research School for Transportation Infrastructure and Logistics. Within both organisations continuous research efforts are made to provide better models of individual driver behaviour and flows in traffic networks, and to provide control systems for traffic flows or vehicle operations that improve traffic flow (see section 1.3.1). Within this context it is not desirable to provide just a single new traffic flow model, but rather a modelling framework is desired that enables one to flexibly include new behavioural models or ADA designs.

The objective of this dissertation thesis is to improve traffic flow theory as a basis for a traffic flow model that:

- provides a realistic representation of traffic dynamics, in particular congested traffic flow operations;
- bridges the gap between macroscopic traffic flow models and microscopic behavioural models of the individual driver;
- is suited for further refinement of the behavioural models for individual driver behaviour (provided these models are available), with the aim of modelling congested traffic operations and dynamics more realistically;

- is suited, among others, for modelling ADA systems and (co-operative vehicle – vehicle, vehicle – infrastructure) traffic control systems and their impact on traffic flow dynamics;
- provides a flexible framework in which alternative behavioural specifications for the driver or the ADA system can be modelled.

1.2.2 Research scope and limitations

We limit the scope of the research in this thesis as follows. First of all only motorised vehicles are considered such as cars, trucks, busses, and vans. Motorcycle, bicycle, or pedestrian traffic is beyond the scope.

Secondly we only consider unidirectional uninterrupted flow, that is: motorway traffic and *not* for instance urban roads with traffic lights or parallel parking, or rural roads with bidirectional traffic.

Thirdly the scope is on ADA systems supporting the longitudinal control task in *normal* traffic conditions. We thus exclude support systems for the more strategic level of control (e.g. route navigation), lateral control (e.g. lane keeping systems), or support limited to emergency manoeuvres (e.g. Anti-lock Brake Systems, Collision Avoidance Systems), unless these systems cause (indirect) changes to the normal longitudinal driving behaviour (Brookhuis et al., 2001). For instance, drivers with ABS may prefer shorter following distances because they feel safer in case of emergency braking (risk homeostasis).

So far, we have set limits to the scope of the system under examination. In addition, we also limit the viewpoint with respect to the system that we examine: the focus is on modelling the interaction between vehicles in traffic flow and the dynamic traffic patterns that emerge. We thus exclude detailed research into the individual models of the driver and of the ADA system itself. We will limit ourselves to formulating hypothetical, intuitive models that need to be refined, justified and validated in specific research efforts on these subjects (see also section 1.3.1). The hypotheses and assumptions on driver and ADA behaviour in this thesis are therefore given as first approximations. They also illustrate the potential mathematical format in which the results of the separate investigations of driver and ADA behaviour could be cast in order to serve as inputs for the traffic flow modelling framework presented in this thesis. Moreover, without attempting to select the best possible numerical solution scheme for our model, we apply a standard numerical scheme that is traditionally applied to macroscopic traffic flow models of this type, albeit with some pragmatic modifications.

Finally, the dependence of the model specifications presented in this thesis on behavioural models that are not yet readily available, prevents us – for the time being – to validate the model extensively. In theory it would be possible to calibrate and validate the model on the macroscopic level (i.e. matching macroscopic traffic flow variables by changing the microscopic models and parameter settings). However, this would render the effort of refining the microscopic foundations of the model a purely

theoretical issue, since validation at the macroscopic level would *not* guarantee that microscopic relationships underlying the model be valid. In that case, less elaborate models exist that are not based on microscopic principles, while being valid at the macroscopic level.

1.3 Research approach

Since, as was indicated in the previous section, the research in this thesis is most meaningful when considered as a building block in the research context to which it contributes, the research approach is first briefly positioned within a broader context (section 1.3.1), after which the approach of the work reported in this thesis is outlined (section 1.3.2).

1.3.1 Complementary research activities

Although the research reported in this thesis forms an autonomous piece of work in itself, it cannot be completely detached from the broader research context in which it was performed. We mention two complementary research activities that were performed in parallel and are still being continued at the time of writing this thesis: the development of quantitative individual driver behaviour models, and the collection of empirical data of individual vehicle operations in congested traffic.

1.3.1.1 Quantitative models of individual driver behaviour

In this thesis a traffic flow model is developed that builds on a quantitative model of individual driver behaviour, taking into account the influence of the behavioural variable *activation level*. In contrast to the variables in classical microscopic or macroscopic models this variable is not a physical quantity that can be directly measured. Yet, models have been formulated by psychologists to describe the driving task with consideration of behavioural aspects like alertness, awareness, activation level, attention level, work load etcetera. Such models are indispensable to provide a sound quantitative foundation for the activation level-based extension of the human-kinetic traffic flow model (chapter 6).

The model development in this thesis has therefore been performed in close co-operation with the Human Factors institute of TNO, where a mode-based system dynamic individual driver model is being developed. This *optimal control model* is an application of an existing behavioural framework that was designed to predict the performance of humans involved in a control task. Previous applications of this framework include a pilot flying a plane and various task definitions for controlling vehicles (Wewerinke, 1989; Wewerinke 1992). The main functions, which the driver has to perform to fulfil the complex driving task, are: perception, estimation (central information processing) and decision making concerning the task mode, and control involved in a given task mode. Within each task mode, the driver/vehicle model can be considered as an optimal control model (see e.g. Kleinman et al., 1970). The optimal control model (OCM) is based on the hypothesis that the driver is well-trained and well-

motivated, in other words, behaves optimally. Optimal here refers to both the assignment of the mental resources over the different driving tasks, and to the control response in steering the vehicle, given inherent limitations of erroneous perception and delayed processing.

The model was applied to certain scenarios typical of congested traffic, such as simple car-following (Wewerinke & Hoedemaeker, 2001; Wewerinke et al., 2002), closing in on slow traffic, stop-and-go driving (Wewerinke & Hogema, 2003) and merging onto a busy motorway (Wewerinke et al., 2003). Also a method was conceived to interface the results of this detailed driver model to the human-kinetic framework presented in this thesis (Tampère et al., 2002). The elaboration of this concept, a further development of the optimal control individual driver model to cover the full width of scenarios in congested traffic flow, and the validation on the individual driver level, are subjects strongly recommended for further research.

1.3.1.2 Empirical data collection of individual vehicle operations in congestion

In order to calibrate and validate the individual driver model discussed in the previous section and the human-kinetic model presented in this thesis, purely macroscopic traffic flow data that are widely available are not sufficient. For that reason a substantial effort has been done to collect empirical data of individual vehicle operations in congestion. For this purpose remote sensing techniques were developed that permit the conversion of accurate digital video images to position, speed and spacing data of all individual vehicles in the flow (Hoogendoorn, 2003). The images are recorded from a remote platform, e.g. a helicopter, flying approximately vertically at heights of ±700 m above congested traffic (Figure 1-1).

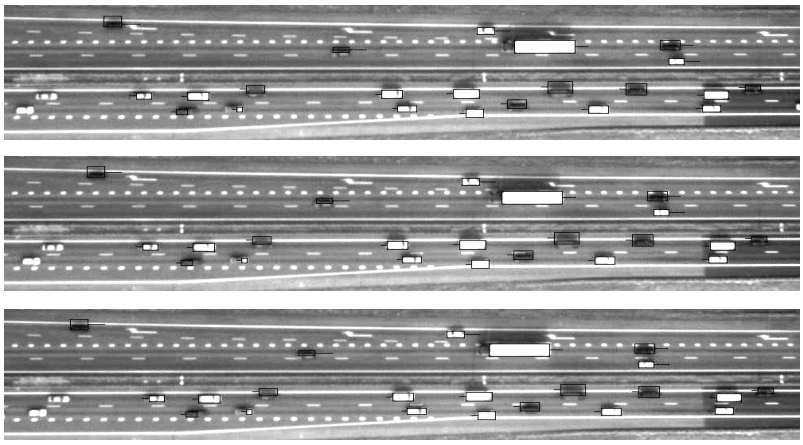


Figure 1-1 Sequence of processed images obtained through remote sensing; recognised vehicles are indicated as boxes, with a 'tail' proportional to the speed

Two data sets have been collected so far. One served as a test set for the development of the conversion algorithms from images to traffic data (Hoogendoorn et al., 2002). During the second helicopter flight, a rich data set for traffic research purposes was acquired from a stretch of 10 km upstream of the Everdingen exit on the A2 motorway in the Netherlands. Research on this data set is ongoing and publications are

forthcoming. Calibration and validation of the human-kinetic traffic flow model based on this data is strongly recommended for further research.

1.3.2 Research approach

Traffic flow theory is a field of research that has been extensively studied. Any contribution in this field should therefore start from and build on the empirical and theoretical findings reported in literature, as well as on a critical review of the numerous models that have been previously developed.

A state-of-the-art overview of empirical findings and theories is given in chapter 2. Care has been taken throughout the thesis to refer to empirical theory where appropriate, so that properties and outputs of the human-kinetic model can be qualitatively compared against empirical facts and theories. In chapter 3 the state-of-the-art of traffic flow modelling and applications to Advanced Driver Assistance systems is discussed. It is also explained how the findings from this review – together with our own experience – have led to the choice of the modelling approach that underlies the remainder of this thesis.

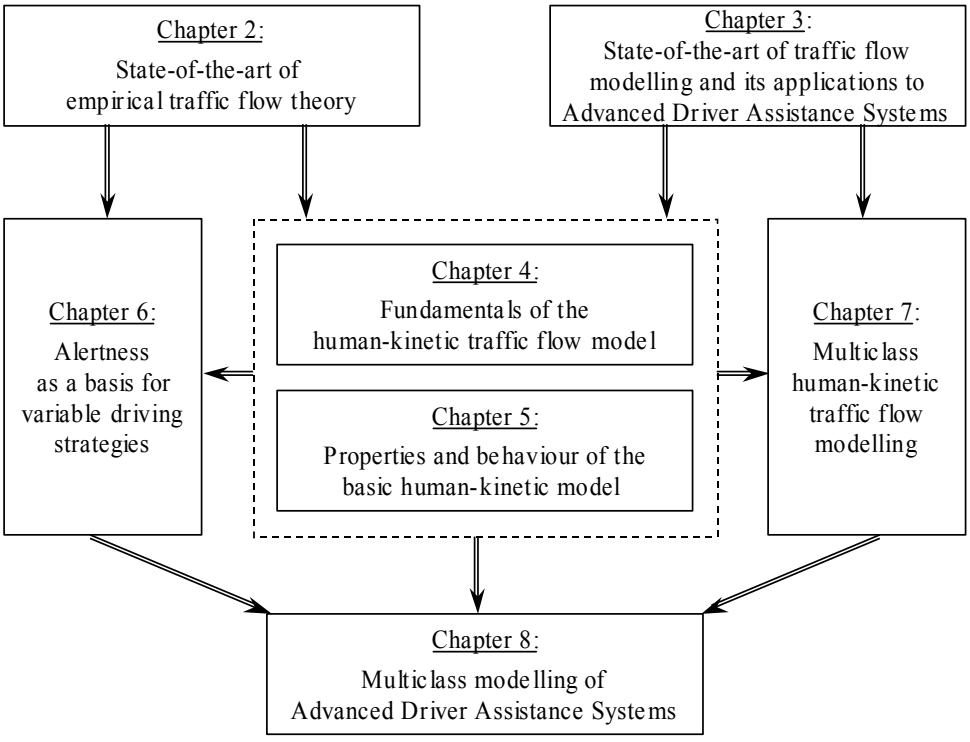


Figure 1-2 Structure of this thesis' main text

From the review in chapter 3, the kinetic modelling approach is selected as the most suitable for our purpose. The main arguments for this choice are: (i) that it inherently provides a relationship between the individual driver behaviour and the overall dynamics of traffic flow (micro-macro link); (ii) the specification of an individual driver is not extremely detailed and non-deterministic, which makes such models suitable for application in an early stage of ADA development, where the main concept and

workings of the system are known, but the exact deterministic specification in all its details is not.

Using the mathematical formalism from the kinetic modelling approach, we develop our own class of kinetic traffic flow models in chapter 4. Since the kinetic theory originated from the description of gas-kinetics, special care has been taken in the derivation in this thesis to avoid gas-like behaviour of the vehicles (which has often been a criticism against the kinetic modelling approach). Starting from a unique formulation of the acceleration behaviour of drivers in the *acceleration integral*, a macroscopic model is developed that builds solely on microscopic specifications of individual driver behaviour. For that reason the modelling approach of chapter 4 is called the *human-kinetic* traffic flow theory. Its properties are mathematically and numerically analysed in chapter 5 that together with chapter 4 and 6 constitutes the core of the thesis. The main results of chapter 5 are the stability analysis and the presentation of various congestion patterns that compare well to those of state-of-the-art models.

With the basic human-kinetic model established and its properties thoroughly analysed, the model is further refined in chapter 6. An inherent limitation of most existing traffic flow models is relaxed: instead of mechanistically modelling drivers as machines responding with exactly the same control actions in all identical traffic conditions, we introduce the *activation level* of drivers as a basis for *variable driving strategies*. The activation level is an additional variable that depends among others on microscopic and macroscopic traffic conditions and is associated with either an active or rather passive driving strategy. Firstly it is explained how this new variable is mathematically treated using the techniques of kinetic modelling. Secondly some example behavioural specifications are elaborated that illustrate how the activation level might depend on traffic conditions, and how it in turn affects driver behaviour. This part of the thesis in particular has a hypothetical character, which needs to be refined and validated in further research (see section 1.3.1). However, we are able to show how consideration of variable driving strategies, related to increases or decreases of the activation level, can contribute to some typical phenomena in congested traffic, like the capacity funnel and capacity drop.

Before applying the refined human-kinetic model to ADA systems, a multiclass version of the human-kinetic theory is developed in chapter 7. The aim is to be able to model mixed traffic flows consisting of ADA-equipped and non-equipped vehicles. Due to the special treatment of acceleration behaviour in the basic model of chapter 4, the state-of-the-art method for distinguishing multiple user classes in kinetic models needs to be tailored to the human-kinetic approach by means of a multiclass version of the acceleration integral.

Finally we illustrate the potential application to ADA analysis of the multiclass human-kinetic traffic flow modelling approach in chapter 8. The ADA system performs a queue-tail warning function based on inter-vehicle communication, comparable to the function that is currently implemented as an infrastructure-based system on Dutch motorways. The application uses all the features of the human-kinetic model that were established in the previous chapters.

1.4 Scientific contributions

The contributions of the research reported in this thesis to traffic flow theory and to the modelling of driver behaviour and ADA systems in kinetic and macroscopic traffic flow models may be summarised as follows:

1. *Establishment of a new class of kinetic traffic flow models (and equivalent macroscopic traffic flow models)*: the human-kinetic traffic flow modelling approach. This approach provides a micro-macro link, since it takes a specification of individual driver behaviour (e.g. a classical car-following model) as a starting point for the kinetic model development. Acceleration and deceleration are treated as continuous processes, so that the model naturally accounts for:
    a. Adaptive driver behaviour as a function of traffic flow conditions, through the stimulus-response structure of the underlying micro model
    b. Anisotropy of driver behaviour: drivers are only influenced by stimuli coming from downstream
    c. Non-locality of stimuli: drivers respond to conditions at a variable distance downstream, depending on the driving speed
    d. Finite space requirements: vehicles are not modelled as point-sized particles, but require a finite space that is a function of driving speed
    e. Limited acceleration and deceleration capability of vehicles
    f. All relevant correlations between speeds of predecessors and followers and the gap between them can in theory be accounted for.
2. Accounting for pure reaction time delay and for anticipation behaviour in macroscopic traffic flow models. The meaning of these terms is consistent with the definition of *reaction time* and *anticipation* at the individual driver level.
3. Insight into the behaviour and into the properties of this model: equilibrium solutions (which are output of this model, no input as is traditionally the case), analytical stability criteria, and numerical solutions. As a result new insights are gained into typical congestion related phenomena like the occurrence of different types of congestion waves and the role of driver behaviour related parameters herein.
4. Inclusion of *activation level* as a behavioural variable influencing actual driving strategy. As a result of this refinement, drivers are no longer modelled as machines responding with exactly the same control actions in all identical traffic conditions. Their driving strategy depends on the activation level, and the activation level depends on traffic conditions.
5. A multiclass model having the same properties as the basic human-kinetic model with activation level dynamics.

6. Modelling of Advanced Driver Assistance systems in a macroscopic traffic flow model, including inter-vehicle communication. This application exploits all the previously mentioned strengths of the activation level-based human-kinetic model.
7. This work was performed within a broader context aimed at closing the gap between individual driver modelling, data collection of individual driver operations in congestion and traffic flow modelling, in which the author actively participated (section 1.3.1).

1.5 Practical relevance

The work presented in this dissertation thesis can contribute to the solution of the following practical problems:

- The human-kinetic model is a flexible framework for exploring and assessing new ADA concepts. Its focus is on the macroscopic level, i.e. the impacts of the ADA system on traffic flow dynamics. Typically this level of detail is of interest when considering system optimum solutions (e.g. benefits for society, throughput), whereas at the level of the individual vehicle, the focus lies on the user optimum (e.g. driver comfort). The human-kinetic approach is therefore typically suited for the design or assessment of ADA systems by stakeholders interested in the system optimum, like road authorities or policy makers.
- In general, the improved description of traffic flow in the human-kinetic model is applicable in many practical estimation and prediction problems where macroscopic traffic flow models are applied. For instance, decision support systems and traffic information systems in traffic management centres use on-line traffic state estimation and prediction tools based on macroscopic models. Traffic management strategies are assessed and optimised off-line using macroscopic traffic flow models. These and similar applications benefit from an improved description of traffic flow.
- The introduction of behavioural models of the individual driver potentially improves the transferability of a validated traffic flow model. Current traffic flow models need to be recalibrated for each application in a new traffic network. This indicates that behaviour of traffic be dependent on local conditions in a way that is *not* described in existing traffic flow models. Consideration of (one or more) behavioural variables like the activation level can help to overcome this problem, so that – once the model is validated – it is transferable to other networks purely by configuring it to the local infrastructure, but without changing the behavioural parameters of the model.

2 STATE-OF-THE-ART OF EMPIRICAL TRAFFIC FLOW THEORY

2.1 Aim and structure of this review

This dissertation deals with modelling traffic flow dynamics, especially in congestion, and the role of human driver behaviour therein. The aim of this chapter is to provide insight into the current knowledge of the process that is modelled: traffic flow, its dynamics, phenomena commonly referred to as *congestion,* and the role of driver behaviour in these phenomena. We review the most important empirical findings in this area and the theories that were formulated based on these observations. As such, this chapter gives concrete form to the first requirement set to traffic flow models for ADA analysis in section 1.1.3: in order to provide a good representation of traffic flow, a model should be able to qualitatively and quantitatively reproduce the empirical observations discussed in this chapter, in particular those related to congested traffic. Throughout the thesis, we will therefore show simulation results obtained through the newly developed model and relate these to the empirical observations and theory discussed in this chapter. A review of the state-of-the-art of traffic flow *models* is given in a separate chapter (chapter 3).

The state-of-the-art review of empirical traffic flow theory is structured as follows. After a brief introduction of macroscopic traffic variables (section 2.2.1), the most common presentation of empirical data is taken as a starting point to discuss some basic characteristics of traffic flow: the fundamental diagram (section 2.2.2). In the remainder

of the chapter, we focus on some phenomena that are commonly referred to as congestion phenomena: traffic breakdown (section 2.3), congestion patterns (section 2.4), and the wave fronts that occur in these patterns: queue tails (section 2.5) and queue heads (section 2.6).

In the discussion of traffic breakdown in section 2.3, the conditions in which the breakdown process can be initiated and the role of spontaneous transitions and bottlenecks are discussed in section 2.3.1. Subsequently we focus on the nature of traffic breakdown, a chain reaction leading eventually to low speeds or standstill (section 2.3.2).

The role of traffic instability in the formation of congestion patterns (section 2.4) is taken as a guideline for the discussion: first we discuss congestion patterns in which instability is absent; then we turn to patterns governed by instability (sections 2.4.1 and 2.4.2 respectively).

Having recognised different congestion patterns, we zoom in on the processes at the queue tail, the upstream front of the congested area (section 2.5). Again the main distinction is made through the role of instability: the propagation or spillback of queue tails characterised by a stable deceleration process in section 2.5.1, and the role of instability in the queue tail in section 2.5.2.

Next the processes at the downstream front or *head* of the congested area are discussed (section 2.6). These processes determine how efficiently traffic can leave the congested region: phenomena like hysteresis and capacity drop are determinant in this respect (section 2.6.1). Again instability plays a decisive role in determining whether the queue head remains at a fixed location or not (section 2.6.2). The properties of a moving queue head are briefly addressed (section 2.6.3). Finally we discuss the location of the queue head as compared to the location of the instability that triggered jam formation (section 2.6.4).

The conclusions of this empirical state-of-the-art overview for traffic flow modelling in general and in this thesis specifically are drawn in section 2.7.

2.2 Macroscopic traffic flow variables and their fundamental correlations

In this section macroscopic traffic flow variables are introduced. The term *macroscopic* refers to the variables describing traffic flow characteristics directly at the level of detail of the flow, as opposed to variables used to described the motion of individual vehicles that are referred to as *microscopic* variables.

2.2.1 The macroscopic traffic flow variables: density, flow rate and average speed

Macroscopic traffic flow analyses have been performed since the 1950's. A traditional way of representing traffic data is based on the relationship between the speed, flow rate, and density. These variables are defined in Figure 2-1, where *trajectories* (i.e.

paths in space and time of individual vehicles) are represented. Traffic *density* $k_{\Delta x}$ is defined as the number of vehicles in a spatial region of width Δx at time t. In Figure 2-1 the density at time t over region Δx equals the number of black squares divided by the length of the region Δx. Note that the density $k_{\Delta x}(t, x, \Delta x)$ depends on the region Δx and may vary with time. Since the number of vehicles is an integer number, the density $k_{\Delta x}$ is a stepwise varying function of time.

Analogously to the definition of the density, the *flow rate* $q_{\Delta t}$ is defined as the number of vehicles in a temporal interval of length Δt at location x. Note that the flow rate $q_{\Delta t}(t, x, \Delta t)$ depends on the choice of the temporal interval Δt and may vary with location. In Figure 2-1 the flow rate at location x over interval Δt is determined by the number of black circles divided by the length of the interval Δt.

Since the number of vehicles is an integer number, the flow rate $q_{\Delta t}$ is a stepwise varying function of time.

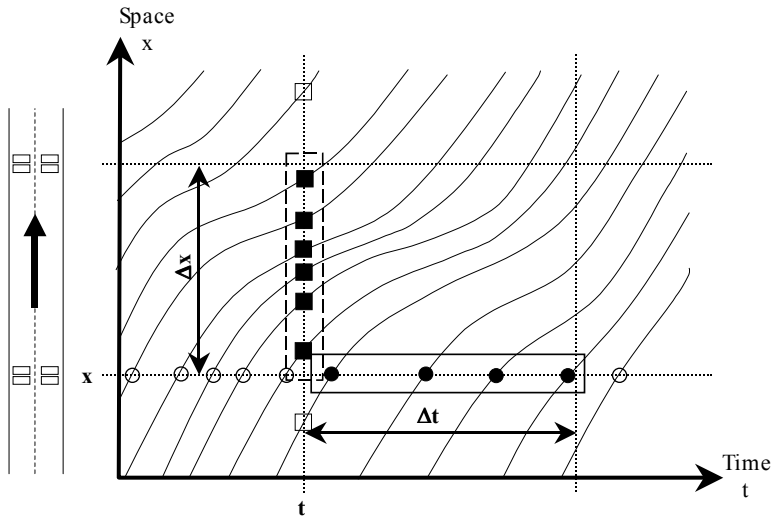


Figure 2-1 Definition of the density $k_{\Delta x}$ and the flow rate $q_{\Delta t}$ from trajectories

Without going into details, we mention that a time-continuous density and flow rate definition $k(t,x)$ and $q(t,x)$ can be obtained as the derivative of a smooth approximation of the cumulative count functions with respect to space or time respectively or from stochastic considerations (see section A.2 in Annex A). These continuous variables are local, instantaneous variables and are also the basis of macroscopic traffic flow models. One can prove that the following relation holds for the *continuous* density and flow rate in stationary conditions (e.g. Leutzbach, 1988):

$$q = k V \tag{2.1}$$

In this equation V stands for the continuous *average speed*, which is herewith implicitly defined (see also section A.2 in Annex A).

2.2.2 *Fundamental diagram*

Our empirical knowledge of traffic flow stems almost entirely from measurements obtained from local traffic detectors, i.e. observations at a cross-section. The aim of this section is to provide the reader who is not familiar with empirical traffic data and terminology related to congested traffic flow theory, with some examples of traffic data that were obtained through local detectors. At the same time we introduce the traditional way of representing these data in fundamental diagrams, along with terminology that will be defined and discussed more precisely in the remainder of this chapter. Finally our intention is to show the inherent limitations of local empirical data: their representation in fundamental diagrams is subject to all kinds of measurement site specific influences that complicate the derivation of any general theory from these measurements. We believe that this is the main cause for much debate among traffic flow theorists up to present.

Most local traffic detection devices – usually intended for daily traffic monitoring in traffic management centres – detect the passage of individual vehicles at a fixed cross-section (longitudinal position) along the road. From these detectors the following direct measurements are obtained:

- passage time of individual vehicles
- occupation time of the detector by individual vehicles

Often these direct measurements are not at the disposal of the researcher, since before transmitting the data to the central traffic information centre, local processing stations transform the single-vehicle data into one or more indirectly derived measurements or time-aggregates of individual vehicle measurements, for instance:

- individual speed (from passage time difference at a closely spaced detector pair),
- estimation of the vehicle length, often as indication of vehicle type (from occupation time and speed of individual vehicle),
- flow rate (see 2.2.1),
- occupancy (% of time a detector was occupied by vehicles during some time interval),
- average speed (average of individual speeds during some time interval, or derived from occupancy and flow rate; e.g. Zwet et al., 2002),
- traffic composition (aggregate of vehicle types).

The density is often estimated from the aggregate flow rates and speeds using equation (2.1), although the reader should be aware that this relationship for continuous variables only approximately holds for measurement data that are by definition discrete (we refer to Leutzbach (1988) for some classical pitfalls of data aggregation and incorrect use of theoretical relations like (2.1)).

A classical fingerprint of traffic flow at a certain measurement location is obtained through the so-called *fundamental diagrams* of average speed against flow rate, speed

against density, or flow rate against density. It is important to realise that such diagram reflects a *local relationship* between the macroscopic traffic variables, which should *not* be perceived as a *causal* relationship.

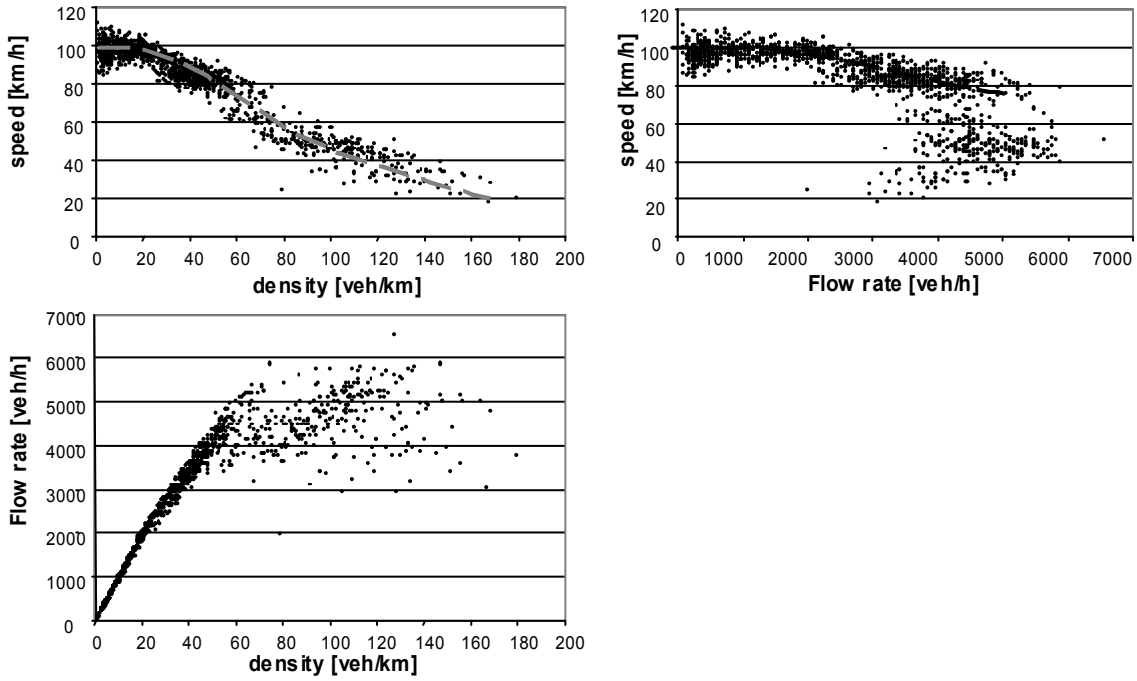


Figure 2-2 Example of empirical fundamental diagrams: 1-minute averaged measurements from 0.5 km upstream of the western entrance of the Kennedy tunnel on the E17, a 3-lane motorway in Antwerp, Belgium (see Figure 2-4)

Figure 2-2 shows an example of an empirical fundamental diagram[1] (obtained through local measurements of flow rate and speed, together with relation (2.1) as an approximation for the density). Although such a diagram is by definition only a representation of local traffic conditions at the location of the detector – and therefore by no means 'generally true' – this example has some characteristics that have also been observed by several researchers worldwide and can be regarded as generally accepted facts (while other characteristics in this specific example are *not* generally true, see below):

- The trend line of the average speed against density (dashed line in Figure 2-2) decreases monotonically. In the example here the decay is exceptionally strong (speed drops to 30 km/h already for 150 veh/km or 50 veh/km/lane), which can be explained from the particular location of the measurement (see Figure 2-4).

[1] The aim of the fundamental diagrams shown in this section is *not* to interpret as correctly as possible the specific examples, but rather to illustrate for the reader not familiar with this type of data, the kind of observations that can be made. Of these observations we indicate in the text which ones are general, in the sense that they are also observed at most other locations, and which ones are rather exceptional. Moreover, we preliminarily introduce a lot of terminology here that is defined later on in the text; see the cross-referenced sections for definitions.

- Under some critical density k_{crit} (here approximately 60 veh/km or 20 veh/km/lane), the relationship between the flow rate and the density is strongly correlated. Data points are grouped closely around a trend line with almost constant slope for low density; the slope decreases when approaching the critical value.
- In the same density range, the average speed is more or less constant for low density (and equal to the free speed, here 100 km/h) and decreases when approaching the critical value (dashed line in Figure 2-2).
- Above the critical density, the data points are scattered, i.e. the correlation between the traffic variables is less pronounced than for densities below the critical value. Would the time aggregation interval be smaller, this scattering would even be larger and vice versa.
- The data points seem to be clustered in groups, in this specific case two point clouds with k_{crit} or V≈70 km/h as an approximate boundary: speeds appear to drop suddenly from values of 80 km/h and higher to speeds of 60 km/h and lower. In general, groups of data points exist indeed, belonging to separate *traffic phases* with relatively sharp *phase transitions* between the phases (see section 2.4), although the corresponding data point clouds will not always be as clearly separated as in the example of Figure 2-2.
- Generally the phase with strongly correlated flows and density ($k < k_{crit}$) is referred to as the *free flow* branch of the fundamental diagram, and the more widely scattered point cloud for $k > k_{crit}$ as the *congestion* branch.
- The flow rate does not exceed some maximum value (here 5800 veh/h). In this specific example the flow rate does not decrease for higher values of the density, which is typical for locations in or downstream of a bottleneck (see below). This maximum value is referred to as the *capacity*, in this case the *queue discharge rate* (see below and section 2.6.1).

Some observations in Figure 2-2 are not found in similar plots of data collected at different locations and vice versa. In general, the shape and characteristics of the fundamental diagrams depends strongly on the measurement location and on the composition of traffic (i.e. mix of vehicle types). Figure 2-3 schematically shows the dependence of the fundamental diagram on the location of the measurement with respect to the *bottleneck* location (see section 2.3.1), upstream of which a queue builds up. The bottleneck C can process flows up to some maximal value, the *capacity* (see section 2.6.1 for definitions of capacity). Traffic demand below the capacity of location C is indicated in Figure 2-3 with dots; traffic demand higher than or equal to the capacity of C with white circles. For locations D downstream of the bottleneck that can theoretically process higher flow rates than bottleneck C, all measurements lie on the *free flow* branch of the fundamental diagram. The same applies for locations A with higher capacity that are so far upstream of bottleneck C, that the queue building up in region B does not spill back to A. Measurements in zone B shortly upstream of the bottleneck segregate into two groups of data points, corresponding to different traffic *phases*: a free flowing phase, corresponding to the period before bottleneck C has been

activated and generated a queue that reached B, and a congested phase with flows lower than or equal to the bottleneck capacity, low speeds, and high density.

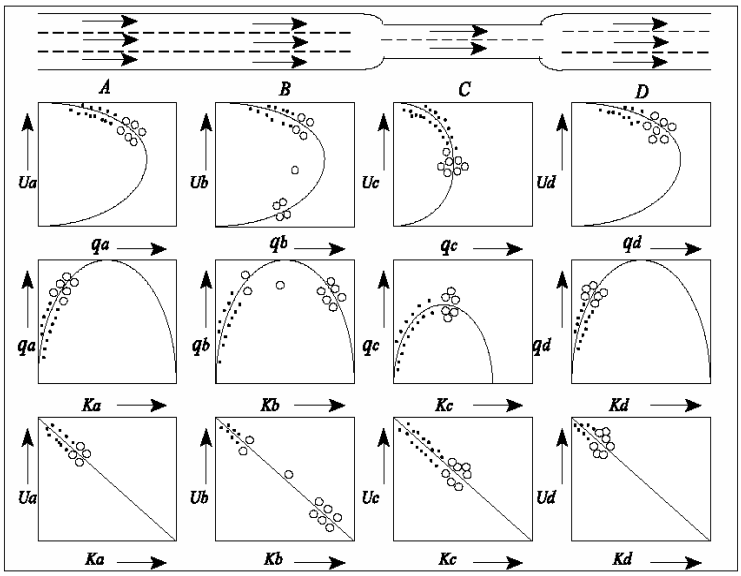


Figure 2-3 Schematic dependence of the fundamental diagram on the measurement location with respect to the bottleneck location (reproduced after May, 1990)

Comparing Figure 2-2 to Figure 2-3, it is no surprise that the measurements of Figure 2-2 were recorded shortly downstream (near location C) of a strong active bottleneck, where a busy on-ramp enters the motorway from the left (whereas in Belgium right-hand driving rules apply). This specific location (Figure 2-4) also explains some peculiar characteristics of the measurements that deviate from the general pattern. The unusually strong decay of speed with density occurs because traffic enters from the left, so that slow vehicles have to traverse 2 lanes to reach the slow lane, which causes strong disturbances. Since in the example of Figure 2-2 the speeds in the low speed phase are clustered around approximately 50 km/h with high flow rates, we conclude with Helbing (2001, fig. 5) that we measure the recovery from congestion to free flowing traffic (zone C, very near to B in Figure 2-3). Similar observations were also reported by Banks (1999) and by Kerner & Rehborn (1996b). The latter author suggests that they are associated with a third traffic phase called *synchronized traffic* (see section 2.4.2).

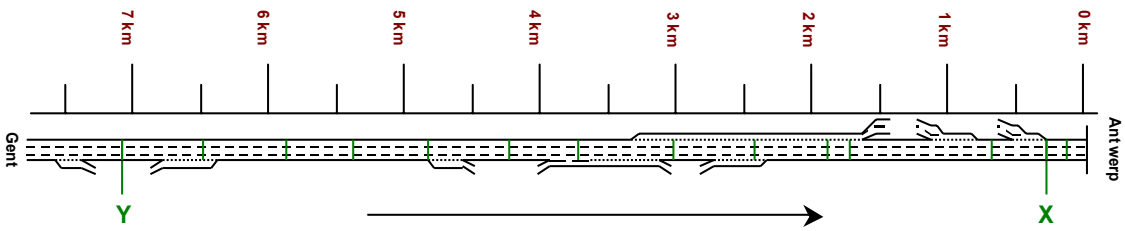


Figure 2-4 Road configuration with indication of the locations of the measurement data shown in Figure 2-2 (X) and in Figure 2-5 (Y)

Typical measurements associated with a location upstream of a bottleneck (location Y in Figure 2-4, comparable to a location of type B in Figure 2-3) are depicted in Figure 2-5. Again this example is only for illustrative purposes, and it is shown here since – apart from local peculiarities – it reveals some additional observations of traffic flow that are agreed upon worldwide:

- The diagram of flow rate versus density consists of two branches: the already mentioned free flowing branch and the congested branch; here it appears clearly that the congested data points are widely scattered around a regression line (dashed line in Figure 2-5) with negative slope: flow rate decreases with the density.
- The free flow branch has a few measurement points at higher flows than those occurring in the congested region, so that, together with the aforementioned observation, the diagram of flow rate against density has a typical *reversed lambda* (λ) shape (Koshi et al., 1983).
- The high flows in the tip of the reversed lambda are maintained only during a short period (flows > 5000 veh/h during 26 minutes only, indicated with arrows in Figure 2-5), after which a transition from free flowing traffic to congested traffic occurs.
- Due to the reversed lambda shape, there is a range of densities (here 40-70 veh/h or 13 to 23 veh/h/lane) for which traffic can be either congested or free flowing, dependent on the history and the traffic state downstream of the measurement location.
- Eventually at very high densities, traffic speed measurements approach zero. The density associated with standstill is called the *jam density*.

The observations of the reversed lambda shape and the short periods of high flows prior to congestion are related to the phenomenon of *capacity drop* and *hysteresis* (see section 2.6.1). Note that Figure 2-5 has a particular property that is *not* typical for fundamental diagrams: for very low flows, the speed decreases. This effect is site-specific and an effect of changing traffic composition: these data points are related to night traffic, which – due to the presence of a large harbour in Antwerp – has a very high fraction of trucks, driving at lower speeds than cars. This site-specific observation shows how the characteristics of the fundamental diagram not only depend on the measurement location with respect to bottlenecks but also on the composition and characteristics of traffic.

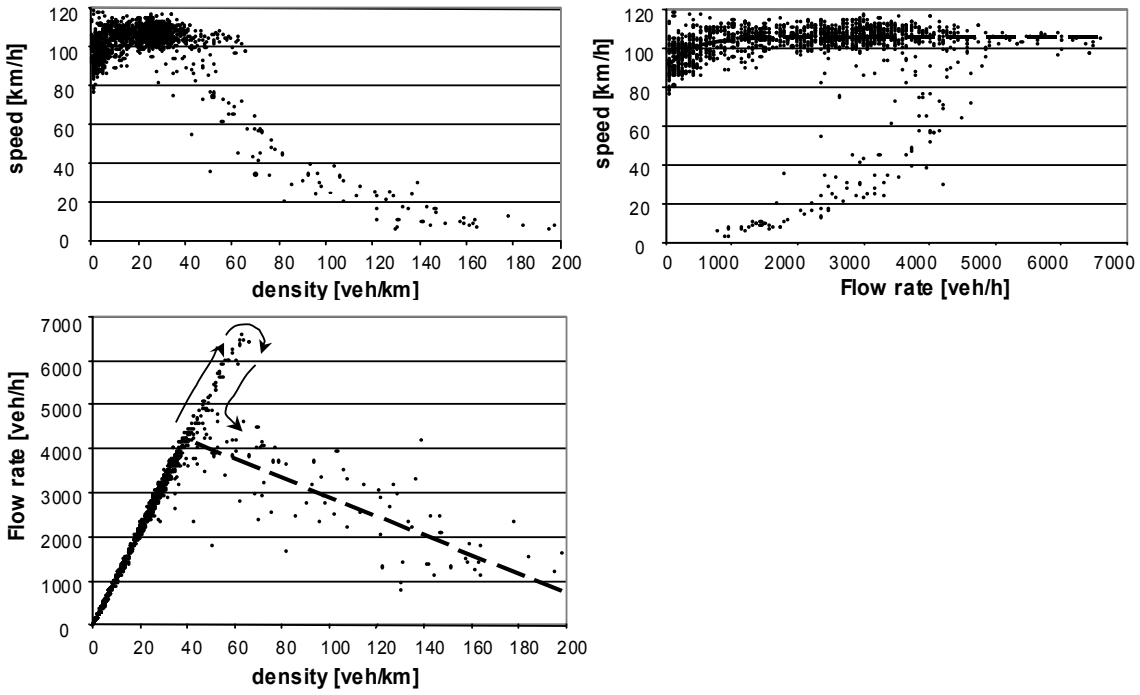


Figure 2-5 Example of empirical fundamental diagrams: 1-minute averaged measurements from 7 km upstream of the western entrance of the Kennedy tunnel on the E17, a 3-lane motorway in Antwerp, Belgium (see Figure 2-4)

For the sake of completeness, we mention that more insight from local measurements in traffic flow dynamics can be gained by combining measurements from different locations. Two methods for representing data from multiple locations are worth mentioning: *rescaled cumulative count* plots and *space-time* plots. The former method plots the cumulative counts of vehicles at different locations in one single graph (Figure 2-6), so that the accumulation of vehicles between the measurement locations is visualised, along with more refined characteristics of the flow (Cassidy & Windover, 1995; Windover, 1998). This representation of data has been the basis of many studies by the 'school of Berkeley' into *first order* behaviour of traffic (see sections 2.5.1 and 3.3.1).

Alternatively, one of the variables flow, speed, or density can be plotted as a function of time t and location x in space-time plots, possibly with interpolated values for t and x values that have not been measured (Treiber & Helbing, 2002). This method especially supports the visual inspection of the propagation and growth of disturbances, and of phase transitions. Therefore, it is the preferred representation for analyses of *second order* behaviour of traffic (see sections 2.5.2 and 2.4.2).

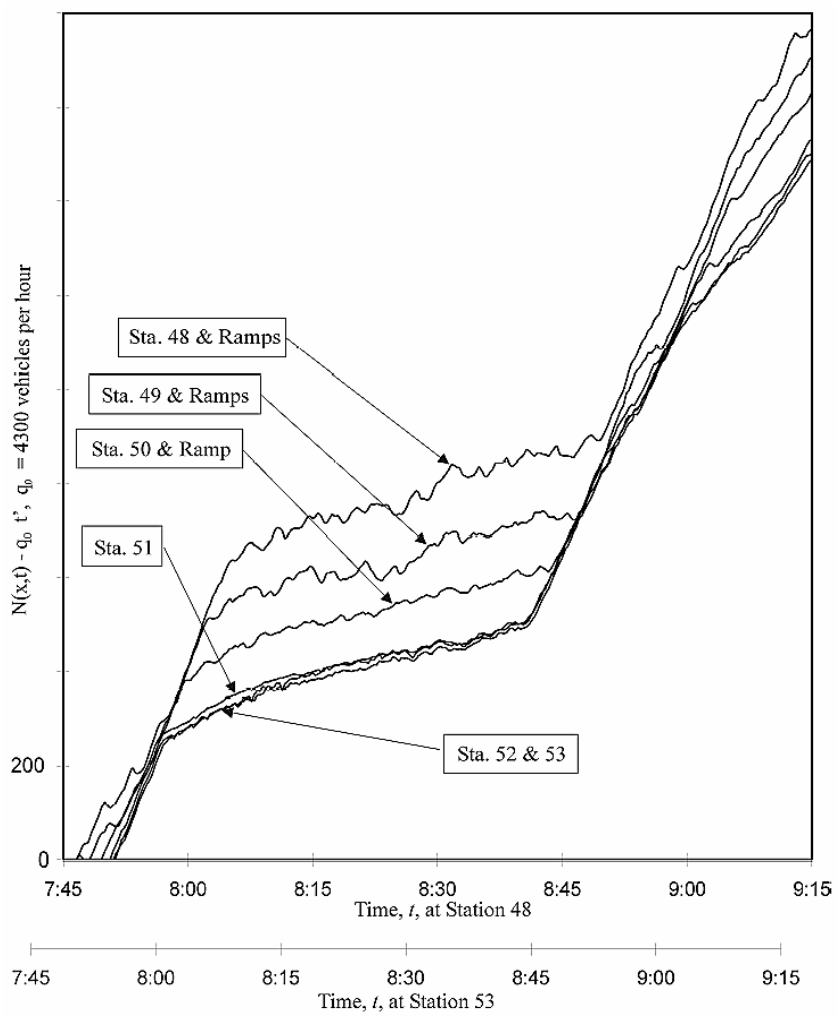


Figure 2-6 Example of traffic flow analysis through cumulative count plots; the figure represents rescaled cumulative counts of vehicles at different detector stations ("Sta.") (reproduced after Mauch, 2002)

Properties of the fundamental diagram in relation to driver behaviour

The observations in the fundamental diagram can also be interpreted from the point of view of individual driver behaviour. The fundamental diagrams of speed against flow or density confirm that drivers have some *desired* maximum speed: even in low densities – when drivers are not influenced by other traffic – the average speed of traffic is limited and nearly constant. When the density increases, the probability that drivers are confronted with slower traffic and cannot immediately overtake increases, which leads to a slight decrease of the average speed with density (up to the critical value). This applies more in situations where drivers are forced to perform more lane changing manoeuvres – as in the case of Figure 2-2; this is in contrast to Figure 2-5 where traffic breaks down *before* this effect can be sensed.

Once faced with congestion, drivers are all driving in a *constrained* instead of *free* mode. This is: their behaviour is (partly) determined by the presence of another vehicle immediately in front. The free flow regime thus corresponds to flow in which drivers are free to choose their own speed during most of the time, or are bound to the desired

speed of other drivers. The congestion regime is associated with *all* drivers driving slower than their desired speed because virtually every driver is constrained.

The declining trend line, around which the congested data points are scattered, reveals that there is a positive correlation between the average driving speed and the average distance between vehicles. The fundamental diagram reveals no causality in either direction. Therefore, from the point of view of the driver two interpretations are possible: either the driver adapts his speed to the available (small) distance to the predecessor, or the driver adapts the following distance to his predecessor to the speed with which traffic is driving.

Conclusions from fundamental diagrams

Traffic has long been analysed based on the representation of cross-section measurement data in fundamental diagrams. It appears already from the examples given in this section that certain factors complicate the analyses based on this representation:

- Measurements are widely scattered, especially data points corresponding to congested traffic and for short aggregation intervals Δt;
- There is no one-on-one relationship between density and flow: certain densities can be associated with more than one flow regime;
- The relationships between the measurements are strongly influenced by the location of the measurement with respect to a bottleneck;
- The relationships between the measurements are strongly influenced by the composition of traffic if the behaviour or physical dimensions of different user groups differ significantly;
- The representation does not indicate how the measurements evolve in time: what is the sequence of the data points and do they represent transient states or stable conditions (also called *equilibrium conditions,* see section 3.3)?

Nevertheless, the representation of cross-section data in a fundamental diagram reveals a few characteristics of driver behaviour. The main distinction in the fundamental diagram between free flow and congested regime, is associated with (part of the) drivers reaching their desired speed or not. However, the diagram only suggests *correlations* between declining speed and decreasing road space per vehicle (i.e. increasing density) without indicating the causality. Neither does the fundamental diagram reveal how the transition between free flowing and congested regimes occurs. The next section therefore summarises empirical knowledge about the process of *traffic breakdown* and the conditions in which it occurs.

2.3 Traffic breakdown

Traffic consists of vehicles, controlled by drivers, who are all assumed to have the intention to travel safely at their desired maximum speed. Then how can a high speed free flowing traffic state change into a low speed congested traffic state? Traffic flow

theorists do not agree on this subject. Surely, the discussion of the fundamental diagram in the previous section showed that the density of traffic is an important indicator determining which traffic state will occur. The debate is whether an excess of some critical density at bottlenecks in the network is the *only* cause for breakdown into a congested state, or that *instability* of traffic flow plays an equally important role. In this section we discuss the role of bottlenecks and of traffic flow instability in section 2.3.1. Having recognised the importance of traffic flow instability, we elaborate on the factors contributing to this phenomenon in section 2.3.2.

2.3.1 Conditions for traffic breakdown: bottleneck-induced or spontaneous breakdown and the role of traffic flow instability

The primary cause of traffic breakdown is clearly that the motorway is too heavily loaded. This causes the high densities in which breakdown becomes highly probable.

Some authors therefore argue that traffic breakdown is always deterministically associated with the presence of a bottleneck (Daganzo et al., 1999). These authors build their theories on the concept of *capacity* of a motorway. For now, we define the capacity as the maximum flow rate that can be processed by a motorway before flow breaks down (we refine this definition in section 2.6.1, where it turns out that the definition given here is the 'prior-congestion' capacity). A *bottleneck* is a location where either (a) traffic enters the motorway so that the sum of the main lane traffic and merging traffic exceeds the capacity or (b) the capacity of the road drops to a lower value than the actual flow rate it is carrying. The former happens exclusively at motorway on-ramps or merges, the latter can be associated with a steep grade, an incident, a change in traffic regulation (speed limit), a weaving area, a tunnel, a bridge, or another visual discontinuity (Janssen et al., 1995). Muñoz & Daganzo (2000) report another type of bottleneck associated with an off-ramp or diverge, where traffic leaving the motorway spills back onto the main lanes due to capacity limitations at the end of the off-ramp. In fact this bottleneck can be regarded as type (b), since the capacity for traffic on the main lanes is reduced due to capacity being partly occupied by exiting traffic that blocks some lane(s).

Kerner (1998) reported the breakdown to congestion in absence of any apparent bottleneck. This seemingly *spontaneous breakdown*, also referred to as *phantom jams*, occurs in critically dense traffic. The theory of Kerner states that in these conditions specific events (like lane changes, other actions by individual drivers) can cause perturbations that exceed some critical threshold, above which the perturbation is amplified and eventually causes congestion. That is, spontaneous breakdown occurs if two conditions are met: (i) due to some deterministic cause (e.g. lane change, control error by a driver), an initial perturbation is generated with enough amplitude to grow (*pinch region*), and (ii) traffic flow in which this happens is unstable, so that the initial perturbation grows into a jam (see section 2.3.2). Apart from spontaneous transitions, which are rather rare phenomena, Kerner also recognises transitions to congestion due to bottlenecks. In fact, Kerner distinguishes two types of congested traffic: synchronised traffic (characterised by medium speeds and high flows) and wide jams (characterised

by speeds and flow rates close to zero). Kerner's theory (1999) states that transitions between free, synchronised or jammed traffic states are all possible. They are caused by local critical perturbations (pinch region) in traffic. Since the critical amplitude of local perturbations needed for a transition from free flow to synchronised flow is much lower than that needed for a free to jammed flow transition, the former are by far more numerous. In fact, spontaneous transitions are always from free flow to synchronised traffic, since due to the absence of external sources of strong perturbations (bottlenecks) the transition to jammed traffic is highly improbable.

Daganzo and Kerner are the most explicit propagators of seemingly different theories on traffic breakdown: the former recognises only bottlenecks as a cause for breakdown, whereas the latter points out that instability is crucial and can even lead to spontaneous congestion in absence of a bottleneck. However, both views are not *entirely* contradictory. In view of the mechanism of instability (explained in section 2.3.2), the breakdown mechanisms near bottlenecks and for spontaneous breakdown are not fundamentally different. Both situations differ in the probabilities for critical perturbations and the susceptibility of the flow for perturbations. In bottlenecks of type (a), the extra traffic causes a high number of large perturbations while at the same time the presence of extra traffic raises the density to such an extent that the probability for amplifying critical perturbation is practically 100%. Therefore, the probability of traffic breakdown at such bottlenecks is very high, and the bottleneck behaves like a deterministic cause of breakdown. In bottlenecks of type (b), the amount of traffic remains constant, but capacity decreases. That is, flow becomes more susceptible for amplifying critical perturbations. This bottleneck theory is confirmed by Krauss et al. (1999), who show that the probability of breakdown is dependent on the traffic conditions (flow rate, velocity variance) and on the length of the considered homogeneous road. This is consistent with the theory, since the probability of a critical perturbation increases with time and thus with the length travelled. Other than the length of the road, the time-to-breakdown may therefore also be a random variable describing the breakdown probability (Evans et al., 2001).

2.3.2 Stability, metastability and instability of traffic flow

Given the bottleneck theory of the previous section, instability of traffic flow is a crucial phenomenon. Moreover, as we discuss later in sections 2.4.2 and 2.5.2, instability is also the key factor determining the type of congestion that occurs after free flowing traffic has broken down. This section reviews empirical findings and theories on the stability of traffic flow.

The stability of traffic flow is determined by the stability of follower-leader pairs in platoons (section 2.3.2.1) and by the density, which among other factors is determinant for the distribution of platoons and inter-platoon gaps in traffic (section 2.3.2.2). These factors determine whether an initial perturbation leads to a chain reaction in which the perturbation is transferred to subsequent vehicles while growing in amplitude, or dampens, so that traffic remains free flowing. Finally section 2.3.2.3 recalls the concept of *metastability*.

2.3.2.1 Local and asymptotic stability of car-following in platoons

A control system is said to be *stable* if disturbances with finite amplitude to an equilibrium state lead to a dynamic response in which the amplitude of the disturbance decreases in time, and the equilibrium state is gradually restored. In line with this definition, the stability of traffic can be analysed on several levels. If a single pair of a leader vehicle and a following vehicle is considered, the stability of the car-following behaviour is referred to as *local stability* (Leutzbach, 1988).

In traffic flow *platoons* of vehicles occur. Platoons are sequences of one leader (who drives unconstrained) followed by a number of constrained following vehicles. The distance between the last vehicle of one platoon and the first vehicle of the next one is the *inter-platoon gap*. A platoon is by definition *asymptotically stable* if a disturbance with finite amplitude in the speed of the platoon leader is transferred to consecutive follower-leader pairs in the platoon with decreasing amplitude (Leutzbach, 1988). Local stability is a necessary but no sufficient requirement for asymptotic stability (which is also referred to as *platoon stability*).

When considered as a feedback control process, the primary influence factors for local and asymptotic stability of car-following behaviour are the pure time delay or reaction time of drivers, and the dynamic relaxation time or feedback control gain.

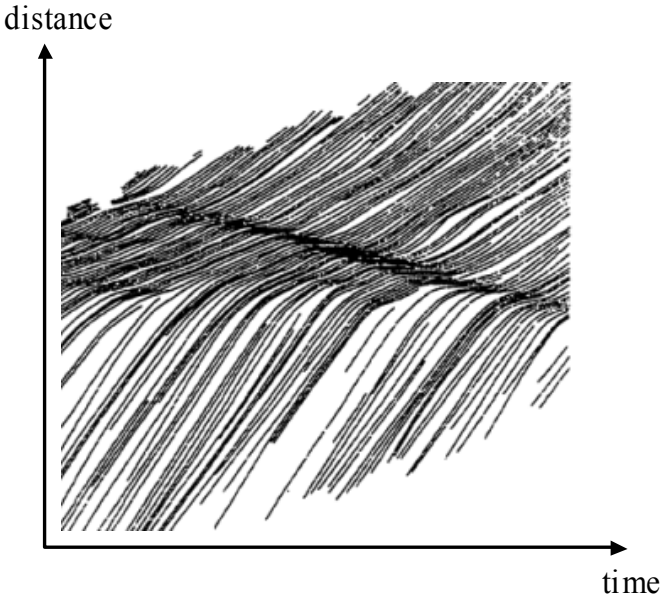


Figure 2-7 Trajectories of vehicles forming a jam (reproduced after Treiterer & Myers, 1974)

The question now is: is real traffic locally or asymptotically stable. This can only be examined using measurements of individual vehicle movements or so-called *trajectory* data. The rare publications discussing this type of data reveal that platoons are – at least in the reported cases – unstable. The famous article by Treiterer & Myers (1974) shows trajectories of a so-called *phantom jam*: the speed disturbance in a long platoon is gradually amplified until vehicles come to a standstill (Figure 2-7). It takes until the occurrence of a large inter-platoon gap for the jam to disappear. Coifman (1997)

publishes trajectories of thirteen shockwaves in platoons. In four of these recordings a small disturbance grows as it travels upstream in the platoon, until vehicles have to halt. In none of these, lane changes play any role, so that the growth of the disturbance's amplitude is purely due to asymptotical instability of the car-following behaviour in a vehicular platoon. Empirical evidence for local instability of car-following behaviour is given by Ranjitkar et al. (2003), who conclude from an experiment on a test track that the average responses of drivers were unstably both locally and asymptotically.

2.3.2.2 Traffic flow instability: determined by platoon instability and platoon structure

If car-following in a platoon is asymptotically unstable, why do initial perturbations then sometimes grow into congestion and sometimes not? That is because stability of traffic flow is not the same as local or platoon stability. Traffic flow must contain enough long platoons combined with short inter-platoon gaps, so that instability within one platoon is transferred in a chain reaction onto the next platoon, where the perturbation continues to grow in amplitude.

It is therefore imaginable that a platoon is *not* stable but traffic flow *is*. Consider an asymptotic unstable platoon undergoing a disturbance. The last vehicle will undergo the disturbance with larger amplitude than the platoon leader. If this vehicle succeeds to accelerate to a speed equal to or higher than the speed of the leader of the next platoon *before* this leader is close enough to react to the disturbance, then the next platoon will remain undisturbed. No matter if the next platoon is asymptotically stable, the initial disturbance is damped and hence traffic flow is stable. It appears that asymptotic instability of long platoons is sufficient to make traffic flow unstable *only* if platoons are long enough, and inter-platoon gaps short enough; the density is therefore a determinant variable. However, also the heterogeneity of traffic and choice of lane by drivers determine the platoon structure. For instance, if fast and slow vehicles are mixed in heterogeneous traffic, the faster vehicles will form platoons behind the slower vehicles.

Summarising, traffic flow stability is determined by:

- The local and asymptotical stability of car-following behaviour;
- The distribution of platoon sizes;
- The distribution of inter-platoon gaps.

Evidence of traffic flow stability versus platoon stability is already present in the single disturbance of Figure 2-7 discussed by Treiterer & Myers (1974). The platoon instability leading to a short jam is not enough to cause traffic flow breakdown within the time frame observed, because a large enough inter-platoon gap prevents the disturbance from being transferred onto the next platoon. Experiments in the Holland Tunnel, New York, reported by Edie & Foote (1961) confirm this theory. The authors observed that single-lane flow without overtaking in the tunnel was disturbed by flow waves (instabilities) occurring in dense traffic. By introducing 4-second gaps in the flow entering the tunnel, they were able to prevent the occurrence of these waves and to slightly increase capacity. They even report an experiment that started with uncontrolled

flow; once the waves occurred they started introducing gaps again, which made the waves dissolve. With these experiments Edie & Foote proved that traffic flow stability can be obtained in spite of platoon instability, if sufficiently large gaps or inter-platoon headways are present.

2.3.2.3 Metastability

The stability of traffic flow as discussed in the previous section is determined by the platoon length, by the length of inter-platoon gaps, and by the amplitude of the disturbance. Both the lengths of the platoons and of the inter-platoon gaps are stochastically distributed in traffic flow and dependent on the density, the distribution of the desired speeds, the speed limit, overtaking opportunities, etcetera.

Therefore, the stability of traffic flow with respect to a given disturbance is not deterministic. It depends on the probability that the disturbance affects a platoon that is long enough, combined with an inter-platoon gap that is short enough for the disturbance to be amplified. In this context traffic flow is stable if this probability is very low and unstable if this probability is high. Traffic flow is called *metastable* if it is stable with respect to disturbances with small amplitude, but unstable with respect to larger disturbances.

Metastability in traffic was empirically confirmed among others by Lee and co-workers (2000), who observed traffic near a bottleneck where different congestion patterns emerged (or not) dependent on the history (like earlier occurrence of perturbations). Similar empirical evidence for metastability and different congestion patterns for seemingly equal boundary conditions are reported by Treiber et al. (2000). Other observations of metastable flow are made by Kerner & Rehborn (1996b) and Kerner (1999).

2.4 Congestion patterns

In the previous section, the breakdown of traffic flow from free flowing regime to congestion was described. Furthermore, instability of traffic flow was discussed and empirical evidence was referenced, showing that traffic flow instability can be observed at least in certain conditions. In this section we discuss congestion patterns that have been reported in literature. The notion of stability and instability is crucial in this respect. Section 2.4.1 reviews traffic flow theory and empirical findings on congested traffic patterns that can be explained *without* considering traffic flow instability; these patterns are called *first order* patterns, named after the first order traffic flow model that is able to capture these aspects of congestion formation and propagation (see also section 3.3). The next section 2.4.2 adds to this, congestion patterns in which instability is crucial, named *higher order* patterns, since a higher order traffic flow model is required to describe the emergence of these patterns (see also section 3.3 and 3.4). This section discusses the congestion patterns as a whole. For a more thorough review of the processes at the up- and downstream end of these patterns, the reader is referred to the next sections 2.5 and 2.6.

2.4.1 'First order' congestion patterns

Daganzo (2002a), Mauch (2002), Cassidy & Bertini (1999a, 1999b), Muñoz & Daganzo (2000, 2001) present a variety of empirical studies of congestion patterns near bottlenecks. These studies are based on data obtained from local detectors that detect the moment of occupation and clearance of a location by every individual vehicle. Common in these studies is the analysis by means of cumulative count curves and transformations of these curves (Cassidy & Windover, 1995; Windover, 1998, see also Figure 2-6). We give a summary of findings using this method, and refer to Daganzo (2002a) for a more elaborate discussion and more references:

- congestion always sets in at deterministic locations in the network, i.e. at bottlenecks, for instance immediately downstream of merges, or upstream of diverges on motorways (Muñoz & Daganzo, 2000);
- more precisely, congestion sets in at or some short (~ 1 km) distance downstream of a bottleneck (capacity funnel, section 2.6.4) after a period of high prior-congestion flow rates (hysteresis, Cassidy & Bertini, 1999a; see section 2.6.1);
- at on-ramps, merging traffic enters the motorway at the cost of traffic on the main lanes – even if the motorway is congested, i.e. upstream of an on-ramp the flow rate is equal to the queue discharge minus the ramp flow, or in other words: congestion does usually not spill back onto the on-ramp because merging traffic takes priority over main lane traffic (even though legislation may say the opposite) (Cassidy & Bertini, 1999a; Cassidy & Mauch, 2001);
- in the absence of on-ramps, oscillations in flow rate, speed, or occupancy propagate regularly (i.e. without instability nor growth in amplitude) in traffic flow, either with the speed of traffic or against the direction of flow with speeds approximating 20 km/h (Cassidy & Bertini, 1999b; Cassidy & Mauch, 2001; Muñoz & Daganzo, 2000);
- large oscillations moving upstream with increasing amplitude have been observed on motorways with on-ramps (Daganzo, 2002a; Mauch, 2002); although the exact cause for this growth cannot be read from fixed location detectors, there is evidence that it is closely related to lane changing effects (mainly those near the on-ramp into and on the main lanes); there is no evidence for growth of disturbances other than through lane changes.

The main conclusion by this school of authors is that a congestion pattern is determined by deterministic sources of perturbations at *bottleneck* locations (Daganzo et al., 1999). These perturbations then move up- or downstream in a *stable* way, i.e. according to the first order kinematic or shock wave theory (see section 3.3.1). Seeming contradictions with this theory are mostly due to differences between lanes and/or between user groups. Therefore, the first order traffic flow theory is complemented by lane changing and accordingly the notion of one- versus two-pipe regimes (homogeneous over lanes versus segregated in slow and passing lanes) and of fast and slow vehicles (rabbits and slugs). Remarkably the authors belonging to this school have never published empirical

research into the mechanisms causing congestion patterns that are attributed to instability of traffic by other authors (e.g. emergence of stop-and-go waves, oscillatory congested traffic, localised clusters, etcetera). Nevertheless, several empirical observations of such patterns are known in literature (see next section 2.4.2). At best, the propagation of these perturbations in moderately dense traffic is considered (Mauch, 2002), or the authors show that phase transitions from free flow to congestion have been caused by bottlenecks, but without explaining the amplitude or oscillating character of the resulting congestion wave(s) (Daganzo et al., 1999).

2.4.2 'Higher order' congestion patterns

Apart from the patterns described in the previous section, other phenomena have been reported that find their origin in instability of traffic flow. Most notable are the phase diagram of congested states presented by Helbing et al. (1999), Lee et al. (2000) and Treiber et al. (2000), and the wide jams and synchronised traffic reported by Kerner and co-workers, which are all discussed in the next sections.

2.4.2.1 Congested phase diagrams of Helbing and Treiber and of Lee and co-workers

In 1998 and 1999, two teams of researchers presented the theory that the congested regime may not be a single dynamic phase but rather a collection of multiple phases, each of which is realised under different conditions. Initially the theory was formulated based on simulation results, but empirical evidence for the theory was published shortly thereafter. Helbing et al. (1999) simulated motorway traffic near a merge, using different combinations of main lane flow and on-ramp flow (Figure 2-8). They found that light combinations yielded instabilities or narrow clusters of congestion that either remained at the location of the on-ramp (Pinned Localised Cluster, PLC) or travelled upstream (Moving Localised Cluster, MLC). For larger flows the occurrence of the clusters repeated itself (through the boomerang effect, section 2.6.4), so that a pattern of Triggered Stop-and-go waves (TSG) emerged. Yet more loaded combinations yielded a continuous pattern of Oscillating Congested Traffic (OCT), and the worst combinations led to Homogeneous Congested Traffic (HCT). These patterns are consistent with the theory provided in section 2.3: the on-ramp flow acts as a source of disturbances (the size of which depends on the flow rate) and limits the capacity for the main lane flow; whereas the main lane flow rate determines the probability that these perturbations are amplified. Therefore, different combinations of both flows yield more or less frequent – and eventually constant – breakdown of traffic near the merge. In later publications Treiber et al. (2000) provided empirical evidence for the existence of all these patterns and even of coexisting patterns on German motorways. Moreover, they were able to simulate the patterns using both microscopic and macroscopic traffic flow models with open boundaries.

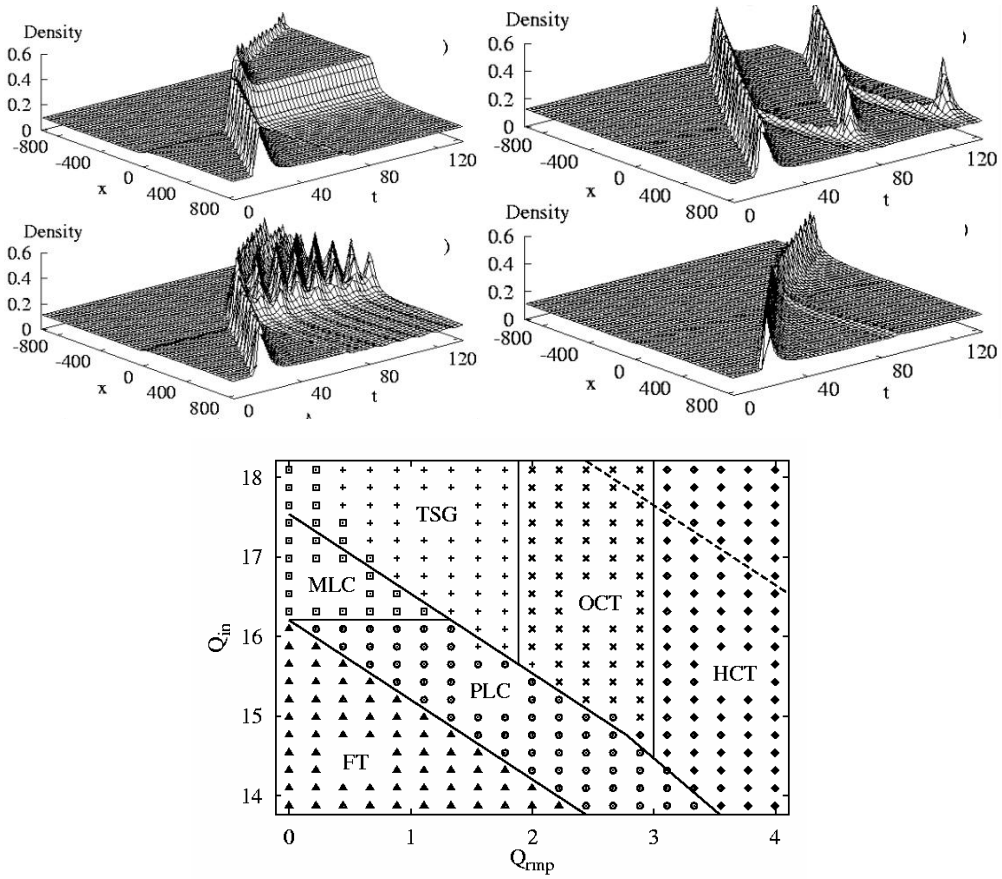


Figure 2-8 (a)* above*: Simulated higher order congestion patterns; (b)* below*: phase diagram of congested states (on-ramp flow Q_{rmp} versus main-lane flor Q_{in}) (reproduced after Helbing et al., 1999)

Almost simultaneously Lee et al. (1998, 1999) presented a similar theory. Using a macroscopic traffic flow model (section 3.3), they were able to show that metastable traffic can adopt several oscillatory or stable congested states. They showed that metastability (see also section 2.3.2.3) implies that either free flow or one or more congested traffic patterns can exist for the same initial traffic density, dependent on the presence and precise features of a temporal and finite perturbation in the potential bottleneck. As a result a jam cluster forms that – dependent on the on-ramp and main-lane flow – remained static near the on-ramp or moved upstream, triggering new jams. Theoretically Lee et al. (1999) were able to predict a phase diagram similar to that of Helbing (Figure 2-8). Empirical confirmation of their hypotheses was published in 2000. Lee et al. (2000) registered three types of congestion patterns at the same on-ramp in South-Korea. They were able to identify these as PLC, OCT and HCT and confirmed that the time-averaged on-ramp flow and main lane flow are characteristic parameters of the congested traffic flow. Since not all potential combinations of traffic demand at the bottleneck occurred, the other congestion patterns were not registered, but their existence is neither falsified. Apart from these correspondences to the findings of Helbing and co-workers, also subtle differences were found. For instance, Helbing found that for a fixed main lane flow an increasing ramp flow determines whether TSG, OCT or

HCT occurs, whereas Lee identifies the main lane flow as the determinant factor, causing similar congestion patterns while the on-ramp flow remains fixed. Probably the different bottleneck configuration explains these minor differences, so that the correspondences between the theories and empirical findings of both research teams provide convincing evidence for a common bottleneck theory.

2.4.2.2 Three-phase theory of Kerner

The work of Helbing and Lee was partially inspired by empirical findings earlier in the second halve of the nineties by the German physicist Kerner and co-workers. In 1996 Kerner & Rehborn published experimental data showing the existence of *wide jams* (speeds and flows nearly zero). In another publication Kerner & Rehborn (1996b) reported the existence of synchronised traffic (speeds lower than free flow, flows comparable to free flow). These three empirically observed traffic phases and transitions between these states have formed the basis of the *three-phase theory* by Kerner (1999). The fact that synchronised traffic can be either homogeneous, homogeneous-in-speed (but not in flow or density) or a complex pattern, suggests that it is a collection of traffic states, similar to those distinguished by Lee or Helbing. Helbing & Treiber (2002) argue that their theory and observations, and those of Kerner and co-workers are in fact closely related and support a common theory. Kerner on the other hand disagrees on this viewpoint and shows that both theories have essentially different solutions and that three phase traffic theory has its own and qualitatively different phase diagram (Kerner et al., 2002). More specifically Kerner denies the existence of a unique equilibrium relationship between density and average speed (or flow rate), as a stable attractor of congested traffic states; instead he argues that there exists a 2-dimensional range in the flow-density plane of equilibrium states that can be maintained for hours, *without converging* to a *single* point on an equilibrium curve. As far as driver behaviour is concerned, this discussion essentially comes to the following issue: do drivers following a predecessor with a stable constant speed have *one* desired equilibrium following distance as a function of the driving speed, or are they satisfied with *any distance within a reasonable range* (Kerner & Klenov, 2003)?

2.5 Processes at the queue tail

In this section we discuss in more detail the processes in the upstream jam front or *queue tail* of jams that occur as parts of the patterns described in the previous section 2.4. In the next section 2.6, processes in the downstream jam front or *queue head* will be more closely examined. Understanding these processes is required to explain how the first or higher order patterns of the previous section can occur and expand.

As in the previous section, the discussion is structured dependent on the role of traffic instability. Section 2.5.1 examines the queue tail movement in absence of traffic instability, whereas in section 2.5.2 the influence of instability in the queue tail is discussed.

2.5.1 *Queue spillback*

In the queue tail, vehicles decelerate from a phase with higher speed to a phase with lower speed. It is agreed upon that the motion of the upstream jam front is determined by the flow rate and density in the fast region upstream of the front, and by the flow rate and density inside the slow congested region. The physical law of conservation of vehicles explains that a jam front where vehicles can be supplied at a higher rate than the rate with which they leave, moves in the direction of the supply (and vice versa) since vehicles are then stacked upstream of the jam front. The density of this packing compared to the original density determines the speed of the jam front motion. The same principle holds for smaller differences between zones with different flow rate and density; we speak of *kinematic waves* referring to the propagation of incremental perturbations, whereas the propagation of a finite perturbation is called a *shock wave*. The propagation of these waves is described by the first-order traffic flow theory discussed in section 3.3.1. Empirical evidence for this theory is provided by the authors mentioned in section 2.4.1.

2.5.2 *Instability in the queue tail*

Besides propagation of the queue tail, its stability is an important property. The stability of the queue tail is determined by two counteracting processes: the retarded adaptation to the (low) speed in the queue versus the ability to anticipate.

Retarded deceleration

Vehicles approaching slower traffic have to decelerate from a high speed level in free flowing traffic to a lower speed level in the queue. This manoeuvre is not instantaneous for two reasons: (i) a pure time delay exists between the perception of the traffic state and the acceleration response and (ii) to annihilate the speed difference takes a finite time, dependent on the deceleration level. Therefore, the speed is not immediately adapted to that in the queue, so that compression to higher density than that of the original queue can occur and the flow is potentially unstable.

Anticipation

On the other hand, drivers have the ability to anticipate: they can see the traffic conditions downstream and are able to extrapolate the current conditions into the near future. This gives them the opportunity to start the deceleration before actually reaching the queue tail, so that the speed difference between free flowing and congested regime is bridged gradually, thereby approaching the queue tail with the appropriate speed.

In theory researchers agree that retarded deceleration and anticipation behaviour exist and that they counteract each other: slower responses can be compensated by further anticipation (see also sections 3.3.2 and 4.3.2). However, researchers disagree whether these aspects are always balanced or not. The former case (anticipation is at least large enough to compensate for delayed response) would lead to stable propagation of jam fronts and traffic that can be adequately described by models that neglect traffic instability (see first order models in section 3.3.1). In the latter case delay sometimes

prevails (which would lead to propagation and growth of amplitude of jam fronts, which can only be described by higher order models, see section 3.3.2). To the best knowledge of the author, no empirical observations specifically of the queue approaching process are reported. Conclusions should therefore be drawn from the empirical studies of congestion patterns as discussed in section 2.4.

Roughly speaking researchers of the University of Berkeley California defend the first order theory, thus implicitly assuming stability of traffic during the approach of the queue tail (e.g.: Daganzo, Newell, Cassidy, Bertini, Muñoz, Mauch; see also section 2.4.1) whereas European researchers, most pronouncedly represented by (German) physicists, defend second or higher order theories that recognise potential instability of queue tails (Kerner, Helbing, Treiber, Hennecke, Papageorgiou, Kuehne, Lee, Hoogendoorn, Aw, Rascle, Klar, Wegener, Zhang; see also section 2.4.2).

2.6 Processes at the queue head

The processes at the downstream front or *head* of the congested area are important since they determine how efficiently traffic can leave the congested region: phenomena like hysteresis and capacity drop are determinant in this respect (section 2.6.1). As in the previous section the stability of traffic during formation of the jam plays a crucial role: it determines whether the downstream congestion front (or queue head) remains at a fixed location or not (section 2.6.2). In section 2.6.3 we review empirical properties of the moving queue head. As a last process at the downstream congestion front, we discuss the location of the queue head as compared to the location of the instability that triggered jam formation (section 2.6.4).

2.6.1 Hysteresis, capacity drop or the two-capacity phenomenon

So far, we have not defined the capacity of a motorway; we have only briefly mentioned that the term capacity refers to the maximum flow rates that can be processed by a road (section 2.2.2). This is however not a sufficient definition. The absolutely highest flow rates are in general observed shortly before traffic breakdown. We will refer to these high flow rates as *prior-congestion* capacity. They can however not be maintained during periods longer than approximately 0 to 40 minutes[2] before traffic breaks down. Another capacity definition is therefore that of the *queue discharge* capacity. This definition refers to the highest flow rates that occur at the head of the queue once traffic has broken down. Queue discharge capacity (also: outflow of the queue) can last for long periods but is in general lower than the prior-congestion flow rates. The difference between maximum prior-congestion and queue discharge flow rates is called the *capacity drop* or the *two-capacity* phenomenon (Banks, 1991). Differences of 5-10% (Cassidy & Bertini, 1999a; Bertini & Cassidy, 2002), over 20% (Smulders et al, 2000) up to 50% (Kerner, 1999) have been reported.

[2] Approximate range derived from the different references in this section.

These widely varying results suggest that different factors can contribute more or less to the capacity drop depending on the specific situation. Let us discuss three potential contributors:

- *Time-to-breakdown*
  Once traffic reaches a critical density, a sufficiently large perturbation is needed before traffic actually breaks down (see also section 2.3); in this period flow rates might be temporarily higher than the flow rate that can be stably sustained afterwards. Because of the unstable character of these overcritical densities, these conditions do usually not last longer than 10-30 minutes, dependent on the demand pattern on the main lanes (Lorenz & Elefteriadou, 2000). This process is comparable to superheating or supercooling of fluids before the transition to vapour or ice, which is only a matter of time before a nucleus initiates the phase transition (Kerner, 1999).

- *Hysteresis in car-following behaviour*

  It is observed that the representation of the traffic state follows a different path or phase trajectory in the fundamental diagram during breakdown to congestion than during recovery from congestion back to free flowing traffic (Figure 2-9); this effect is called *hysteresis*[3] (e.g. Treiterer & Myers, 1974; Zhang, 1999). Different aspects of car-following behaviour contribute to this effect. Drivers respond with a certain delay (reaction time + dynamic relaxation of the response; see also section 2.5.2) to actions of the predecessor, leading to temporary deviations from the desired following distance. During braking, gaps are therefore smaller and during acceleration larger than during stable following; this is even more the case when the control response of drivers during braking is stronger than during acceleration (since during decelerations control errors or delays are more dangerous with respect to rear-end collisions). As a result of this compression and expansion, maximum flow rates in accelerating traffic (=queue discharge) are lower than those in stable or decelerating flows. Zhang (1999) presents a theory on driver behaviour that explains hysteresis. Consistently with an earlier theory by Newell (1965), he distinguishes an *anticipation-dominant* phase, a *relaxation-dominant* phase and a balanced anticipation and relaxation phase in between these two extremes. The essential difference is that in dilute traffic conditions drivers anticipate to disturbances ahead, but in denser traffic drivers cannot see the disturbance until it reaches them, so that relaxation to prevailing traffic conditions becomes dominant. Drivers at the head of the queue are constrained and cannot anticipate to conditions ahead. Therefore, the velocity – and hence the flow rate – is retarded which causes capacity drop.

[3] *Hysteresis* in general, refers to the fact that a system follows a different trajectory in the state space, dependent on the direction of the transition.

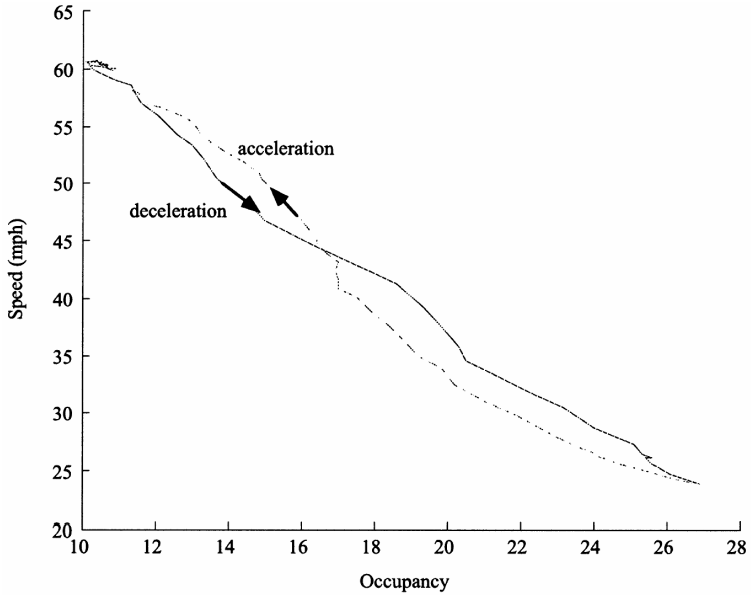


Figure 2-9 Phase trajectory with hysteresis (reproduced after Zhang, 1999)

- *Changes in driver behaviour*

  The above mentioned aspects cannot explain *large* capacity drops. Some authors therefore suggest that drivers accept shorter following distances prior to congestion than after traffic breakdown (Smulders et al., 2000). Several explanations are possible here. Drivers might be more motivated to follow closely when speeds are higher than when speeds have collapsed and the gain of taking more risk disappears (Van Toorenburg, 1983a, 1983b; Daganzo, 2002a; Dijker et al., 1999). This behaviour leads to an increase of the already mentioned hysteresis: longer distances between vehicles at the head of the queue and hence a queue discharge rate lower than without this less motivated behaviour. Variable driver behaviour is further discussed in section 6.2.

2.6.2 Fixed or moving location of the queue head

The queue head is the downstream jam front of a congested region. In this front traffic accelerates from a lower to a higher speed regime (usually free flowing traffic). Apart from the spontaneous breakdown reported by Kerner (see section 2.3.1) the jam front is initially (i.e. immediately after breakdown to congestion) always at the location of a bottleneck. Dependent on the role of traffic instability, the acceleration front or queue head remains stably at the location of the bottleneck or moves upstream, after which a new breakdown can occur. We discuss both possibilities based on Figure 2-10.

A bottleneck is considered, caused by on-ramp traffic entering a motorway. Suppose that the stationary demand is A on the main lanes and B on the on-ramp. Assume further that the maximal flow that can be carried by the motorway both before and after the merge is C, and that the sum of A and B exceeds C. For simplicity suppose that the merging flow B takes priority over the main lane traffic. Therefore, traffic breaks down near the merge and a queue builds up on the main lanes.

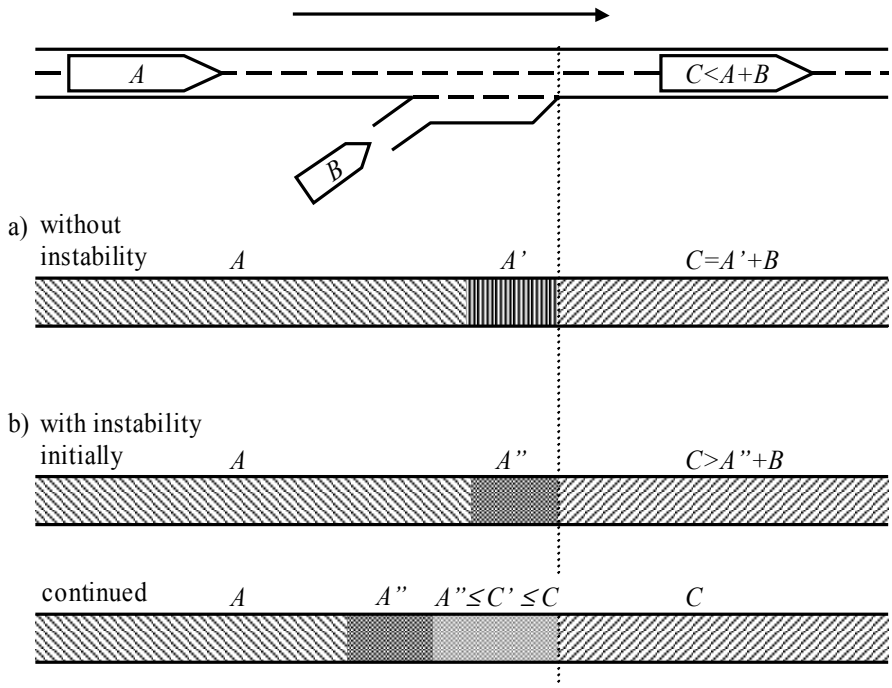


Figure 2-10 Flows determining the dynamic behaviour of the queue head at a bottleneck

Stable decelerating jam front

First consider the scenario of Figure 2-10a without instability in the decelerating jam front (upstream). Due to the capacity constraint in the bottleneck and the assumption of merging traffic taking priority over main lane traffic, the flow in the jammed area is limited to $A'=C-B$. Therefore, the region A' has higher density and lower speed than the region with flow rate A. The deceleration of vehicles crossing the upstream jam front between region A and A' is stable and hence also the flow rate inside the region A'.

Now let us examine what happens at the downstream end of the congested region A' as a result of this stable flow. Consider the accelerating jam front (dotted line): at the left hand side of this front (upstream side), a total flow arrives equal to $B+A'$, whereas at the right hand side of this front (downstream side), traffic is evacuated at a flow rate C. Because these flows are equal, traffic *leaves* the accelerating jam front at the same rate than the rate with which it is *supplied*, so that this front remains *stationary* at the location of the merge.

Unstable decelerating jam front

Now consider the scenario of Figure 2-10b with instability in the decelerating jam front (see the remainder of the article for more insight into the conditions for traffic instability). As in the previous case, traffic initially breaks down due to the merging towards a flow rate A' and the associated lower speed. However, instability implies that due to *overshoot* the speed inside the jam temporarily drops lower (and density

increases) and hence also the flow rate drops below A' to A"<A' (in the worst case, traffic comes to a standstill and the flow A" would temporarily become zero).

Let us again examine what happens at the downstream end (accelerating jam front) of the congested region at the instant when flow has dropped to A". At the left hand side of the accelerating jam front, the sum of the flows is B+A", whereas at the right hand side, traffic can be evacuated at a rate C>A"+B. Since more traffic is evacuated by flow rate C than can be supplied by A", the downstream end of the densely packed zone A" is 'consumed', in other words: the accelerating jam front starts moving in the upstream direction.

However, as time elapses (see Figure 2-10b, continued), the accelerating jam front arrives upstream of the merging area, the flow B is no longer present, and the discharge rate of the accelerating jam front (which is indicated as C') can gradually increase from A" to the maximal flow rate on the main lanes, which was previously assumed to be C. After some time this gradually increasing flow C', combined with flow B of the on-ramp again exceeds the maximal throughput C, so that, at the location of the merge, traffic breaks down again, and the process that was just discussed repeats itself.

This discussion has shown that, dependent on the stability of traffic flow during breakdown at a bottleneck, two scenarios can occur. Either traffic is stable during breakdown, so that the downstream (or accelerating) jam front remains fixed at the bottleneck location. Or traffic during breakdown is unstable, so that a repetitive process is initiated: a jam is formed, but the accelerating jam front of that jam travels upstream, so that flow at the breakdown location is restored; this causes a restore of the flows to such an extent that a new breakdown occurs, with a new jam that travels upstream and so on. This illustrates that instability is a necessary condition for the occurrence of oscillating (or periodic) congestion patterns like stop and go waves or oscillating congested traffic.

2.6.3 Properties of a moving queue head

The analysis of the previous section has illustrated how moving queue heads can be generated at bottlenecks with unstable traffic. In this section some empirical studies of moving queue heads and their special properties are mentioned.

Kerner & Rehborn (1996a) present data and theories stating that the downstream front of wide jams (see section 2.4.2) has fixed and 'universal' properties: propagation velocity (upstream), density upstream and downstream of the queue head, and outflow rate (about 50% to 2/3 of prior-congestion capacities). The fixed properties only depend on the weather, composition of traffic, vehicle length, and accepted safe time clearance, but not on traffic conditions. These findings of stationary properties of jam outflow have been (qualitatively) confirmed by Treiber et al. (2000).

The more or less constant outflow rate of wide jams has been confirmed in a research of Dutch motorways (Kruisbergen, 2001). A number of moving jams on different two-lane motorways was tracked and the in- and outflow rates and wave speeds of both upstream and downstream jam front were measured. The outflow rate turned out to be

approximately constant (in person-car equivalents; about 70% of prior-congestion capacities). The speed of the queue head however, varied between 3 and 6.5 m/s. No determinant factors, like percentage of trucks or weather influence, were tested as potential explicatory variables.

2.6.4 Capacity funnel

According to the discussion in section 2.6.2, queue heads would be generated exactly at the location of bottlenecks, for instance where traffic merges. In this section some empirical evidence is discussed that refines this image: the first onset of congestion often occurs some distance downstream of the actual bottleneck that initiates the jam formation.

The first reference to this phenomenon was made by Buckley & Yagar (1974). They used the name *capacity funnel* to indicate this phenomenon, since it seems like the capacity is only gradually reduced in the direction downstream from the location of the bottleneck. Bertini & Cassidy (2002) and Cassidy & Bertini (1999a) discuss a capacity funnel that was predictably reproduced on several weekdays. Traffic breakdown occurred always between 500 and 1000 m downstream of a merging lane and remained fixed there. In an attempt to confirm these findings, Smulders et al. (2000) found that the effect was less predictable than reported by Cassidy & Bertini. On some days the onset and stable queue head location of congestion coincided spatially with the location of a merging area, on other days the onset and queue head of congestion were located more than one kilometre downstream of the merging, and yet on other occasions the onset was downstream of the merge, after which the queue head settled stably at the merging location.

Little theories are available explaining these differences. Helbing (2001) shows simulations where congestion sets in approximately 1 km downstream of a merging area. In other publications Helbing & Treiber (1998) show a *boomerang* effect: a disturbance at a motorway entry is first carried in the direction of the flow, while in the meantime its amplitude increases. Having gained enough amplitude, the direction of the disturbance inverts, and it travels upstream back to the on-ramp. Upon passing the location of the merging, a new perturbation is generated that initially travels downstream, and the process repeats itself.

From the point of view of driving behaviour, the postponed – and therefore spatially shifted – onset of congestion could be explained by considering the motivation of drivers. Daganzo (2002a) suggests that drivers can drive with shorter average headways if they are motivated to do so. Other empirical evidence of temporary higher flow rates due to motivated driving was provided by Muñoz & Daganzo (2001, see also section 6.2.1). Near merging zones several motivations are conceivable, like:

- merge-give way behaviour (drivers temporarily accept shorter headways for vehicles entering the motorway)
- the opposite behaviour: drivers following closely to prevent vehicles from cutting in in the gap in front of them

- anticipation by lane changing: in order to facilitate the merging drivers change lanes to the passing lanes, where temporarily higher densities are accepted.

We will explicitly consider the capacity funnel effect in this thesis and make assumptions on driver behaviour in these conditions (see section 6.5).

2.7 Conclusions

Since the earliest empirical and theoretical studies in the 1950's, the insight into traffic flow dynamics has increased gradually. The so-called first order behaviour of traffic flow is nowadays widely accepted, and is able to describe the occurrence of queues, the propagation of jam wave fronts and the relation between average speed, density and flow rate. The vast majority of empirical data this theory was based upon comes from traffic detectors at fixed locations (cross-section data), which has long-time been analysed mainly through its representation in fundamental diagrams. In the second halve of the 1990's, more refined aspects of congestion formation and propagation have gained broad attention. On the one hand, researchers have tried to extend the first order theory to explain specific congested traffic flow characteristics near bottlenecks. On the other hand, traffic instability – that has been studied since the 1970's already – has gained renewed attention, since it forms the basis of a wide variety of so-called higher order congested traffic patterns, including stop-and-go traffic. From the point of view of driver behaviour, aspects of longitudinal behaviour are the main factors influencing traffic dynamics, like: distance keeping with respect to the predecessor, anticipation to speed changes, and retarded adaptation of the speed to changing traffic conditions.

Although traffic instability has been empirically studied and theoretically analysed, it remains unclear to which extent and in which conditions it influences the formation and propagation of congestion in real traffic flows. Two schools of traffic flow theorists vigorously defend different views on this issue, the first one denying the role of instability in congestion formation or propagation, the second one focussing on traffic features and corresponding traffic flow models (see next chapter) with instability as a core mechanism. Meanwhile a third so-called 'three-phase' traffic flow theory based on according empirical observations has been proposed. This theory also builds on instability mechanisms, but adheres to a different driver behaviour model, which is less deterministic and mechanistic than in previous theories.

We conclude from this review that – despite lots of well-accepted facts and theories on (congested) traffic flow – not all features of congested traffic have been unquestionably observed or described, nor satisfactorily explained. Most importantly, whereas the mechanism behind congestion wave propagation is widely accepted, the *formation* of congestion waves and the role of instability therein is the main issue of discussion. However, the amount of empirical observations and analyses of oscillating congestion patterns and a variety of congested traffic phases that has emerged since the second halve of the 1990's forms strong evidence that instability indeed plays a significant role in processes governing congestion dynamics.

Because on the one hand traffic flow models have helped researchers to understand certain empirical features better, and on the other hand driver behaviour is crucial in all these issues, we proceed in this thesis by building a traffic flow model that is inextricably linked to individual driver behaviour. This allows us to relate behaviour of traffic flow (e.g. instability) to characteristics of the driver. This development might enable future researchers to look for answers to the unresolved issues of congested traffic in observations both at the level of traffic flow and at the level of individual driver behaviour.

Before proceeding to the model development, the next chapter gives a review of existing traffic flow modelling approaches, with special attention for the representation of individual driver behaviour and the potential for modelling ADA systems.

3 STATE-OF-THE-ART OF TRAFFIC FLOW MODELLING AND ITS APPLICATIONS TO ADVANCED DRIVER ASSISTANCE SYSTEMS

3.1 Aim and structure of this review

The aim of the state-of-the-art review in this thesis is threefold:

- to give an overview of modelling approaches that have been proposed to describe and explain the empirical facts and theories discussed in chapter 2;
- to provide insight into the role of individual driver behaviour in these models;
- to give an overview of how the influence of ADA systems on traffic flow has been analysed using these modelling approaches.

The review is structured using the well-known distinction based on the level-of-detail of traffic flow description on the one hand, and of traffic flow behaviour on the other hand. Therefore, three main types of models are reviewed:

- *microscopic* traffic flow models: the description of traffic flow is vehicle-based, with the behaviour of the vehicle-driver combinations specified also on the individual or microscopic level (section 3.2);

- *macroscopic* traffic flow modes: the description of traffic flow is flow-based (or macroscopic, i.e. without distinction of individual vehicles), with the behaviour of the flow specified also on the level of traffic flow (section 3.3);
- *kinetic* traffic flow models: the description of traffic flow is macroscopic, with the behaviour of the flow derived from individual or microscopic behavioural specifications (section 3.4).

We do not strive for an exhaustive review of all models that have been proposed so far, but refer to Hoogendoorn & Bovy (2001c) and Helbing (2001) among others for readers desiring such broad overview. Rather the intent is to provide for each model type:

- a concise introduction to the model principle, based on a discussion of the most representative model(s);
- a selection of some representative models of that type, illustrating the progress that was made in the past in describing empirical phenomena, driver modelling etcetera;
- a discussion of the role of individual driver behaviour in that modelling approach;
- an overview of typical ADA analyses using that modelling approach (Tampère, 2001).

Finally in section 3.5, the chapter is concluded with some critical discussion and conclusions on traffic flow modelling, motivating and positioning the work performed in this thesis.

3.2 Microscopic traffic flow models: state-of-the-art and ADA applications

Traffic flow models are called *microscopic* if they describe traffic flow on the level of individual vehicles. Therefore, and in contrast to macroscopic models, microscopic traffic flow models build by definition on individual driver behaviour specifications. Microscopic models are briefly reviewed in this section, since they serve as a basis for the kinetic model development in this thesis. For a more elaborate review of microscopic models, the reader is referred to Minderhoud (1999).

Generally speaking the aim of a microscopic traffic flow model is to mimic the way in which a driver controls his vehicle, and as a result of that describe traffic flow operations. Several behavioural approaches and models for this control task are available (see Hoedemaeker (1999) for an overview). Often a distinction is made between subtasks; the most basic distinction of subtasks is that of the longitudinal and lateral control tasks. Models for these two subtasks are reviewed in sections 3.2.1 and 3.2.2 respectively. Section 3.2.3 briefly addresses microsimulation tools, since some of these form the basis for microscopic model analyses of ADA systems, which are discussed in section 3.2.4.

3.2.1 *Longitudinal control models for car-following*

From the discussion on traffic flow stability (section 2.3.2), we recall that longitudinal behaviour plays a dominant role in the formation and propagation of congestion. It is therefore not surprising that longitudinal vehicle control is the most studied aspect of driver behaviour. The longitudinal control task consists of:

- *speed and acceleration choice during free or unconstrained driving*: the choice of speed and acceleration when the motion of the vehicle is not hindered by the presence of other traffic; this is mainly governed by the 'desired speed' of a driver.
- *distance, speed and acceleration in constrained driving* (*car-following*): determines the equilibrium and transient acceleration behaviour when approaching or following a predecessor in the same lane.

When studying the congestion regime, unconstrained driving is less important, so our discussion is limited to constrained driving or car-following models.

Classical car-following theory

Car-following theory relates the acceleration of a vehicle-driver unit to motivational or perceived stimuli such as: desired speed, speed difference and distance to the predecessor. For an extensive overview of car-following models, we refer to Rothery (1999), Minderhoud (1999), Ahmed (1999) or Brackstone & McDonald (1999). Car-following theory has been developed and gradually improved since the 1950's by authors like: Pipes (1953), Chandler et al. (1958), Helly (1961), Gazis et al. (1959, 1962), Todosiev & Barbosa (1964), May & Keller (1967), Wiedemann (1974), and many others. Generally speaking, such models give the acceleration $\dot{v}_j$ of a vehicle *j* that follows a leader *j-1* at time *t* as:

$$\dot{v}_j(t) = f_{CF}\left(v_j, v_{j-1}, s_j, \dot{v}_{j-1}\right) \tag{3.1}$$

The function f_{CF} is a function of the stimuli such as: speed, speed v_{j-1} of and distance s_j to the predecessor, acceleration $\dot{v}_{j-1}$ of the predecessor, but possibly also the speed of the pre-predecessor, etcetera. When the limitations on perception by drivers are taken into account in car-following models, we speak of *psycho-spacing* models, like for instance the model of Wiedemann (1974). This approach represents the driver as a satisfising controller rather than an exact controller: a control action is only triggered if the stimuli exceed certain thresholds. For instance, small speed differences with the predecessor or small deviations from the desired following distance are accepted without correcting action whereas large deviations are not.

In the second halve of the 1990's, car-following theory attracted renewed attention, when physicists discovered several congestion patterns (see section 2.4.2) and tried to formulate appropriate car-following models explaining the new empirical observations. In this context the works of Bando et al. (1995), Helbing, Treiber and co-workers, Kerner & Klenov (2002, 2003) and a special class of car-following models called *cellular automata* are worth mentioning.

Optimal velocity model

The model of Bando et al. (1995) was one of the first models covering both unconstrained and constrained driving while at the same time exhibiting a realistic form of instability. The essence of their model, also indicated as the *optimal velocity* model, is as follows. Driver response of the follower is not directly coupled to the speed of the leader, but depends on the deviation with respect to the speed that fits best to the actual following distance (the 'optimal velocity'). The authors show that under certain conditions small perturbations are amplified and grow into jams, which allows the study of stop-and-go patterns.

Intelligent Driver Model

Helbing, Treiber and co-workers (Treiber et al., 2000; Helbing et al., 2002) proposed a microscopic car-following rule called Intelligent Driver Model (IDM) that refines the optimal velocity model. The concept of *optimal velocity* is now defined as a function of the speed difference *and* following distance with respect to the predecessor instead of a function of the following distance only. The IDM model shows good accordance with empirical data of different types of congestion. Especially metastability and the phase transitions between non-congested and congested conditions and between different types of congested traffic are reproduced well. Moreover, the authors propose a macroscopic generalisation of the IDM, and are able to interface both aggregation levels seamlessly.

Three-phase theory microsimulation

Recently, Kerner & Klenov (2002, 2003) proposed a microscopic model supporting the three-phase traffic theory of Kerner (free, synchronised and jammed traffic). The essential difference with traditional microsimulation models is twofold: (i) instead of assuming a 1-to-1 relationship between the driving speed and the preferred following distance, Kerner & Klenov discern two distance thresholds: one (speed dependent) minimal safety distance and one *synchronisation* distance; if the former is exceeded the driver response is the usual: braking until (at least) the safe distance is restored, however, the driver is *indifferent* with respect to distances between the former and latter threshold: the only response is to eliminate potential speed differences; (ii) when eliminating speed differences with the predecessor there exists always a probability of errors: drivers who are assumed to brake temporarily accelerate by error. Assumption (ii) is a variant of rather accepted stochastic perturbation terms used in cellular automaton models (see below). Its presence in the model is essential, since it is responsible for phase transitions of synchronised flow over highly compressed synchronised flow (called 'pinch region' by the authors) to wide moving jams. The first assumption (i) is more fundamental, since as a result there exists a broad range of steady state or equilibrium densities for a given speed. Therefore, this model has no fundamental diagram or equilibrium speed curve for congested traffic states, but rather a 2-dimensional range of equilibrium solutions. This property corresponds to the main experimental feature of synchronised traffic reported by Kerner & Rehborn (1996b) or Kerner (1999) (see also section 2.4.2).

Cellular automata models

To conclude this section, a special class of discrete microsimulation models is discussed: the *cellular automata* (*CA*) models. This approach reduces the microscopic car-following behaviour to a minimal set of simple driving rules that still (qualitatively) reproduce the dynamics of traffic flow including transitions to congestion. However, although the CA operationalisation is microscopic (distinguishing individual vehicles), the behavioural specification is more comparable to that of kinetic models (see section 3.4). In CA models space and time – and therefore also speed – are divided into discrete intervals. Vehicles are particles hopping from one cell to another. The speed is maximal with the limitation that collisions with the predecessor be prevented. Crucial for the occurrence of instabilities is the introduction of a certain probability for accelerating less strongly than possible. After the initial introduction of the CA-model in traffic flow theory by Nagel & Schreckenberg (1992), a variety of CA-models have been tested. The CA-models have been further developed including multi-lane variants by among others Helbing & Schreckenberg (1999), Emmerich et al. (1998), Wu & Brilon (1999). Although CA models are based upon trial and error rather than upon scientific driver behaviour models, the output of the CA-models resembles macroscopic properties of traffic flow quite realistically. This raises an intriguing theoretical hypothesis about the essence of traffic instability and congestion formation: are CA-models the extreme simplification of human driving that in spite of the simplicity still captures the essential aspects of congestion formation and propagation? This question is even more challenging since Nagel (1998) proved that the continuous limit of the CA-model corresponds to the well-known LWR first order macroscopic model (see section 3.3.1). In the last years, the trend in CA modelling has been towards further refinement of the original, rather coarse model assumptions (e.g. Chowdhury et al., 1999). As a consequence, the difference between CA and traditional microscopic car-following rules is fading. For instance the microscopic model of Kerner & Klenov (2002) has characteristics of both approaches, and the authors have also proposed a CA version (Kerner et al., 2002) using the same underlying theory and yielding qualitatively the same solutions as the continuous microscopic version.

3.2.2 Lateral control models for lane-changing and merging

Apart from lane-keeping, which is beyond our scope, the lateral control task consists of:

- (*spontaneous*) *lane-changing behaviour*: balances the wish versus the opportunity to change lanes on a multi-lane highway for speed gain, comfort, individual preferences, or because of legislation (e.g.: European driving rules).
- *merging behaviour*: describes how VDU's merge into their target lane when *forced* to leave their present lane (e.g. on-ramp, approach of a lane-drop, restricted lane use (truck lanes,…)).

The role of lane changing and merging in the formation or propagation of congestion in multilane traffic is less important than that of longitudinal control, but not negligible. First of all, merging manoeuvres near on-ramp or lane-drop bottlenecks are the direct cause for overloading of certain lanes and thus traffic breakdown (e.g. Cassidy &

Bertini, 1999a; Daganzo, 2002b). Apart from merging, the role of spontaneous lane changing during the transition to congestion ('one-pipe regime') was among others stressed by Daganzo (2002a). Also, spontaneous lane changing can be the origin of perturbations that could grow into jams in unstable traffic. In contrast to longitudinal control models, the number of lane changing models is fairly limited. For an overview of lane changing models, we refer to Ahmed et al. (1996). For the sake of completeness of the review, we limit the discussion in this section to the most representative models only.

The best known lane change model is that of Wiedemann (1974). The rationale behind this model is that drivers try to avoid discomfort on their own lane (car-following instead of free driving) and seek speed advantage in other lanes. This lane change desire only leads to a lane change manoeuvre if the gap in the target lane is sufficient to perform a safe manoeuvre. Helbing (1997) proposes a lane-changing model that considers repulsive and attractive forces acting on a driver. Stimuli, like the hindrance by other vehicles in the same lane and speed gain on the target lane, push or attract drivers to leave their lane, while the absence of appropriate gaps in the target lane may act as a repulsive force, keeping drivers in their actual lane.

It can be concluded that the theory is more or less agreed upon (the stimuli and safety considerations), but has different operational variants. However, all variants introduce a significant number of new parameters in the description of traffic flow; these parameters have significant influence on the simulation results (e.g. bottleneck capacity) but are not directly measurable and thus hard to validate. Ahmed et al. (1996) is among the rare contributions describing extensive validation efforts for lane-changing models using individual vehicle data.

3.2.3 Microsimulation tools

Before proceeding to the discussion of ADA analyses with microscopic traffic flow models, some considerations are devoted to microsimulation tools (often the basis of the ADA analysis). Microsimulation tools are software packages for commercial or research purposes, based on a combination of some car-following logic with a set of lane-changing rules. Moreover, these tools contain route choice models, node models, traffic control measures, etcetera. Many of these models have been extended with more detailed heuristic rules for specific circumstances experienced during driving. For a review of microsimulation tools, the reader is referred to Minderhoud (1999).

Most of these models are used as commercial tools or as research tools. Examples are: SIMONE (Minderhoud & Bovy, 1998), MIXIC (Van Arem et al., 1997b), PELOPS (Wallentowitz et al., 1998), AIMSUN (Barcelo et al., 2002), VISSIM (PTV, 2003), PARAMICS (Quadstone, 2002), FOSIM (Fosim, 2004), FLOWSIM (Wu et al., 2000), MITSIM (Yang et al., 1999) and many others. Some of these models (like the first three in the list above) include ADA Systems as an option. Adaptive Cruise Control is the most common ADAS in microsimulation. To the author's knowledge, only FOSIM (Fosim, 2004) and SIMONE (Minderhoud & Bovy, 1998) have been adapted to include alternative car-following in congested conditions.

A problem of a detailed microsimulation tool with many unobservable parameters is the large number of degrees of freedom. This implies the risk that the same apparent traffic flow behaviour can be generated through different internal parameter sets. This may seem harmless, but it means that the model can 'explain' the same phenomena in different ways, and many different parameter sets can be tuned to yield a seemingly correct reproduction of the validation data. This raises questions about what is the real mechanism? This is especially important when applying variations to a so-called 'validated' reference situation, e.g. when estimating the effect of ADA Systems that have not been tested in practice. Note that this problem is typical of models with many parameters, like for instance neural networks. With the latter type of model, the risk of *over fitting* can be reduced by the introduction of a regulisation term in the error function that is minimised during calibration. Such regulisation term penalises large parameter adaptations during calibration, since these might improve the fit of the model to the calibration data, while rendering it less suited for extrapolations to new data (e.g. Lint, 2004).

3.2.4 Microscopic model analysis of ADA systems

Microscopic models are well suited and easily adapted for simulating and analysing ADA Systems. The representation of the driver and his vehicle in the model relates to the corresponding elements in the real world in a straightforward and obvious way. This also makes communication about the simulation set-up and results comprehensible and easily accessible to outsiders. We give some examples here, and refer to Minderhoud (1999) for a more elaborate overview.

In the mid 1990's, Van Arem et al. (1996, 1998) analysed Adaptive Cruise Control (ACC) with the MIXIC microsimulation tool. The model was especially developed for ADA analysis on motorway stretches and produces detailed output on motorway efficiency, safety and emissions of noise and exhaust gases. The detailed driver and vehicle model are typical of a submicroscopic model. An even more detailed vehicle model is the basis of the Pelops simulator by Ludmann & Weilkes (1995, 1997). Pelops was developed in the frame of the European research program PROMETHEUS and also applies to ACC on motorways and urban areas. ACC was also the subject of research of Minderhoud (1999); here the driver model is more detailed. It was based on findings from behavioural experiments in a driving simulator, where the driving behaviour with ACC was examined (Hoedemaeker, 1999). Besides ACC, other ADA systems like Intelligent Speed Adaptation (ISA) have been examined as well with microsimulation (Hoogendoorn & Minderhoud, 2001). Also in the United States microsimulation tools have been widely applied for the ex-ante assessment of ADA systems. The Shift programming language – a framework for hierarchical modelling of dynamic control systems – of the PATH research group (Deshpande et al., 1997) offers microsimulation, together with the network- or link-level control layer simulation for fully Automated Highway Systems (e.g. Horowitz & Varaiya, 2000). The simulation tool has also been applied to ADA systems with the driver in the loop (VanderWerf et al., 2001). For that purpose the model has been extended with an extensive behavioural model for driver behaviour (Delorme & Song, 2001).

Our own experience in earlier research, reported by Alkim et al. (2000), suggests that microsimulation with its high level of detail is not always the most suited tool for ADA assessment in the early explorative phase of ADA design, where the strengths and weaknesses of ADA support and control *concepts* are studied. Alkim et al. (2000) explored the conceptual advantages of vehicle-to-vehicle communication and vehicle-to-infrastructure communication in mixed manual and automated traffic. It turned out that the ADA system did not lead to the expected improvement of traffic flow. However, due to the high level of detail of the microsimulation tool – and hence also of the ADA specification – it was virtually impossible to conclude whether some detail in the design or parameter setting of the ADA System led to the unexpected deterioration of traffic flow conditions, or the ADA concept itself was responsible. One recommendation was therefore to design and analyse the ADA control concept on a more abstract level of detail first. Only when this analysis leads to a proof-of-concept, should one proceed to the detailed microscopic simulation tool to refine and analyse the ADA design.

Examples of such less detailed (but still microscopic) proof-of-concept analyses are those by Treiber & Helbing (2001) who use the IDM model with different parameter settings to represent ADA or by Kerner (2004) using the cellular automaton-like microscopic model of Kerner & Klenov (2002) in combination with a specific ACC control logic.

3.3 Macroscopic traffic flow models: state-of-the-art, behavioural foundation and ADA applications

Macroscopic traffic flow models describe the dynamics of traffic flow on an aggregate level, i.e. as a continuous medium in terms of flow rate, average speed and density, and without distinguishing individual vehicles. The dynamic behaviour of traffic is purely defined on the macroscopic level, for instance based on variations in time or space of the macro traffic variables and of a priori known *equilibrium* conditions: the value towards which the macroscopic variables would converge in absence of variations in space and time. Therefore, the relation with individual driver behaviour in these models is indirect.

The core of each macroscopic traffic flow model is the *continuum equation*, expressing the conservation of vehicles. Macroscopic models differ from each other depending on the assumptions that supplement the continuum equation. Models that assume a static empirical equilibrium relation between the average speed and the density are called *first order models*. A *second order model* assumes a dynamic equation for the speed as a function of time and of spatial derivatives of density and/or speed.

In the remainder of this section on macroscopic models, the first order model is discussed (section 3.3.1), then the model of Payne as the prototype for second order models (section 3.3.2), after which we review how both model types have been refined in order to increase consistency with individual driver behaviour (section 3.3.3). Finally

section 3.3.4 deals with ADA applications modelled through macroscopic traffic flow models.

3.3.1 The first order model of Lighthill, Whitham and Richards

Analogies with hydrodynamic flow were used by Lighthill & Whitham (1955) and independently by Richards (1956) to establish a dynamic flow model for traffic, commonly referred to as the LWR model after the initials of the authors. It consists of three basic equations: the continuity law that describes the time and spatial relation between traffic density k and flow q, an instantaneous or fundamental relation between flow q, speed V and density k (see also equation (2.1)) and a theoretical or empirical relation between the so-called equilibrium speed V^e and the density k:

$$\frac{\partial k(t,x)}{\partial t} + \frac{\partial q(t,x)}{\partial x} = 0 \tag{3.2}$$

$$q(t,x) = k(t,x)V(t,x) \tag{3.3}$$

$$V(t,x) = V^e(k(t,x)) \tag{3.4}$$

The most appealing property of the LWR model is its simplicity that makes it possible to solve the differential equation numerically, analytically or graphically. It can be shown that fluctuations in the density propagate through traffic (kinematic waves) with a density dependent speed that is never larger than the equilibrium average speed V^e of traffic. According to the LWR theory, the amplitude of these waves does not grow. However, as time elapses the continuous density gradient gradually transforms into a discontinuous jump, separating regions with different densities (shock waves).

Apart from the attractive property of providing easy insight into the propagation of disturbances (information flow), the LWR model has also some drawbacks:

- The discontinuous nature of shock waves is unrealistic from the microscopic point of view since it assumes that individual vehicles accelerate or decelerate instantaneously when flowing past a jam front; Lebacque (2003) has formulated a variant of LWR where this approximation has been remedied, in so doing bridging the gap with second order macro models (see section 3.3.2);
- The assumption that traffic cannot deviate from equilibrium conditions is falsified by empirical observations. These show a wide scatter of data points around the equilibrium curve. These deviations are of both stochastic and of structural nature (e.g. the hysteresis effect when arriving at or accelerating from jammed regions, section 2.6.1);
- Since the amplitude of perturbations remains constant, the model does not account for instability, which is crucial for modelling the formation of density clusters growing into stop-and-go waves or other phase transitions. These transitions do occur in empirical data, but are not accounted for by the first order theory that therefore denies some fundamental behaviour of unstable traffic flow (see section 2.4.2).

Still, when instability is irrelevant for some application (e.g. traffic flow with signalised intersections) the first order models are usually sufficient. Moreover, for these applications elegant numerical solution schemes for the LWR model have been proposed, e.g. the (lagged) cell transmission model by Daganzo (1994, 1995b) and the Godunov-scheme that was applied by Lebacque (1996). Daganzo and his co-workers at the University of Berkeley (Newell, Cassidy, Mauch) stress the applicability of the LWR model to explain a wide variety of phenomena observed in empirical studies of bottlenecks (section 2.4.1). It appears that consideration of multiple user-classes in the LWR model increases considerably the descriptive power of the first order approach (Daganzo, 1997; Zhang & Jin, 2002; Chanut & Buisson, 2003; Logghe, 2003; Zhu et al., 2003). Crucial is the distinction between fast and slow user-classes ('rabbits and slugs'), and an appropriate choice of the fundamental speed curve (3.4), as was pointed out by Daganzo (2002a, 2002b). Other extensions to the LWR model have been formulated by Zhang (1999; see section 3.3.3.3). Despite this progress, the absence of an instability mechanism in the LWR model remains a theoretical obstacle, which led to the development of a variety of *higher order models*.

3.3.2 *Reaction time and anticipation in the Payne model: the basis of higher order models*

Payne (1971) was the first author to recognise that delayed response and anticipation behaviour of drivers are fundamental in macroscopic traffic flow models if one wishes to account for traffic instability. Payne approximated individual driver behaviour as follows:

$$v\left(t+T, x\left(t+T\right)\right) = V^e\left(k\left(t, x\left(t\right)+\Delta x_a\right)\right) \tag{3.5}$$

This equation expresses that the speed of an individual vehicle after a reaction time T is equal to the equilibrium speed corresponding to the density some distance Δx_a downstream. After repeatedly applying linear Taylor approximations to equation (3.5) and approximation of $\Delta x_a \approx \frac{1}{k(t,x)}$ the following dynamic equation for the average speed is obtained:

$$\frac{\partial V\left(t,x\right)}{\partial t} + \underbrace{V\left(t,x\right)\frac{\partial V\left(t,x\right)}{\partial x}}_{convection} = \underbrace{\frac{V\left(t,x\right)-V^e\left(k\left(t,x\right)\right)}{T}}_{relaxation} \underbrace{-\frac{c_0^2\left(t,x\right)}{k\left(t,x\right)}\frac{\partial k\left(t,x\right)}{\partial x}}_{anticipation} \tag{3.6}$$

$$c_0^2\left(t,x\right) = -\frac{1}{T}\frac{dV^e\left(k\left(t,x\right)\right)}{dk} \tag{3.7}$$

The left-hand side of this equation is the total time derivative of the average speed, that is: the time-evolution of the average velocity as it would be seen by an observer travelling with V. The interpretation of the different terms is:

- Convection: accounts for the change in average speed at a location, due to vehicles leaving or arriving at the location with different speeds; in other words: the transportation of velocity along with the flow;
- Relaxation: it is assumed that an equilibrium velocity exists: $V^e(k)$, but the traffic state can deviate from equilibrium. When other influences, reflected by the convection and anticipation terms, are small, traffic tends to relax to the equilibrium relation;
- Anticipation: traffic approaching a spatial change of traffic conditions (reflected by the spatial derivative of density) adapts its speed to these conditions (see also section 4.3.2).

This modelling approach solves two drawbacks of the LWR model. Firstly, the relaxation term allows traffic to deviate from the equilibrium situation when the spatial conditions are not homogeneous. When driving into or out of congestion, this mechanism results in average velocities (and therefore also flow rates) that deviate substantially from the equilibrium state. Secondly, the anticipation term solves the theoretical 'shock wave problem' of the LWR model: since drivers tend to react to the spatial variation of the density, discontinuous behaviour is avoided. As a result, queue fronts have a finite width (i.e. fronts are spread over some finite x-range) and the acceleration or deceleration of traffic need not be instantaneous.

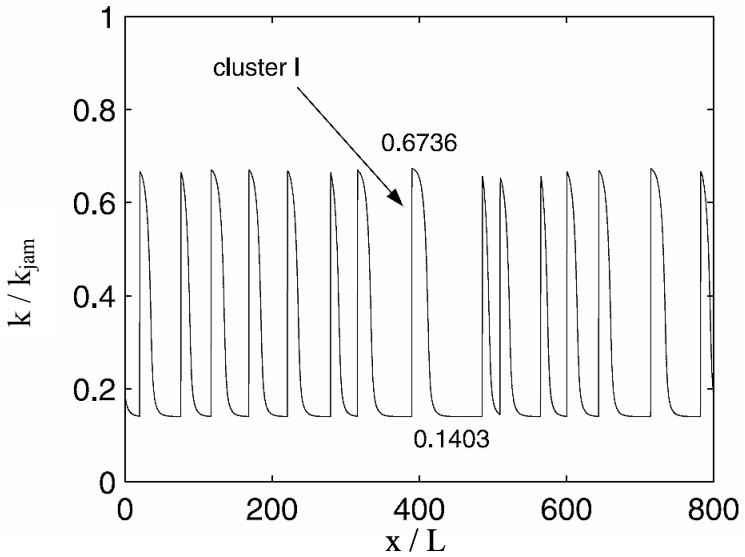


Figure 3-1 Jam formation (clusters of the density) due to instability in the Payne model (Jin & Zhang, 2003)

It can be shown that in a certain range of density, perturbations of finite amplitude can grow into jams (see section A.6.2 in Annex A and Figure 3-1). This instability property implies the potential of this model to account for empirically observed stop-and-go traffic. Another property of the Payne model (and other similar models) is that small perturbations are transported along two (curved) *characteristic lines* (one moving faster and one moving slower than average vehicular speed), which is unlike the LWR model that contains one (straight) characteristic. However, this means that in Payne-type models information is transported downstream *at a speed higher than traffic velocity*.

This is a violation of the *anisotropic* property of traffic flow that refers to the fact that drivers only respond to traffic conditions ahead. This was one of the arguments for Daganzo (1995a) to reject the Payne-type models. In a reaction to this criticism, Aw & Rascle (2000) and Zhang (2000, 2003a) proposed variants of the Payne model that avoid this flaw (see also section 4.3.2 for a discussion related to this issue).

From the point of view of driver behaviour modelling, the merit of the Payne model is that he introduced anticipation and delayed reaction by drivers. Just like in microscopic traffic flow models these aspects determine the stability of traffic flow. However, despite the car-following model that was taken as a basis for the development of the velocity equation, the behavioural foundation is poor. Firstly, the equilibrium speed that plays a crucial role in (3.5) and (3.6) can only be indirectly linked to individual driving behaviour parameters. Secondly, the direct but coarse linear Taylor approximation used in the derivation does *not* maintain the microscopic character of the underlying individual driver model. As an illustration let us consider the reaction time of drivers (typically in the order of 1 s) that is introduced in the simple 'microscopic' model (3.5). This parameter emerges in the macroscopic velocity dynamic equation as the relaxation time, which stands for a different process and therefore must adopt substantially larger values than normal drivers' reaction times in order to fit empirical data (typically in the order of 10-20 s).

Together with the LWR model, the model of Payne is the most widely applied traffic flow model. It is the core of the METANET model (Messmer & Papageorgiou, 1990; Papageorgiou, 1989) that not only offers a numerical solution scheme for the model of Payne, it also provides tools to build networks and to apply various sorts of control, like ramp metering and route diversion. Some other authors that use slightly modified versions of the Payne model are: Cremer & May (1985) and Smulders (1990).

3.3.3 Refinements to driver modelling in macroscopic traffic flow models

Four main approaches have been proposed in literature to further bridge the gap between individual behaviour specification and macroscopic traffic flow behaviour:

- extensions in analogy to hydrodynamics (section 3.3.3.1)
- micro-macro links in analogy to Payne (section 3.3.3.2)
- refinement of the equilibrium speed (section 3.3.3.3)
- (gas-) kinetic approach (see section 3.4.1)

In the next subsections the former three approaches are reviewed. The latter one is the kinetic or mesoscopic approach that is discussed in a separate section 3.4.1.

3.3.3.1 Extensions in analogy to hydrodynamics

The dynamic speed equation (3.6) of Payne has a structure similar to hydrodynamic convection-diffusion models of compressible fluids. The anticipation term is therefore also often called 'pressure' term, because of a similar term in hydrodynamic models. However, in these models the friction between particles or molecules mutually, and

between particles and boundaries is responsible for another contribution to the dynamic equation, which is not present in the model of Payne. This 'viscosity' or 'diffusion' term in hydrodynamic models causes transitions to dynamic regimes such as turbulence. Therefore, it has inspired researchers to include a similar term in traffic flow models, in order to capture dynamic regimes like stop-and-go traffic. Many of these models can be cast into the following generalised form of the Payne model:

$$\frac{\partial V}{\partial t} + V\frac{\partial V}{\partial x} = \frac{V^e(k) - V}{\tau(k)} - \frac{1}{k}\frac{\partial P^e(k)}{\partial x} + \nu(k)\frac{\partial^2 V}{\partial x^2} \tag{3.8}$$

The anticipation term now consists of two contributions, the *traffic pressure P* and the *viscosity* ν. With the definitions $\tau(k) = \tau$, $\nu(k) = 0$ and $P^e = c_0^2 k$, the model of Payne is obtained. Another member of the Payne family of models is the model proposed by Kühne (1984). He proposes an equilibrium velocity – density relation derived from a car-following model. Furthermore he specifies: $\tau(k) = \tau$, $\nu(k) = \nu$, and $P^e = \Theta_0 k$. The only substantial difference with Payne, other than the fundamental diagram, is the constant viscosity coefficient ν. This model behaves like the model of Payne with an instability region, but the viscosity term smoothens out sharp changes in the speed profile. Kerner and Konhäuser (1993), and later Lee et al. (1998) adopted the model of Kühne, with a density dependent viscosity coefficient: $\nu(k) = \frac{\eta_0}{k}$. This model – sometimes referred to as the KKKL model after the authors – shows the emergence of density clusters within a certain instability range of the density. These clusters can grow into jams that propagate with different speeds, so that after some time these waves can catch up with each other and combine to form wide traffic jams that propagate with a constant characteristic velocity. This process is well-known from empirical data (see section 2.4.2). Moreover, Kerner and Konhäuser show that their model is meta-stable within a certain density range, another empirical characteristic of congested traffic that was previously not captured in macro models (see section 2.3.2.3). This means that perturbations of subcritical amplitude are damped, while larger perturbations grow to form jams. Other variants of Payne-type models have been proposed by among others: Liu et al. (1998), Helbing (1997), Zhang (1998), and Hoogendoorn (1999).

From the point of view of driver behaviour, authors who advocate viscous traffic flow models identify the viscosity with a 'higher order tendency' of traffic to go along with the flow. The second order derivative of the speed indicates whether the first order speed gradient will increase or decrease in space. This causes drivers to respond firmer to gradients that will increase even further. In other words: they anticipate to the fact that traffic immediately ahead will change speed soon. Helbing (1997) has motivated the introduction of the viscosity term by assuming a change of driver state from brisk to careful driving. Although these interpretations might be reasonable, the corresponding term in the macroscopic model is often introduced ad hoc, to improve the potential for representing dynamics traffic flow patterns, or even for the sake of numerical robustness. Only Zhang (2003b) was able to mathematically derive the occurrence of the viscosity term from an individual driver model (see next section).

3.3.3.2 Micro-macro links in analogy to Payne

Recently, progress was made in establishing macroscopic traffic flow models from specifications of microscopic models for individual driving behaviour. We mention here the works by Zhang and Hennecke & Helbing.

Similar to the approach of Payne, Zhang (2003b) used Taylor approximations of a microscopic model. His microscopic model is similar to that of Payne:

$$\begin{aligned} v_j(t+T) &= F\left(x_{j-1}(t)-x_j(t)\right)+G\left(v_{j-1}(t)-v_j(t)\right) \\ &= V^e\left(k\left(t,x(t)+\Delta x_a\right)\right)+\alpha\left(\delta_1\frac{\partial V(t,x+\delta_2)}{\partial x}\right) \qquad with \quad 0<\delta_2<\delta_1 \end{aligned} \tag{3.9}$$

The first line of equation (3.9) is the purely microscopic one. It says that the speed of a vehicle *j* is adapted after some 'adaptation time' *T* as a function *F* of the distance, to and a function *G* of the speed difference with the predecessor *j-1*. In the second line, the transition to a spatially continuous spacing function and speed function is made, and the functions *F* and *G* are specified. The function *F* is related to the equilibrium speed V^e and the function *G* is a simple linear damping of the speed difference. Comparing this equation to (3.5), the speed damping term introduces a spatial derivative of the speed that was not present in the model of Payne. Applying the Taylor approximation in the same manner as Payne did, yields a viscous traffic flow model with the structure of equation (3.8). The spatial derivative of the speed in (3.9) is in this manner transformed into the second order viscosity term of (3.8) .

Hennecke et al. (2000), and Helbing et al. (2002), generalising an idea originally proposed in Helbing (1998) on a specific microscopic model presented by Helbing & Tilch (1998) (see section 3.4.2), use a comparable approach in which they allow a generic form of the microscopic model:

$$\frac{dv_j(t)}{dt} = f\left(x_{j-1}(t)-x_j(t),\ v_j(t),\ v_{j-1}(t)-v_j(t)\right) \tag{3.10}$$

By an appropriate definition of the average speed *V* as a linear interpolation of individual vehicle speeds v_j, and analogously of the average acceleration *A* as a linear interpolation of accelerations according to (3.10), they build a macroscopic equivalent of the microscopic model. Furthermore, Hennecke and Helbing avoid Taylor approximations by explicitly allowing non-local terms in the macroscopic models. These are terms that influence the average speed at location *x* while being numerically evaluated at location $x+\Delta x_a$, i.e. at some anticipation distance downstream. This is not only an elegant method with respect to numerical solution of the resulting traffic flow model, it also solves an issue raised by Daganzo (1995a), who criticised second and higher order models for not being anisotropic (see also section 3.3.2). *Anisotropy* indeed is a fundamental difference between driver behaviour and the physical systems that have inspired traffic flow modellers. It means that behaviour of drivers is influenced only by information coming from downstream. In traffic flow models that contain spatial derivatives of *V* or *k,* the behaviour at *x* is equally influenced by traffic conditions immediately downstream or upstream, since the derivative has no 'direction'. As a result of not being anisotropic, Daganzo has shown that in certain realistic starting

conditions, traffic can move backwards in second and higher order models, which is clearly unrealistic. The explicit non-local approach of Hennecke (further described in Treiber et al., 1999 and Helbing et al., 2001) avoids this pitfall.

From the point of view of driver behaviour, the approach of Zhang is a justification of already existing viscous traffic flow models, but does not add any new links to individual driver behaviour, since his specification of functions F and G is such that it still builds explicitly on the macroscopically defined equilibrium speed. The approach of Hennecke, Treiber and Helbing allows more flexibility in the definition of macro models based on a microscopic specification. Moreover, the non-locality is solved in a way that guarantees anisotropy of driver behaviour.

3.3.3.3 Refinement of the equilibrium speed

Except for the approach of Hennecke, Treiber and Helbing that is more generic in the sense that it theoretically allows any specification of individual driver behaviour, all previously mentioned models have introduced ways to allow the traffic state to deviate temporarily from the empirical equilibrium speed. This was desired, since in the LWR first order model, the restriction of the traffic state to the equilibrium speed led to a model that neglects hysteresis, instability, metastability, and stop-and-go waves.

However, there is a different school in traffic flow modelling, where the same issue is solved differently. Traffic is still restricted to an equilibrium state (as in the LWR model). But the definition of this equilibrium speed is refined, so that the resulting (quasi-) first order model reflects some desired refinement in traffic flow dynamics. As in the case of the second and higher order models, the refinements are motivated by appropriate assumptions about the 'gross' or 'average' behaviour of drivers.

Newell (1965) and Zhang (1999) presented quasi-first order models with so-called 'phase curves'. Phase curves are a set of relations between the average speed and the density, with the actual flow type determining which phase curve has to be chosen as complement to the first order equations (3.2) and (3.4). For instance, the assumption that drivers respond delayed to accelerating or decelerating speed profiles leads to different phase curves. For the same density, the corresponding speed on the deceleration phase curve is higher than for the acceleration phase curve. The result is a compression of traffic flow in upstream jam fronts where speeds drop sharply, and an extension in the downstream jam front where speeds recover from low values (hysteresis, see 2.6.1). Zhang & Kim (2001) present an individual driver behaviour model with hysteresis, i.e. a model in which drivers choose a different preferred distance to the predecessor dependent on whether traffic flow is free flowing or has settled in a congestion regime. Although this individual model might serve as a basis for different phase curves, no macroscopic equivalent of this model has been presented so far.

Recently, Daganzo (2002a, 2002b) has proposed another LWR-type model with a refined equilibrium curve (Figure 3-2). Not only the specific reversed-lambda shape of the equilibrium flow is essential in this model (for modelling the capacity drop, section 2.6.1), also the assumption that (a) traffic consists of two types of drivers: 'rabbits' and 'slugs', each with their own equilibrium flow and (b) a multilane road is divided into

fast and slow lanes with different equilibrium flows in certain speed regions. The model is motivated by assumptions on (the gross effect of) driver behaviour of the rabbits and slugs, so that the proposed equilibrium flows (a) and (b) are consistent. For instance: slugs always drive on the slow lanes, while rabbits drive on the fast lanes in free flowing traffic ('two-pipe regime') and spread over both fast and slow lanes in congested traffic, in so doing annihilating the speed difference between both types of lanes ('one-pipe regime', separated from the two-pipe regime in Figure 3-2 by the dotted line). Interestingly, the behaviour of the rabbits changes dramatically in the transition between the two regimes, so that rabbits, who originally drive close to each other at high speeds in a *motivated frame of mind* in the two-pipe regime, need more free space at the same driving speed in the one-pipe regime due to a *loss of motivation*.

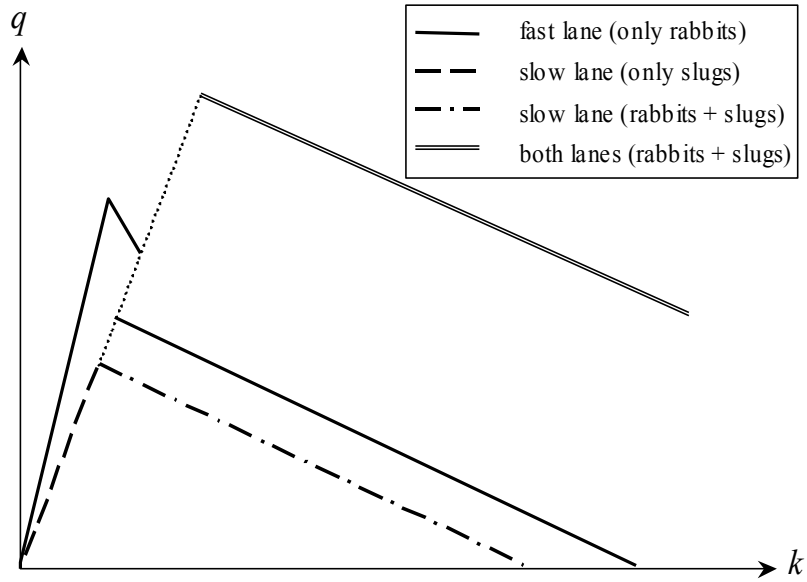


Figure 3-2 Refined equilibrium curve of flow against density for LWR model with* rabbits *and* slugs *(Daganzo, 2002a)

The theory of Daganzo was meant as a 'caricature' of traffic flow (which is more or less confirmed by partial validation, see Banks (2002), Banks & Amin (2003)), and might be criticised for being oversimplified and not founded on behavioural research. However, we regard it as an important first attempt to use changes in driving behaviour as an explanation for traffic flow phenomena near or in congestion. Moreover, the first order macroscopic modelling framework used by the author allows easy and illustrative solution and interpretation of the model outcome in a variety of traffic flow conditions. The model of Daganzo is valuable since it synthesises – albeit a caricature – the minimal set of driver and traffic flow behaviour needed to reproduce a rich variety of traffic flow dynamics. However, the first order modelling framework is still not suited to establish an explicit and quantitative relationship between individual driver behaviour specification and macroscopic traffic flow models.

3.3.4 Macroscopic model analysis of ADA systems

Model applications of macroscopic traffic flow models to ADA systems are rare. For such studies either an empirical measurement of the fundamental diagram in the presence of the ADA system is needed, or a direct relationship between individual driver behaviour and the dynamic speed equation. Since ADA systems are still in the experimental or market introduction stage, empirical fundamental diagrams are not available and can only be estimated for fully-automated and hence deterministically predictable ADA systems.

A comprehensive example of a macroscopic model application to vehicular traffic control is given by Karaaslan et al. (1990). The authors use a time and space discretisation of the Payne model (section 3.3.2). They derive a control law for an automated highway, assuming that vehicle speeds can be controlled externally. The idea is that congestion starts with inhomogeneities in the traffic stream; the control law avoids the growth of clusters by controlling the vehicle speeds. For that purpose the anticipation term of the Payne model is replaced by a control term that depends on the density along the freeway. Density is homogenised by moving some of the load in the congested regions to less dense regions. The speeds computed with the control law are the mean speeds, which are macroscopic variables. They may be regarded as target speeds in each section that have to be obeyed by an appropriate control policy at the microscopic level of individual vehicles. Such a two level control policy may be applied using communication and automatic cruise control methods. Alternatively, a similar effect may be achieved if the control speeds are displayed as advisory speeds on Variable Message Signs (VMS); this was studied by Lenz using the KKKL model (section 3.3.3) and the model of Helbing et al. (2001), or by Bellemans (2002, 2003) and Hegyi (2004) using the METANET model of Papageorgiou (1989).

Swaroop & Rajagopal (1998) present a design approach for the Adaptive Cruise Control strategy of vehicles in an Automated Highway System based on macroscopic traffic flow stability analysis. The authors do their analysis in two steps: first they examine the effect of the spacing policy employed by the Intelligent Cruise Control (ICC) systems on traffic flow stability; in a second step they show how the dynamics of ICC system affects traffic flow stability. In their first step, Swaroop and Rajagopal use the conservation of vehicle principle to show that for idealised ICC systems (no spacing error and immediate regulation of following distance) with any spacing policy, density perturbations lead to travelling waves in traffic flow. With this approach they prove that – even under these idealised circumstances – density disturbances propagate upstream and unattenuated for any ICC controller with a constant time headway policy. For the analysis of the ICC control law with consideration of the ICC system dynamics (i.e. with spacing error), the authors use a macroscopic model with a speed equation that is derived from the microscopic ICC control law. This model clearly differs from the Payne model: the anticipation term is absent because the control actions of the ICC depend solely on the behaviour of the vehicle immediately preceding it. However, if one were to model the effectiveness of feeding back information from the roadside devices into the cruise control law (e.g., VMS), an anticipatory term will be present in the equation for the dynamics of traffic speed. Finally, the authors propose a macroscopic

speed equation with which ICC control logic can be designed, starting from a speed equation showing the 'desired' speed dynamics of the controlled traffic flow. Swaroop shows that the ICC controller then must depend on macroscopic properties of the flow (spatial derivatives of density and speed). He suggests that this knowledge be estimated in an operational system by roadside systems and then be transmitted to the vehicles.

Demir (2001) presents an analysis of ACC traffic using the Payne macroscopic model. The fundamental diagram is first 'calibrated' to the simulation results of a microsimulation model with ACC vehicles. Therefore, the validity of the simulation results is doubtful. First of all, the results combine the model uncertainties of two simulation models; secondly, only the equilibrium solution of the Payne model was calibrated, so that it remains unsure whether the dynamics of the ACC traffic are captured correctly.

3.4 Kinetic traffic flow models: state-of-the-art, behavioural foundation and ADA applications

In this section the so-called (gas-)kinetic traffic flow models are reviewed. The basic concepts and theory are only briefly described. For a precise insight into the theory, we refer to Annex A where this approach is extensively elaborated. The review in this section gives – apart from the general approach – a brief history of the development of (gas-) kinetic models with special attention for the representation of driver behaviour.

The kinetic approach combines principles from microscopic and macroscopic modelling, in that the description of traffic is at the flow level, but with traffic behaviour specified on the level of individual drivers. Using the mathematical framework borrowed from statistical mechanics and gas kinetics, traffic is described with the *generalised density function* ρ. The generalised density ρ is the probability distribution of finding a vehicle at location x at time t in a certain state S, for example with the state S defined as the individual speed v. Also other characteristics of the vehicle or its driver like its desired speed and user-class, can be distinguished. Microscopic behaviour is translated into equivalent changes in the generalised density ρ. For example, a vehicle that brakes to avoid a collision with its predecessor causes a decrease of the probability density ρ at high speeds and a corresponding increase of ρ at lower speeds.

The kinetic description of traffic in terms of a generalised (probability) density of individual vehicle state is a level-of-detail in between microscopic and macroscopic traffic flow models, therefore the name *mesoscopic* is also customary to indicate kinetic models. The consideration of microscopic (interaction) processes is especially helpful for the specification of the model. For practical applications or simulations, the kinetic level of description is less suited. Therefore, the kinetic model formulation is mathematically transformed in a macroscopic formulation, for which numerical solution schemes and analysis techniques (e.g. stability analysis) are available.

In the next section 3.4.1, the basic kinetic traffic flow model as originally proposed in the 1960's is briefly addressed. Then, in section 3.4.2 the discussion is focussed on

refinements to kinetic models with respect to driver behaviour. Finally in section 3.4.3, some applications of kinetic models to ADA systems are reviewed.

3.4.1 The basic kinetic model by Prigogine, Herman and Paveri-Fontana

The basic structure and processes considered in the very first kinetic traffic flow model by Prigogine (1961) are still the basis of modern kinetic traffic flow models. Prigogine introduced the concept of the generalised density ρ (t,x,v), the probability density function at time t and location x of vehicles with speed v. He considered three processes causing changes to ρ:

- Convection: due to the movement of the vehicles, the generalised density ρ at a fixed location changes in time;
- Continuous change of the speed due to acceleration towards the desired speed;
- Discrete[4] change of the speed due to interactions of faster vehicles with slower traffic.

Prigogine derived the following dynamic equation for the (probability) distribution function ρ (t,x,v):

$$\frac{\partial \rho(t,x,v)}{\partial t} + v\frac{\partial \rho(t,x,v)}{\partial x} = \left(\frac{\partial \rho(t,x,v)}{\partial t}\right)_{continuous} + \left(\frac{\partial \rho(t,x,v)}{\partial t}\right)_{discrete} \tag{3.11}$$

The continuous acceleration towards the desired speed is specified as an exponential relaxation process and depends on the desired speed distribution w(t,x,v) of vehicles with speed v:

$$\left(\frac{\partial \rho}{\partial t}\right)_{continuous} = -\frac{\partial}{\partial v}\left(\rho\frac{w(t,x,v)-v}{\tau}\right) \tag{3.12}$$

For the discrete deceleration process, Prigogine & Herman (1971) specify the following. Fast vehicles catching up with slower traffic have a probability p of changing lanes and continue their journey without braking. With probability (*1-p*) the fast vehicle has to brake. In this interaction the fast vehicle adopts the speed of the slow predecessor, whose speed remains unaffected by the manoeuvre. The change in probability density for speed v at location x is then the balance of an inflow into state v by vehicles with an original higher speed v_0^f, who brake to speed v and an outflow of state v by vehicles who had speed v, but brake to adopt a lower speed v_1^f (= the original speed of the slower predecessor v_0^s):

$$\left(\frac{\partial \rho}{\partial t}\right)_{discrete} = (1-p)\left(\int_{v_0^f > v}\left(v_0^f - v\right)\varphi\left(t,x,v_0^f,x',v\right)dv_0^f - \int_{v > v_0^s}\left(v - v_0^s\right)\varphi\left(t,x,v,x',v_0^s\right)dv_0^s\right) \tag{3.13}$$

[4] Indicated in formulae by $(\ldots)_{event}$ or $(\ldots)_{discrete}$ or by mentioning explicitly the nature of the discrete event, e.g. $(\ldots)_{merge}$

In this equation the *pair-wise vehicle distribution* occurs. It denotes the joint probability distribution of finding a fast vehicle at location x with speed v_0^f, while at the same time a second (slower) vehicle is present at location x' having speed v_0^s. Prigogine simplifies the pair-wise vehicle distribution by (a) neglecting the finite length and space requirement of vehicles and (b) assuming *vehicle chaos*. The former assumption considers vehicle interactions only of vehicles at the same location x, whereas with the latter assumption all correlations between the speeds of predecessor and follower are neglected:

$$\begin{aligned}\varphi\left(t,x,v_0^f,x',v_0^s\right) &\overset{(a)}{=} \varphi\left(t,x,v_0^f,x,v_0^s\right) \\ &\overset{(b)}{=} \rho\left(t,x,v_0^f\right)\rho\left(t,x,v_0^s\right)\end{aligned} \tag{3.14}$$

Paveri-Fontana (1975) improved the statistical treatment of the desired speed in the original formulation of Prigogine (1961). Driver behaviour is the same as the rules described above.

3.4.2 Refinement of driver behaviour in kinetic models

Essentially the driver behaviour specification in kinetic traffic flow models has undergone relatively little evolution or refinement in three decades since the works of Prigogine & Herman (1971) and Paveri-Fontana (1975). The contributions by among others Nelson (1995, 1997), Helbing (1997), Klar & Wegener (1999), Hoogendoorn (1999), Hoogendoorn & Bovy (2000), and Helbing et al. (2001) are mainly focussed on a better *statistical* description of the traffic process and to a lesser extent of driver behaviour, with considerably improved macroscopic traffic flow models as a result.

3.4.2.1 Multi-lane models, finite space requirements, anisotropy, speed correlations

Helbing (1997), Hoogendoorn (1999), and Hoogendoorn & Bovy (2000) propose multi-lane extensions of the kinetic traffic flow model. For that purpose, the individual driver model is extended with lane changing rules. Hoogendoorn replaces the lane changing model of Prigogine (with overtaking probability $p(k)$) by a more refined one that takes into account free lane changes, postponed lane changes, mandatory lane changes and gap acceptance of drivers who want to perform a lane change. Both authors also refine the longitudinal driving rules. Helbing accounts for the fact that fast drivers with speed v_0^f who find slower vehicles with speed v_0^s on their path, might overreact and decelerate to a speed $v_1^f \leq v_0^s$. This deceleration is still considered instantaneous.

Helbing introduces the *finite space requirement* of drivers. This means that instead of only decelerating if vehicles have the same x-positions (assumption (a) by Prigogine in equation (3.14)), braking interactions already take place between vehicles with a finite, speed-dependent headway (after Jepsen (1998)). The effect of this specification is an increased interaction rate and hence a lower average speed at high densities; therefore the model is especially suited for traffic flows in and near congestion. Hoogendoorn

adds to this correction by distinguishing constrained and unconstrained drivers (or: platoon followers and leaders). He allows constrained drivers to accelerate only to the minimum of the own desired speed and that of the platoon leader. Moreover, he distinguishes different user-classes, each with their own parameters (e.g. desired speed, relaxation time and space requirement).

In later work of Helbing together with Treiber (1999), non-local interactions are introduced (see also the work of Helbing et al. (2001)). By doing so they account for the fact that drivers only respond to information coming from downstream, while being unaffected by traffic conditions upstream (anisotropy of driver behaviour, see also sections 3.3.2 and 3.3.3).

Finally Shvetsov & Helbing (1999) explicitly take correlations between speeds of successive vehicles into account. This is in contrast to the vehicular chaos assumption (assumption (b) in equation (3.14)) that underlies all previously mentioned models.

3.4.2.2 The first steps towards inclusion of individual driver models

Nelson and co-workers (1995, 1997), Helbing (1998), Klar & Wegener (1999), and Waldeer (2000, 2001, 2004a, 2004b) proposed kinetic models with refinements to the longitudinal individual driver behaviour. We regard these approaches as the first steps towards inclusion of individual driver models, because they are the first experiments to refine the instantaneous interaction model by Prigogine. Moreover, the model of Klar & Wegener is amongst the few (together with Shvetsov & Helbing (1999)) that consider correlations in the speeds of successive vehicles.

Nelson (1995) and Nelson et al. (1997) formulated a 'mechanical model' that treats acceleration towards the desired speed and deceleration due to slower traffic in a similar fashion: both as discrete processes (whereas other authors follow Prigogine in the specification of acceleration as a continuous process). Dependent on the headway and the speed difference with the predecessor, a driver does one of the following:

- instantaneously accelerate towards the own desired speed;
- instantaneously accelerate towards the speed of the predecessor;
- maintain the own current speed;
- instantaneously decelerate towards the speed of the predecessor;
- instantaneously decelerate towards standstill (speed zero).

Klar & Wegener (1999) follow an approach similar to Nelson, but treat longitudinal and lane changing behaviour simultaneously. Moreover, in addition to the above mentioned five possible driver responses in case of an event, they allow drivers to adopt also intermediate speeds (i.e. between the own speed and that of the predecessor in case of an interaction, or between the own speed and the own desired speed in case of being 'released' from a car-following situation). Moreover, in theory events are possible in which drivers change lanes and speed simultaneously (and instantaneously).

Apart from the speed correlations and improved individual driver modelling, the work by Klar & Wegener (1997) is interesting for another reason. To our knowledge it is one

of the rare examples (see also: Hoogendoorn & Bovy, 2001b) where the kinetic model is numerically evaluated in its kinetic formulation to obtain dynamic solutions, whereas most other kinetic models are first cast into their macroscopic equivalent formulation before being numerically solved.

Helbing (1998) proposed a method for obtaining a kinetic (and corresponding macroscopic) traffic flow model based on a specific individual driving rule: the general forced model of Helbing & Tilch (1998). The individual driving rule, containing both acceleration to the desired speed and deceleration for slower predecessors, is considered as a continuous process, replacing the acceleration term in the original model formulation by Prigogine. After some simplifying assumptions, a macroscopic equivalent model is obtained that behaves similarly to the microscopic model of Bando et al. (1995). Helbing however, abandoned this path of research and proposed another direct (not kinetic) relation between micro and macro model layers (Helbing et al., 2002; see also section 3.3.3).

Recently Waldeer (2000, 2001, 2004a, 2004b) proposed a kinetic model in which the description of vehicle states is generalised to include also the current acceleration of a vehicle (instead of only the location, the speed, and eventually the desired speed in traditional approaches). In this approach finite acceleration and deceleration times are taken into account, acceleration and braking are treated on the same time scale, and the approach allows in theory to incorporate any car-following relation. Moreover, the model of Waldeer recognises that the main control of drivers over their vehicle is through the acceleration (gas or brake pedal). Another theoretical advantage is that by taking the acceleration as a state variable, the *history* of individual control actions is better described, thus avoiding illogical state transition sequences (in a Markov-like interpretation of the kinetic model, see Waldeer (2000) and section 7.4.3). A disadvantage is the fact that another dimension is added to the multivariate probability density ρ, for which theoretically correlations with the other variables (speed, gap, speed of predecessor) have to be established; and hence another correlation that is violated when *vehicular chaos* (neglecting correlations, see equation (3.14), section 3.4.1) is assumed. For now, it is unclear to us which approach is the better one (if any). The fact that Waldeer only considers equilibrium solutions to the kinetic (mesoscopic) level of description and not to the equivalent macroscopic model makes it difficult to compare this approach to the other models mentioned in this section. Interestingly, Waldeer (2002a, 2002b, 2004c) uses Monte Carlo simulations to solve the kinetic system, a valuable alternative to the direct integration described by Klar & Wegener (1999).

To conclude our review of kinetic models, we mention the work by Nelson & Raney (1999). These authors investigate theoretically the interpretation of kinetic traffic flow models. They consider two possible interpretations of the probability density function ρ:

- single-instance sampling: ρ describes the probabilities of possible vehicle states of all vehicles present on a specified road at a specified designated time, or

- ensemble sampling: ρ describes the probabilities of possible vehicle states of all vehicles present at a designated time on one of an ensemble of identical roads (or: ρ determines the histogram of all instances of macroscopically equivalent sets of individual vehicle states on different days).

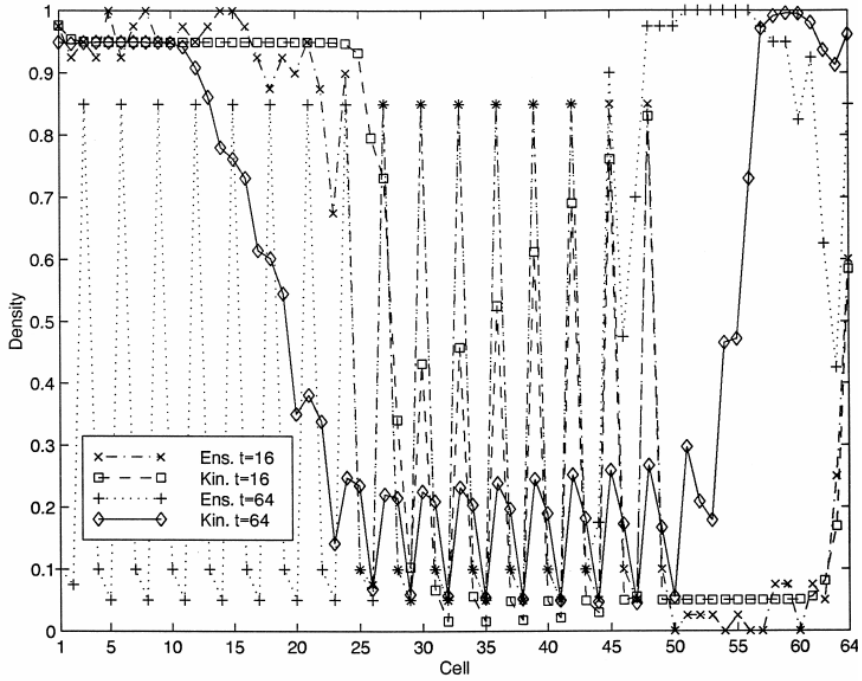


Figure 3-3 Differences between the average density of an ensemble of microsimulations with a CA model (indicated 'Ens.') and the density in an equivalent kinetic model (indicated 'Kin.') at two different time steps (Nelson & Raney, 1999)

Whereas the authors advocate the second interpretation, they show that the model output for an ensemble of Cellular Automata (CA) simulations and the corresponding kinetic model differ significantly (Figure 3-3), thereby questioning the validity of kinetic traffic flow models in general. A closer look at the CA under consideration shows that it has strongly asymptotical unstable properties, which cause traffic flow instability in the metastable density range (see section 2.3.1). We interpret these results as a warning that stable or homogeneous-in-space solutions of the kinetic model should be interpreted with care: the solution might represent only the 'optimistic' subset of a wider (metastable) range of solutions. Stability of a solution should therefore only be 'believed' if the superposition on the solution of (finite but large enough) perturbations converges again to the same stable solution. In fact this is consistent with the findings of Lee et al. (1998) and Helbing et al. (1999) for their metastable models. Or differently: metastability of a macro or kinetic model might be regarded as an indication for platoon instability of the (implicitly or explicitly) underlying microscopic car-following model.

3.4.3 Kinetic model analysis of ADA systems

To the authors' knowledge, kinetic models as such have not been applied to the problem of vehicle control. There are only some examples of authors who used the macroscopic transcription of kinetic models and apply it to the vehicle or traffic control problem.

We briefly mention the works by Lenz (1999) and Hoogendoorn & Bovy (2001a). Both authors apply the macroscopic formulation of a kinetic model to variable speed control using Variable Message Signs. The principle applied by Lenz is very similar to that exploited by Karaaslan et al. (1990, see section 3.3.4). Lenz applies a sliding mode controller to the model of Helbing et al. (2001) that is transformed into moving coordinates for this purpose. Stop-and-go waves are damped because the controller takes care of increasing the flow above some density and by decreasing the flow below this density, which is mimicked by a suitable change of speed limits on the VMS. Hoogendoorn & Bovy (2001a) consider a similar application, but with various control algorithms, among others the homogenising control strategy as applied on Dutch motorways. VMS warn drivers of slower traffic ahead. Hoogendoorn assumes that drivers decrease their desired velocity when control is active and implements the variable *average* desired speed in the *macroscopic* implementation of his model. As a result the occurrence of *phantom traffic jams* (induced by a temporary truck density perturbation) is prevented in the situation with control, as long as the speed adaptation by drivers is sufficiently large.

It appears that in neither of the two examples, the effect of traffic flow control was specified on the level of the microscopic behavioural rules that underlie the kinetic model. Rather the macroscopic transcription was used, and the effect of traffic flow control measures on the macroscopic behaviour is presupposed. The advantage of using kinetic based models rather than purely macroscopic models is that the former have more parameters with a direct relation to driver behaviour that can be influenced by the control policy (e.g. desired speed distribution). In the applications presented here, this appeared to be enough to analyse the application under consideration.

We conclude that the use of kinetic model principles for the design or assessment of ADA systems is a new field of research that has not been explored yet. The challenge is to implement the effects of traffic flow control and ADA systems into these models by flexibly adapting the microscopic model foundations, after which the behavioural changes can be transcribed into a macroscopic model formulation.

3.5 Conclusions and discussion

After reading a review of microscopic traffic flow models, the behavioural researcher Boer (1999) concluded that driver behaviour models in traffic flow theory and traffic engineering have been mainly inspired by mechanistic approaches while insights from behavioural research have been neglected. We conclude from our review that this is even more so for the behavioural assumptions behind kinetic (mesoscopic) and macroscopic traffic flow models. Assumptions on individual driver behaviour are only

loosely or indirectly related to specific terms in the dynamic equations in second or higher order macro models, or to a specific shape of the equilibrium relation or phase curves in the first order or LWR-type models. In kinetic models individual driver behaviour is taken as a starting point, but the behavioural assumptions are rather coarse. So far, the development of these models has been focussed on (necessary) improvements of the statistical representations and assumptions. Recently however, there is renewed attention for the behavioural foundations of macroscopic traffic flow models. With the works of Zhang (2003b) and Hennecke et al. (2000) in the Payne-type models, Daganzo (2002a) in the LWR-type models, and Nelson (1995), Nelson et al. (1997), Klar & Wegener (1999), and Waldeer (2000, 2001, 2004a, 2004b) in the kinetic models, we see a clear tendency to make macroscopic traffic flow models more consistent with realistic individual driver behaviour.

Within the kinetic theory, authors have strived for models explaining all available empirical knowledge on traffic flow dynamics, *without* assuming driver behavioural changes as an explanation. Although we recognise the value of these efforts from the point of view that "the simplest model explaining all relevant observations is the best", we proceed along a different path, where the role of *variable driver behaviour* or *variable driving strategies* is recognised as a potential (alternative or additional) explanation for certain empirical facts. After all we can – from a theoretical point of view – not deny the fact that driver behaviour is highly adaptive under the influence of all kinds of factors, like: weather, car equipment (e.g. ADAS or cellular phones), traffic legislation and enforcement, distracting factors inside and outside the car, road lay-out, tunnels etcetera. Since from the perspective of the individual driver, different traffic regimes are experienced quite differently (e.g. free driving, oscillatory or stationary congestion patterns), it is hardly imaginable that driver behaviour be constant in these widely varying conditions. However, up to now little attention has been given by traffic flow model developers (mainly engineers, mathematicians or physicists) to providing appropriate entries in their models where driver psychology can be plugged in. Therefore, the work in this thesis is aimed at introducing as many degrees of freedom as possible to include various behavioural hypotheses in macroscopic traffic flow models.

We believe that a better representation of driver behaviour and a link with individual behavioural models is not only necessary from the theoretical point of view. With Advanced Driver Assistance systems still in its infancy, the question is now raised how traffic flows will behave once ADA systems are more widely deployed. At best, empirical observations from individual equipped test vehicles or driving simulator experiments are available, on the basis of which individual behaviour models are built. Traffic flow models are the only means of extrapolating these individual models to the macroscopic level of traffic flows.

We reviewed some example applications of traffic flow analyses with ADA systems. Especially the (sub-) microscopic and macroscopic levels of detail have been successfully applied. However, both types of models are not equally suited for all types of traffic flow analyses with ADA equipped vehicles. The (sub-) microscopic approach is appealing, because the high level of detail allows virtually all ADA systems to be implemented. Moreover, the direct correspondence with the real world makes results

and model specifications comprehensible to outsiders. However, microscopic models are not necessarily the best option for exploratory studies of ADA concepts. In that stage of development, detailed specifications of the ADA systems are not precisely known, and the microscopic approach might be too detailed. In these cases the macroscopic approach might be the better choice, since it directly shows the collective dynamics of traffic flow without detailed behavioural specifications. However, here the problem is how to include the behavioural changes of the drivers and vehicles without a priori limiting the model's outcome to a predictable solution. Explorative analyses have therefore been limited to relatively simple strategies with influence on those few parameters that occur in the macroscopic model (like average speed at equilibrium). External speed control using advisory or compulsory speed limits on VMS signs or by directly determining the speed of automated vehicles (e.g. set speed of an ICC controller) is a common case study reported in literature using macroscopic models.

Kinetic models, with a level of detail between micro and macro approaches, theoretically offer more flexibility for implementing ADA systems than macroscopic models, but do not require too detailed specifications like the more complex (sub-) microscopic modelling tools. However, ADA applications in mesoscopic models have not been reported in literature so far. We therefore strongly belief that explorative analysis of ADA systems in kinetic (mesoscopic) models is worthwhile, both from a scientific and practical point of view. In the next chapter, the foundations of a kinetic model are laid that builds explicitly on a refined representation of individual driver behaviour in a most flexible way. This property allows simulating alternative individual behaviour, such as behaviour with ADA systems, as illustrated in the model application in chapter 8.

4 FUNDAMENTALS OF THE HUMAN-KINETIC TRAFFIC FLOW MODEL

4.1 Introduction

The purpose of the model development within this thesis is twofold. On the one hand, we aim for a model that is capable of accurately describing and explaining traffic flow dynamics in all regimes, including congestion patterns. On the other hand, a model is desired to explore the influence of Advanced Driver Assistance (ADA) systems and corresponding changes of the behaviour of the driver and/or the vehicle on traffic flow dynamics.

The former purpose is a prerequisite: a model aimed at exploring the influence of alternative behavioural specifications should at least reproduce the reference situation correctly. In our case the reference situation is current traffic. The minimal requirement to the model is therefore that it describes traffic flow dynamics at least as good as the best existing traffic flow models. For now, the development is focussed on single or mixed-lane traffic (instead of distinguishing each lane separately, which is referred to as *multi-lane* modelling), with special attention for congested traffic and the transitions between free flowing and congested traffic

The latter purpose requires that the model buils only on specifications of individual vehicle and driver behaviour, and as such provides maximal flexibility for exploring alternative specifications of individual behaviour of the driver and/or the vehicle. For instance, if new insights into individual driver behaviour in congested traffic become

available, these may form the basis of a novel, more refined individual driver model; our model should then be capable of predicting the influence of this refined behaviour on the macroscopic or traffic flow level. Also the introduction of Advanced Driver Assistance systems in traffic may change the behaviour of the driver and/or that of the vehicle. The proposed model should contain the most important aspects of individual behaviour of the driver and the vehicle, so that behavioural changes due to the presence of ADA systems in the vehicle can be accommodated. These requirements to the model can be summarised as follows: the model should bridge the gap between the individual or microscopic level of description of driver and vehicle behaviour, and the macroscopic level of traffic flows.

From the previous chapter, we concluded that kinetic traffic flow models are the best starting point for our purposes for three reasons:

(i) the model's output and formulation is directly at the macroscopic level, the same level of description as the collective traffic flow dynamics in which we are interested;

(ii) the representation of traffic flow dynamics – especially that of congested traffic patterns – is among the best available to date;

(iii) the basic characteristic of the kinetic approach is that it relates the motion of individual vehicles to the dynamics of traffic flow.

The model development in this and the next chapters is thus based on the kinetic theory of traffic flows. We refer to Annex A for an introduction to the existing basic theory of kinetic vehicular traffic flow modelling. Our own contribution to this theory concentrates on the representation of individual driver modelling; we have therefore chosen to call our approach of traffic flow modelling *human-kinetic* modelling. This name expresses that the individual driver model, that was inspired by interactions of gas particles in traditional *gas-kinetic* traffic flow models, is more inspired by the human driver in our approach (see section 3.4).

The remainder of this chapter presents the starting point of our contribution to kinetic traffic flow modelling by providing a critical discussion of time scales of the processes in traffic flow (section 4.1.1). Section 4.1.2 summarises the individual behavioural model that underlies the development of the human-kinetic traffic flow model in this and the next chapters. The human-kinetic modelling approach is developed in sections 4.2 and 4.3. The next section 4.4 indicates how the human-kinetic traffic flow model can be numerically evaluated. Finally section 4.5 concludes the chapter with conclusions and an outlook to the next chapter.

4.1.1 Motivation for modelling adaptive individual driver behaviour as a continuous process: discussion of time scales in kinetic traffic flow models

In this section a key argument is developed that underlies the approach used throughout this thesis. A fundamental choice in this approach is to model both acceleration and deceleration of traffic in the same way, that is: as smooth or continuous processes with a

finite duration. Although different from traditional gas-kinetic traffic flow modelling approaches, this choice is justified based on arguments of *time scales.*

The basic argument is the following. Suppose that a many particle system such as gas, granular, or vehicular flow is considered. Then in principle the duration of all interaction processes of the particles influencing the dynamics of that system should be modelled, *unless* the duration of the interaction is *sufficiently small*. The latter is the case only if:

- the duration (time scale) of the interaction process is much shorter than the time scale of the overall dynamics of the system, *and*
- once an interaction between a number of particles has started, the probability that another particle – initially not involved in the interaction – influences this interaction (or is influenced by it) before the end of the interaction process, is sufficiently small.

Only if these two conditions are fulfilled, can the duration of the interaction be neglected (Jeans, 1982). It is then sufficient to consider an instantaneous transition from the particle states prior to the interaction, into a modified state after the interaction, without taking into account the transient states during the interaction process.

As an illustration let us consider the kinetic theory of gasses, where interactions between particles are often considered instantaneous. This is justified by looking at typical time scales of the processes involved. For the sound and other waves in air for example, the speed is in the order of 300 m/*s* with wave lengths in the order of 10^{-2} to 10 meters (i.e. sounds of 20000-30 Hz, Kinsler et al., 1999). This means that pressure, density and other macroscopic variables for gas flows change with a time scale in the order of $10^{-2} – 10^{-4}$ seconds. On the other hand, the duration of interactions between gas particles is many orders of magnitude smaller ($\sim 10^{-12}$ seconds, Dettmann et al., 1999). Moreover, the probability that during the interaction between two gas particles an interaction with another (third or more) particle occurs is negligibly small. Therefore, when analyzing macroscopic flows of gasses, the interaction time between molecules is indeed negligible.

Now do the same arguments hold for traffic flows? In traffic, typical (macroscopic) wave speeds lie in the order of 10-20 m/*s* with wavelengths in the order of 100 – 500 meters, so that a typical time scale for changes of the macroscopic traffic flow variables like density is 10 – 50 seconds. For comparison, consider the duration of individual acceleration: a powerful sports car can accelerate from 0 to 30 m/*s* in approximately 5 seconds, but normal person cars (and a fortiori trucks) rather approach 50 – 100 seconds in normal traffic flow. This is the same time scale as the macroscopic wave speed. Already from the argument of time scales, it can be concluded that acceleration should not be considered an instantaneous process. Also the second condition is not fulfilled: when a vehicle is accelerating, one cannot be sure that it reaches its target speed before a braking interaction with a newly perceived slower vehicle starts. For these reasons acceleration is treated as a *continuous* process in traditional gas-kinetic traffic flow models (see section 3.4) and in the human-kinetic model developed in this thesis.

Consider now the deceleration or braking interactions. Such interaction usually takes approximately 1 – 5 seconds. Compared to the macroscopic time scale of 10 – 50 seconds, the difference is hardly 1 order of magnitude, much smaller than in gas kinetics (where it was approximately 8 orders of magnitude). Moreover, it is not self-evident to neglect the probability that a decelerating vehicle would need to increase its braking force, for instance because the slower predecessor approaches another even slower car. Authors of existing gas-kinetic traffic flow models have appartently accepted the small time scale difference as an argument to neglect the duration of deceleration interactions – albeit rather implicitly by adopting the analogy with kinetic models for gas flows. As a result, gas-kinetic traffic flow models contain a *continuous* acceleration term that is balanced by an *instantaneous* deceleration term[5].

The model development in this chapter however, proceeds along a different path. It is felt that the conditions for neglecting the duration of interactions are fulfilled neither for acceleration nor for deceleration. The approach outlined in this chapter therefore models acceleration and deceleration in a similar way, i.e. by taking the finite duration of both processes explicitly into account.

The approach consists of determining the expected continuous acceleration of individual vehicles in the so-called acceleration integral. The acceleration integral acts upon the macroscopic speed in a continuous way, thus influencing macroscopic traffic behaviour (see section 4.2.2).

4.1.2 The model of individual driver behaviour in the human-kinetic traffic flow model

In this section we give an overview of assumptions on individual driver behaviour that are used throughout this thesis. The aspects mentioned here are elaborated in the next chapters. Therefore, the next section only outlines the steps in the model development, and refers to the sections in this or the next chapters where each step is motivated and detailed.

The human-kinetic traffic flow model models traffic flow behaviour on motorways. This thesis mainly elaborates the longitudinal driver behaviour, whereas lane changes are only briefly discussed. With respect to lane changing, free lane changes and mandatory lane changing at on-ramps (merging) are considered. The individual driver behaviour is not necessarily characterised by constant control parameters throughout the journey; we refer to variable driver behaviour or *various driving strategies* during the journey. As a first approximation, the current driving strategy is assumed to be a function of one single parameter, the *activation level*, that can change due to events in the vehicle, in traffic surrounding the vehicle, or even in the environment of the road. Finally the influence is considered of in-vehicle equipment that assists the driver in his task (Advanced Driver Assistance or ADA).

---

[5] Actually, the model of Klar & Wegener (1999) is an exception that treats both acceleration and deceleration as instantaneous processes.

Note that the term *activation level* is used here, but that in behavioural psychology also the terms alertness, attention, awareness, motivation etcetera are used, which all have a rather different meaning. The term activation level is used here not in contrast to these other terms, but as a more generic behavioural parameter that determines the current driving strategy (see section 6.1). Eventually, when the behavioural model is further refined in future work, it may be necessary to refine this behavioural variables into different behavioural components as well.

The different aspects of driver behaviour are discussed in the next sections. Distinction is made between *basic* and *variable* individual behaviour. *Basic* behaviour refers to the behavioural model of the driver (both longitudinal and lane changing behaviour) *without* consideration of temporal variations of the driving strategy due to activation level variations. In theory all parameters describing basic individual driver behaviour can be assumed to be *variable* and dependent on the activation level. Those parameters that will be considered variable further in this thesis, are marked in the summary below as *"variable"*, otherwise the label *"invariable"* is used.

Basic individual longitudinal driving behaviour

The longitudinal behaviour is described by the car-following model underlying the human-kinetic traffic flow model. This is a deterministic feedback control loop where the driver controls the acceleration of the vehicle, based on perceived audio-visual cues. Although in principle alternative car-following models can be considered to form alternative human-kinetic traffic flow models, the model in this thesis builds on a car-following rule that is loosely based on the model of Helly (1961) and that forms the foundation of the Mixic simulation model (Van Arem et al., 1997b).

The longitudinal driver behaviour has the following characteristics:

- Vehicles have a finite vehicle length (no point-sized vehicles as in many gas-kinetic models).
- Drivers only react to traffic conditions in front of them (downstream). This behaviour is also called *anisotropic* behaviour, in analogy to anisotropic physical systems, for which the response (e.g. elastic counterforce) depends on the direction in which the stimulus is applied.
- Drivers control their vehicle longitudinally through the acceleration, that is: by applying either the gas pedal or the brakes. No distinction is made between using the brakes and using the engine brake; we assume that the driver is able to apply exactly the control needed to achieve the desired acceleration. A fixed time delay is considered between the perception of the stimuli influencing the desired acceleration and achieving the actual acceleration. This *reaction time* (*invariable*) is a lumped delay consisting of delay in perception, in taking decisions, in applying the controls, and the delay in vehicle response.
- The acceleration of the vehicle is physically limited between a fixed minimal and maximal acceleration (*invariable*). The minimal acceleration (maximal braking) is stronger than the maximal acceleration, so that car-following behaviour is potentially asymmetric.

- Other than the limitations to acceleration and deceleration and the delayed vehicle response, implicitly modelled in the lumped reaction time, no dynamic vehicle model is taken into account.
- Drivers drive either in *unconstrained* (also: *free flow*) mode or in *constrained* (also: *car-following*) mode; these modes are distinguished at each time instant, dependent on which mode actually requires the lowest acceleration; in other words: a driver accelerates according to free flow mode unless the presence of other traffic downstream requires a lower acceleration; the driver then switches to car-following mode until the lower acceleration for the predecessor(s) is not longer required.
- In free flowing mode, the aim of the driver is to reach or maintain his *desired speed* (*invariable*).
- In car-following mode, the aim of the driver is (a) to reach or maintain a *desired velocity-dependent distance* with respect to the predecessor, and (b) to adapt the own speed to that of the predecessor. The desired distance is equal to the own vehicle length, increased by: a constant minimal safety margin (*invariable*) and a speed dependent safety margin (*variable*); the latter is linear or quadratic with the own driving speed.
- In car-following mode, the driver reacts to the state of the direct predecessor on the own lane only. However, we also model *anticipation behaviour* (*variable*), that is: not only the current state of the predecessor is used; the driver also extrapolates the state of the direct predecessor in the near future, based on the perceived spatial speed profile that the predecessor will inevitably follow.

The model development is considerably simplified by assuming all drivers identical except for the desired speed. For all other parameters related to the individual driver behaviour, only one value is considered that is characteristic for some *average driver*. Of course in reality every driver is different. Therefore, important differences between groups of drivers can be included in the model by distinguishing *user-classes*. In theory this distinction is not only suited to discern passenger cars from trucks and busses, or ADA-equipped vehicles from non-equipped vehicles, one could also consider driver types with considerably different parameter sets or even behaviour. Since in theory the number of user-classes is not limited, one could theoretically consider as many user-classes as there are drivers. In practice however, this is not a realistic and efficient option.

The basic longitudinal behaviour is further specified in sections 4.2 to 4.3.2 of this chapter. The distinction between user-classes is discussed separately in chapter 7.

Basic individual lane change behaviour (invariant)

Within this thesis, distinction is made between free lane changing and mandatory lane changing at on-ramps. The following assumptions are made with respect to *free lane changing*:

- A lane change is performed if the driver desires to change lanes and if it is possible to do so (gap acceptance).
- A driver desires a free lane change only when he is in following mode and has to brake for a slower predecessor. If the predecessor accelerates the driver stays in his lane.
- A driver who desires to change lanes is able to do so with a certain probability that a suitable gap is available on the target lane. We do not further specify this probability, but do some suggestions for further development.
- The model considers mixed-lane traffic, i.e. flows are not explicitly modeled separately for each lane (as in multi-lane models). This is however no theoretical limitation to the model, but a result of priorities that have been set during model development. It is therefore not necessary to specify the target lane (right or left).
- The effect of a free lane change on the distribution of traffic over different lanes of a motorway is not taken into account; only the longitudinal effect is considered: a driver who changes lanes continues his journey without loss of speed, whereas a driver who is not able to change lanes has to brake.
- The duration of a lane change is not taken into account: lane changes happen instantaneously[6].

The following assumptions are made on individual driving behaviour with respect to *mandatory lane changing for merging* onto the motorway:

- Merging behaviour is not modelled individually but only at the macroscopic level. The assumptions discussed here are implicit assumptions behind the macroscopic specification.
- We consider a time dependent total inflow of traffic onto the motorway; this traffic effectively enters the motorway within a merging zone (part of the on-ramp where lane-changing onto the motorway is possible).
- Every driver who desires to merge forces himself onto the main lanes, unless there is physically no space to enter the main lanes.
- Not every driver merges at the same longitudinal position. It is not specified individually how this choice is made; it is assumed that in the merge zone the merging intensity is maximum at a given point, and that the merging intensity decreases linearly up- and downstream of this point to the beginning and end of the merging zone.
- Each driver merging onto the motorway has adapted his speed to match that of the main lanes, prior to performing the merge.

[6] Strictly speaking the conditions on *time scales* (see section 4.1.1) are not fulfilled for lane changes. It would therefore be more correct to model these as continuous processes instead of instantaneous interactions. The assumption made here is made for convenience and in analogy to existing approaches. The refinement as continuous process is recommended for future research.

The basic lane change behaviour is further specified in section 4.3.3 of this chapter. In this thesis *no variability of lane changing parameters* is considered, although in theory this could be done in the same way as is illustrated for variable longitudinal behavioural parameters.

Variable individual longitudinal behaviour (summary)[7]

The current behaviour as described in the previous section is referred to as the *driving strategy*. The driving strategy may vary during the journey. Within this thesis the simplified assumption is used that one behavioural variable sufficiently characterises the current driving strategy: this variable is denoted as the *activation level* (see section 6.1). The activation level may vary due to events in the vehicle, in traffic surrounding the vehicle, or even in the environment of the road. In the following, only a very limited number of behavioural assumptions with respect to the activation level is considered, mainly as an illustration of how variable driving behaviour can be modelled and analysed.

These are the assumptions concerning the variability of the activation level:

- The activation level varies between certain minimal and maximal boundaries.
- Each driver has some *comfortable* or *normal* activation level, with which he preferably drives.
- When the activation level deviates from the normal level due to some cause, the driver will gradually return to the normal activation level after the cause for a higher or lower activation level has disappeared.
- Temporary deviations from the normal activation level are possible in both senses: a driver may temporarily drive more active or can temporarily be distracted, so that less mental resources are available for the driving tasks. We only consider some example specifications:
    - The activation level of a driver suddenly rises to a high level when a vehicle cuts in in front of the vehicle.
    - The activation level rises to a high level when merging onto the highway.
    - The activation level gradually decreases if a driver is caught in prolonged congestion, i.e. when driving speed remains low for a longer period.

The activation level would not affect traffic dynamics if it had no effect on individual driving behaviour. Among the many potential hypotheses and model specifications for the influence of activation level on driving, the following assumptions are elaborated as an illustration:

- With a higher than normal activation level, the driver in car-following mode is able to follow the predecessor at a shorter distance than the normal desired following distance. The driver is indifferent with respect to a range of distances

[7] This section provides only a summary of the assumptions on variable individual longitudinal behaviour; for further details and a motivation of the assumptions the reader is referred to chapter 6.

from a minimal acceptable distance (which decreases with increasing activation level), to the normal desired following distance. This means that as long as the actual following distance is within this *indifference band*, the driver does not brake nor accelerate (at least not with the aim of changing the following distance; possibly he does accelerate or brake to minimise speed difference with the predecessor).

- With a lower than normal activation level, the desired time headway increases; this means that a driver can drive less active, if this is compensated by a longer following distance.
- With a higher than normal activation level, the driver is able to anticipate better. That is: a more active driver weighs the extrapolation of the predecessor's state based on the downstream spatial speed profile stronger in his car-following response.

To avoid making too many changes at a time, the above-mentioned behavioural reactions to increased or reduced activation level are isolated in separate case studies. Hence, the combination of both behavioural aspects is not considered for the sake of comprehensibility, although such combinations are plausible in practice.

Variable individual driver behaviour is specified and motivated separately in chapter 6.

Individual driver behaviour with ADA system[8]

As an example for the applicability of the human-kinetic model developed in this thesis to the analysis of ADA systems, a queue-tail warning system is considered. This is an ADA system that automatically sends warnings upstream if the vehicle brakes sharply. ADA-equipped vehicles receiving such warnings provide a warning to the driver for the upcoming braking manoeuvre. The warning is stronger as the warnings received from downstream are more numerous. The behavioural assumptions with respect to this system are:

- If a driver receives a warning, he immediately adopts a higher activation level than the current level. The change of activation level is proportional to the strength of the warning.
- As a result the driver anticipates stronger to upcoming speed drops (as discussed in the previous section).

The modelling of ADA-systems and the according individual driver behaviour are specified separately in chapter 8.

[8] This section provides only a summary of the assumptions on individual driver behaviour with ADA; for further details and a motivation of the assumptions the reader is referred to chapter 8.

4.2 Continuous adaptive individual driver behaviour in the human-kinetic macroscopic traffic flow model

From the discussion on time scales in section 4.1.1, we concluded that it is theoretically better to build kinetic traffic flow models with acceleration and deceleration both modelled as smooth processes. Starting from generic results from the kinetic theory, summarised in section 4.2.1, the smooth *acceleration integral* is developed in section 4.2.2. The acceleration integral relates the microscopic specification of individual driver behaviour to the macroscopic level of traffic flow dynamics through the use of *microscopic car-following* models.

4.2.1 Generalised kinetic continuity equation and the generic macroscopic traffic flow equations

The starting point of the kinetic model development is the definition of the non-local and local vehicle state vector S' and S of an individual vehicle and the probability density function ρ of individual vehicle states $\rho(t,S')=\rho(t,x,S)$. The latter function is defined so that $\rho(t,x,S)\ dx\ dS$ defines the expected number of vehicles having state $S\in[S,S+dS)$ at time t and location $x\in[x,x+dx)$. For a more precise definition of these traditional kinetic variables the reader is referred to section A.2, Annex A, where the fundamentals of the kinetic traffic flow theory are summarised. Moreover, it is explained there how the generalised density ρ relates to the traditional vehicular density $k(t,x)$, the average speed $V(t,x)$, the flow rate $Q(t,x)$, and the speed variance $\Theta(t,x)$.

For the generalised density $\rho(t,x,S)$, a continuity equation in the state space can be derived (Hoogendoorn ,1999; Leutzbach, 1988):

$$\frac{\partial \rho}{\partial t}+\nabla_{S'}\cdot\left(\rho\frac{dS'}{dt}\right)=\left(\frac{d\rho}{dt}\right)_{discrete} \tag{4.1}$$

In this equation the dependency of ρ on (t,S') or (t,x,S) has been omitted for notational simplicity; the product operator $\cdot$ denotes the inner product and the Nabla operator $\nabla_{S'}$ for the non-local state vector S' with dimensions $(x,s_1,s_2,\ldots,s_n)$ is defined by:

$$\nabla_{S'}\equiv\left(\frac{\partial}{\partial x},\frac{\partial}{\partial s_1},\frac{\partial}{\partial s_2},\ \ldots\ ,\frac{\partial}{\partial s_n}\right) \tag{4.2}$$

The interpretation of equation (4.1) is as follows. In the LHS it states that changes in time of the generalised density are caused by changes in time of the state variables s_i. These changes are larger as the expected number of vehicles in that state is larger (hence the multiplication by ρ) and as the sensitivity of ρ for changes in the state variable is larger (hence the derivative with respect to s_i in the Nabla operator). On the other hand, some discrete event can cause a sudden increase or decrease of ρ for a certain state S. For instance, new vehicles with state S can suddenly enter the system at position x, causing a discrete increase of $\rho(t,x,s)$. These discrete contributions to the change of ρ are accounted for in the term in the RHS of equation (4.1).

Traditionally, the *kinetic* continuity equation is cast into an equivalent system of coupled differential equations for the *moments* of the generalised density function ρ. The *method of moments* is traditionally used to transform the kinetic equation (4.1) into its equivalent macroscopic formulation (for a description of this method, see section A.3.2, Annex A). Typical of the macroscopic system derived from (4.1) is that the n^{th} equation in the macroscopic system describes the dynamics of the $(n-1)^{th}$ order moment of ρ, in terms of the first $n+1$ moments (i.e. moments 0 until n). Therefore, the macroscopic system is recursive, and in order to obtain a *closed* system of n equations (i.e. for all moments up to order $(n-1)$), some a priori assumption needs to be made for the n^{th} moment.

For the macroscopic equivalent of the human-kinetic traffic flow model, we choose to close the macroscopic system with the second-order moment of ρ. This means that the density k and the average speed V are independent macroscopic variables, the second-order moment Θ is a priori considered as an equilibrium variable that is a function of k and V, and higher-order moments are assumed to be zero. The motivation for this choice is the following:

- in first order models (closed by assuming an equilibrium value for the speed V as a function of the density k) instabilities do not occur; therefore this choice would neglect some essential empirical features of congested traffic (see sections 2.4.2 and 3.3.2);
- third order models (Helbing, 1997; Hoogendoorn, 1999) have theoretical benefits (e.g. Hoogendoorn shows that such models have three characteristics moving with speeds slower than, equal to, and faster than the average speed of traffic). These models however have not incontrovertibly proven their surplus value in describing actual traffic flow dynamics, which has made Helbing decide to return to second order models in later publications (for example Helbing et al., 2001; see also section 2.4.2.1).

After application of the method of moments to (4.1) and closing the system with the second order moment of ρ, the following equations form the basis for the human-kinetic traffic flow model (for the mathematical derivation, see A.3.2.2, Annex A):

$$\frac{\partial k}{\partial t} + \frac{\partial kV}{\partial x} = \left(\frac{dk}{dt}\right)_{discrete} \tag{4.3}$$

$$\frac{\partial kV}{\partial t} + \underbrace{\frac{\partial kV^2}{\partial x}}_{convection} + \underbrace{\frac{\partial k\Theta}{\partial x}}_{pressure} = \underbrace{k\left\langle \frac{dv}{dt} \right\rangle_v}_{smooth\ \ acceleration} + \underbrace{\int_v v\left(\frac{d\rho}{dt}\right)_{discrete} dv}_{discrete\ \ acceleration\ 1} \tag{4.4}$$

Equation (4.3) is the macroscopic continuity equation for the vehicular density k. The term $\left(\frac{dk}{dt}\right)_{discrete}$ in the RHS stands for flow of vehicles into or out off the road section under consideration. Equation (4.4) describes the dynamics of the flow rate $Q=kV$, also called the speed momentum equation or conservative speed equation. The primitive formulation or speed equation is mathematically equivalent, and reads:

$$\frac{\partial V}{\partial t} + \underbrace{V\frac{\partial V}{\partial x}}_{convection} + \underbrace{\frac{1}{k}\frac{\partial k\Theta}{\partial x}}_{pressure} = \underbrace{\left\langle \frac{dv}{dt} \right\rangle_v}_{smooth\ \ acceleration} + \frac{1}{k}\left(\underbrace{\int_v v\left(\frac{d\rho}{dt}\right)_{discrete} dv}_{discrete\ \ acceleration\ 1} - \underbrace{V\left(\frac{dk}{dt}\right)_{discrete}}_{discrete\ \ acceleration\ 2} \right) \quad (4.5)$$

Before introducing the new developments of this thesis, let us recall the physical interpretation of the different terms in equations (4.4) and (4.5):

- convection term: change of the average speed V due to a spatial speed gradient that is carried with the speed of the flow V;
- pressure term: change of the average speed V due to individual vehicles that travel with speed $v < V$ or $v > V$; the former remain 'longer' at location x, whereas the latter leave x 'earlier'; for a discussion on the physical meaning of the pressure term in this model equation and other macroscopic models, the reader is referred to Annex A; in the terminology developed there, the pressure term in equation (4.5) is typified as pure *speed variance pressure*;
- smooth acceleration: change of the average speed V due to an unbalance in smooth individual accelerations;
- discrete acceleration 1: change of the average speed V due to any event causing a discrete change of $\rho(t,x,v)$, the expected number of vehicles with speed v;
- discrete acceleration 2: change of the average speed V due to the discrete in- or outflow $\left(\frac{dk}{dt}\right)_{discrete}$ of vehicles into or out off the road section under consideration; the term expresses that the total speed momentum kV has to be redistributed over a suddenly changed number of vehicles.

The generic results of this section are used as a basis for our own modelling approach in the next section.

4.2.2 Continuous adaptive individual driver behaviour in the human-kinetic macroscopic traffic flow model: micro-macro link

Where traditionally the acceleration of vehicles is specified in the 'smooth acceleration'-term of equation (4.4) or (4.5) and decelerations as interactions in the 'discrete acceleration 1'-term (see A.4, Annex A), the model development in this thesis now proceeds along a different path and treats both processes simultaneously in the 'smooth acceleration'-term, as motivated in section 4.1.1.

The second order macroscopic traffic flow model established by equations (4.3) and (4.4) or (4.5) contains no specification of the behaviour of drivers and vehicles so far, other than the assumption that the local state $S(t,x)$ of an individual vehicle-driver unit is fully determined by the speed v of the vehicle only.

For simplicity and without loss of generality, let us assume in this section that the discrete in- or outflow of traffic $\left(\frac{dk}{dt}\right)_{discrete} = 0$. Therefore, the 'discrete acceleration 2'-term is also zero. Since deceleration is considered as a smooth process, there are no discrete transitions from one speed to another in the flow, since all such transitions are considered smooth, therefore also the 'discrete acceleration 1'-term is zero in the human-kinetic traffic flow model. The human-kinetic traffic flow model (in primitive formulation) is now described by:

$$\frac{\partial k}{\partial t} + \frac{\partial kV}{\partial x} = 0 \tag{4.6}$$

$$\frac{\partial V}{\partial t} + V\frac{\partial V}{\partial x} + \frac{1}{k}\frac{\partial k\Theta}{\partial x} = \left\langle \frac{dv}{dt} \right\rangle_v \tag{4.7}$$

The discussion concentrates on the specification of the acceleration and deceleration behaviour in the 'smooth acceleration'-term in the RHS of equation (4.7). This term is defined as the expected value of smooth individual vehicle accelerations or decelerations $\dot{v}$:

$$\left\langle \frac{dv}{dt} \right\rangle_v \equiv \int_v p_v\left(v|t,x\right)\dot{v}\,dv \tag{4.8}$$

As shown in the review of microscopic traffic flow models (3.2.1), numerous models exist that describe the microscopic acceleration $\dot{v}$ for an individual vehicle. These microscopic car-following relations give the individual acceleration of a vehicle/driver combination j as a function of various stimuli. The most common stimuli in car-following models are:

- own speed v_j of vehicle/driver combination j
- presence of a predecessor j-1[9]
- speed v_{j-1} of the predecessor
- acceleration $\dot{v}_{j-1}$ of the predecessor
- net distance s_j to the predecessor (gap between own front bumper and rear bumper of the predecessor minus a static safety margin, see section 4.2.3).

Some car-following models also take the potential presence of a pre-predecessor into account, in that case additional stimuli are:

- the presence of a pre-predecessor j-2
- the speed v_{j-2}, the acceleration $\dot{v}_{j-2}$ and the net distance s_j^{pp} to the pre-predecessor.

A general car-following relation is then defined as:

[9] the index j is used here as a counter for individual vehicles, starting from downstream to upstream

$$\dot{v}_j\left(v_j, s_j, v_{j-1}, \dot{v}_{j-1}, s_j^{pp}, v_{j-2}, \dot{v}_{j-2}\right) \tag{4.9}$$

For the sake of notational convenience, but without loss of generality, the discussion is restricted to car-following models that take into account only one predecessor. However, analogously to the procedure described here one could also elaborate acceleration integrals that consider pre-predecessors as well (or even a deliberate number of other vehicles acting as stimuli, possibly including vehicles in adjacent lanes, or vehicles elsewhere in the flow, the position and speed of which are transmitted e.g. through inter-vehicle communication). Moreover, the analysis is further restricted to a subclass of car-following models that is independent of the predecessor's *s* acceleration. Therefore, in the remainder of this thesis we choose to discuss only car-following models of the form:

$$\dot{v}_j\left(v_j, s_j, v_{j-1}\right) \tag{4.10}$$

In order to evaluate equation (4.10), a specification is needed for the speed v_j of the follower, for the gap s_j, between the follower and the predecessor, and for the predecessor's speed v_{j-1} as inputs for the function f_{CF}.

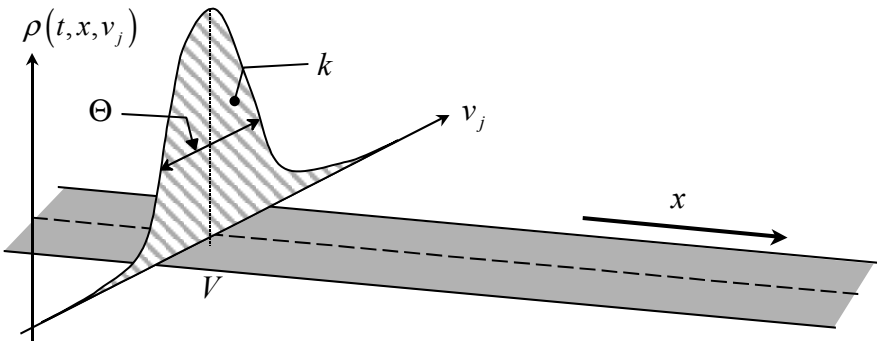


Figure 4-1 Probability density of individual vehicle speed if the macroscopic speed, speed variance and density are known

In contrast to a microscopic model implementation of the car-following model (4.10), the stimuli are not deterministically known in the kinetic model, but stochastic variables obeying the probability density function ρ. Figure 4-1 illustrates this for the speed v_j of the follower. If there is a vehicle present at location x, its speed is not known with certainty; we only know that the probability density ρ at that location describes the expected number of vehicles at t and x having speed v_j. Actually in equation (4.10) all three stimuli are stochastic, possibly correlated variables. Therefore, a three-variable correlated distribution has to be considered, which is schematically represented in Figure 4-2.

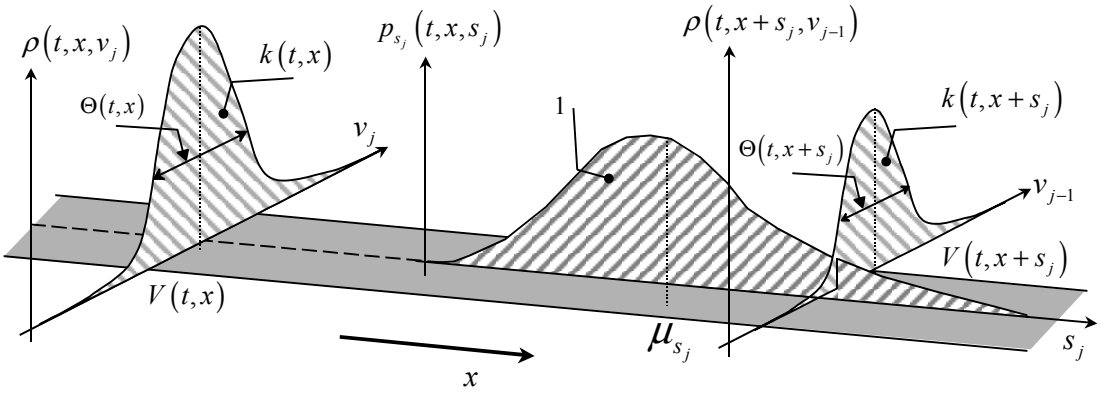


Figure 4-2 Simultaneous distribution of three variables: follower's speed, gap and leader's speed

The kinetic equivalent acceleration law (4.8) of the microscopic deterministic function (4.10) thus becomes:

$$\left\langle \frac{dv}{dt} \right\rangle_v \equiv \int_{v_j} p_v\left(v_j\right) \int_{s_j} p_s\left(s_j\right) \int_{v_{j-1}} p_v\left(v_{j-1}\right) \dot{v}_j\left(v_j, s_j, v_{j-1}\right) dv_{j-1}\, ds_j\, dv_j \tag{4.11}$$

with:

$$\begin{aligned} \left\langle \frac{dv}{dt} \right\rangle_v &= \left\langle \frac{dv}{dt}(t,x) \right\rangle_v \\ p_v\left(v_j\right) &= p_v\left(v_j \middle| t,x\right) \\ p_s\left(s_j\right) &= p_s\left(s_j \middle| t,x,v_j\right) \\ p_v\left(v_{j-1}\right) &= p_v\left(v_{j-1} \middle| t,x+s_j^0+s_j,v_j,s_j\right) \end{aligned} \tag{4.12}$$

Equations (4.11) and (4.12) are the most general formulation of the *acceleration integral* in the human-kinetic model for car-following functions of the form given in equation (4.10). The acceleration integral expresses that – if there is a vehicle *j* present at *t* and *x* – the expected smooth acceleration in the human-kinetic traffic flow model is found by considering all potential microscopic states at *t* and *x* in combination with all potential locations and speeds of the predecessor *j-1*; for all these potential car-following situations, the deterministic acceleration is calculated according to equation (4.10) and multiplied by the probability that this car-following situation occurs. The latter probability is a combination of the probabilities $p_v\left(v_j\right)$, $p_s\left(s_j\right)$, and $p_v\left(v_{j-1}\right)$ of occurrence of the follower's speed v_j, of the gap between the follower and leader s_j, and of the leader's speed v_{j-1}. The correlation between these probabilities is expressed in the conditions in equation (4.12): one of the three correlated variables can be chosen freely, here the follower's speed, the probability for the gap is conditional to the fact that there is a follower at *x* with speed v_j, and finally the speed of the predecessor is conditional to the fact that it is driving at a distance $s_j^0 + s_j$ in front of a follower at *x* with speed v_j.

Note that the gross distance between position x_j of the follower and position x_{j-1} of the leader is equal to the vehicle length (plus safety margin) s_j^0 of the follower, plus the actual gap s_j and is thus related to the density k (see next section for precise definitions).

With equation (4.11) an implicit crucial choice is made, in that the integral is expressed in terms of the gap s_j between a vehicle at x and his predecessor at x', and *not* in terms of the (unconditional) probability of having a vehicle at location x'. Indeed, since the car-following relation (4.10) is defined with respect to the predecessor, it would not be correct to consider the unconditional probability of finding any vehicle at x'. This probability would only be interesting if no vehicle were present between x and x'. Instead, by considering the gaps, we express the probability of finding a vehicle at location $x'=x+s_j$, *knowing that this is the first vehicle downstream of vehicle j at x*, which is consistent with the definition of a predecessor, and hence with the car-following relation (4.10).

Note that it follows from equations (4.11) and (4.12), that any traffic flow model that builds on this definition of the acceleration integral is by definition anisotropic and non-local. It is *anisotropic*, which means that traffic is only influenced by conditions downstream, because in the second integral only gaps are present with respect to vehicles in the driving direction (downstream). Hence, only traffic downstream of x affects the state of traffic at x. It is *non-local* since we look at interactions of vehicles having speed v_j at location x with vehicles having speed v_{j-1} at location $x+s_j^0+s_j$.

The human-kinetic model has to be further specified by making specific choices for the car-following relation (4.10) and for the probabilities in equation (4.12). These are given in the next section.

4.2.3 Specification of the acceleration integral

This section further specifies the acceleration integral of equations (4.11) and (4.12). It starts with the probabilities $p_v(v_j)$, $p_s(s_j)$ and $p_v(v_{j-1})$ of occurrence of the follower's speed v_j, of the gap between the follower and leader s_j, and of the leader's speed v_{j-1}.

4.2.3.1 Specification of probability distributions

In theory $p_v(v_j)$, $p_s(s_j)$, and $p_v(v_{j-1})$ are all correlated probabilities. They are implicitly defined by the generalised density function $\rho(t,x,v)$ and some additional spatial correlation functions, which is sometimes summarised by considering a two-vehicle non-local probability density $\rho_2(t,x,v_j,x+s_j,v_{j-1})$ (Helbing, 1997). Now two fundamental choices are made:

(i) as in most traditional kinetic traffic flow model, the assumption traditionally called *vehicular chaos* (see equation (3.14)) is adopted. This means that the correlations between the speeds of the follower and the leader and the gap are neglected;

(ii) all specifications of the probability distributions $p_v(v_j)$, $p_s(s_j)$, and $p_v(v_{j-1})$ are related to the moments of ρ ; i.e. specific types of distribution functions are chosen and their moments are calculated from the moments of ρ.

The former approximation is a pragmatic one. It is motivated by two arguments: (a) model development becomes much more complex when taking into account the correlations, so it is first analysed whether the model performs well without this complication; (b) those efforts in literature where correlations have been taken into account (see section 3.4.2) have so far been pure theoretical ones: eventually the researchers had to conclude that not enough data is available to calibrate the covariance matrices that occur in the model, and decided to neglect the correlations after all. In section 7.4.3 we reflect upon the consequences for the human-kinetic model of this simplification.

The latter choice is unavoidably related to the macroscopic character of the human-kinetic traffic flow model: since the macroscopic system was closed with the second-order moment of ρ only the first three moments of ρ are known at the macroscopic level: the total number $k(t,x)$ of vehicles expected at t and x, the expected speed $V(t,x)$ of these vehicles, and the variance $\Theta(t,x)$ of the speed. Therefore, these variables are the only information available to specify the probability distributions.

The remainder of this section gives the specifications for the speed and gap distributions. Gaussian or normal distributions are adopted for the speeds, and a lognormal distribution for the gaps. Both distributions are fully characterised by two moments: the average and the standard deviation. We show how these relate to the moments k, V, and Θ of ρ.

Speed distribution of the follower

Following Helbing (1997) and Hoogendoorn (1999), who examined empirical single vehicle data on motorways, a normal distribution is taken for the speeds. Note that the generalised density, given a time t and a location x then also is normally distributed, since the local speed distribution and the local generalised density function are coupled through:

$$p_v(v_j|t,x) = \frac{\rho(t,x,v_j)}{k(t,x)} \tag{4.13}$$

Therefore, by definition the average value of this distribution is the macroscopic speed $V(t,x)$; the standard deviation of the distribution is equal to the square root of the macroscopic speed variance $\Theta(t,x)$, that was earlier identified with an equilibrium value (sections 4.2.1 and A.3.2.4). The probability (4.13) then reads:

$$p_v(v_j|t,x) \equiv N(v_j, \mu_{v_j}, \sigma_{v_j}) \tag{4.14}$$

with:

$$
\begin{aligned}
N\left(y,\mu,\sigma\right) &\equiv \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(y-\mu)^2}{2\sigma^2}} \\
\mu_{v_j} &\equiv V\left(t,x\right) \\
\sigma_{v_j} &\equiv \sqrt{\Theta\left(t,x\right)} \\
&\equiv \sqrt{\Theta^e\left(k\left(t,x\right),V\left(t,x\right)\right)}
\end{aligned}
\tag{4.15}
$$

Note that with this choice the occurrence of negative speeds is theoretically not excluded. However, an appropriate choice of the speed variance can prevent negative speeds from having significant probabilities, e.g. $3\sqrt{\Theta} \le V$ leaves less than 1% probability for speeds <0. When the model is solved numerically, the remaining small error can be corrected by clipping the normal distribution at its physical boundary, and by normalising the total probability to 1 again.

The *equilibrium* speed variance $\Theta(t,x)$ is assumed to be a function of the density and speed only. From the empirical analyses by Helbing (1997) or Hoogendoorn (1999), it appears that for k approaching zero, the standard deviation $\sqrt{\Theta}$ is maximal and in the order of 4 m/*s* and declines monotonically with the density. We adopt the approximation by Shvetsov & Helbing (1999) that defines:

$$
\begin{aligned}
\sigma_v^2 &\equiv \Theta^e\left(k\left(t,x\right),V\left(t,x\right)\right) \\
&\equiv V^2\left(t,x\right)\, B\left(k\left(t,x\right)\right) \\
&\equiv V^2\left(t,x\right)\left(b_0 + b_1\left(1+e^{-\frac{k(t,x)-b_2\, k_{jam}}{b_3\, k_{jam}}}\right)^{-1}\right)
\end{aligned}
\tag{4.16}
$$

From this definition it is obvious that the shape function $B(k)$ is a monotonically increasing and strictly positive function of the density, if the parameters b_0 and b_1 are strictly positive.

Speed distribution of the predecessor

Consistent with the distribution function for v_j, the speeds v_{j-1} for the predecessor are also assumed normally distributed. Here the average of the distribution is the non-local macroscopic speed $V\left(t,x+s_j^0+s_j\right)$, since it should be evaluated at the location $x+s_j^0+s_j$ of the predecessor; the standard deviation equals the square root of the non-local macroscopic equilibrium speed variance:

$$
p_v\left(v_{j-1}\middle|t,x+s_j^0+s_j\right) \equiv N\left(v_{j-1},\mu_{v_{j-1}},\sigma_{v_{j-1}}\right) \tag{4.17}
$$

with:

$$\begin{aligned}\mu_{v_{j-1}} &\equiv V\left(t, x + s_j^0 + s_j\right) \\ \sigma_{v_{j-1}} &\equiv \sqrt{\Theta\left(t, x + s_j^0 + s_j\right)} \\ &\equiv \sqrt{\Theta^e\left(k\left(t, x + s_j^0 + s_j\right), V\left(t, x + s_j^0 + s_j\right)\right)}\end{aligned} \tag{4.18}$$

Distribution of the gap

Before proceeding to the specification of the distribution function for the gap s_j between the follower and the leader, first some gaps and related distances are defined, and it is shown how the average of the gap is related to the density k.

We choose to take the rear bumper of a vehicle as a reference for the longitudinal position x. The following distances can then be defined (Figure 4-3):

- the *headway* h_j is the distance between the own rear-bumper and that of the predecessor;
- the *gross vehicle length* $s_j^0 \equiv$ vehicle length l_j + the static safety margin $s_j^{0\prime}$;
- the *net gap* s_j = the headway h_j – the *gross vehicle length* s_j^0.

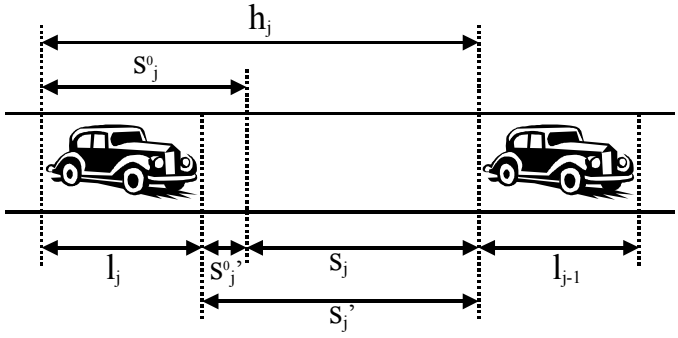


Figure 4-3 Definition of the headway, gross vehicle length and net gap

By definition the following relations hold (Figure 4-3):

$$h_j \equiv s_j^0 + s_j \tag{4.19}$$

When the gross vehicle length is constant for all vehicles, then by definition it is equal to the inverse of the jam density k_{jam}:

$$s_j^0 \equiv \overline{s_j^0} = \frac{1}{k_{jam}} \tag{4.20}$$

The expected value of the headway is by definition:

$$\overline{h} \equiv \frac{1}{k} \tag{4.21}$$

The expected headway h is thus obtained by taking the inverse of the density k. The question now is: at which x-location should the density k be evaluated? This is not trivial, since $k(t,x)$ relates to a specific location x and varies along the x-axis, whereas

the headway h relates to a range of x-values – a range over which the density $k(t,x)$ is – generally speaking – not constant. A consistent choice would be to relate the average headway $\overline{h}$ for a vehicle located at x to the density halfway of the interval $[x, x+\overline{h}]$, which yields the implicit definition:

$$\overline{h}(t,x) \equiv \frac{1}{k\left(t, x+\frac{\overline{h}(t,x)}{2}\right)} \tag{4.22}$$

Solving this implicit definition can be avoided by the following explicit approximation:

$$\overline{h}(t,x) \approx \frac{1}{k\left(t, x+\frac{1}{2\,k(t,x)}\right)} \tag{4.23}$$

The data analysis of individual vehicle passages by Hoogendoorn (1999), shows that the gap distribution can be approximated by a lognormal distribution, with average and variance dependent on the vehicular density. After a thorough study of traffic data at high volumes, Sadeghhosseini & Benekohal (2002) also recommended using a (shifted) lognormal model for modelling the headway distribution for any given flow rate (and thus for the gap distribution, since headway and gap only differ by the speed as a scaling factor). The gaps are therefore assumed to be lognormally distributed. The lognormal (net) gap distribution s_j is then specified:

$$p_s\left(s_j \middle| t,x,v_j\right) \equiv LgN\left(s_j, \mu_{s_j}, \sigma_{s_j}\right) \tag{4.24}$$

with:

$$\begin{aligned} LgN\left(y,\mu_{s_j},\sigma_{s_j}\right) &\equiv \frac{1}{\sigma'_{s_j}\sqrt{2\pi}} e^{-\frac{\left(\ln(y)-\mu'_{s_j}\right)^2}{2\sigma'^{\,2}_{s_j}}} \\ \mu'_{s_j} &\equiv \ln\left(\mu_{s_j}\right) - \frac{1}{2}\ln\left(\frac{\sigma^2_{s_j}}{\mu^2_{s_j}}+1\right) \\ \sigma'_{s_j} &\equiv \sqrt{\ln\left(\frac{\sigma^2_{s_j}}{\mu^2_{s_j}}+1\right)} \end{aligned} \tag{4.25}$$

and:

$$\begin{aligned} \mu_{s_j} &\equiv \overline{h}(t,x) - \overline{s_j^0} \\ &= \frac{1}{k\left(t, x+\frac{1}{2\;\;k(t,x)}\right)} - \frac{1}{k_{jam}} \\ \sigma_{s_j} &\equiv \mu^b_{s_j} \end{aligned} \tag{4.26}$$

In this specification the standard deviation of the gaps is chosen to be a monotonically increasing function of the average gap. This is justified by the following argument. At high densities (and hence small net gaps), the fraction of traffic that is constrained or following is large, yielding small standard deviation of the gaps; at low density or large gaps some vehicles are still constrained with small gaps and others are unconstrained with large gaps, so the standard deviation is large. At intermediate densities, the number of constrained vehicles increases with the density, so the standard deviation decreases monotonically. The exponent in the standard deviation is therefore b > 0 and is normally not larger than 1.

Note that with the specifications in this section, potential correlations have been neglected between the distributions for v_j, s_j, and v_{j-1}, following the vehicular chaos assumption (equation (3.14)). However the term 'chaos' is somewhat pessimistic, since through the macroscopic variables k, V, and Θ, and the non-local relation between the averages for v_j and v_{j-1}, the *averages* of the distributions are always logically coupled, but the *deviations* from the averages are not.

4.2.3.2 Specification of the car-following relationship

Lastly, this section specifies the car-following rule $\dot{v}_j(t) = f_{CF}(v_j, v_{j-1}, s_j)$. Any car-following model can in principle be transformed into a human-kinetic macroscopic variant. The approach is illustrated with the car-following rule described in this section.

In our choice of a car-following model we have not strived for the *best* model. We believe that this selection may vary dependent on the types of vehicles one is modelling, the local driving behaviour and legislation, the ADA system under examination, etcetera. We adopt here the core of the longitudinal model for the microscopic simulators with which we have most experience: a car-following model that is an extension of Helly (1961), and that forms the core of the MIXIC and FOSIM microscopic simulation tools used at our research institutes (Van Arem et al., 1997b; Dijker et al., 1999, Fosim, 2004). The model reads:

$$\dot{v}(v_j, v_{j-1}, s_j) \equiv \min\left(\frac{w_j - v_j}{\tau_w} \quad ; \quad \frac{s_j - s^d(v_j)}{\tau_s} + \frac{v_{j-1} - v_j}{\tau_v} \right) \tag{4.27}$$

with:

$$\begin{aligned} acc_{min} \le \dot{v} \le acc_{max} \\ s^d(v) \equiv s_1^d v + s_2^d v^2 \end{aligned} \tag{4.28}$$

This car-following behaviour states that a driver accelerates towards his desired speed w_j, unless the gap s_j or speed difference with the predecessor urges him to brake. It defines the driver to be in *free flowing* mode if the acceleration to the desired speed is the most restrictive in the minimal operator, and in *constrained* or *car-following* mode otherwise. In the constrained mode, the driver aims at maintaining a desired gap s^d (a function of the speed, defined in equation (4.28)) on the one hand, and at minimising the speed difference with the predecessor on the other hand. In all three cases, errors

with respect to the desired state are controlled using a linear feedback control gain $1/\tau$, so that – if the other terms were zero – each term models an exponential adaptive process with relaxation time τ. Finally the acceleration is bounded within physically realistic limits acc_{min} and acc_{max}.

Note that in general, the desired speed w_j differs among the driver population. Hoogendoorn (1999) suggests a normal distribution for the desired speeds, with an average W and variance σ_w^2 equal to those of the individual speeds at very low density. With equation (4.16) the variance of the desired speed becomes:

$$\sigma_w^2 \approx W^2 \, b_0 \tag{4.29}$$

As a consequence the expected acceleration is obtained through an additional integration of equation (4.11) over the desired speeds:

$$\left\langle \frac{dv}{dt} \right\rangle_v = \int_{w_j} p_w\left(w_j\right) \int_{v_j} p_v\left(v_j\right) \int_{s_j} p_s\left(s_j\right) \int_{v_{j-1}} p_v\left(v_{j-1}\right) \dot{v}_j\left(w_j, v_j, s_j, v_{j-1}\right) dv_{j-1}\, ds_j\, dv_j dw_j \tag{4.30}$$

Since in this thesis a multi user-class version of the human-kinetic model is considered (see chapter 7), the desired speed variance in traffic flow can also be approximated by considering multiple user-classes, each with a different average desired speed, but no within-class variance. For the sake of computational efficiency, all simulations in this thesis neglect the desired speed variance, so that equation (4.11) in stead of (4.30) can be used.

It is important to notice that the car-following relation (4.27) - (4.28) identifies a single desired following distance with every speed (or vice versa). Any deviation of the driving speed or the following distance results in a corrective action by the driver. Note that this assumption is incompatible with the three-phase traffic flow theory of Kerner (sections 2.4.2.2 and 3.2.1). It is left as an interesting subject for further research to analyse whether the empirical observations made by Kerner and co-workers (synchronised flow) could be reproduced in the human-kinetic modelling framework using an alternative car-following relation, similar to that of Kerner & Klenov (2002).

4.3 Completing the basic human-kinetic model: introduction of finite reaction time, anticipation, lane-changing and merging

Because the model specified in the previous sections builds completely on individual microscopic driver behaviour, the model is devoid of typical gas-like behaviour – like the instantaneous speed changes and infinitely small vehicle length, for which the original gas-kinetic based traffic flow models are sometimes criticised. Moreover, it is possible to add some specific aspects of driver behaviour that may influence the dynamics of traffic flow. Following the specification of driver behaviour in section 4.1.2, *finite reaction times* and explicit *anticipation* behaviour are now introduced. *Lane changing* is not considered in the model so far. We here show briefly how lane changes can be modelled, but do not further detail this aspect within this thesis. Finally,

this section specifies how merging traffic at a motorway entry is modelled, so that realistic scenario's can be modelled.

4.3.1 Finite reaction times

It is well known that a time delay T exists between the perception of stimuli by the driver and his response, the control action, which in the car-following theory is an acceleration. It is custom practice in microscopic traffic modelling to explicitly account for this time delay in the car-following rule (see section 3.2.1). A typical car-following rule can therefore in general be written as:

$$control\ action(t) = sensitivity \cdot stimulus(t - T) \tag{4.31}$$

or equivalently:

$$control\ action(t + T) = sensitivity \cdot stimulus(t) \tag{4.32}$$

In the case of equation (4.11), this means that it has to be rewritten as:

$$\left\langle \frac{dv}{dt}\left(t+T, x\left(t+T\right)\right)\right\rangle_v = \int_{v_j} p_v \int_{s_j} p_s \int_{v_{j-1}} p_v \dot{v}_j\left(v_j, s_j, v_{j-1}\right) dv_{j-1}\, ds_j\, dv_j \tag{4.33}$$

with:

$$\begin{aligned} p_v &= p_v\left(v_j \middle| t, x(t)\right) \\ p_s &= p_s\left(s_j \middle| t, x(t), v_j\right) \\ p_v &= p_v\left(v_{j-1} \middle| t, x(t) + s_j^0 + s_j, v_j, s_j\right) \end{aligned} \tag{4.34}$$

Equations (4.33) and (4.34) say that the acceleration integral calculated at time t and location x act upon the traffic flow only T seconds later at location $x(t+T)$. Although in reality the reaction time of drivers is statistically distributed, it is approximated here as a fixed value for the entire driver population. In a linear approximation (justified since T is small), the location $x(t+T)$ can be written as:

$$x(t+T) \approx x(t) + T \cdot V(t, x(t)) \tag{4.35}$$

One option now is to approximate equation (4.33) using Taylor approximations, and rewrite it in terms of variables in t and x only. This is the way in which the models of Payne (1971) and Zhang (2003b) are derived, and the approximation we use in the linear stability analysis of section 5.2.1. It is preferable however, to use as little approximations as possible. No approximations are needed for the numerical evaluation of (4.33) if *non-temporality* is introduced in the numerical solution approach. This means that upon updating the numerical solution to the macroscopic traffic flow equations at time t and location x, the previously calculated (and temporarily stored) solution to equation (4.33) at time t-T and at location $x(t$-$T)$ is used.

The effect on model behaviour of taking the reaction time into account is discussed in the stability analysis of section 5.2.1.

4.3.2 Explicit anticipation behaviour

Following the specification of driver behaviour in section 4.1.2, we look for a way of modelling anticipation behaviour of drivers explicitly. Prior to elaborating the model specifications, the next section first discusses briefly the essence of anticipation behaviour.

4.3.2.1 Driver anticipation

In this thesis distinction is made between *non-local behaviour* and *anticipation behaviour*. Non-local behaviour was defined in sections 4.2.2 and 4.2.3. It refers to the fact that drivers behave according to a stimulus-response control law, with some of the stimuli perceived at a different location than the own. Anticipation behaviour refers to the fact that drivers do not only respond to the actually perceived stimuli, but also to the expected evolution of these stimuli in the near future. They can do so because humans have the ability to predict the behaviour of their immediate predecessor (e.g. speed change, lane change), based on the state of traffic further downstream of the predecessor. Figure 4-4 shows that different information in the flow is used for the direct (mixed local and non-local) stimulus-response behaviour, compared to information used for anticipation[10].

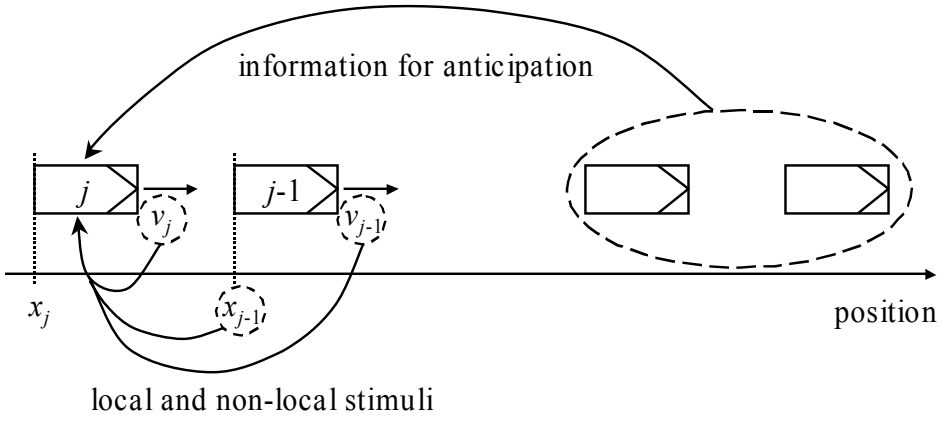


Figure 4-4 Non-local behaviour versus anticipation behaviour

The rationale behind the specification of anticipation in the human-kinetic model is the following. Since the main stimulus in our model is the speed of the predecessor – who will be obliged to adapt his speed to conditions further downstream – we identify the 'expected' evolution in time of the stimulus with the macroscopic trend in space of the average speed. A driver is assumed to respond to the *anticipated* speed of the predecessor, rather than to the actual speed of the predecessor. The anticipated speed is

[10] It is worth noting that – using the definitions given here – the so-called 'anticipation' term in the model of Payne (1971, see section 3.3.2) would be identified with non-local behaviour rather than with anticipation behaviour, because this term is the macroscopic equivalent of a non-local stimulus in the underlying microscopic model.

then equal to the actual speed, adjusted more or less (dependent on the *anticipation strength*) with the trend in the macroscopic speeds, as perceived within a reasonable range downstream of the current position.

Note that anticipation only applies for an anticipated deceleration or negative trend of the speeds. Only there does the driver have the freedom to possibly decelerate stronger or earlier than his predecessor. In accelerating traffic the driver may want to respond to conditions further downstream (where speeds are higher), but the physical presence of the predecessor prevents him from doing so *before* the predecessor has. This is why the anticipation model in the human-kinetic traffic flow model is specified *asymmetrically* with respect to deceleration or acceleration.

4.3.2.2 Specification of anticipation behaviour

So far, the speed v_{j-1} of the predecessor of a vehicle j at position x was modelled as a Gaussian distributed speed with average $V\left(t, x + s_j^0 + s_j\right)$ (section 4.2.3.1). Physically this is the correct transcription of the microscopic car-following relation (4.27). With this specification drivers are assumed to be aware of the exact spatial evolution of the speeds downstream, up to the position of the immediate predecessor only. As a consequence in busy traffic, when the probability of small gaps is high and thus the predecessor is likely to be close ahead, drivers would not be aware of the evolution of speeds further downstream, for instance sharp speed drops.

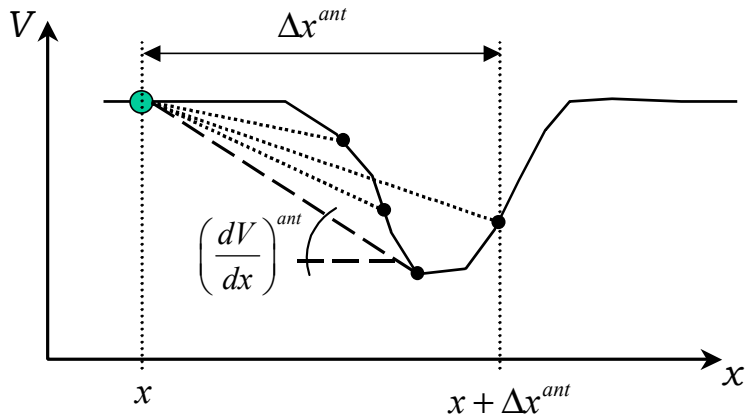


Figure 4-5 Illustration of the anticipation definition, based on preview of the macroscopic spatial speed gradient

The new element in the specification of anticipation is the assumption that drivers respond to the *anticipated* non-local speed of the predecessor, rather than to the actual speed. This is specified by replacing the physically determined average speed $\mu_{v_{j-1}} = V\left(t, x + s_j^0 + s_j\right)$ in equation (4.17) with an *anticipated* non-local speed, defined by:

$$\begin{aligned}\mu_{v_{j-1}} &\equiv V^{ant}\left(t, x + s_j^0 + s_j\right) \\ &\equiv V\left(t,x\right) + f^{ant} \cdot \left(s_j^0 + s_j\right) \left(\frac{dV}{dx}\right)^{ant}\end{aligned} \tag{4.36}$$

with:

$$\left(\frac{dV}{dx}\right)^{ant}(t,x) \equiv \min_{\Delta x \in (0, \Delta x^{ant}]} \left(\frac{V\left(t, x+\Delta x\right) - V\left(t,x\right)}{\Delta x}\right) \tag{4.37}$$

Equation (4.36) expresses that the anticipated non-local speed depends on the anticipated instead of the physical spatial gradient of the spatial speed profile and on a factor f^{ant}, the anticipation strength. The *anticipated spatial speed gradient* is equal to or lower than the physical speed gradient. If drivers can estimate speeds in a range Δx_{ant} downstream, the anticipated gradient of the spatial speed profile $\left(\frac{dV}{dx}\right)^{ant}$ is then defined as the minimum physical speed gradient in this range (Figure 4-5).

The minimum operator in equation (4.37) makes the anticipation behaviour asymmetrically sensitivity for accelerating or decelerating spatial speed profiles. If speeds are lower at the maximum anticipation range than immediately in front, then the driver will anticipate to this trend. If speeds far ahead are higher but the immediate predecessors have not yet accelerated, then neither will do the driver, since the minimum speed gradient is then the one over the nearby range $\Delta x << \Delta x^{ant}$. If $s_j^0 + s_j > \Delta x^{ant}$ then the driver does not perceive the actual speeds at that distance. He is – in other words – only aware of speed gradients within his anticipation range. Moreover, with $\Delta x^{ant} \rightarrow 0$ we obtain the local speed gradient and the perceived speed converges to the actual (nearby) speed gradient (albeit the *forward* difference, so that the anisotropy property is not violated in the case of shocks!).

4.3.3 Lane changing and merging

Two types of lane changing are important: free lane changing and mandatory lane changing. The former occurs when a faster vehicle meets a slower vehicle on its lane and sees opportunity to change lanes to a faster lane. The latter occurs when the current travelling lane will soon be unavailable for the driver, for instance because the lane ends within some distance downstream or because the lane does not lead to the desired destination while another lane does.

Although no multi-lane motorways are considered in the case studies in this thesis – and therefore the free lane change specification is not used in any simulation results shown - for the sake of completeness, section 4.3.3.1 shows briefly how the classical free lane change modelling of kinetic models can be applied and refined, without however detailing the free lane change specification. Section 4.3.3.2 discusses the only type of mandatory lane change that is used in this thesis: merging of traffic at on-ramps onto a

motorway. Also this behaviour is, for the time being, restricted to single-lane traffic, that is: merging onto a one-lane motorway.

4.3.3.1 Free lane changing

Although free lane changing by drivers is not discussed in much detail in this thesis, this section is an indication of how more or less detailed specifications for free lane changing can be established.

The behavioural assumptions of the simplified free lane changing model suggested in this section, are borrowed from the traditional way of modelling lane changes in kinetic traffic flow models, like for instance proposed by Prigogine (1961). It is assumed that if a faster vehicle is impeded by a slower vehicle (i.e. it would have to brake if it remains on the current lane), it has two possibilities: either it is able to change lanes, or it adapts its speed according to the car-following relation. If it is able to change lanes, it continues its journey with the same speed at which it was driving, so that $\dot{v} = 0$ in that case. The way in which the overtaking probability $p_{overtake}$ is defined can yield a variety of different, more or less refined models.

In terms of model equations, the free lane change model affects the calculation of the individual acceleration according to the car-following model of the acceleration integral in the speed equation. Remember that according to equation (4.27), the driver is either in the free flowing mode or in car-following mode, dependent on which mode yields the minimal acceleration. Equation (4.27) is rewriten as follows:

$$\dot{v}\left(v_j, v_{j-1}, s_j\right) \equiv \min\left(\dot{v}_{free}\ ;\ \dot{v}_{cf}\right) \tag{4.38}$$

with

$$\begin{aligned} \dot{v}_{free} &\equiv \frac{w_j - v_j}{\tau_w} \\ \dot{v}_{cf} &\equiv \frac{s_j - s^d\left(v_j\right)}{\tau_s} + \frac{v_{j-1} - v_j}{\tau_v} \end{aligned} \tag{4.39}$$

A driver only considers lane changing if he is in car-following mode *and* if he has to brake (that is: if he is able to accelerate along with the predecessor he does not consider lane changing). If the driver desires to change lanes, then with the probability of immediate overtaking $p_{overtake}$, he is able to do so, and the acceleration $\dot{v}_{cf}$ is reset to zero. In the other case, with probability $(1 - p_{overtake})$, a lane change is not possible, and the acceleration $\dot{v}_{cf}$ takes the value according to equation (4.39). These specifications can be summarised by modifying equation (4.38) as follows:

$$\dot{v}\left(v_j, v_{j-1}, s_j\right) \equiv \min\left(\dot{v}_{free}\ ;\ \max\left(\dot{v}_{cf}\,;\,\left(1 - p_{overtake}\right)\dot{v}_{cf}\right)\right) \tag{4.40}$$

The definitions of the free flow and car-following acceleration remain the same as in equation (4.39).

The free lane-changing model should be completed by a suitable specification of the overtaking probability $p_{overtake}$. We do not proceed into refined model specifications and just mention some options:

- apply a constant value for $p_{overtake}$,
- apply a density-dependent value $p_{overtake}(k)$,
- relate $p_{overtake}$ to the gap distribution: $p_{overtake}$ equals the probability that a gap is present larger than some (possibly speed-dependent) desired safe gap (see section 3.2.2); this probability is easily calculated using the cumulative probability density function for the gaps,
- the same as the previous option in a multi-lane model, but then using the gap distribution in the adjacent lane(s).

Note that with the simple specifications given here, the lane changing is modelled as an instantaneous process. Actually, this is not consistent with the time scale considerations made earlier for the acceleration and braking processes (section 4.1.1). Theoretically, it would be more elegant to consider the finite duration of the lane change in some way. This issue is left for further research. Especially when considering the influence of different free lane changing strategies on traffic flow dynamics or the influence of ADA systems supporting free lane changing, a more detailed and continuous instead of discrete specification of lane changing behaviour is advised.

As mentioned earlier the focus in this thesis is on longitudinal behaviour and free lane changing has not been applied in any of the case studies presented, which is equivalent by taking $p_{overtake}$ in the above mentioned simple specification constant and equal to zero.

4.3.3.2 Merging

So far, no in- or outflow of traffic was considered into the road section under consideration: $\left(\frac{dk}{dt}\right)_{discrete} = 0$. This section specifies the model behaviour with traffic inflow at a motorway on-ramp (Figure 4-6).

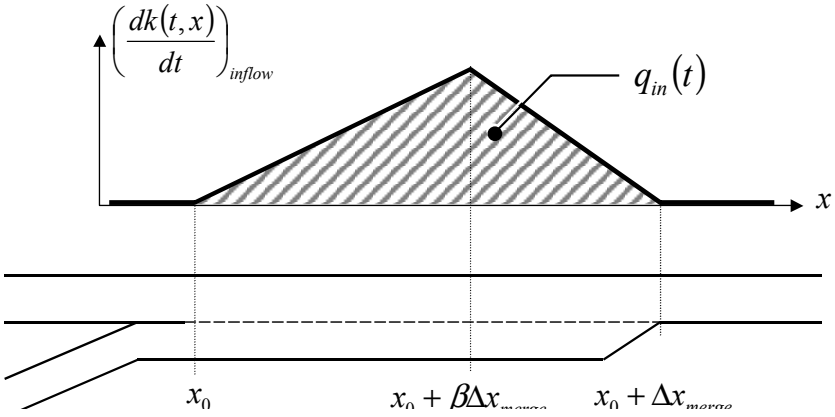


Figure 4-6 Spatial spreading of merging intensity

Figure 4-6 introduces the notation $q_{in}(t)$ for traffic demand per unit time at the on-ramp and Δx_{merge} the length of the slip lane that starts at x_0. The merging intensity is maximal at location $x_0+\beta\Delta x_{merge}$ ($0\leq \beta \leq 1$). Then the merging intensity or inflow as a function of x is a triangular function (Figure 4-6):

$$\left(\frac{dk(t,x)}{dt}\right)_{inflow} \equiv f(x)\ \ q_{in}(t) \tag{4.41}$$

with:

$$f(x) \equiv \begin{cases} \dfrac{x-x_0}{2\beta\Delta x_{merge}^2} & x_0 \leq x < x_0+\beta\Delta x_{merge} \\ \dfrac{\Delta x_{merge}+x_0-x}{2(1-\beta)\Delta x_{merge}^2} & x_0+\beta\Delta x_{merge} \leq x < x_0+\Delta x_{merge} \\ 0 & elsewhere \end{cases} \tag{4.42}$$

In order to avoid that equation (4.41) leads to exceeding of the jam density on the main lanes, $f(x)$ is set equal to zero if the sum of $\left(\frac{dk(t,x)}{dt}\right)_{inflow}$ and k exceeds jam density at any point x. With this specification, merging traffic always enters the motorway (unless there is physically no space). In fact, this is not consistent with traffic legislation – that prescribes that main lane traffic has priority over merging traffic. It is however consistent with observed practical behaviour, where merging traffic takes priority over main lane flow, as is shown by Cassidy & Bertini (1999b) and Mauch (2002).

The fact that $\left(\frac{dk}{dt}\right)_{in/outflow} \neq 0$ has also consequences for the speed equation. In the conservative speed equation (4.4), the 'discrete acceleration 1'-term accounts for the speed momentum that flows into the road. The momentum equation is thus:

$$\frac{\partial kV}{\partial t}+\frac{\partial kV^2}{\partial x}+\frac{\partial k\Theta}{\partial x} = k\left\langle\frac{dv}{dt}\right\rangle_v + \underbrace{\int_v v\left(\frac{dk}{dt}\right)_{inflow} dv}_{discrete\ \ acceleration\ 1} \tag{4.43}$$

The primitive formulation is then:

$$\frac{\partial V}{\partial t}+V\frac{\partial V}{\partial x}+\frac{1}{k}\frac{\partial k\Theta}{\partial x} = \left\langle\frac{dv}{dt}\right\rangle_v + \frac{1}{k}\left(\underbrace{\int_v v\left(\frac{dk}{dt}\right)_{inflow} dv}_{discrete\ \ acceleration\ 1} - \underbrace{V\left(\frac{dk}{dt}\right)_{inflow}}_{discrete\ \ acceleration\ 2}\right) \tag{4.44}$$

In the special case that merging traffic always manages to enter the motorway with its speed adapted perfectly to the average speed on the main lane, the following simplified speed equations are obtained:

$$\frac{\partial kV}{\partial t}+\frac{\partial kV^2}{\partial x}+\frac{\partial k\Theta}{\partial x}=k\left\langle\frac{dv}{dt}\right\rangle_v+V\left(\frac{dk}{dt}\right)_{inflow} \tag{4.45}$$

The primitive formulation in that case remains unchanged with respect to equation (4.7) :

$$\frac{\partial V}{\partial t}+V\frac{\partial V}{\partial x}+\frac{1}{k}\frac{\partial k\Theta}{\partial x}=\left\langle\frac{dv}{dt}\right\rangle_v \tag{4.46}$$

4.4 Numerical evaluation

4.4.1 Numerical evaluation of the acceleration integral

The acceleration integral is approximated in our numerical implementations using the standard method known as the *extended trapezium rule* (e.g. Evans, 1993). Because the acceleration integral is an integral in three variables, this operation is computationally rather expensive, and it would be preferable to utilise more efficient quadrature techniques. However, since computational speed is not the main concern of this thesis and the simple trapezium rule yields sufficiently accurate calculations, we leave efficiency improvements for further research.

4.4.2 Numerical solution of the dynamic system

The system of partial differential equations that define the conservative macroscopic human-kinetic traffic flow model is:

$$\frac{\partial k}{\partial t}+\frac{\partial kV}{\partial x}=\left(\frac{dk}{dt}\right)_{inflow} \tag{4.47}$$

$$\frac{\partial kV}{\partial t}+\frac{\partial k\left(V^2+\Theta^e\right)}{\partial x}=k\left\langle\frac{dv}{dt}\right\rangle_v+\int_v v\left(\frac{d\rho}{dt}\right)_{inflow}dv \tag{4.48}$$

These equations correspond to the general structure of the second order or Payne-type traffic flow models, the numerical evaluation of which is discussed in section A.5. The general structure of this class of models is:

$$\frac{\partial}{\partial t}\mathbf{U}+\frac{\partial}{\partial x}\mathbf{F}(\mathbf{U})=\mathbf{G}(\mathbf{U}) \tag{4.49}$$

In the case of the human-kinetic model:

$$\mathbf{U}\equiv\begin{pmatrix}u_1\\u_2\end{pmatrix}=\begin{pmatrix}k\\kV\end{pmatrix} \tag{4.50}$$

$$\mathbf{F}(\mathbf{U}) \equiv \begin{pmatrix} u_2 \\ \frac{u_2^2}{u_1} + u_1 \quad \Theta^e(u_1, u_2) \end{pmatrix} \tag{4.51}$$

$$\mathbf{G}(\mathbf{U}) \equiv \begin{pmatrix} \left(\frac{du_1}{dt}\right)_{inflow} \\ u_1 \left\langle \frac{dv}{dt} \right\rangle_v + \int_v v \left(\frac{d\rho}{dt}\right)_{inflow} dv \end{pmatrix} \tag{4.52}$$

Given the definition of the equilibrium speed variance of equation (4.16), the eigenvalues of the linearised system (see A.6) are:

$$\lambda_{1,2} = \frac{u_2}{u_1}\left(1 + B(u_1) \pm \sqrt{B^2(u_1) + B(u_1) + u_1 \frac{dB(u_1)}{du_1}}\right) \tag{4.53}$$

Since $B(k)$ is a monotonically increasing and strictly positive function, the eigenvalues are always real. Moreover, none of the eigenvalues is always zero, except maybe for some specific combinations of u_1 and u_2. Therefore, the system is hyperbolic, which means that it has wave-like solutions (Hirsch, 1990), and that the numerical solution schemes for Payne-type models are applicable (see Annex A, section A.5). In this thesis the extended HLLE scheme presented by Ngoduy et al. (2004) is adopted, since the authors concluded convincingly that this scheme is superior to some other well-known schemes for Payne-type models.

4.4.3 The pressure term in the human-kinetic model and its consequence for the numerical evaluation

The pressure term is important from a numerical point of view. In Annex A, the discussion in section A.5.2 shows in general the importance and different physical meanings of the pressure term in various Payne-type models. It turns out that the definition and parameters of the pressure term determine the sign of the eigenvalues of the Jacobian matrix $\frac{\partial \mathbf{F}}{\partial \mathbf{U}}$ of $\mathbf{F}(\mathbf{U})$ in equations (4.49) and (4.51). These eigenvalues play a central role in Godunov-type numerical solution schemes for Payne-type models (like the extended HLLE scheme applied in this thesis). The intent of the current section is to show that the eigenvalues of the human-kinetic model are always positive, which raises specific problems in the numerical solution scheme (section 4.4.3.1). For now, this issue is remedied by making some pragmatic modifications to the numerical scheme (section 4.4.3.2), after which some directions for a more fundamental treatment are suggested (section 4.4.3.3).

4.4.3.1 The nature of the pressure term in the human-kinetic model: pure speed variance pressure and corresponding numerical issues

The pressure term in the speed equation (4.48) of the human-kinetic traffic flow model contains solely the speed variance. It is therefore identified as *pure speed variance*

pressure (in contrast to *non-local* or *anticipation* pressure, terminology introduced for pressure terms in existing models with a different physical origin, see the discussion in section A.5.2, Annex A). As a consequence, the eigenvalues in the human-kinetic traffic flow model are always positive. This can be verified from equation (4.53) (except maybe when $\frac{dB(u_1)}{du_1}$ is very large) and becomes obvious from the approximation of the eigenvalues:

$$\lambda_{1,2} \approx V \pm \sqrt{\Theta^e} \tag{4.54}$$

For Gaussian distributed speeds this always yields positive eigenvalues.

Convection and acceleration at the upstream interface of a jammed cell in the case of positive eigenvalues

Let us investigate the behaviour of the numerical solution scheme in the high-density regime for such model with positive eigenvalues. Consider the numerical discretisation in Figure 4-7 (this is the same case as the one examined for existing Payne-type models in section A.5.2.2, Annex A). The specific situation of Figure 4-7 is characterised by a jammed cell *j+1* (i.e. densely packed traffic at standstill), preceded by a densely occupied but not-yet jammed cell *j*. When the model has positive eigenvalues only, the convective flux into cell *j+1* over cell interface $j+\frac{1}{2}$ is exclusively determined by the traffic state upstream of this interface (see also section A.5.2.2, Annex A), i.e. by cell *j*. However, traffic in that cell is still flowing, since it is not yet jammed. Therefore, so will traffic at the cell interface $j+\frac{1}{2}$. The conclusion is that the flux across this interface and into the jammed cell is positive and the jammed cell is further filled due to convection of traffic from upstream.

Convection may have this undesired effect *unless* the flow *inside* the upstream cell *j* itself stops, and as a result also convection across the interface $j+\frac{1}{2}$. In the human-kinetic model, the non-local interaction and anticipation (modelled in the acceleration integral in the RHS or *numerical source term* of the speed (momentum) equation) have this function: in the case of Figure 4-7, the acceleration integral for cell *j* yields a braking response for traffic approaching the jammed cell (given that the car-following model that forms the core of the acceleration integral prevents rear-end collisions). This should prevent the jammed cell from being further filled. However, as will be shown in the next example, this interplay between convection and acceleration (in the numerical source term) is *not* modelled correctly in the *discrete* solution when using the widely applied Godunov-type numerical schemes for Payne-type models (with a fixed spatial discretisation Δx and a fixed time increment Δt).

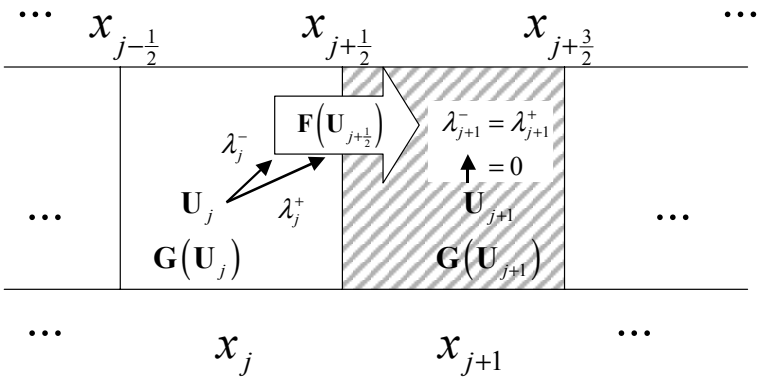


Figure 4-7 Determination of numerical fluxes at the boundary between not-jammed and jammed cells

Continuous human-kinetic model solution for approaching jammed traffic

Let us examine further the example of Figure 4-7. The downstream cell j+1 is homogeneously jammed, with traffic standing still at jam density; the upstream cell j is flowing. If the car-following model of the acceleration integral is indeed collision-free, then the speed and average acceleration in the upstream cell j are such, that vehicles halt at last just before the interface with the jammed cell. Therefore, according to the continuous version of the human-kinetic model, the outflow of the upstream cell j is zero (the inflow is irrelevant for the discussion here). Let us suppose that vehicles enter cell j with the maximum speed v_0 that allows stopping just before the interface with the jammed cell with a *constant* deceleration d. The following kinematic relations then hold between entry speed v_0 and the distance Δx travelled during this deceleration manoeuvre (that takes $\Delta t'$ seconds to complete):

$$0 = v_0 - d\Delta t' \tag{4.55}$$

$$\Delta x = v_0 \Delta t' - \frac{d}{2}\Delta t'^2 = \frac{v_0^2}{2d} \tag{4.56}$$

For the upstream cell j, the *average* deceleration is equal to the constant *individual* acceleration $-d$ and the average speed V_j equals ½ v_0.

Discrete human-kinetic model solution for approaching jammed traffic

Let us now investigate how the discrete solution scheme would handle this situation. Suppose that the calculation of the average acceleration is exact, so $\left\langle \frac{dv}{dt}\left(x_j\right)\right\rangle_v = -d$.

Because the eigenvalues are positive, the fluxes over the upstream and downstream interfaces of the jammed cell j+1 are (only the density fluxes are shown here):

$$\begin{aligned} F_1\left(k_{j+\frac{1}{2}}\right) &= k_j V_j > 0 \\ F_1\left(k_{j+\frac{3}{2}}\right) &= k_{j+1} V_{j+1} = k_{jam}\, 0 = 0 \end{aligned} \tag{4.57}$$

Balancing the net inflow $F_1\left(k_{j+\frac{3}{2}}\right)-F_1\left(k_{j+\frac{1}{2}}\right)$, the conclusion is that the jammed cell is further filled with traffic. This is physically incorrect and not consistent with the continuous solution just derived.

Moreover, this incorrect solution lasts until the speed V_j in the upstream cell drops to zero. However, the discussion in the remainder of this section shows that this can not happen immediately.

The speed in the upstream cell j is adapted by two processes: (i) traffic energy flows into the cell via the upstream interface $j-\frac{1}{2}$ on the one hand (convection) and (ii) traffic inside the cell decelerates with deceleration d (numerical source term). We do not detail the convective energy flow; it is sufficient here to point out that in a decelerating speed profile, speeds in upstream cells are higher than in downstream cells, so that the energy convection turns V_{j-1} into V'_{j-1} (according to equation (A.56), Annex A), and the following inequality holds:

$$V'_{j-1} \geq V_{j-1} = \tfrac{1}{2} v_0 \tag{4.58}$$

Now consider the acceleration inside cell j. The speed is further adapted according to equation (A.57) in Annex A:

$$V''_{j-1} = V'_{j-1} - d\Delta t \tag{4.59}$$

This equation reveals that the speed in the upstream cell can only immediately drop to zero (i.e. within one time step Δt) if:

$$\Delta t = \frac{V'_{j-1}}{d} \tag{4.60}$$

However, the fixed time increment Δt has been chosen such, that numerical stability is guaranteed. The latter is only possible if no vehicle – regardless of its speed – can traverse a cell with size Δx within one time step Δt, which is the case if:

$$\Delta t < \frac{\Delta x}{\max(V)} = \frac{\Delta x}{\max(w)} \tag{4.61}$$

with $\max(w)$ the maximum desired speed of a driver. This is clearly incompatible with equations (4.56), (4.58) and (4.60) since it would yield:

$$\begin{aligned} \frac{V'_{j-1}}{d} &< \frac{\Delta x}{\max(w)} \\ \frac{V'_{j-1}}{d} &< \frac{v_0^2}{2d\max(w)} \\ \max(w) &< \frac{v_0^2}{2V'_{j-1}} \leq v_0 \end{aligned} \tag{4.62}$$

The latter inconsistency demonstrates that the speed in cell *j* can impossibly drop to zero within one time step.

Conclusion

The analyses in this section have shown that for any Payne-type model with only positive eigenvalues (and hence for our human-kinetic model), and using widely applied Godunov-type numerical solution schemes for Payne-type models:

(a) the outflow of a cell preceding a jammed cell is modelled incorrectly in the *discretised* numerical model, since the numerical scheme assumes a non-zero outflow for cases in which traffic in reality decelerates until standstill before reaching the downstream interface of that cell;

(b) this error is sustained for at least more than one time step, since constraints to the cell size and time increment prevent that the acceleration integral calculated inside the upstream cell reduces the speed to zero within one time step.

The interested reader can verify in section A.5.2.2, Annex A that existing Payne-type models have both positive and negative eigenvalues, so that for such models the flow at cell interfaces in congestion is determined by both adjacent cells, thus preventing the issue of 'numerically' over-filling jammed cells. It is said that in such models *information* flows *convectively* in both upstream and downstream directions, according to speeds corresponding to the negative and positive eigenvalues respectively[11].

The next sections propose a pragmatic solution to this numerical issue in the human-kinetic traffic flow model.

4.4.3.2 Modifications to the numerical scheme of the human-kinetic model

Due to the positive eigenvalues in the human-kinetic model, some modifications to the numerical solution scheme are needed to avoid the problems mentioned in the previous section. Since the aim of this thesis is primarily to refine the behavioural models underlying the interactions between vehicles in kinetic models, and not to refine numerical solution schemes, some pragmatic modifications to standard solution schemes for Payne-type models are proposed that, for now, solve the issue. The main concern is to avoid making concessions to the behavioural model for the sake of easy numerical implementation. Therefore, since we are facing a numerical problem, this problem is circumvented by changes to the numerical scheme only, and the search for theoretically better solution schemes (for instance the suggestions in the next section) is left as an issue for further research.

Two modifications are made to the extended HLLE scheme presented by Ngoduy et al. (2004). Firstly, the order of calculation of convective and source terms is changed,

[11] Note that this *not* mean that the human-kinetic model does not account for negatively directed information flows. However, it has no negatively directed flow in the *convective* part of the model (the LHS of the flow equations); negative information flow – which also occurs in reality and should therefore be modelled in one way or another – is accounted for in the source term in the RHS (acceleration integral).

which reduces the problem, but does not avoid it entirely. Secondly, the calculation of the eigenvalues is modified so as to artificially force backwardly directed convective information flow, albeit in the limit of high-density traffic only.

Changed order of calculation

The rationale behind this first modification is the following. With the original order of calculation in Godunov-type solution schemes (described in section A.5.2.2, Annex A), the acceleration of traffic in cell *j*, calculated in the acceleration integral of term $\mathbf{G}\left(\mathbf{U}_j\right)$, is only partially imposed on the vehicles in cell *j*, since it allows traffic to flow out of the cell *before* the variables $\mathbf{U}_j$ are actually changed according to the source term **G**. Let us again illustrate this with an example.

Consider the speed and flow in two consecutive cells *j* and *j+1*. Suppose that the initial speed is the same in both cells and that the acceleration term in **G** is such that in cell *j* the decision is to brake until standstill, whereas in cell *j+1* drivers decide to not change their speeds. In Figure 4-8 the effect of the original order of calculation (first convection, then source terms) is schematically presented. It appears that a part of the drivers who were originally in cell *j* and decided to brake until standstill, has left the cell and entered cell *j+1 before* the braking has started (due to first accounting for convection). In this new cell these drivers are imposed the behaviour of that cell (when the source term is processed), which is to drive on without braking. This part of traffic thus escapes unaffected by its own decision to brake.

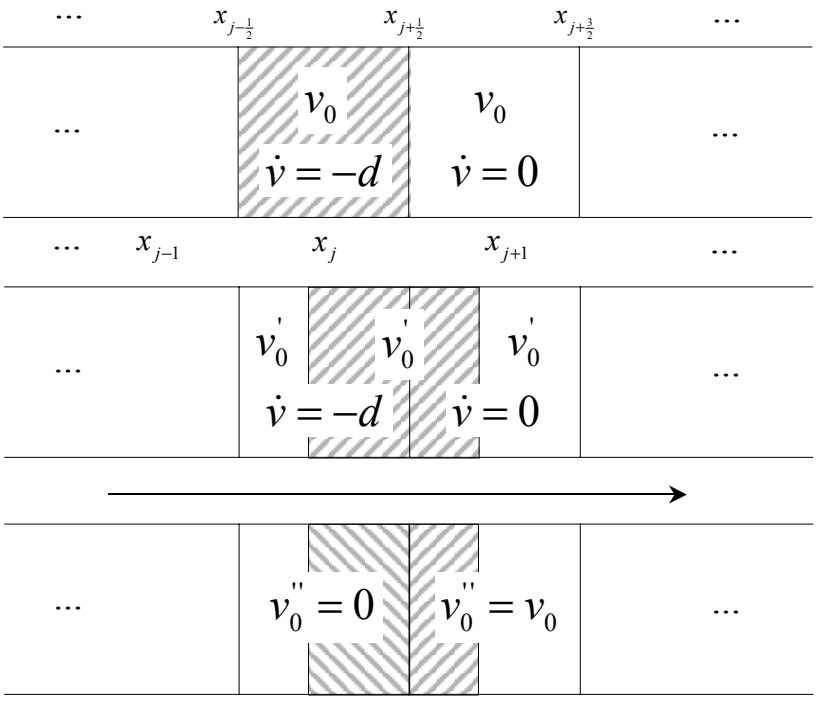


Figure 4-8 Effect of the traditional order of calculation for the numerical source term $\mathbf{G}\left(\mathbf{U}_j\right)$

One might decide to invert the order of calculation, and first impose the accelerations calculated in **G** *before* convectively moving traffic. However, this neglects the fact that during the finite duration of the acceleration process vehicles travel some finite distance. We therefore propose the following order of calculation:

1. Account *partially* (for a fraction $p_{prior} \in [0,1]$) for the change of $\mathbf{U}_j$ during the time interval Δt due to internal and external forces contained in the source term $\mathbf{G}\left(\mathbf{U}_j\right)$:

$$\mathbf{U}'_j \equiv \mathbf{U}_j + p_{prior} \Delta t \, \mathbf{G}\left(\mathbf{U}_j\right) \tag{4.63}$$

2. Dependent on the partially changed conditions $\mathbf{U}'_{j-1}$, $\mathbf{U}'_j$, and $\mathbf{U}'_{j+1}$ in the cell x_j under consideration, and in the neighbouring cells x_{j-1} and x_{j+1}, determine the numerical fluxes over the boundaries $\mathbf{F}\left(\mathbf{U}'_{j-\frac{1}{2}}\right)$ and $\mathbf{F}\left(\mathbf{U}'_{j+\frac{1}{2}}\right)$.

3. The change of $\mathbf{U}'_j$ during the time interval Δt due to numerical fluxes is then:

$$\mathbf{U}''_j \equiv \mathbf{U}'_j - \Delta t \frac{\mathbf{F}\left(\mathbf{U}'_{j+\frac{1}{2}}\right) - \mathbf{F}\left(\mathbf{U}'_{j-\frac{1}{2}}\right)}{\Delta x} \tag{4.64}$$

4. Finally further adapt $\mathbf{U}''_j$ by superimposing the unused fraction $(1-p_{prior})$ of the source term $\mathbf{G}\left(\mathbf{U}_j\right)$ on the previous result:

$$\mathbf{U}'''_j \equiv \mathbf{U}''_j + \left(1 - p_{prior}\right) \Delta t \, \mathbf{G}\left(\mathbf{U}_j\right) \tag{4.65}$$

Theoretically the best value for p_{prior} is ½. To understand this, consider the distance Δs travelled during Δt. From basic kinematics it appears that:

$$\Delta s = v \, \Delta t + \dot{v} \frac{\Delta t^2}{2} \tag{4.66}$$

In the new order of calculation the speed v is first changed into a speed v':

$$v' = v + p_{prior} \, \dot{v} \, \Delta t \tag{4.67}$$

after which traffic is convectively moved:

$$\begin{aligned} \Delta s &= v' \, \Delta t \\ &= v \, \Delta t + \dot{v} \, p_{prior} \Delta t^2 \end{aligned} \tag{4.68}$$

The latter result is equal to equation (4.66) only in the case of p_{prior} = ½. However, keeping in mind the overestimation of the numerical fluxes discussed in section 4.4.3.1, choosing a larger value of p_{prior} might be justified. Note that this procedure has some resemblance with the second order numerical scheme of MacCormack (Hirsch, 1990) and other predictor-corrector schemes, but that we avoid evaluating the source terms $\mathbf{G}\left(\mathbf{U}_j\right)$ twice, which – in the case of the human-kinetic model – is a numerically

expensive operation due to the necessity to numerically evaluate the acceleration integral.

Forced inversion of information flow

The analysis in section 4.4.3.1 has shown that in a model with only positive eigenvalues, blocking the outflow of a cell j by a sufficiently large deceleration inside that cell (so that vehicles would stand still before reaching the downstream cell interface) can never be achieved within one time step Δt in a Godunov-type numerical procedure with fixed time increment Δt and cell size Δx. Briefly speaking, the reason is that the flow in the cell centre is assumed to be homogeneously valid over the entire cell, hence also over the downstream interface, whereas in reality the flow varies with x over the dimension of the cell. The change of order of calculation just proposed can therefore *reduce* the numerical issues discussed in 4.4.3.1 (because the flow is (partially) reduced *before* instead of *after* convectively moving traffic), but cannot *entirely* solve the issue.

A second pragmatic modification to the numerical scheme is therefore proposed. The rationale behind this modification is the following: as the occurrence of a negative eigenvalue in Payne-type models prevents the problem described in section 4.4.3.1 (see section A.5.2.2, Annex A), a negative eigenvalue is artificially introduced in slowly moving traffic (i.e. only then when the problem would otherwise occur).

In the calculation of the eigenvalues on the cell interfaces, a constant contribution c_0^2 is added to the speed variance that is used to calculate the smallest eigenvalue (and *only* there). This means that the eigenvalues are calculated with the following equations instead of equation (4.53):

$$\lambda_1 = \frac{u_2}{u_1}\left(1 + B(u_1) + \sqrt{B^2(u_1) + B(u_1) + u_1 \frac{dB(u_1)}{du_1}}\right) \tag{4.69}$$

$$\lambda_2 = \frac{u_2}{u_1}\left(1 + B(u_1) - \sqrt{B^2(u_1) + B(u_1) + u_1 \frac{dB(u_1)}{du_1} + \frac{u_1^2}{u_2^2} c_0^2}\right) \tag{4.70}$$

Note that the largest eigenvalue remains unchanged by this operation. Therefore, we do *not* introduce information flow faster than the fastest vehicle in our model (a flaw that is present in existing approaches described in section A.5.2.3 of Annex A, and that was already criticised for Payne-type models (Daganzo, 1995a)). Moreover, the extra contribution c_0^2 is *not* added to the speed variance when calculating the *physical* flux over the cell interfaces according to equation (4.51). Therefore, this correction has no influence on the energy flux between cells, a flaw that is present in the existing approaches by Klar & Wegener (1999) and by Helbing (1997). Their correction term appears in the convective model as a *non-local* traffic pressure, and therefore contributes to the *physical* flux of energy, although the non-local component of the traffic pressure only represents the flow of *information* (see section A.5.2.3 of Annex A).

Finally let us remark that the artificial contribution c_0^2 is added only to force the 'freezing' of flow in the case of low speed conditions. Therefore, a correction c_0 in the order of 3 to 4 m/*s* is sufficient. This means that the correction only has an effect if the vehicular speed has already dropped below c_0 due to anticipation and braking as specified in sections 4.2 and 4.3.

Conclusion

We conclude that, within the context of this thesis, the extended HLLE scheme presented by Ngoduy et al. (2004) with a fixed time and spatial discretisation can be used. However, in order to avoid purely numerical problems that occur in this and similar solution schemes because of the non-negativity of the eigenvalues of the human-kinetic model, two pragmatic modifications were introduced:

(a) change the order of calculation of convective flow and source terms;

(b) add a constant (small) contribution to the calculation of the smallest eigenvalue of the human-kinetic model (and *only* there) to force the inversion of information flow in the model during the evaluation of the convective terms in low speed conditions (near standstill).

It is important to realise that the latter correction purely fixes a *numerical deficiency* by a *numerical modification*; the real information flow for anticipation and non-local (or anisotropic) behaviour is accounted for elsewhere (in the acceleration integral in the source term) and is *unaffected* by the modification.

4.4.3.3 Some suggestions for further research on the numerical solution of the human-kinetic traffic flow model

Further research is needed to replace the pragmatic approach of the previous section by a more fundamentally justified numerical solution. One option might be to use so-called *staggered grid* discretisations in combination with numerical solution schemes that treat convective and source-terms *simultaneously*. Some examples of such schemes can be found in Versteeg & Malalasekera (1995) and other specialised literature on computational fluid dynamics.

The principle of a staggered grid is that scalar variables (like the density k and the activation level A (see chapter 6) in the case of the human-kinetic model) and speed components of the flow are evaluated on different discretisation points (see Figure 4-9).

This has the advantage over a traditional grid that the speed values in the approximation:

$$\left(\frac{\partial kV}{\partial x}\right)_j \approx \frac{k_{j+\frac{1}{2}}V_{j+\frac{1}{2}} - k_{j-\frac{1}{2}}V_{j-\frac{1}{2}}}{\Delta x} \tag{4.71}$$

for the convective flux gradient of the continuity equation (4.47) for the density k_j in the calculation cell j around x_j need *not* to be approximated, since the speeds are already defined on the cell interfaces $j-\frac{1}{2}$ and $j+\frac{1}{2}$.

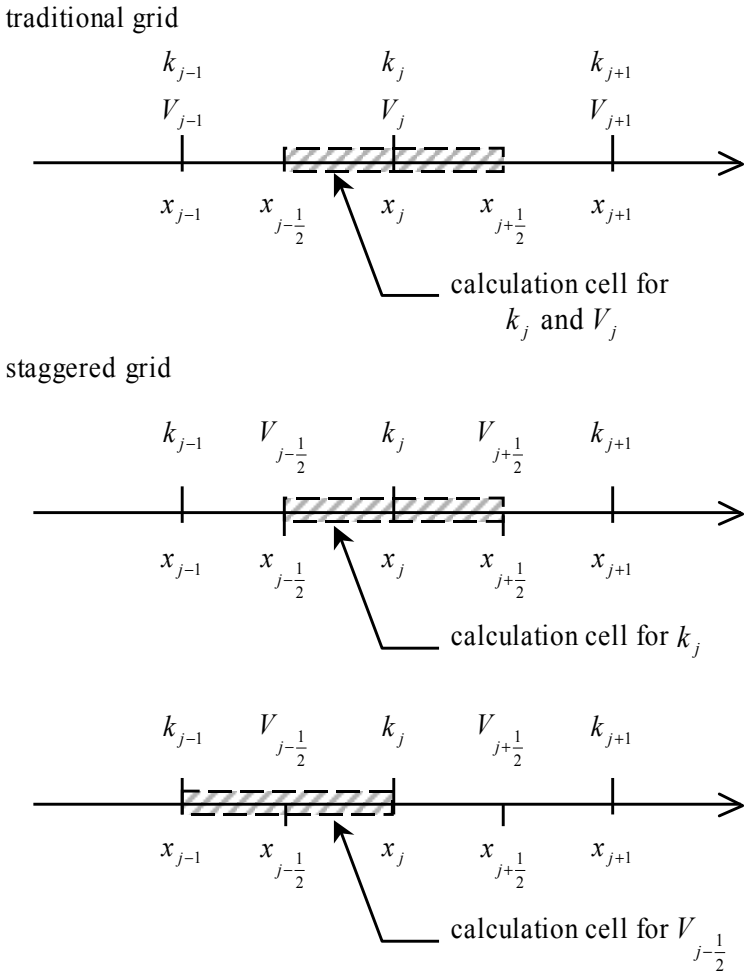


Figure 4-9 Traditional versus staggered discretisation grid and the definition of the calculation cells

This staggered discretisation would solve the issue of section 4.4.3.1 in the following way. Let us recall the example of the jammed cell (Figure 4-7). The speed at the upstream boundary of the jammed cell would approach zero (since due to anisotropy of the acceleration integral only information of the jammed cell is used to calculate the acceleration behaviour, and traffic in that cell is standing still). Therefore, also the numerical fluxes over this interface into the jammed cell would be zero. Notice that on the other hand, the densities on the cell interfaces still need to be approximated, but this introduces smaller errors, especially in nearly jammed traffic, since multiplication with the speed (that would approach zero in those conditions) freezes flow anyway.

Approximations would also be needed for the speed values on the interfaces of the staggered calculation cells for the speed (see Figure 4-9), for instance V_j and $V_{j\text{-}1}$ in the calculation of the energy flux over the interfaces j-1 and j of the calculation cell $j-\frac{1}{2}$ for $V_{j-\frac{1}{2}}$. One suggestion would be to approximate these speeds according to the kinematic law of uniformly accelerated motion applied between staggered grid points, which would yield:

$$V_j \approx \sqrt{V_{j-\frac{1}{2}}^2 + 2\left\langle \frac{dv}{dt} \right\rangle_{j-\frac{1}{2}} \left(x_j - x_{j-\frac{1}{2}} \right)} \tag{4.72}$$

This would solve the issue of the order of calculation of convective and source terms (Figure 4-8), since the numerical approximation of convection now implicitly accounts for simultaneous acceleration and convection. However, also other more refined approaches are available in computational fluid dynamics that treat the issue of simultaneous calculation of convection and source terms in a correct way, for instance the PISO scheme (Issa, 1986).

More research is needed, however, to elaborate the techniques suggested in this section for macroscopic traffic flow models – including the human-kinetic model. For instance, the potential occurrence of shocks (or at least steep density and speed gradients) in traffic flow requires special care in the numerical solution, as well as the definition of boundary conditions on staggered grids. For now, we leave this issue for further research.

4.5 Conclusions

In this chapter the fundamentals of the human-kinetic traffic flow models were described and motivated. The behavioural model that is used throughout this thesis was summarised, after which the mathematical basics were outlined. The starting points are the terminology and generalised continuity equation of traditional kinetic theory. However, motivated by a discussion on time scales, no distinction is made between the mathematical treatment of acceleration and deceleration, as is traditionally done in kinetic traffic flow models. The new approach consists of modelling acceleration and deceleration together as continuous processes in the so-called *acceleration integral*.

The core of the human-kinetic traffic flow model, the acceleration integral, links the microscopic and macroscopic levels of description together, through consideration of a car-following relationship of individual driver behaviour with stochastically distributed input variables: speeds of the follower and predecessor, and the gap between them. The individual gap distribution is related to the non-local macroscopic traffic density shortly downstream of the follower. The speed distributions of follower and leader are derived from local and non-local macroscopic variables: the average speed and speed variance. The latter macro variable is considered to be an equilibrium function of the speed and density, so that the macroscopic equivalent of the human-kinetic traffic flow model is characterised mathematically as a *second order* model.

Due to the micro-macro link, further refinements could be made to the model: delayed reaction by drivers, anticipation to nearby future evolution of the predecessor's speed, a simple lane changing model, and a merging model near on-ramps are introduced. Eventually this leads to a model with the following desirable properties:

- Adaptive driver behaviour as a function of traffic flow conditions, through the stimulus-response structure of the underlying micro model

- Anisotropy of driver behaviour: drivers are only influenced by stimuli coming from downstream
- Non-locality of stimuli: drivers respond to conditions at a variable distance downstream, depending on the driving speed
- Finite space requirements: vehicles are not modelled as point-sized particles, but require a finite space that is a function of driving speed
- Limited acceleration and deceleration capability of vehicles
- In theory, the approach can account for all relevant correlations between speeds of predecessors and followers, and the gap between them.

Albeit not the main issue in this thesis, the numerical solution of the human-kinetic traffic flow model equations is important. It turns out that, although the system of partial differential equations is hyperbolic as are most existing macroscopic traffic models, its pressure term differs from other models. The treatment of non-local aspects and anticipation in the acceleration integral (which is mathematically a *source term*), instead of isolating these from the source term and treating them together with the convective part of the equation, causes flaws in the numerical solutions, when widely applied Godunov-type solution schemes would be used. Some pragmatic remedies for this issue are proposed, as well as suggestions for a more fundamental treatment, which is however left as an issue for future research.

With the model specifications in this chapter, the fundamentals of the human-kinetic model are laid. Although the derivation is equally applicable to alternative behavioural specifications and refinements (some examples of which are presented in the remainder of this thesis), it is interesting to examine the behaviour of this basic version of the human-kinetic model in more detail. In the next chapter, a number of analyses are presented, both focussing on theoretical issues like stability and consistency between microscopic model foundations and macroscopic formulation, and on practical issues such as the representation of congestion patterns and the sensitivity of the model solutions to various parameter settings. In the remainder of the thesis, the human-kinetic traffic flow model is then further applied to variable driver behaviour, multi user-class traffic, and mixed non-supported and ADA-equipped traffic.

5 PROPERTIES AND BEHAVIOUR OF THE BASIC HUMAN-KINETIC MODEL

In this chapter the properties and behaviour of the basic human-kinetic traffic flow model, with the specifications given in chapter 4, are examined. The behaviour of the model is explored in increasing order of complexity: first equilibrium behaviour, then dynamic behaviour. The behaviour is discussed based on mathematical analyses and simulated results. Mathematical (or theoretical) analyses are performed only on the model without speed or gap variances. We call this the *deterministic* model (as the opposite of the *stochastic* model with variances), since in that case the individual speeds and gaps are deterministically related to the macroscopic speed V and the density k.

The outline of the chapter is as follows. Distinction is made between static or equilibrium behaviour (section 5.1), and dynamic behaviour of the model (section 5.2). It turns out that – even without the numerical solution techniques discussed in the previous section – equilibrium solutions can be calculated that are homogeneous in time and in space (section 5.1.1). These solutions are the fundamental diagrams corresponding to the microscopic specifications of the model. Section 5.1.2 examines the sensitivity of the fundamental diagram for the parameters of the model. An important aspect determining the dynamic behaviour is the stability of equilibrium solutions. A linear stability analysis is performed, and the role of parameters determining the stability of the model is investigated (section 5.2.1). The influence of other dynamic parameters is illustrated in the dynamic sensitivity analysis of section 5.2.2). In section 5.3, the previous results are taken as arguments for a partial proof of equivalence between the microscopic and macroscopic model formulation. Finally in

section 5.4, the human-kinetic traffic flow model is applied to a hypothetical on-ramp and different congestion patterns are examined.

5.1 Equilibrium model behaviour

In this section solutions of the human-kinetic model are considered that are homogeneous in space and time. This means that – once the model has reached such a time and spatially homogeneous state – this macroscopic state does not change, unless some external influence enters the model (through the boundary conditions). These states are therefore also referred to as *equilibrium* solutions. A homogeneous traffic state requires that the acceleration on the macroscopic level is zero, so that the equilibrium solutions can be directly derived from the definition of the acceleration integral. This procedure, as well as a sensitivity analysis of equilibrium solutions for changes to the model's parameters, is presented in this section.

5.1.1 Equilibrium solutions: theoretical equilibrium speed

The definition of the acceleration integral (4.33) with (4.36) is only dependent on the macroscopic state variables $k(t,x)$ and $V(t,x)$ on location x and some distance downstream of x. Formally the acceleration integral can thus be written as:

$$\left\langle \frac{dv}{dt}\left(t+T,x\left(t+T\right)\right)\right\rangle_v = f\left(k\left(t,x\right),k\left(t,x+\frac{1}{2\,k\left(t,x\right)}\right),V\left(t,y\right)\right) \qquad y\in\left[x,x+\Delta x_{ant}\right] \tag{5.1}$$

Assume a solution for the dynamic density and speed equations that is homogeneous in x, that is:

$$\begin{aligned} k\left(t,x\right) &= k\left(t\right) \quad \forall\, x \\ V\left(t,x\right) &= V\left(t\right) \quad \forall\, x \end{aligned} \tag{5.2}$$

Then:

$$\left\langle \frac{dv}{dt}\left(t+T,x\left(t+T\right)\right)\right\rangle_v = f\left(k\left(t\right),V\left(t\right)\right) \tag{5.3}$$

If there exist combinations of $k^e(t)$ and $V^e(t)$ for which the acceleration integral equals zero, this would mean that the homogeneous macroscopic speed V^e would remain unchanged in time, and therefore also the density k^e. These are by definition macroscopic equilibrium solutions of the human-kinetic model. Omitting the dependency of t the implicit definition for the equilibrium condition (k^e,V^e) or $V^e(k^e)$ can be written:

$$\begin{aligned} \left\langle \frac{dv}{dt}\right\rangle_v &= \int_{v_j} N\left(v_j\right)\int_{s_j} LgN\left(s_j\right)\int_{v_{j-1}} N\left(v_{j-1}\right)\dot{v}_j\left(v_j,s_j,v_{j-1}\right)dv_{j-1}\,ds_j\,dv_j \\ &= f\left(k^e,V^e\right) \\ &= 0 \end{aligned} \tag{5.4}$$

with:

$$\begin{aligned} N\left(v_j\right) &= N\left(v_j, \mu_{v_j}\left(k^{\mathrm{e}}, V^{\mathrm{e}}\right), \sigma_{v_j}\left(k^{\mathrm{e}}, V^{\mathrm{e}}\right)\right) \\ LgN\left(s_j\right) &= LgN\left(s_j, \mu_{s_j}\left(k^{e}\right), \sigma_{s_j}\left(k^{e}\right)\right) \\ N\left(v_{j-1}\right) &= N\left(v_{j-1}, \mu_{v_{j-1}}\left(k^{\mathrm{e}}, V^{\mathrm{e}}\right), \sigma_{v_{-1}j}\left(k^{\mathrm{e}}, V^{\mathrm{e}}\right)\right) \end{aligned} \tag{5.5}$$

Due to the general form of the car-following rule $\dot{v}_j$, it is not possible to explicitly derive a general closed expression for the equilibrium relation $V^e(k^e)$. Given a fixed set of parameters for the car-following model and for the distribution functions in equation (5.5), the equilibrium condition can be solved numerically. A typical result is given in Figure 5-1a that depicts the value of the acceleration integral as a function of k and V. The equilibrium condition is met at the intersection of this acceleration integral function with the plane $\left\langle \frac{dv}{dt} \right\rangle_v = 0$. Figure 5-1b shows the equilibrium relation $V^e(k^e)$ for the same parameter set.

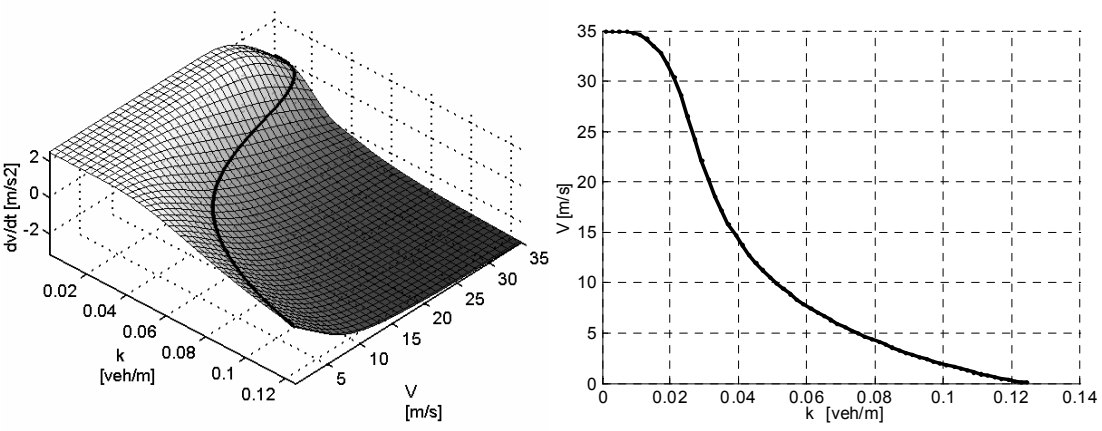


Figure 5-1 (a)* left*: acceleration integral as a function of the density and the speed; intersection with $\left\langle \frac{dv}{dt} \right\rangle_v = 0$ ***; (b)* right*: equilibrium speed-density relationship***

Figure 5-1b has the typical shape of equilibrium speeds or fundamental diagrams that are used in macroscopic traffic flow models, like the model of Payne (1971) or the LWR model.

It turns out that the acceleration integral that is solely based on an individual driver model and specifications of distribution functions for the speed and gaps, implicitly defines an equilibrium relation between speed and density. This relation is similar to traditional empirically derived fundamental diagrams. However, the acceleration integral also defines the expected acceleration when actual traffic conditions deviate from the equilibrium.

5.1.2 *Sensitivity of the equilibrium solutions of the human-kinetic traffic flow model* (*fundamental diagram*)

The human-kinetic traffic flow model contains a considerable number of parameters. The aim of this section and section 5.2.2 is to provide insight into the sensitivity of model solutions for changes in these parameters. In this section the sensitivity analysis focuses on those parameters that influence the equilibrium solution. In section 5.2.2 the influence of parameter settings on dynamic model solutions is examined.

The sensitivity analysis of the equilibrium solutions is split into three parameter groups:

(i) parameters with influence on the deterministic equilibrium solutions (i.e. with speed and gap variances = 0)

(ii) the speed and gap variances

(iii) parameters with no influence on the deterministic equilibrium solutions, but with influence on stochastic equilibrium solutions

The parameters analysed in this section, the group to which they belong, and the default value are listed in Table 5-1.

Table 5-1 Parameters and default values for the equilibrium sensitivity analysis

	Para-meter	Description	Default value	Introduced in equation
(i)	k_{jam}	jam density; inverse of gross vehicle length	0.125 veh/m	(4.20)
	s_1^d	linear factor for desired gap	1.0 s	(4.28)
	s_2^d	quadratic factor for desired gap	0.0 s^2/m	(4.28)
	W	average desired speed	35 m/s	(4.28)
(ii)	b	exponent determining the gap variance	0.8	(4.25)
	B	shape factor determining the speed variance[12]	0.012 / 0.028 (low / large k)	(4.16)
(iii)	acc_{min}	minimal acceleration	-5 m/s^2	(4.28)
	acc_{max}	maximal acceleration	3.5 m/s^2	(4.28)
	τ_w	relaxation constant for the desired speed	2.5 s^{-1}	(4.27)
	τ_s	relaxation constant for the desired distance	3.333 s^{-2}	(4.27)
	τ_v	relaxation constant for the speed difference	1.111 s^{-1}	(4.27)

[12] Actually, the shape factor is a function of the density (see equation (4.16)); this function changes rather abruptly from an almost constant value for low density, to another virtually constant value for high density. For that reason we treat the shape factor as a constant in the sensitivity analysis.

5.1.3 *Parameter sensitivity of the deterministic equilibrium solution*

With the variance of speeds and gaps zero, the equilibrium solutions are simply solvable from the condition (5.4). The solution exists of two branches: the *free flow* branch (for which the free flow acceleration in the car-following rule (4.27) is determinant in the minimum operator) and the *car-following* or *congestion* branch (otherwise). The free flow branch is defined by:

$$V^e = W \qquad \forall k^e < k_{crit} \tag{5.6}$$

The congestion branch follows from:

$$\frac{1}{k^e} = \frac{1}{k_{jam}} + s^d\left(V^e\right) \quad \forall k^e \geq k_{crit} \tag{5.7}$$

These two conditions implicitly define the critical density as the intersectin of the free flowing and congestion branches:

$$\frac{1}{k_{crit}} = \frac{1}{k_{jam}} + s^d\left(W\right) \tag{5.8}$$

From equations (5.6) – (5.8), the sensitivity of the equilibrium solutions can be directly read. As an illustration some variants to the reference parameter set of Table 5-1 are plotted in Figure 5-2, which shows the fundamental diagram of density against flow rate.

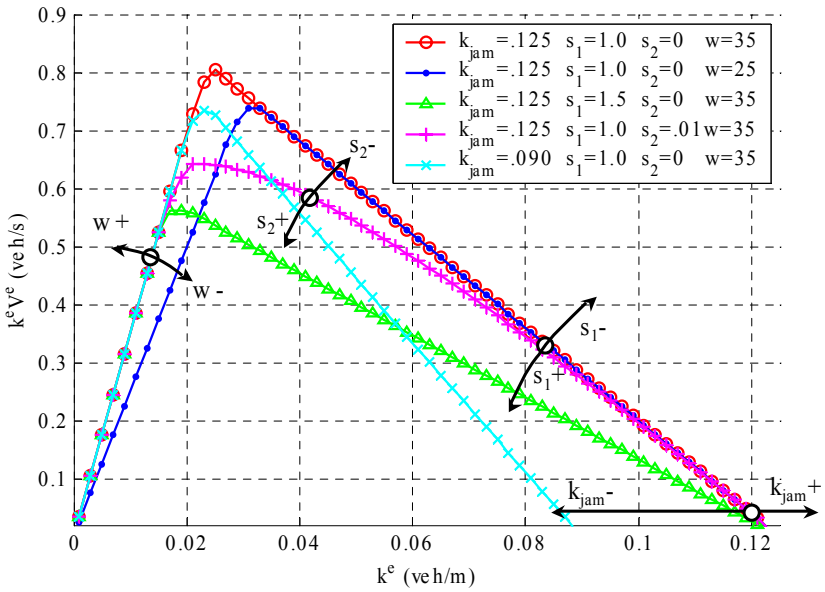


Figure 5-2 Parameter sensitivity of the deterministic equilibrium solution

The effect of a lower desired speed W is a lower slope of the free flow branch, and therefore a shift of the critical density to higher density, corresponding to a lower maximal flow rate. A larger linear desired gap factor s_1^d leads to a lower maximal flow

rate at a lower critical density. Introducing a quadratic factor s_2^d different from zero in the desired gap definition, changes the straight congestion branch into a curved, concave one with lower equilibrium flow rates around the critical density. The difference of the latter two cases with the reference variant disappears for large density, since the equilibrium speed near jam density approaches zero, so that the linear or quadratic factors have no effect anymore. Finally, a lower jam density leads by definition to standstill at a lower maximal density, while the maximal flow rate changes only marginally. Therefore, the slope of the congested branch is steeper. Let us remark that all conclusions in this paragraph or course are also valid vice versa, i.e. changing the parameter in the other sense leads to the inverse effect as described here.

5.1.4 *Sensitivity of the equilibrium solutions for speed and gap variance (stochastic equilibrium solutions)*

This section examines the effect of introducing speed and gap variance in the equilibrium solution. As a reference we use the first curve (red – circle markers) in Figure 5-2 and Figure 5-3. The figures have the same scales to allow easy comparison. A priori one can expect that introducing variance leads to increased vehicular interactions and hence lower equilibrium flow rates and speeds for the same density. Moreover, this effect will be dominant for the medium density range, since for low density there is enough space on the road to drive at the desired speed, and for near-jam density the average gaps and speeds become so small and therefore also the variance.

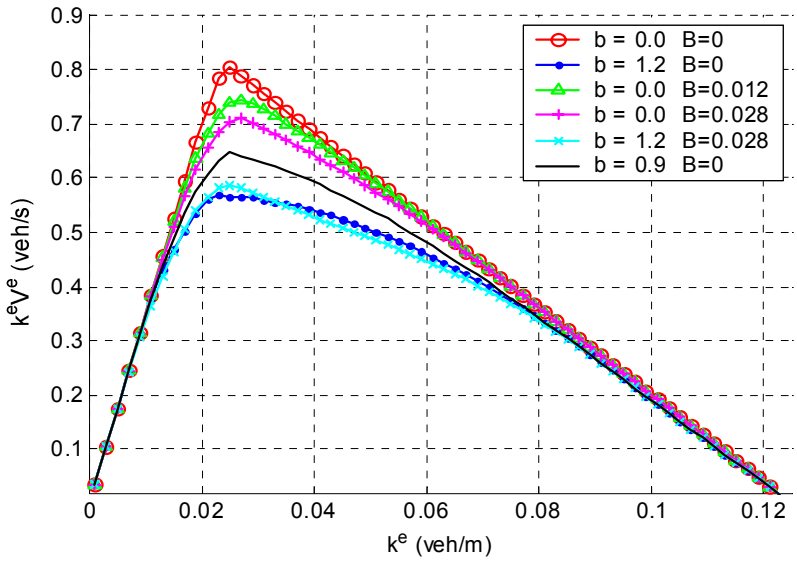


Figure 5-3 Sensitivity of the equilibrium solutions for speed and gap variance

These expectations are confirmed in Figure 5-3. Any variance or combination of variances leads to lower equilibrium flow rates compared to the deterministic reference case. The effect of the speed variance – with a realistic order of magnitude of the parameters obtained from Dutch motorway data (Helbing, 1997) – is rather small. Helbing proposes an equilibrium variance proportional to the squared equilibrium speed

(equation (4.16)), with the shape factor B varying between the extremes shown in Figure 5-3. The equilibrium solutions are almost equal, with a maximal flow rate in the order of 10% lower than the deterministic reference. The effect of b, the gap variance parameter, is strong. The curve for b=0.8 is not shown since it coincides with that of B=0.028; this shows that the gap and speed variance act upon the equilibrium relation in a similar fashion: an increased interaction frequency between vehicles reduces the equilibrium speed with respect to the deterministic case, especially for the medium density range. A further increase of b to 0.9 and 1.2 leads to substantially lower capacity. The curve for b=1.2 and that for b=1.2, B=0.028 almost coincide, which shows that the gap variance is dominant over the speed variance for larger values of b. This strong sensitivity suggests that a more subtle definition of the gap variance than equation (4.25) might be better; more research into the gap distribution is needed.

5.1.5 *Parameter sensitivity of the stochastic equilibrium solution*

The third class of parameters has no effect on the deterministic equilibrium solution, but once individual speed and gap differences are introduced, they influence the fundamental diagram. For example, if differences between the individual speeds exist, an influence of τ_v, the relaxation constant for the speed difference is probable.

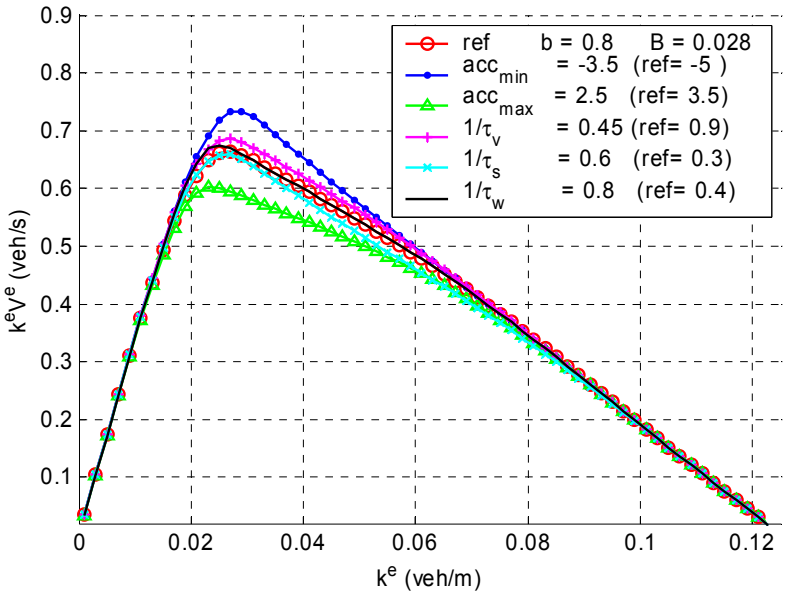


Figure 5-4 Parameter sensitivity of the stochastic equilibrium solution

Figure 5-4 confirms this expectation. It turns out that the equilibrium solution is little sensitive to the relaxation constants τ_v, τ_s and τ_w: the solutions almost coincide with the reference curve. The minimal and maximal accelerations acc_{min} and acc_{max} have a more pronounced influence. A lower maximal or lower minimal acceleration shifts the fundamental diagram towards lower equilibrium speeds (e.g. the curve for a decreased acc_{max}) and vice versa (e.g. the curve for an increased acc_{min}). This effect is strongest in the medium density region where the variance is most pronounced. The explanation is

logical: the low accelerations in the acceleration integral are stronger (with lower acc_{min}) or less compensated for by high accelerations (with lower acc_{max}), which shifts the result of the acceleration integral calculation towards lower average accelerations, and hence the equilibrium to lower speeds.

We conclude that parameters of group (i) have the strongest influence on the equilibrium solution. Moreover, their influence affects the entire density range. The variance of speeds and gaps (group (ii)) is also important, especially that of the gap variance. The effect of the parameters of group (iii) is indirect, through the existence of speed and gap variance. Of this group only the extreme accelerations have a substantial influence on the equilibrium solution.

It is interesting to note that the fundamental diagram of flow rate against density is essentially triangular (in the deterministic case with linear desired gap function). The influence of the parameters either changes the shape of the triangle (slope of the free flow or congestion branches) or smoothens the tip of the triangle. In all cases the fundamental diagram is concave. In light of the kinetic wave theory, this means that illogical first order model behaviour corresponding to convex fundamental diagrams (Zhang, 2000) does not occur in the human-kinetic traffic flow model.

5.2 Dynamic model behaviour

If the initial conditions to the model are not in equilibrium (see previous section), or some boundary condition generates disturbances to an initial traffic state in equilibrium, the solution of the human-kinetic model becomes dependent on space and time. The analysis of the dynamic behaviour is partially mathematical (or theoretical), and partially based on numerical simulation results. The theoretical analysis (section 5.2.1) involves a stability analysis of the equilibrium solutions, calculated in the previous section. The question here is: if an equilibrium solution is slightly disturbed, does the dynamic behaviour restore the original equilibrium solution or does the amplitude of the disturbance grow in time? This analysis yields insight into the stability of the model and the influence of parameter settings on stability. These insights are completed with a numerical sensitivity analysis (section 5.2.2), in which the influence of parameter settings on non-equilibrium starting conditions is examined.

5.2.1 Linear stability analysis of the dynamic deterministic solution

In section A.6.2 of Annex A, we derive a general linear stability criterion for any Payne-type model that can be cast into the form:

$$\frac{\partial k}{\partial t}+\frac{\partial kV}{\partial x}=0 \tag{5.9}$$

$$\frac{\partial V}{\partial t}+V\frac{\partial V}{\partial x}+\frac{1}{k}\frac{\partial P}{\partial x}=f\left(k,V\right) \tag{5.10}$$

The stability condition is:

$$\left(\frac{\partial f}{\partial V}\right)^2 \frac{\partial P}{\partial k} - k^{*2}\left(\frac{\partial f}{\partial k}\right)^2 - \frac{\partial P}{\partial V}\frac{\partial f}{\partial V}\frac{\partial f}{\partial k} > 0 \tag{5.11}$$

It is shown below, that the human-kinetic model with the speed- and gap-variance set equal to zero can be cast into the form of equations (5.9) and (5.10) using linear Taylor approximations, which allows to apply the generic linear stability condition (5.11) to our model. In order to increase the insight into how the different behavioural aspects contribute to the linear stability of the model, the analysis is performed in three steps:

(i) the human-kinetic model without reaction time nor anticipation;

(ii) the human-kinetic model with reaction time but without anticipation;

(iii) the human-kinetic model with reaction time and anticipation.

5.2.1.1 Without reaction time nor anticipation

The primitive speed equation (without in- or outflow of traffic into the modelled road section) contains the acceleration integral in the RHS. In absence of speed- and gap-variance the integration reduces to:

$$\left\langle \frac{dv}{dt}(k,V) \right\rangle_v = \min\left(\frac{w_j - v_j}{\tau_w} \quad ; \quad \frac{s_j - s^d\left(v_j\right)}{\tau_s} + \frac{v_{j-1} - v_j}{\tau_v} \right) \tag{5.12}$$

with:

$$\begin{aligned} s^d\left(v_j\right) &= s^d\left(V\right) \\ &= s_1^d V + s_2^d V^2 \end{aligned} \tag{5.13}$$

$$\begin{aligned} v_j &= V\left(t,x\right) \\ v_{j-1} &= V\left(t, x + h\left(t,x\right)\right) \end{aligned} \tag{5.14}$$

$$s_j = h\left(t,x\right) - s_j^0 \tag{5.15}$$

Remember the approximation for the average headway $h(t,x)$ of equation (4.23):

$$h\left(t,x\right) = \frac{1}{k\left(t, x + \dfrac{1}{2k\left(t,x\right)}\right)} \tag{5.16}$$

First of all, let us remark that it is only necessary to consider the second case in the minimum operator in equation (5.12). In the other case, the density is apparently so low that the vehicles do not feel each other's presence and are all free to accelerate towards their desired speed w. That low-density limit is always stable, since any small deviation from an equilibrium would return to the equilibrium: the situation that all vehicles travel at speed w with gaps larger than or equal to $s^d\left(w\right)$. This stable region therefore contains all situations where:

$$k < \frac{1}{s_j^0 + s^d\left(w\right)} \tag{5.17}$$

The analysis can thus be focussed on the car-following case, i.e. the second case in the minimum operator in equation (5.12), which is dominant in all situations that do not satisfy (5.17). For the identification of the human-kinetic speed equation (with its non-local terms (5.14) and (5.15)) with the general form of the local Payne-type speed equation (5.10), it needs to be rewritten it in a local form, i.e. approximate the non-localities first using Taylor approximations and neglect all higher order terms (2nd or higher derivatives or products of derivatives).

Let us start with equation (5.16):

$$\begin{aligned} h\left(t,x\right) &\approx \frac{1}{k\left(t,x\right) + \frac{1}{2k\left(t,x\right)}\frac{\partial k}{\partial x}} \\ &= \frac{1}{k}\left(\frac{1}{1+\frac{1}{2k^2}\frac{\partial k}{\partial x}}\right) \\ &\approx \frac{1}{k}\left(1-\frac{1}{2k^2}\frac{\partial k}{\partial x}\right) \end{aligned} \tag{5.18}$$

Substituting this result in (5.14), yields for the speed of the predecessor:

$$\begin{aligned} v_{j-1} &= V\left(t, x + h\left(t,x\right)\right) \\ &\approx V + h\frac{\partial V}{\partial x} \\ &\approx V + \frac{1}{k}\left(1-\frac{1}{2k^2}\frac{\partial k}{\partial x}\right)\frac{\partial V}{\partial x} \\ &\approx V + \frac{1}{k}\frac{\partial V}{\partial x} \end{aligned} \tag{5.19}$$

If the terms with *t*- or *x*-derivatives are collected in the LHS, and the rest in the RHS, the local approximation of the non-local speed equation for the human-kinetic model reads:

$$\frac{\partial V}{\partial t} + V\frac{\partial V}{\partial x} + \frac{1}{k}\left(\frac{\partial k\Theta^e}{\partial x} + \frac{1}{2\tau_s k^2}\frac{\partial k}{\partial x} - \frac{1}{\tau_v}\frac{\partial V}{\partial x}\right) = \frac{1}{\tau_s}\left(\frac{1}{k} - s_j^0 - s^d\left(V\right)\right) \tag{5.20}$$

Since the speed variance is assumed = 0, the equilibrium variance Θ^e also equals zero. However, since we consider a small numerical contribution c_0^2 to the eigenvalues in our model, Θ is withheld in the derivation as a constant equal to c_0^2 and its influence is examined later. With the substitutions:

$$\tilde{P} \equiv kc_0^2 - \frac{1}{2\tau_s k} - \frac{V}{\tau_v} \tag{5.21}$$

$$f(k,V) \equiv \frac{1}{\tau_s}\left(\frac{1}{k} - s_j^0 - s^d(V)\right) \tag{5.22}$$

equation (5.20) is cast into the desired form:

$$\frac{\partial V}{\partial t} + V\frac{\partial V}{\partial x} + \frac{1}{k}\frac{\partial \tilde{P}}{\partial x} = f(k,V) \tag{5.23}$$

Therefore, the linear stability condition is given by (5.11) with equations (5.21) and (5.22) substituting P and f respectively. The linear stability condition now becomes:

$$\frac{1}{\tau_s^2}\left(k^2\left(s_1^d + 2s_2^d V\right)^2\left(c_0^2 + \frac{1}{2\tau_s k^2}\right) + \frac{s_1^d + 2s_2^d V}{\tau_v} - 1\right) > 0 \tag{5.24}$$

From this result it appears that if $\tau_s \to \infty$, the model is at the limit of instability independent of the values of k, V or any other parameter. Otherwise, distinction is made between the results for c_0 equal to zero or not.

If $c_0 > 0$, the stability conditions are:

$$1 - \frac{s_1^d + 2s_2^d V}{\tau_v} - \frac{\left(s_1^d + 2s_2^d V\right)^2}{2\tau_s} < 0$$

$$or \tag{5.25}$$

$$k > \frac{1}{c_0\left(s_1^d + 2s_2^d V\right)}\sqrt{1 - \frac{s_1^d + 2s_2^d V}{\tau_v} - \frac{\left(s_1^d + 2s_2^d V\right)^2}{2\tau_s}}$$

This result reveals that the following parameters have a stabilising effect (the greater the parameter, the more stable is traffic flow): c_0, s_1^d, and s_2^d. Large relaxation times τ_s or τ_v destabilise traffic flow. Furthermore, stability is reached for high speed V (if $s_2^d \neq 0$, otherwise the speed has no influence on stability) or high density k. This is consistent with the general finding that traffic is stable for low density (equation (5.17)) or very high density (equation (5.25)), and can have an unstable region for intermediate density. The latter is however not necessarily the case: it depends on the parameter settings whether or not the model has an unstable medium density range, as is shown by condition (5.25).

If $c_0 = 0$ the stability condition is independent of the density k:

$$1 - \frac{\left(s_1^d + 2s_2^d V\right)^2}{2\tau_s} + \frac{s_1^d + 2s_2^d V}{\tau_v} < 0 \tag{5.26}$$

In the specific case that $s_2^d = 0$ (the parameter choice for all numerical simulations with the human-kinetic model shown in this thesis), this further reduces to a condition that is also independent on V:

$$1-\frac{\left(s_1^d\right)^2}{2\tau_s}+\frac{s_1^d}{\tau_v}<0 \tag{5.27}$$

The latter result is of special interest with respect to the interpretation of simulation results shown in this thesis. This equation, together with equation (5.17), says that free flowing traffic is always linearly stable, but once the critical density given (5.17) is exceeded and traffic enters the congested regime, congestion is either always linearly stable, or always linearly unstable, dependent on whether or not the parameters satisfy condition (5.27). At least, these results hold for a model without speed- or gap-variance, and neither reaction time nor anticipation. The latter restrictions are released in the next two cases.

5.2.1.2 Including reaction time

For the modelling of reaction time, the RHS and LHS of equation (5.12) are evaluated with a time lag T, so at time t-T and t or at time t and t+T respectively. We now make similar first-order Taylor approximations with respect to the time and the space increment, e.g.:

$$\begin{aligned} &V\left(t-T,x\left(t-T\right)+h\left(t-T,x\left(t-T\right)\right)\right) \\ &\quad\approx V-T\left(\frac{\partial V}{\partial t}+\frac{dx}{dt}\frac{\partial V}{\partial x}\right)+h\left(t-T,x\left(t-T\right)\right)\frac{\partial V}{\partial x} \\ &\quad=V-T\frac{\partial V}{\partial t}-TV\frac{\partial V}{\partial x}+\frac{1}{k}\left(1+\frac{T}{k}\frac{\partial k}{\partial t}-\frac{1-Tk V}{k^2}\frac{\partial k}{\partial x}\right)\frac{\partial V}{\partial x} \\ &\quad\approx V-T\frac{\partial V}{\partial t}-TV\frac{\partial V}{\partial x}+\frac{1}{k}\frac{\partial V}{\partial x} \end{aligned} \tag{5.28}$$

For the non-local and non-temporal approximation of k, the continuity equation (5.9) is used to simplify the expression:

$$\begin{aligned} k\left(t-T,x\left(t-T,x+\Delta x\right)\right)&\approx k-T\left(\frac{\partial k}{\partial t}+V\frac{\partial k}{\partial x}\right)+\Delta x\frac{\partial k}{\partial x} \\ &=k-T\left(-k\frac{\partial V}{\partial x}\right)+\Delta x\frac{\partial k}{\partial x} \end{aligned} \tag{5.29}$$

After approximating all non-local and non-temporal terms in the speed equation and collecting those with time or space derivatives in the LHS and the rest in the RHS, one obtains:

$$\frac{\partial V}{\partial t}+V\frac{\partial V}{\partial x}+\frac{1}{k}\frac{\partial \tilde{P}}{\partial x}=f\left(k,V\right) \tag{5.30}$$

with:

$$\tilde{P}\equiv\tilde{f}\left(kc_0^2+\left(\frac{T}{\tau_s}-\frac{1}{\tau_v}\right)V-\frac{1}{2\tau_s k}\right) \tag{5.31}$$

$$f(k,V) \equiv \tilde{f}\frac{1}{\tau_s}\left(\frac{1}{k} - s^d(V)\right) \tag{5.32}$$

$$\tilde{f} \equiv \frac{1}{1 - \frac{T}{\tau_s}\frac{ds^d}{dV}} \tag{5.33}$$

For the sake of notational simplicity, the desired distance function $s^d(V)$ is assumed to be linear in the speed V, so that the factor $\tilde{f} = \frac{1}{1 - \frac{T}{\tau_s}s_1^d}$ is independent of k and V.

Because the speed equation (5.30) is again in the general form of equation (5.10), the general linear stability condition (5.11) holds again with the definitions (5.31) – (5.33). The stability condition can be simplified (if $\tilde{f} \neq 0$) to yield:

$$\frac{1}{\tau_s^2}\left(k^2\left(s_1^d\right)^2\left(c_0^2 + \frac{1}{2\tau_s k^2}\right) + \frac{s_1^d}{\tau_v} - 1\right) > 0 \tag{5.34}$$

which is exactly the same condition as (5.24) if the desired distance is linear in V ($s_2^d = 0$). This result is rather surprising, since the linear stability turns out independent of the reaction time T.

Note that this not necessarily means that the reaction time has no effect at all on stability, but the effect is non-linear. Also this analysis does not exclude that for large perturbations the reaction time might affect the stability (since the analysis performed here builds on *linear* Taylor approximations, which are only valid for small perturbations). The most important lesson learned from this analysis is that the parameters in equation (5.34) all have a stronger effect on stability than the reaction time. In other words: if traffic flow is adequately modelled by the equations used in this section, and if one intends to improve traffic stability using traffic management measures or ADA systems, these systems should aim at influencing the parameters s_1^d, τ_s, or τ_v, rather than reducing reaction time, since stability is much more sensitive to these parameters than to the reaction time T.

5.2.1.3 Including reaction time and anticipation

For this analysis the dependence of the reaction time can be omitted, since the previous analysis has shown that it does not influence the linear stability analysis. The anticipation defined by equations (4.36) and (4.37) adds two supplementary parameters to the previous equations. The main difference is in the definition of the predecessor's speed, approximated in the previous sections with equation (5.19) and (5.28). It is now modified as follows:

$$v_{j-1} = V(t,x) + h(t,x)\,f^{ant}\left(\frac{dV}{dx}\right)^{ant} \tag{5.35}$$

Now with definition (4.37) for the anticipated spatial speed profile, in the most extreme cases information of $\Delta x \approx 0$ or $\Delta x = \Delta x^{ant}$ downstream is used. Therefore, the following approximation holds with $0 \le \Delta x \le \Delta x^{ant}$:

$$\left(\frac{dV}{dx}\right)^{ant} \approx \frac{\partial V}{\partial x} + \Delta x \frac{\partial^2 V}{\partial x^2} \tag{5.36}$$

Then the speed equation becomes:

$$\frac{\partial V}{\partial t} + V\frac{\partial V}{\partial x} + \frac{1}{k}\frac{\partial \tilde{P}}{\partial x} - \frac{f^{ant}\Delta x}{\tau_v k}\frac{\partial^2 V}{\partial x^2} = f\left(k,V\right) \tag{5.37}$$

with:

$$\tilde{P} \equiv kc_0^2 - \frac{1}{2\tau_s k} - f^{ant}\frac{V}{\tau_v} \tag{5.38}$$

$$f\left(k,V\right) \equiv \frac{1}{\tau_s}\left(\frac{1}{k} - s_j^0 - s^d\left(V\right)\right) \tag{5.39}$$

These equations reveal that the anticipation distance Δx occurs as a factor for a second order derivative. Therefore, its influence is *non-linear only* and the linear stability analysis is unchanged, i.e. Δx has little effect on the stability for *small* perturbations (for which the linear Taylor approximations are accurate), but can have a more important effect on large perturbations in a non-linear way. The anticipation factor f^{ant} occurs as a modification of the relaxation time τ_v. Therefore, the results of the first analysis hold, with the substitution $\frac{1}{\tau_v} \to \frac{f^{ant}}{\tau_v}$: the system is linearly stable if:

$$\frac{1}{\tau_s^2}\left(k^2\left(s_1^d + 2s_2^d V\right)^2\left(c_0^2 + \frac{1}{2\tau_s k^2}\right) + f^{ant}\frac{s_1^d + 2s_2^d V}{\tau_v} - 1\right) > 0 \tag{5.40}$$

The conclusions on linear stability of the first case (neither reaction time nor anticipation) are therefore also valid for the case with anticipation with one modification: the influence of anticipation is felt as a smaller relaxation τ_v with respect to speed differences (for $f^{ant} > 1$), which has a beneficial influence on linear stability of the system. The anticipation distance has no influence on linear stability, but acts upon stability of the system in a non-linear way.

5.2.2 Sensitivity of the dynamic solutions of the human-kinetic traffic flow model

This section discusses the sensitivity of traffic dynamics in the human-kinetic model, based on the sensitivity analysis results presented in Annex B.

Three groups of parameters are considered:

(i) parameters of the car-following model

(ii) parameters of the anticipation model

(iii) parameters of the numerical scheme

For the reference case, the parameter settings of Table 5-2 are used. In each case in Annex B, the sensitivity of the model solution for a change to only one of the parameters is examined (unless stated otherwise).

For the sensitivity analysis, a 6 km long single lane road with open boundaries up- and downstream is considered. Initially, a wide jam (terminology after Kerner, see section 2.4.2.2 and also 5.4.2 and 6.6.3.2) is located between x = 4.5 km and x = 5.5 km. Within the jam the density is equal to the jam density and the initial average speed equals zero. Outside the jammed region, traffic density equals 21 veh/km with an average speed of 30 m/s.

Table 5-2 Parameters involved in the dynamic sensitivity analysis and their default value

	Para-meter	Description	Default value	Introduced in equation
fix	k_{jam}	jam density; inverse of gross vehicle length	0.125 veh/m	(4.20)
	W	average desired speed	35 m/s	(4.28)
	b	exponent determining the gap variance	0.8	(4.25)
	B	shape factor determining the speed variance[13]	0.012 / 0.028 (low / large k)	(4.16)
(i)	T	reaction time	0 s	(4.33)
	τ_w	relaxation constant for the desired speed	2.5 s^{-1}	(4.27)
	s_1^d	linear factor for desired gap	1.0 s	(4.28)
	s_2^d	quadratic factor for desired gap	0.0 s^2/m	(4.28)
	τ_s	relaxation constant for the desired distance	3.333 s^{-2}	(4.27)
	τ_v	relaxation constant for the speed difference	1.111 s^{-1}	(4.27)
	acc_{min}	minimal acceleration	-5 m/s^2	(4.28)
	acc_{max}	maximal acceleration	3.5 m/s^2	(4.28)

[13] Actually, the shape factor is a function of the density (see equation (4.16)); this function changes rather abruptly from an almost constant value for low density to another virtually constant value for high density. For that reason we treat the shape factor as a constant in the sensitivity analysis.

	Para-meter	Description	Default value	Introduced in equation
(ii)	Δx^{ant}	anticipation distance	150 m	(4.37)
	f^{ant}	anticipation strength factor	1	(4.36)
	asym	anticipation asymmetry factor	0	
(iii)	c_0	threshold for numerical information flow inversion	6	(4.70)
	p_{prior}	fraction of numerical source term accounted for before convective part of the traffic model	75%	(4.63)
	$(\Delta t, \Delta x)$	numerical discretisation	(1s,120m)	(4.63)

5.2.2.1 Wide jam propagation and phase trajectory in the reference case

As an illustration of the model behaviour of the basic human-kinetic model, Figure 5-5a presents the evolution in space and time of the density[14] for the reference case (parameters of Table 5-2). The other cases are discussed in Annex B. From this figure the evolution of the shape of the initial jam can be traced, and the wave speed of the jam fronts is visualised. In this case the wave speed of the up- and downstream jam front are constant, with the downstream (accelerating) jam front travelling faster than the upstream (decelerating) jam front. Figure 5-5b represents the evolution in time of the traffic state in the density – flow plane at location x = 3900 m, with the equilibrium relationship plotted in dashed line as a reference. This location is initially free flowing, just upstream of the jam. Because the jam propagates upstream, it crosses the virtual detector, and the traffic state can be traced during and after the passage of the jam. The decelerating curve, corresponding to the passage of the upstream jam front (or queue tail), is marked with circles (⊖), whereas the accelerating curve, corresponding to the passage of the downstream jam front (or queue head), is marked with triangles (⧋). Further in Figure 5-5b, the initial flow - density state is marked with a white square (□), and the final state after passage of the jam with a white star (☆). The latter point is of special interest, since it corresponds to the jam outflow or queue discharge rate (see section 2.6). Also, the relative position of these points determines whether the jammed region grows (in the sense that the total number of vehicles inside the jammed region increases; this is the case when inflow is larger than outflow, i.e. square at higher flow value than star), shrinks (vice versa) or remains unchanged (flow rates coincide). Note that in either case the *shape* of the jam fronts and the *average density* inside the jammed region can change; the rate of in- and outflow only determines *the total number of vehicles* inside the jammed area.

[14] Notice how, during the first 200 seconds for $x > 4$ km, the model transforms the initial condition (which is not necessary compatible with the model) into a solution that is compatible to the model. In this case the jam is compressed to higher density and therefore becomes shorter. The transient solution (for $t < 200$ s and $x > 4000$) should therefore be omitted from the interpretations.

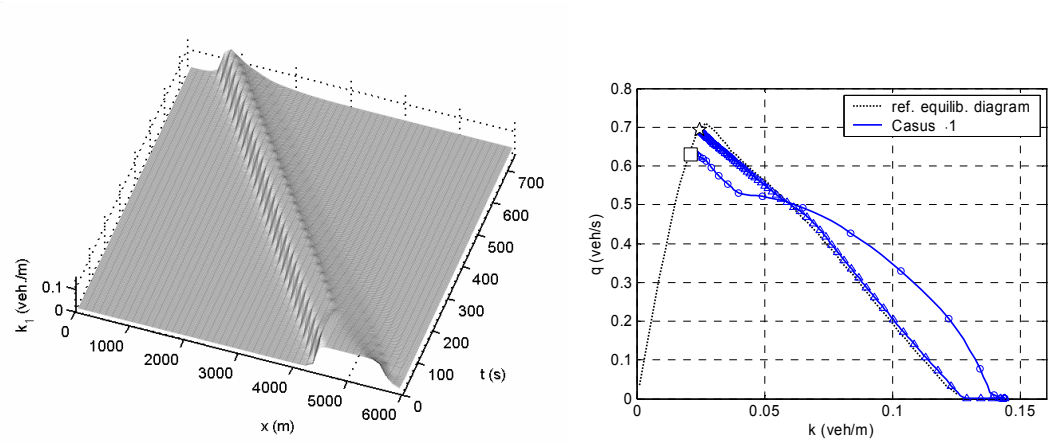


Figure 5-5 Reference run with parameters from Table 5-2; (a)* left*: space-time evolution of the density, (b)* right*: phase trajectory in the flow-density plane

The wave speeds can be determined by taking the slope of the lines connecting begin and end points of the corresponding accelerating and decelerating branches in the flow-density plane. In this case we find approximate values of -5.3 m/s (- 19 km/u) for the decelerating front and -6.8 m/s (-25 km/u) for the accelerating front, which confirms the visual inspection of Figure 5-5a, from which it can be concluded that the downstream wave travels faster than the upstream wave.

The combined deceleration and acceleration branches in the flow – density plane will be referred to as the *phase trajectory*. In the case of Figure 5-5, the phase trajectory has the shape of a stretched *8*. This shape emerges here since the deceleration to jammed traffic deviates substantially from the equilibrium flow curve, whereas the traffic state accelerates again to a free flowing state via the equilibrium curve. The non-equilibrium of the deceleration is characterised by an *anticipation-dominant* phase followed by a *relaxation-dominant* phase (terminology after: Zhang, 1999). In the anticipation-dominant phase (deceleration curve lower than acceleration curve in flow-density plane), the density is still relatively low when the drivers perceive the sharp jam front and adapt their speeds in advance (which is necessary due to their limited acceleration potential and delayed response). As a result of braking, the density increases, but the reduction of the speed is faster than in equilibrium conditions (braking is due to the speed profile downstream, and not due to a higher-than-equilibrium density, hence the name ‘anticipation-dominant’ phase). As a result, the flow rate decreases and remains under the equilibrium flow. Drivers then reach the steepest part of the jam front, where the density increases rapidly, and the relaxation-dominant phase starts (acceleration curve lower than deceleration curve in flow-density plane). The density apparently increases so rapidly now that it exceeds the equilibrium value. Deceleration is now predominantly determined by the urge to annihilate the change rate of the gaps, which are already shorter-than-equilibrium: drivers now adapt their speed to the conditions in which they drive *themselves* (and less by the conditions further downstream, hence the name ‘relaxation-dominant’ phase). Note that the shape and behavioural explanation of the 8-shaped phase trajectory are entirely consistent with the empirical observations and theory of Zhang (1999) and the observations by Treiterer & Myers (1974).

5.2.2.2 Sensitivity to behavioural and numerical parameter settings

We summarise the results of the sensitivity analysis in Annex B according to the three groups of parameters: for car-following, for anticipation, and for the numerical scheme.

The conclusions for the sensitivity to the car-following parameters are summarised as follows:

- The reaction time T has little influence on wide jam propagation in the dynamic (non-linear) simulations of traffic flow, which confirms the conclusions of the linear stability analysis. Also, the effect of the relaxation constant for the desired speed τ_w is marginal.
- The dynamic effect of the desired headway factors s_1^d and s_2^d is mainly sensed through the capacity and stability, as discussed earlier.
- Fiercer control of the desired gap (lower τ_s) leads to less overshoot of the equilibrium speed: actually, traffic flow behaviour is close to that of a first order traffic flow model, although in essence the sequence of anticipation-dominant followed by a relaxation dominant phase during deceleration remains unchanged. The only difference is the smoothness of the upstream jam front, caused by the anticipation behaviour. Fiercer control of the speed difference (lower τ_v) makes traffic more stable, because of a strong anticipation-dominant character. However, the speed assimilation (speed differences are fiercely eliminated) is so strong that the jam smoothens out, which points at exaggerated speed difference control.
- Varying the minimal and maximal accelerations acc_{min} and acc_{max} mainly affects the deceleration and acceleration jam fronts respectively, as could logically be expected. Reduced braking capability makes traffic run into the jam with unrealistically high density; increased acceleration capability lets traffic leave the jam more efficiently (higher discharge rate).
- If the minimal acceleration (strongest deceleration) is moderated, the deceleration curve shifts toward higher speeds (and hence flow rates), while the acceleration curve remains more or less unchanged. Because the braking capacity is limited the density inside the jam also increases to (unrealistically) higher values.

The conclusions for the sensitivity to the anticipation related parameters are summarised as follows:

- In general, the maximal anticipation distance Δx^{ant} has a subtle effect, whereas the anticipation strength f^{ant} is dominant.
- The maximal anticipation distance Δx^{ant} does not change discharge rate nor maximal density inside the jam. Its main effect is a changed chape of the deceleration front, with smoother acceleration while approaching the jam as the anticipation distance increases. This effect is comparable to that of a viscosity term in classical macroscopic traffic flow models.

- Increasing the anticipation strength f^{ant} obviously leads to a more pronounced anticipation-dominant phase and a less important relaxation phase in the approach to a jam, an effect very similarly to that of the relaxation constant τ_v for speed differences. Quiet differently however – and probably less realistically – the outflow from the jam is more efficient (because density inside the jam is lower and anticipation also acts on acceleration behaviour, albeit less strongly).
- A test simulation with a potential remedy to this issue, based on increasing the asymmetry of anticipation for acceleration and deceleration, did not sufficiently solve the issue of overestimated queue discharge so that it would justify the increased complexity of the model. We therefore choose an alternative way of handling this issue through consideration of a variable driving strategy (see section 6.6).

The conclusions for the sensitivity to numerical parameters are summarised as follows:

- The speed threshold c_0 below which negative eigenvalues are introduced artificially in the numerical solution scheme (section 4.4.3.2), confirms the effectiveness of this pragmatic numerical modification, precisely and only where it is needed. Small values actually suppress the mechanism which leads to unrealistic simulations, whereas too large values influence simulation results also at medium density for which it was not needed.
- A similar conclusion holds for the other pragmatic modification: the changed order of calculation controlled by p_{prior}. A smaller parameter value actually suppresses the mechanism, thus yielding unrealistic simulations and proving that the proposed correction indeed (partially) remedies the problems for which it was intended (section 4.4.3.2).
- Finally, the influence of the discretisation step for space and time ($\Delta t, \Delta x$) is considered. With a smaller step size, the simulation results are almost identical. The conclusion is that the discretisation grid needs not to be finer than in the reference case.

5.3 Equivalence of the microscopic and macroscopic model formulations

A very important question about the micro-macro link in the human-kinetic approach is, whether or not the macroscopic model is indeed the correct equivalent of the microscopic model. We do not prove the equivalence in general. However, for the deterministic case (variance of speeds and gaps = 0), the equilibrium solutions and the stability analysis of the previous sections turn out to be valid for both the microscopic and macroscopic models. This section shows that the two model implementations indeed produce equal results, i.e. their equilibrium solutions and stability conditions are equal and the numerical implementation yields perfectly comparable simulation results for some selected examples. We now discuss the arguments that lead us to conclude that microscopic and macroscopic model formulations are equivalent.

The first argument is a theoretical one. It is clear from the speed equation (4.48) with the definition of the acceleration integral (4.33) and all variances equal to zero, that the RHS of the macroscopic model is equal to the microscopic car-following relationship (4.27), the only difference being the continuous character of the speed V and density k in the macroscopic model, whereas these variables are only defined on discrete x-points corresponding to rear bumpers of vehicles (Figure 5-6) in the microscopic formulation. In the microscopic model the same acceleration as in the macroscopic model is thus calculated, albeit only at some discrete points on the continuous speed curve $V(x)$ that is considered in the macroscopic model.

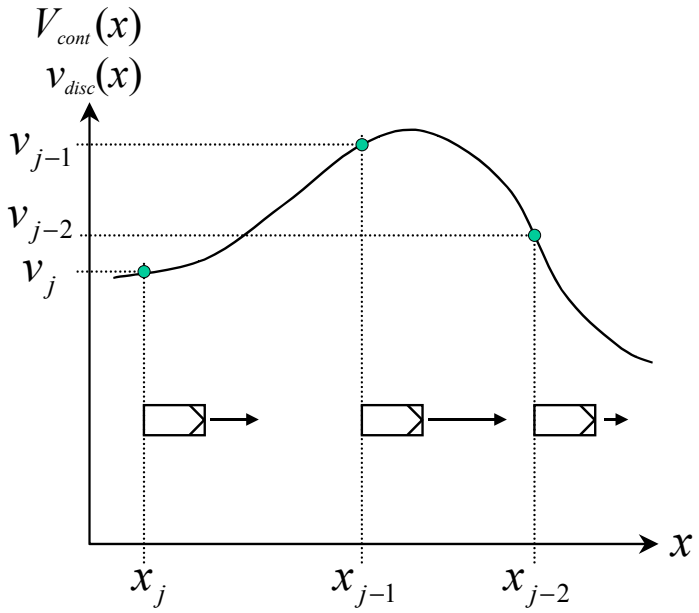


Figure 5-6 Correspondence of macroscopic continuous speed function $V_{cont}(x)$ and the microscopic discrete speed of individual vehicles $v_{disc}(x)$

The difference in model formulation is therefore mainly in the LHS. However, there is no conceptual difference between the microscopic or macroscopic model here: the difference in formulation is due to the frame of reference in which the model is defined. The microscopic model is defined in a moving framework: position and speed are solved for a moving vehicle j, whereas the macroscopic model is defined in a standing framework: speed and density are solved at fixed locations x.

A second argument builds on the equilibrium solutions of the deterministic micro and macro models. For the deterministic case, equation (5.4) reduces to:

$$\begin{aligned}\left\langle \frac{dv}{dt}(k,V) \right\rangle_v &= \dot{v}_j\left(v_j, s_j, v_{j-1}\right) \\ &= \min\left(\frac{w_j - V^e}{\tau_w} \;;\; \frac{\frac{1}{k^e} - s^d\left(V^e\right)}{\tau_s} + \frac{V^e - V^e}{\tau_v} \right) \\ &= 0\end{aligned} \tag{5.41}$$

This equilibrium condition is exactly the same as for the microscopic model. Therefore, the conclusion is justified that the equilibrium solutions for the deterministic microscopic and macroscopic models are the same.

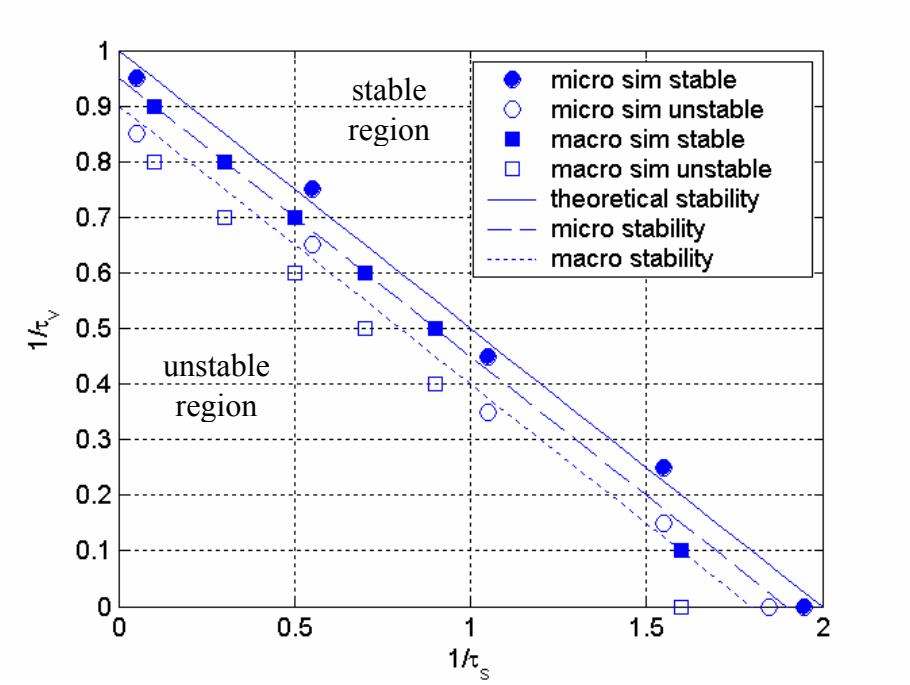


Figure 5-7 Theoretical and simulated stability of equilibrium traffic states for different combinations of τ_s and τ_v; indicated are simulated parameter combinations and whether they produced stable traffic in micro or macro simulations

A third argument is based on the stability of numerical solutions of the models. The stability of numerical solutions of the micro and macro models was tested and compared to the stability condition (5.24). Figure 5-7 and Figure 5-8 give the results of these simulations. It turns out that the theoretical linear stability condition is a good approximation of the stability in the simulations for both the micro and the macro model. For the microscopic model, the simulations are only slightly more stable than predicted theoretically. For the macroscopic model, this difference is larger: the stability of the macroscopic simulations follows the same trend as theoretically predicted, but the simulation is more stable. We also found that the difference becomes larger as the space and time discretisation steps of the numerical solution scheme become larger and disappears when the discretisation becomes finer (Figure 5-8). Probably, *numerical dispersion* (Hirsch, 1990) in the macroscopic solution scheme smoothens out the solution, and therefore counteracts instability mechanisms. Indeed, numerical dispersion becomes larger as the discretisation steps increase. Figure 5-9 reveals another difference between the micro and macro implementation that supports this conclusion. The figure shows micro and macro simulations of the same theoretically unstable situation. Both implementations predict the stability of the solution correctly, but the macroscopic solution is smoother than the microscopic one. Note that with a finer discretisation – that reduces numerical dispersion – also the difference between macro and micro solutions decreases.

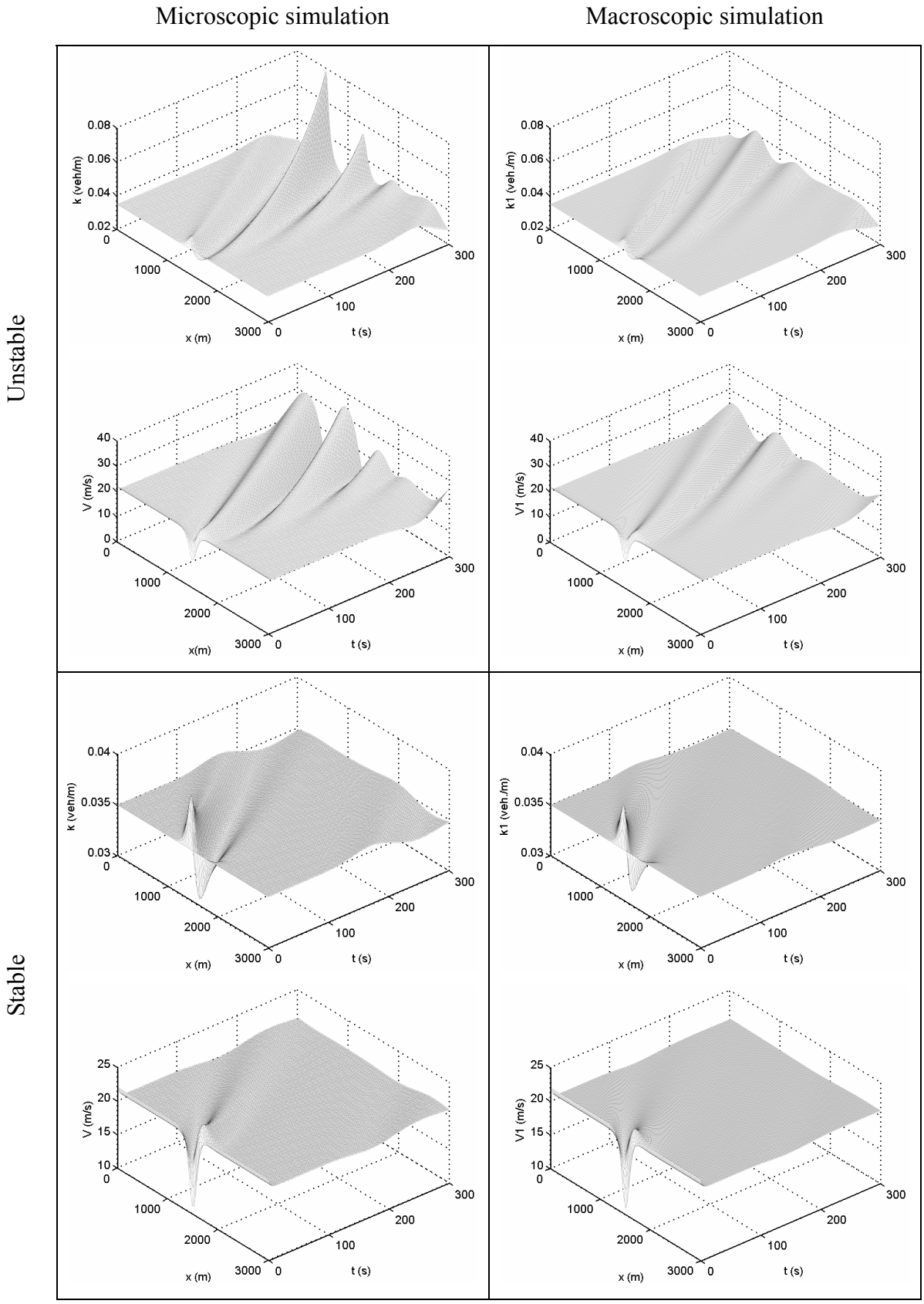


Figure 5-8 Stability and instability of microscopic and macroscopic implementation; shown are the evolution of density (upper graph of each set) and

speed (lower graph of each set) in space and time for the same initial conditions, but stable or unstable parameter settings

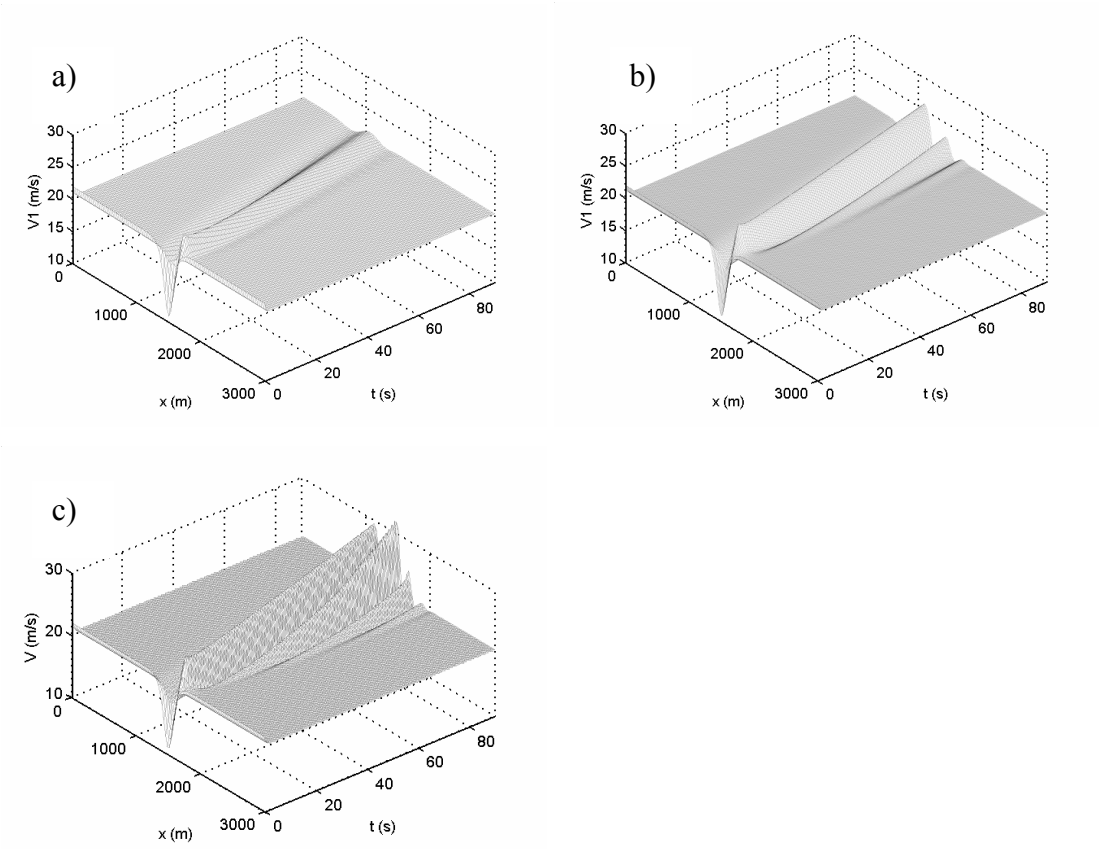


Figure 5-9 (a) macroscopic implementation of unstable traffic flow with an initial speed disturbance, larger time and space discretisation (0.5 s ; 30 m); (b) same case with smaller time and space discretisation (0.125 s ; 7.5 m); (c) microscopic implementation of the same case

Although no full proof was given, we conjecture that the deterministic microscopic car-following model and the deterministic macroscopic human-kinetic model are mathematically equivalent. However, in the numerical solution methods, differences between micro and macro solutions emerge: the macroscopic model is smoother, and therefore more stable than the theoretical prediction and the micro model. Since this difference disappears with smaller discretisation steps in the macro numerical solution scheme, the conclusion must be that this is a *numerical* effect, probably caused by numerical dispersion. Therefore, the conclusions are that:

(i) the microscopic and macroscopic models are theoretically equivalent,

(ii) numerical macroscopic solutions are optimistic: stable solutions near the border of instability might in reality be unstable, and instabilities in the numerical model might in reality be stronger than in simulation,

(iii) better macroscopic numerical solution schemes are desired and

(iv) the numerical macroscopic model is suited for sensitivity analysis and explorative research of different alternative control actions on traffic flow, since the sensitivity of the model for its parameters is qualitatively correct, albeit slightly distorted towards more smooth and stable solutions.

5.4 Theoretical case study

This section examines the behaviour of the human-kinetic traffic flow model near a bottleneck caused by merging traffic at an on-ramp. A variety of congestion patterns can occur, depending on the main lane flow and the on-ramp flow.

The section starts with a definition of the simulations: a description of the domain, the parameters, the initial and boundary conditions, and the discretisation of the numerical scheme. Then the simulation results are shown and discussed.

5.4.1 Simulation set-up

Let us consider a 9 km long single-lane motorway with traffic flowing from 0 to 9000 m. Centred on x=6000m, is a 250m long merging zone. The parameters used in these simulations are those of Table 5-1 and Table 5-2. The starting conditions on the main lanes are homogeneous in space with the initial density k_{main} and flow rate Q_{main} (approximately) in equilibrium. The boundary conditions on the upstream side are defined as a continuous inflow, equal to the initial equilibrium flow; the downstream boundary is free and determined by the cells immediately upstream of the boundary. From the start of the simulation, the total flow rate entering the road through the on-ramp increases from 0 to q_{in} veh/s during the 50 first seconds of the simulation (in order to avoid unrealistic discontinuities). Some congestion patterns do not emerge spontaneously. These so-called *metastable* conditions produce free flowing traffic or congested traffic dependent on the history (state evolved smoothly or transition to congestion triggered by a disturbance). A temporary increase of the inflow is used as the disturbance that triggers the transition from free flowing traffic to the congestion pattern in those cases. The scenarios discussed in the remainder of this section are summarised in Table 5-3.

The numerical solution scheme is the extended HLLE scheme described by Ngoduy et al. (2004), with the modifications of section 4.4.3.2. We use 150 discretisation points (cell length = 60m) and a time step of 1 second. The parameter c_0 that triggers a negative eigenvalue in the case of low speeds (see equation (4.70)) is 4 m/s (this is lower than the value suggested in section B.4; a lower value of c_0 is sufficient here due to the smaller cell length). The fraction p_{prior} of the numerical source term that is accounted for prior to the convective flows (see equation (4.63)) is 75%.

Table 5-3 Definition of the merging bottleneck scenarios

Case	k_{main} [veh/km]	Q_{main} [veh/h]	q_{in} [veh/h]	congestion trigger
1	15	1800	450	200 % q_{in} between t = 40 s and t = 100 s
2	23	2420	100	200 % q_{in} between t = 40 s and t = 100 s
3	23	2420	150	200 % q_{in} between t = 40 s and t = 100 s
4	23	2420	200	200 % q_{in} between t = 40 s and t = 100 s
5	20	2240	400	spontaneous
6	18	2090	600	spontaneous
7	18	2090	900	spontaneous
8	18	2090	1200	spontaneous
9	18	2090	1500	spontaneous

5.4.2 Simulation results: various congestion patterns

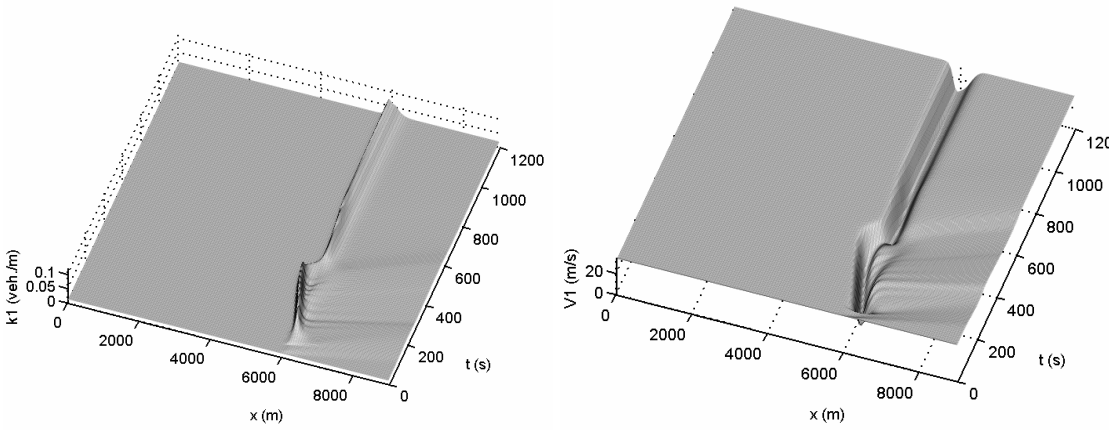


Figure 5-10 Case 1: Q_{main} = 1800 veh/h, q_{in} = 450 veh/h Pinned Localised Cluster; (a)* left*: density in space and time, (b)* right*: speed in space and time

Figure 5-10 shows the evolution of density and speed in space and time. Because the sum of main flow and on-ramp flow is by far lower than the maximal equilibrium flow rate (or *equilibrium capacity,* which equals 2550 veh/h with the parameters used in these simulations) congestion does not occur spontaneously. However, due to a temporary peak in the on-ramp flow, a stop-and-go wave is triggered that initially starts travelling upstream. Since main lane flow is not high enough to maintain this wave, it dissolves quickly, after which a localised speed drop and density cluster stabilises at the location of merging. In the terminology of Helbing et al. (1999, see also section 2.4.2), this phenomenon is indicated as a *Pinned Localised Cluster* (*PLC*). It appears when a relatively large merging flow joins a relatively small main lane flow: the flow rate on

the main lane is not high enough to carry the disturbance upstream, whereas the inflow at the merge is too high to let the cluster dissolve.

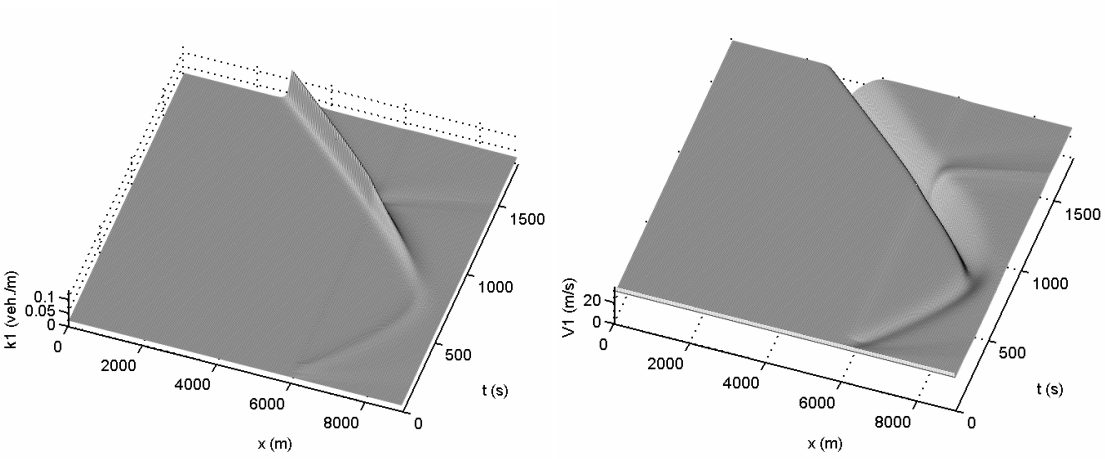


Figure 5-11 Case 2: Q_{main} = 2420 veh/h, q_{in} = 100 veh/h Moving Localised Cluster after 'boomerang'; (a)* left*: density in space and time, (b)* right*: speed in space and time

The second case is the formation of a moving density cluster, shown in Figure 5-11. The main lane flow is substantially higher than in the previous case, so that a stop-and-go wave is maintained and travels upstream, once it is triggered by some breakdown mechanism. In the terminology of Helbing, this phenomenon is called *Moving Localised Cluster* (*MLC*). The MLC is triggered by a temporary increase of the inflow at the on-ramp, which is very low otherwise. Note how, due to the temporary excess of the equilibrium capacity, a density cluster is formed that first travels downstream with traffic, and how it grows while doing so. Once the perturbation has grown enough, the direction of propagation reverses, so that finally the moving cluster travels stably upstream. This process, also reported by Helbing & Treiber (1998), is referred to as the *boomerang* effect. It is a possible explanation for the capacity funnel that was discussed in section 2.6.4. Let us remark here that the conditions in which a single MLC could be triggered had to be carefully chosen, so that in the human-kinetic model (and with the parameters used here) MLC is a rare phenomenon. In section 6.5 an alternative simulation of a capacity funnel is presented that builds on variable driver behaviour. In that case the capacity funnel can be much more reliably reproduced.

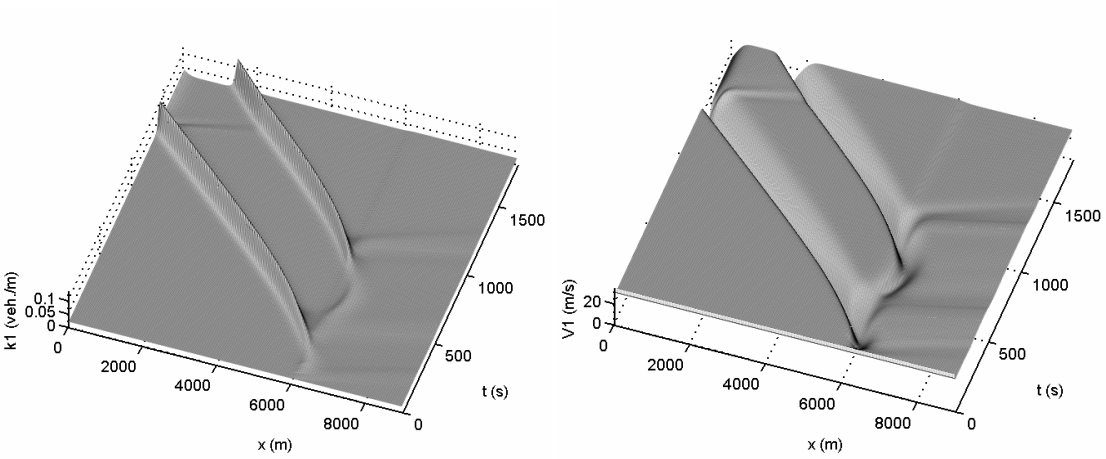


Figure 5-12 Case 3: $\mathbf{Q_{main}}$ ***= 2420 veh/h,*** $\mathbf{q_{in}}$ ***= 150 veh/h Two Moving Localised Clusters with 'boomerang'; (a)*** **left*****: density in space and time, (b)*** **right*****: speed in space and time***

As an illustration of the sensitivity of the MLC of Figure 5-11, simulation results for a slightly increased on-ramp flow (150 instead of 100 veh/h) are shown in Figure 5-12. Jam formation is triggered in the same way as the previous case, which results here in a MLC without the boomerang effect. However, when the MLC starts moving upstream of the on-ramp, it generates another density cluster that eventually grows into a second MLC. After the second MLC has emerged, the process is not repeated anymore, which is in contrast with the next case, for which the on-ramp flow is again slightly increased (to 200 veh/h) with respect to Cases 2 and 3.

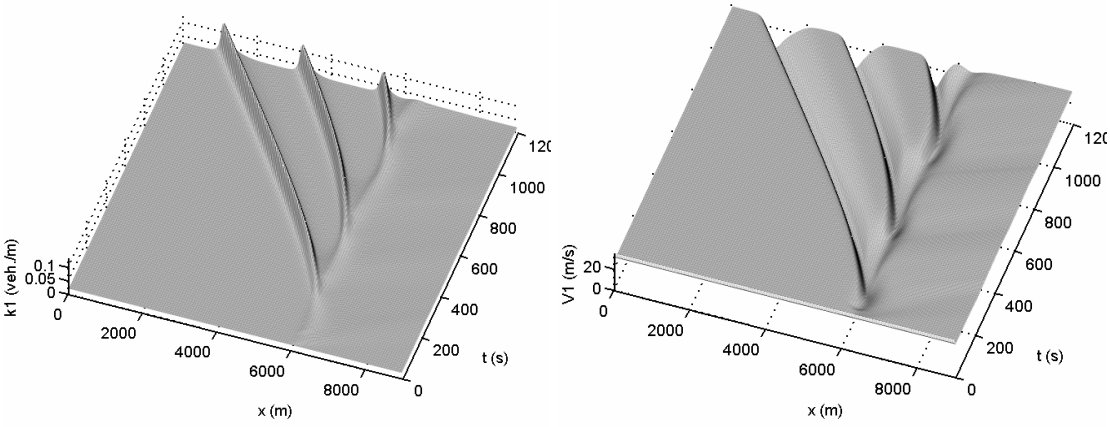


Figure 5-13 Case 4: $\mathbf{Q_{main}}$ ***= 2420 veh/h,*** $\mathbf{q_{in}}$ ***= 200 veh/h Triggered Stop-and-go waves; (a)*** **left*****: density in space and time, (b)*** **right*****: speed in space and time***

In contrast to the previous two cases, the simulation of Figure 5-13 exhibits a periodic generation of stop-and-go waves. The temporary increase of the on-ramp flow now initiates a process in which density clusters repeatedly grow into jams. Each jam travels upstream, and this process is repeated infinitely if the boundary conditions remain unchanged. In the terminology of Helbing, these are *Triggered Stop-and Go waves* (*TSG*).

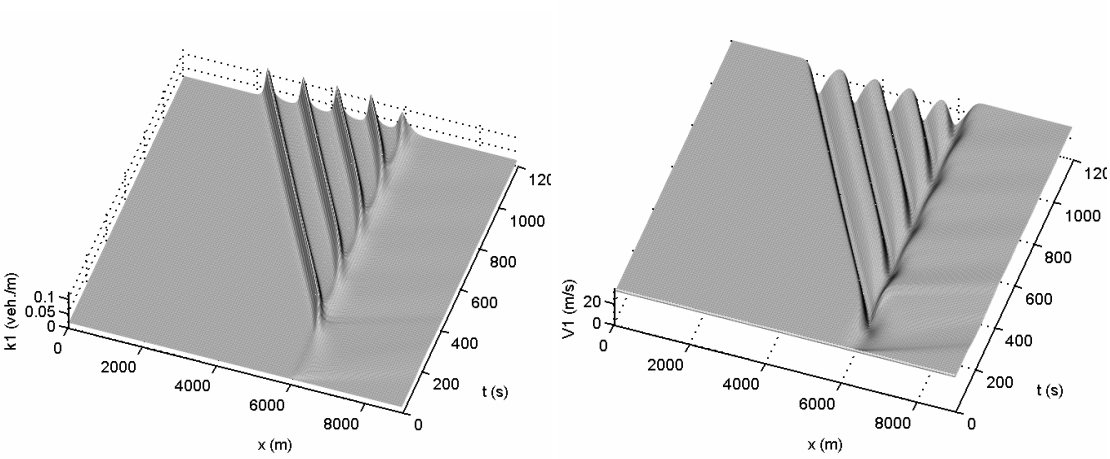


Figure 5-14 Case 5: Q_{main} = 2240 veh/h, q_{in} = 400 veh/h Triggered Stop-and-go waves shorter interval; (a)* left*: density in space and time, (b)* right*: speed in space and time

A further increase of the on-ramp flow (to 400 veh/h), shown in Figure 5-14, generates a similar TSG pattern, but the frequency of the repeated instability increases. The main lane flow has been decreased, so that the total demand for the bottleneck is more or less the same as in the previous examples. Therefore, the wave propagation speed is also lower than in Figure 5-13.

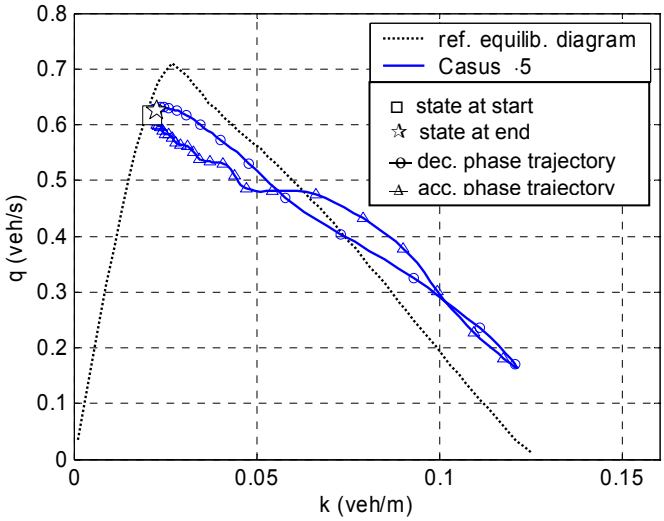


Figure 5-15 Phase trajectory in flow-density plane of one stop-and-go wave of case 5 (TSG)

Note that, in contrast to the *wide jam* (terminology after Kerner, 1999) that was studied in section 5.2.2, the decelerating and accelerating fronts of stop-and-go waves in cases 2 to 5 always travel at the same speed. This is also visible in the phase trajectory in the flow-density plane (Figure 5-15, to be compared to Figure 5-5b). As in previous plots of this type the initial state in the flow-density plane is marked with a square (□), the end state (which we have chosen here after two periods of the oscillation for clarity of the figure) is marked with a star (☆); a decelerating phase trajectory is marked with circles

(⊖) and acceleration with triangles (△). The wave speeds of up- and downstream jam fronts are essentially different in a wide jam compared to a stop-and-go wave: because traffic does not come to a standstill (or only very briefly) in a stop-and-go wave, it accelerates again starting from the maximum density in the wave *without* returning to the equilibrium density first. The acceleration branch therefore grossly follows the same, *non-equilibrium* path in the flow-density plane as the decelerating branch, which means that the wave speeds of the corresponding fronts is equal. In a wide jam these branches can have different average slopes (i.e. different wave speeds). This is possible because the deceleration wave speed depends on the upstream flow and the maximum density at standstill, whereas before accelerating drivers first wait until their minimal desired distance is restored (gross vehicle length of section 4.2.3.1). The acceleration wave therefore starts (in the flow-density plane) from the *equilibrium* jam density and ends at the queue discharge rate (basically following the path of the equilibrium flow, or *J-line* in the theory of Kerner (1999) who does not recognise the existence of an equilibrium diagram). The properties of the downstream (accelerating) jam front of a wide jam are therefore almost exclusively determined by the equilibrium curve, and thus almost independent of traffic conditions upstream of the jam. This conclusion is consistent with the empirical findings of Kerner & Rehborn (1996a) and Treiber et al. (2000), that the outflow of a wide jam has fixed properties, only related to driver behaviour and not to traffic conditions (see also section 2.6.2).

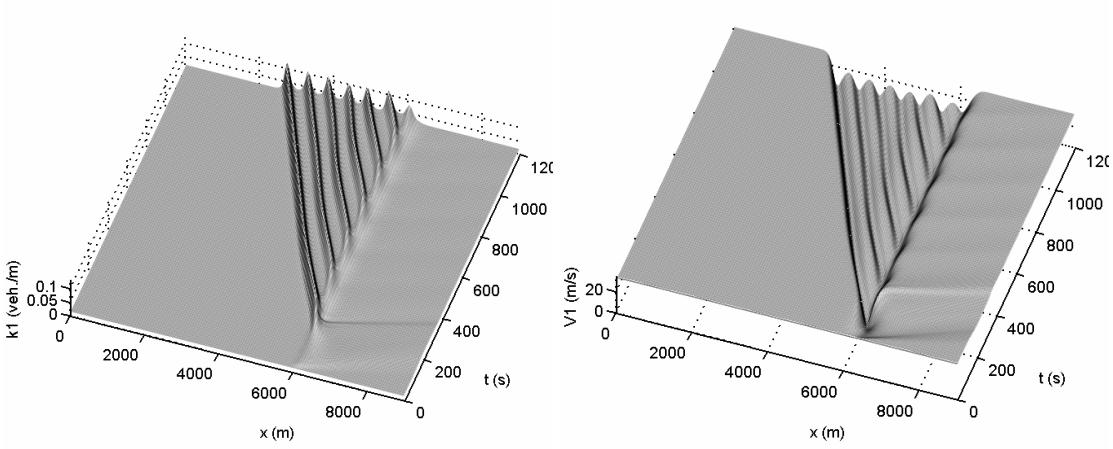


Figure 5-16 Case 6: Q_{main} = 2090 veh/h, q_{in} = 600 veh/h Transition of Triggered Stop-and-go waves to Oscillating Congested Traffic; (a)* left*: density in space and time, (b)* right*: speed in space and time

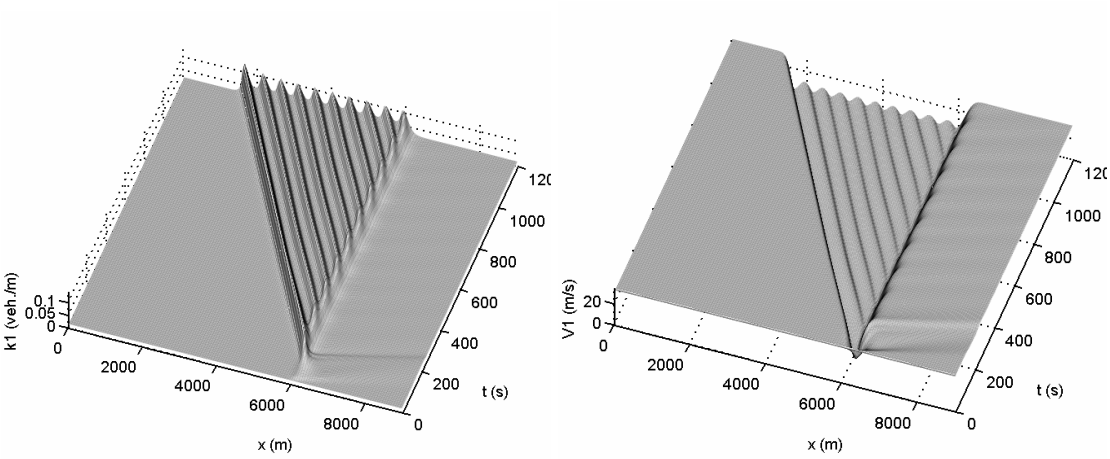


Figure 5-17 Case 7: $\mathbf{Q_{main}}$ ***= 2090 veh/h,*** $\mathbf{q_{in}}$ ***= 900 veh/h Oscillating Congested Traffic; (a)*** **left*****: density in space and time, (b)*** **right*****: speed in space and time***

With still busier traffic at the on-ramp the congestion pattern changes from TSG to *Oscillating Congested Traffic*. The difference with the TSG pattern is that when traffic drives out off a density cluster, its speed does not reach the free flow level again before arriving at the next density cluster. The on-ramp flow determines the maximum speed level inside the congested area, as can be seen by comparing Figure 5-16 with Figure 5-17.

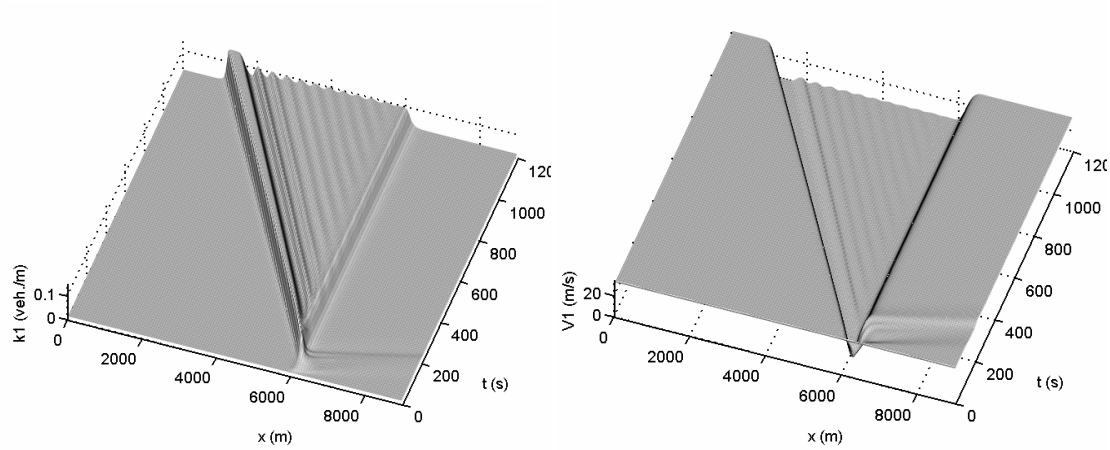


Figure 5-18 Case 8: $\mathbf{Q_{main}}$ ***= 2090 veh/h,*** $\mathbf{q_{in}}$ ***= 1200 veh/h Transition Oscillating Congested Traffic to Homogeneous Congested Traffic; (a)*** **left*****: density in space and time, (b)*** **right*****: speed in space and time***

Finally, a still further increase of the on-ramp flow rate decreases the maximum speed inside the congested area to such extent that the oscillations disappear (Figure 5-18 and Figure 5-19). The latter case can be identified in the terminology of Helbing et al. (1999) with *Homogeneous Congested Traffic* (*HCT*). Other than the simulations shown by this author, the human-kinetic model exhibits two regions of HCT: traffic entering HCT first comes to a standstill at high, non-equilibrium jam density, after which it changes to equilibrium jam density and starts moving slowly. This cycle is clearly visible in the flow-density plot of Figure 5-20. Due to the different wave speeds of the

decelerating branch and the (partial) acceleration branch, the region of totally jammed traffic grows in time into a wide jam.

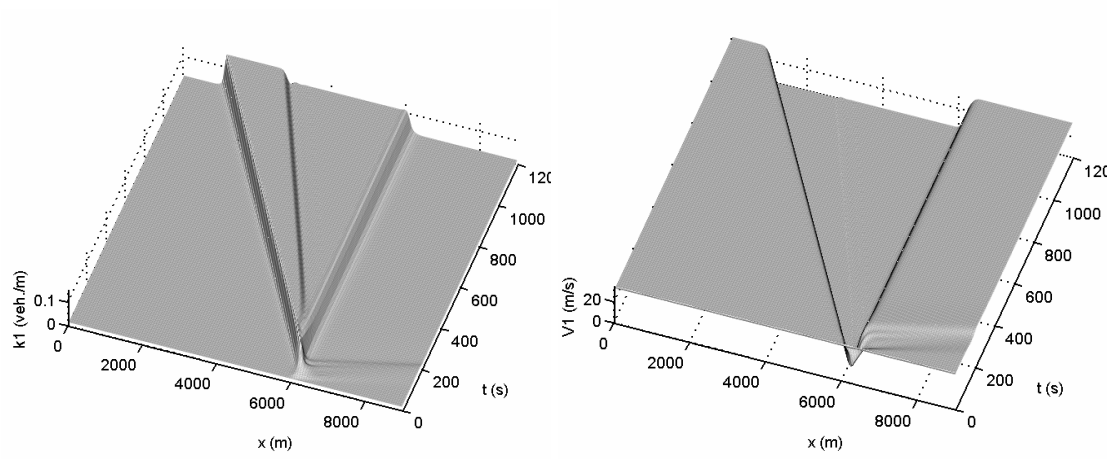


Figure 5-19 Case 9: Q_{main} = 2090 veh/h, q_{in} = 1500 veh/h Homogeneous Congested Traffic; (a)* left*: density in space and time, (b)* right*: speed in space and time

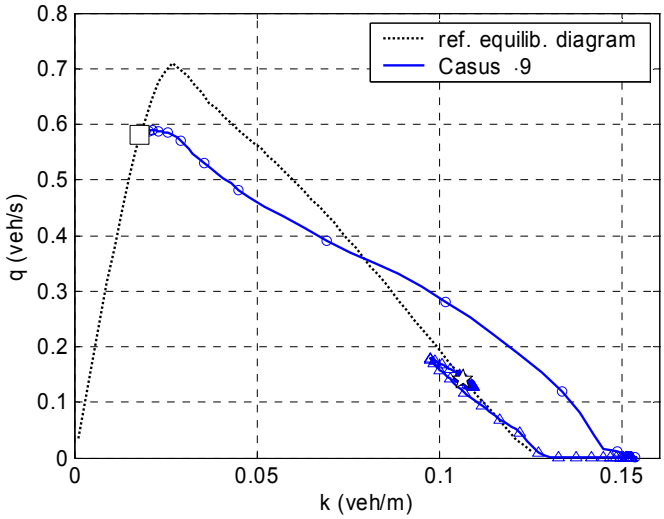


Figure 5-20 Homogeneous Congested Traffic of Case 9 in the flow-density plane (partial phase trajectory)

One aspect that is not visible from the simulation results for each case separately is the evolution of the bottleneck flow. As the on-ramp flow increases and the congestion pattern upstream of the bottleneck becomes more severe, traffic flow apparently becomes less efficient. The bottleneck flow rate is monotonically decreasing with increasing on-ramp flow (Figure 5-21). With an equilibrium capacity of 0.710 veh/s (2550 veh/h) the capacity drop increases from 2.6% in case 2 to 21% in the most severe case 9.

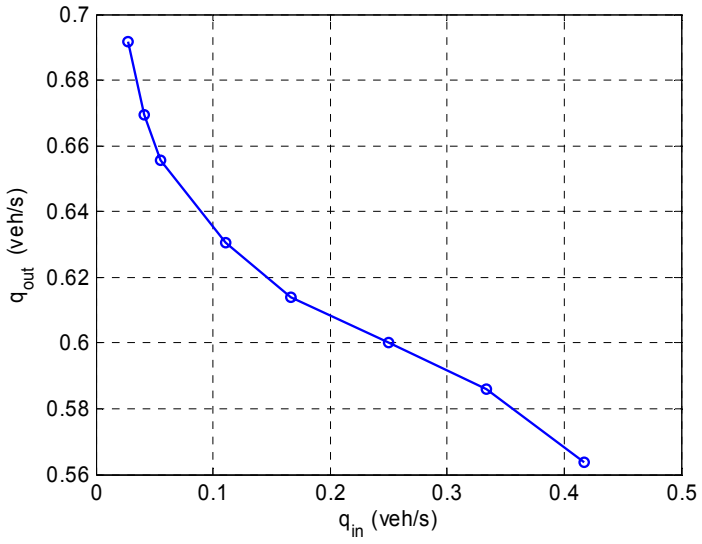


Figure 5-21 Bottleneck flow q_{out} as a function of the on-ramp flow q_{in}

5.5 Conclusions

This chapter examined the behaviour of the basic human-kinetic traffic flow model that was presented in chapter 4.

From its microscopic specification, equilibrium solutions of the human-kinetic model can be obtained, which are equivalent to the fundamental diagram or equilibrium relation in traditional traffic flow models. The fundamental diagram of the human-kinetic traffic flow model is basically triangular; the influence of most parameters causes this triangle to be smoothened to obtain a strictly concave diagram. The most determinant parameters affecting the shape of the equilibrium solution – and therefore of the first order dynamic behaviour of the model – are the car-following parameters for the desired speed, desired gap, vehicle length and minimal and maximal accelerations. The gap variance, and to a lesser extent the speed variance, further smoothen the fundamental diagram.

For non-equilibrium conditions and a model without gap or speed variance, an analytical stability criterion could be derived, which shows that apart from free-flowing traffic that is always stable, the linear stability of congested traffic states depends on parameter settings only, but not on the density itself. Moreover, the reaction time has neither effect on linear stability, and very little on jam propagation or hysteresis. Anticipation on the other hand, plays a dominant role on stability, jam propagation and hysteresis. Especially the anticipation strength – and not the anticipation distance, which has only a second order, viscosity-like effect – has strong influence, which is almost equivalent to the relaxation constant for speed difference in car-following. Interestingly the phase trajectories in the fundamental diagram correspond closely to theoretical and empirical phase trajectories (also often called: *hysteresis loops*) reported in literature: during deceleration in the tail of a jam the anticipation to traffic conditions downstream is dominant initially, after which traffic compresses to densities higher than the

equilibrium values and driver behaviour is dominated by relaxation to the immediately surrounding conditions.

Furthermore, we conclude from theoretical considerations, as well as from numerical simulations, that basically the behaviour of the deterministic microscopic car-following model and the deterministic macroscopic human-kinetic model are mathematically the same. However, although the modifications to the numerical scheme proposed in chapter 4 have the desired effect, numerical differences between micro and macro implementations could not be avoided: the macroscopic model is smoother and therefore more stable than the theoretical prediction and the micro model, probably due to *numerical dispersion*. The conclusion in this respect is that:

(i) the microscopic and macroscopic models are theoretically equivalent,

(i) numerical macroscopic solutions are optimistic: stable solutions near the border of instability might in reality be unstable, and instabilities in the numerical model might in reality be worse than in simulation,

(ii) better macroscopic numerical solution schemes are desired and

(iii) the numerical macroscopic model is suited for sensitivity analysis and explorative research of different alternative control actions on traffic flow, since the sensitivity of the model for its parameters is qualitatively correct, albeit slightly distorted towards more smooth and stable solutions.

Finally, model predictions of congestion patterns near a bottleneck caused by an on-ramp are consistent with best-practice in literature. The ratio of inflow through the on-ramp over demand on the main lanes determines whether solitary localised or moving jams, oscillating patterns, or homogeneous patterns occur. It appears that with constant inflows for the main lane and the on-ramp, both upstream travelling stop-and-go waves and wide jams can be triggered. However, the latter only occur in the most upstream part of severely congested patterns. In section 6.6, it is shown that after introducing variable driving behaviour (caused by loss of motivation at low speeds), wide jams can occur in other, less severely congested patterns as well.

In general, the results presented in this chapter (and in the parameter sensitivity analysis of Annex B) show that the model behaviour is governed to a large extent by a subset of the model parameters, with the other parameters refining the results, sometimes rather subtly. Furthermore, both the equilibrium and non-equilibrium solutions are consistent with the microscopic model foundation. Moreover, dynamic first order and higher order behaviour like jam propagation, hysteresis analysis, jam patterns and capacity drop is comparable to the best models available in literature. However, due to the differences in mathematical structure that were pointed out in the previous chapter, the numerical evaluation of the human-kinetic model requires improvement. As a result, numerical simulation results are qualitatively sound, but slightly smoother – and therefore possibly more stable – than the theoretical continuous solution. In conclusion, the results of this chapter show that the basic human-kinetic traffic flow model of chapter 4 is a sound basis for further model development in the next chapters.

6 ACTIVATION LEVEL AS A BASIS FOR VARIABLE DRIVING STRATEGIES

6.1 Introduction

In the derivation of the basic human-kinetic traffic flow model in chapter 4, a number of parameters have been introduced that relate to driver behaviour. So far, the implicit assumption is that none of these parameters is variable in time or space. However, this is a static and mechanistic approximation of reality: probably behavioural parameters like reaction time, anticipation strength, perception distance, desired following gap, and feedback control strength are all subject to changes over time and along the road, dependent (among other factors) on traffic conditions. The micro-economic theory of Van Toorenburg (1983b) postulates that drivers can be motivated to adopt a more active driving strategy (shorter gaps, sharper acceleration and deceleration,…) if this leads to a gain in efficiency or safety and vice versa. For instance, when driving in prolonged congestion, a driver might change to a driving strategy with slacker control feedback, compensated by a larger desired gap. Or when approaching a regular bottleneck, a familiar driver might expect the onset of congestion and increase his anticipation strength, especially when a queue tail warning system warns for speed breakdown. Modelling such variable driving strategies is not only a theoretical issue. The fact that the behavioural parameters of existing traffic flow models (that do not account for variable driving strategies) need to be recalibrated for every application to a new traffic network, indicates that behaviour of traffic be dependent on local conditions in a way

that is not described in existing traffic flow models so far. Extension of a traffic flow model with variable driver behaviour specifications is therefore also of practical importance, since it paves the way towards a practical model based on behaviour that is more generically valid and thus more easily transferable to other locations.

Thanks to the microscopic foundation of the human-kinetic traffic flow model, all behavioural aspects are present as individual behavioural parameters or specifications. In this chapter an additional microscopic state variable is introduced – and a corresponding macroscopic variable – that allows modelling variations of (specific aspects of) driving dependent on traffic conditions. For now, *one single* behavioural parameter is considered: the *activation level* of the driver (see discussion below). The intention of this chapter is to show how behavioural parameters and variable driver behaviour – that have been neglected so far in the development of traffic flow models by engineers or physicists – can be accounted for in macroscopic traffic flow models. Futhermore, this chapter intends to illustrate the potential influence of variable driver behaviour on traffic flow dynamics.

Choice of the behavioural variable 'activation level' and its potential alternatives

The intention of this chapter is to illustrate the added value of considering variable driver behaviour in traffic flow models, and how this can be achieved by characterising the internal state of the driver by means of a behavioural variable. The choice of a relevant behavioural variable (or a combination of more than one variable) requires a separate study and knowledge of behavioural psychology. It is not the ambition of this thesis to perform such study.

Instead, a one-dimensional behavioural variable is chosen and denoted as the *activation level*, due to its loose resemblance to the well-established term *activation* (see for example De Waard, 1996). Note that *activation level* here is considered as a more generic concept:

> *Within this thesis the term activation level denotes the one-dimensional (driver) state variable that explains temporary (short-term), traffic condition dependent, deviations of the current driving strategy with respect to the average driving strategy of the same subject.*

Note that we herewith exclude differences between subjects caused by, for example, age, physical condition, and gender, as well as longer-term variations within subjects, such as effects of time of day, impairement due to fatigue, drug or alcohol abuse, etcetera: the scope is limited to short-term variations caused by traffic conditions or alerts within the traffic system, like those from in-car systems, traffic control devices (e.g. VMS), or traffic signs.

Within the context of behavioural psychology, it might in some specific cases be more correct to use the terms *alertness, attention, awareness, motivation*, or other terminology, instead of *activation level*. However, rather than suggesting a background in behavioural psychology which is beyond our specialism and scope, we adhere to the just-mentioned and intentionally vague definition of the activation level.

Moreover, the internal state of the driver can in reality probably not be described by a *single* behavioural parameter, but rather by a vector of multiple behavioural parameters. For now, however, it is sufficient to illustrate our method and its impact on traffic flow dynamics for a behavioural variable with one dimension. The human-kinetic theory described in this chapter can in principle be extended to any required dimension.

Structure of the chapter

The chapter is structured as follows. First, a brief review is given of empirical and experimental evidence for traffic condition induced variations of driver behaviour (section 6.2). Section 6.3 continues with the mathematical description of variable macroscopic traffic flow behaviour. This section describes in generic terms how activation level is a property of drivers that flows with traffic, whereas in section 6.4 the role of activation level in variable individual driver behaviour is specified. These specifications are applied in sections 6.5 and 6.6, where two case studies are shown: one of the capacity funnel and one of the capacity drop. Finally, some conclusions are drawn about this chapter in section 6.7.

6.2 Empirical and experimental indications for traffic condition induced variations of driver behaviour

In literature, evidence is present that driving behaviour is subject to variations. The intention of this section is to indicate which aspects of driver behaviour are most likely to vary during driving. Although an exhaustive review is not pursued, it should convince the reader that driver behaviour throughout the journey is in reality indeed *not* characterised by constant behavioural rules and parameters. For more details and references, the reader is referred to the authors mentioned in this section, who more elaborately treat this subject.

6.2.1 Indirect indications: empirical traffic flow features with variable driver behaviour as most plausible cause

In this section, *indirect* indications for variable driving behaviour are discussed. In these studies the behaviour of individual drivers is not directly examined; the authors made empirical observations at the traffic flow level. They suggested variable driver behaviour as the cause underlying the macroscopically observed phenomena. Most indirect observations suggest that the following distance maintained by drivers in dense traffic is variable. Four examples of such observations and corresponding behavioural explanations are given in this section.

Muñoz & Daganzo (2001) report the observation of a moving bottleneck at different successive locations on a two-lane motorway. A slow moving truck partially obstructed the right lane, and faster traffic queued up after this moving bottleneck. Whereas the speed of the truck, and hence of the queue behind it, was constant, the flow rates in the queue gradually decreased from initially high values to lower values as the queue grew longer. This can only be explained by a decrease of the density (and therefore an

increase of the headways) inside the queue in the course of time. The authors suggest that the first drivers caught behind the slow truck were motivated to follow their predecessor closely, since they saw the cause of the queue and knew they would be able to overtake the truck soon. As the queue grew longer, drivers further upstream were less motivated and adopted longer headways.

Another piece of indirect evidence for temporary acceptance of shorter headways is the occurrence of the capacity funnel, observed for instance by Cassidy & Bertini (1999a). These authors found that the onset of congestion occurred approximately 1 km downstream of a busy on-ramp. Temporary acceptance of shorter headways is a potential explanation for this phenomenon, although other explanations are possible (see section 2.6.4). Both the drivers entering the motorway through the on-ramp and the drivers on the main lanes might be motivated to temporarily accept shorter headways. The former can more easily find a gap to merge into, whereas the latter were more or less forced onto the medium and leftmost lanes by merging traffic at the on-ramp, and drivers had to accept short headways to keep traffic flowing.

The capacity drop and hysteresis effect (see section 2.6.1) have also been the inspiration for behavioural hypotheses. Daganzo (2002a) suggests that the transition from high prior-congestion flows to lower queue discharge rates corresponds to a loss of motivation. The author argues that before the transition to congestion multilane traffic flow essentially splits into a faster and slower flow (*rabbits* versus *slugs* or *two-pipe* regime), with the rabbits driving on the leftmost lanes. When traffic becomes denser, the speed difference disappears because the frequency with which some faster slugs change lanes to the fast 'rabbit' lanes increases, the speed in the fast lane decreases, and therefore the rabbits loose their motivation to follow each other closely. Once the speed difference has disappeared, the initially faster rabbits now occupy both the faster and the slower lanes and congestion is maintained during a longer period. Essential in this theory – which is based on numerous macroscopic empirical observations – is the loss of motivation and according expansion (longer headways) of the initially densely packed faster vehicles. Although using different terminology, Zhang & Kim (2001) propose a similar behavioural theory in an attempt to reconstruct and explain the reversed-lambda shape and hysteresis of the fundamental diagram. They propose a multiphase relationship between the *response time* (equivalent to the desired time headway) and the available gap. As traffic gets denser and the available gap decreases, the corresponding headway becomes smaller without loss of speed. As long as the speed is maintained and density is increased only by traffic merging into the densely packed lane, drivers accept ever decreasing headways. However, once the speed drops (congestion sets in), drivers loose their motivation and the desired headway jumps to a higher value. Moreover, an even larger headway is required before drivers accelerate again, which causes hysteresis in the system. Again, these authors need assumptions on driver behaviour to obtain macroscopic traffic behaviour consistent with observations; the minimal set of assumptions that yields realistic macroscopic behaviour includes *multiphase* behaviour, i.e. driving strategy dependent on traffic conditions.

The traffic theory of Kerner (1999; see sections 2.4.2 and 3.2.1) is based on the empirical observation of synchronized traffic. The characteristic property of this regime

is that it covers a two dimensional area in the flow-density plane. This is equivalent on the individual driver level to saying that drivers accept any car-following distance within some range for a given speed. This is also reflected in the microscopic model proposed by the author that assumes to a certain extent *indifference* of drivers with respect to the headway in the synchronized traffic regime. Again, this author recognizes that macroscopic empirical observations are an indirect observation of variable driver behaviour that depends on traffic conditions.

The conclusion from this short review is that, for macroscopic traffic behaviour to be consistent with empirical observations, variable individual driver behaviour is a prerequisite. In some cases, like for instance the capacity funnel or capacity drop, variable driver behaviour is one of multiple possible explanations for empirical traffic flow observations. Nevertheless, attempts to account for this have been scarce and mainly explorative.

6.2.2 *Direct indications: experiments, theories and measurement of variable driver behaviour*

In this section some studies are referenced that form *direct* evidence for variable driving behaviour. The discussion is limited to longitudinal behaviour, more specifically the choice of (desired) headway and anticipation behaviour.

Van Winsum & Heino (1996) report that individual drivers maintain a time headway that is independent of vehicle speed (i.e. they confirm the desired distance relation of equation (4.28) with only the linear term). In addition, Van Winsum (1998) points out that the headway also depends on the driver state, the visual conditions, the mental effort, and attention the driver is willing to pay to the lead vehicle.

Dijker et al. (1999) analysed differences in car-following behaviour between congested and non-congested flow with data from two sites on Dutch motorways. They found that – at the same speeds – passenger car drivers follow with smaller headways in non-congested than in congested conditions. Car-following of truck drivers does not show differences between regimes. The authors suggest that the driver's motivation to follow closely or take over slower cars is lower in congested conditions, causing a different driving strategy.

On a shorter time-scale, car-following behaviour is variable in transient conditions when vehicles cut in ahead of a driver. Driver response with respect to the predecessor is governed by different models and parameters as during undisturbed car-following. Sultan et al. (2002) studied car-following behaviour on motorways immediately after cutting in ahead of the subject. Depending on the criticality of the cut-in situation drivers restore more or less fiercely their desired following distance. On average it takes approximately 20 seconds to restore the desired following distance, indicating that the average response to cut-in manoeuvres is smoother than when the car-following parameters during undisturbed car-following would apply. These findings confirm in real-life situations the earlier findings of Hogema (1995), who reported relaxation times in the order of 14 seconds in similar experiments in a controlled driving simulator experiment.

The experiments of Chang et al. (2001), conducted in a merging area on a busy motorway, suggest that these transient car-following situations can be characterized by increased psycho-physiological load for the driver. In the experiments the vigilance of merging drivers was traced starting from the approach of the on-ramp until leaving the merging zone after completion of the merging. Drivers showed highest vigilance levels during acceleration and merging, after which this level decreased but maintained significantly higher than normal for (at least) 4 seconds after entry on the main lanes (where the measurement ended, further evolution is thus not examined). Moreover, the increase of the vigilance level was strongly related to the traffic flow rate on the outermost main lane: the busier traffic, the higher the psycho-physiological load of the driver. Unfortunately, the authors do not give information on car-following behaviour parameters (such as following distance) during the experiment. However, in light of the study by Sultan et al. (2002) and Hogema (1995), a correlation seems probable between temporarily shorter headways and increased vigilance and driving effort.

Van der Hulst and co-workers conducted behavioural experiments on longitudinal behaviour that put the previous findings in a broader perspective. In a first driving simulator experiment, Van der Hulst (1999) was able to show that drivers adjusted the timing and accuracy of their control actions according to a strategy that resulted in equal minimum time headways for abrupt and gradual decelerations. The response deceleration levels were thus very well adjusted to the criticality in deceleration scenarios. Drivers were able to do so only if the decelerations of the lead car was predictable (through preview of conditions further downstream), which allows to build expectations about possible future threads (anticipation). In another experiment, Van der Hulst et al. (1998) studied driver behaviour in reduced visibility (fog). The results suggested that difficulty of anticipation (lack of preview, further than direct predecessor) was the dominant motivation to change behaviour. Interestingly, dependent on the time constraints given to the subjects, the behavioural adaptation differed: drivers who were imposed a fixed time schedule did *not* increase their headway, but reacted very accurately to decelerations of the lead car (at the cost of a higher driving effort), whereas without time constraints, drivers compensated the lack of preview by larger headways. The latter finding, that lack of preview is preferably compensated by larger safety margins, is also reported by other authors (e.g. Fuller, 1984; Saad, 1996). Van der Hulst et al. (1998) generalized the findings of their experiments, stating that drivers apply a hierarchy of adaptive strategies, which are aimed at the control of time pressure (stress) in driving. In normal conditions, drivers adopt an anticipatory driving strategy. When the possibilities for anticipation are reduced, drivers compensate by speed reductions and increased time headway. When this compensatory strategy is impossible or undesirable, drivers compensate by higher alertness to react accurately to unpredictable events. However, drivers try to avoid increased driving effort costs when they have the opportunity to do so.

Although far from exhaustive the references presented in this section suggest that:

- drivers try to keep driving effort at an acceptably low level;
- drivers therefore preferably choose a car-following distance so that a combination of preview information (anticipation) and perceived information

about the behaviour of the lead vehicle is enough to respond safely to decelerations;

- drivers can change driving strategies, for instance accept temporary reductions of the normal following distance while still being able to respond accurately to deceleration manoeuvres, at the cost of a higher alertness, activation level or vigilance;
- drivers preferably restore the temporary reductions of the acceptable headway, because they prefer driving with lower mental effort when they have the opportunity;
- these temporary variations can be short, for instance due to merging manoeuvres or cutting in of vehicles from an on-ramp or adjacent lane;
- drivers can maintain driving strategies with higher mental effort during longer periods if it is impossible (e.g. due to visibility) or undesirable (e.g. tight time schedule) to increase the headway to 'comfortable' values; an example is traffic prior to congestion, where time constraints and/or the risk of other traffic cutting into a longer (but normally more comfortable) gap make it undesirable to relax to longer headways; this motivation disappears when congestion sets in.

In this chapter assumptions will be made reflecting all (except for the last) aspects of this list and the effects on traffic flow dynamics are explored.

6.3 Variable macroscopic traffic flow behaviour: advection equation for the activation level

6.3.1 Formal mathematical derivation of the advection equation for the activation level

6.3.1.1 Generalisation of the state of a vehicle-driver combination

Let us recall the generic definition of the generalised density $\rho(t,x,S)$. So far, the local vehicle state S was identified with the individual speed v only. Now, the state of an individual vehicle-driver combination is defined by the combination of the speed v and the activation level a of the driver:

$$S \equiv (v,a) \qquad S' \equiv (x,v,a) \qquad \rho(t,S') \equiv \rho(t,x,v,a) \tag{6.1}$$

Within this thesis, the activation level a is considered as a one-dimensional numerical variable. However, following the approach outlined in this chapter, in theory an unlimited number of behavioural variables could be included in the model. With these definitions, the generalised density $\rho(t,x,v,a)\ dx\ dv\ da$ defines the expected number of vehicles with speed $v \in [v,v+dv)$ and activation level $a \in [a,a+da)$ at time t and location $x \in [x,x+dx)$.

Using the generic definition of the moments given by equation (A.14), the 1st order activation level moment or total activation level of drivers in the flow at time t and location x is defined as:

$$\begin{aligned} M_a^1(t,x) &\equiv \int_a \int_v a\,\rho(t,x,v,a)\,dv\,da \\ &= k(t,x)\,A(t,x) \end{aligned} \tag{6.2}$$

The average activation level $A(t,x)$ is defined:

$$A(t,x) \equiv \frac{1}{k(t,x)} \int_a \int_v a\,\rho(t,x,v,a)\,dv\,da \tag{6.3}$$

In general, the generalised density function $\rho(t,x,v,a)$ can be any type of probability density. Just like the speeds and gaps of successive vehicles, correlations between the activation level a and the driving speed v are conceivable. It would, for instance, not be illogical to assume a positive correlation between faster and more active driving and adopt a bivariate correlated distribution function for $\rho(t,x,v,a)$. As a first approximation however, we extend the so-called vehicular chaos assumption (see sections 3.4.1 and 4.2.3.1), and assume that the activation level probability distribution is *independent* of the speed distribution, that is:

$$\begin{aligned} p_a(a|t,x,v) &= \frac{\rho(t,x,v,a)}{\int_a \rho(t,x,v,a)\,da} \\ &= p_a(a|t,x) \end{aligned} \tag{6.4}$$

Moreover, a second simplification is made, purely for notational convenience: the variance of the activation level among the driver population is assumed to be always zero, so that the activation level of individual drivers always equals the average value of the flow:

$$p_a(a|t,x) = \delta(a - A(t,x)) \tag{6.5}$$

In this equation the notation δ is used for the Kronecker delta. The sole purpose of this approximation is to avoid unnecessarily complex notations. For instance, in general the acceleration integral of chapter 4 would be calculated according to the following extension of equation (4.11):

$$\left\langle \frac{dv}{dt} \right\rangle_{v,a} = \int_{a_j} p_a(a_j) \int_{v_j} p_v(v_j) \int_{s_j} p_s(s_j) \int_{v_{j-1}} p_v(v_{j-1})\,\dot{v}_j(a_j,v_j,s_j,v_{j-1})\,dv_{j-1}\,ds_j\,dv_j da_j \tag{6.6}$$

which is now simplified because of the assumption (6.5) to:

$$\left\langle \frac{dv}{dt} \right\rangle_{v,a} = \int_{v_j} p_v(v_j) \int_{s_j} p_s(s_j) \int_{v_{j-1}} p_v(v_{j-1})\,\dot{v}_j(A,v_j,s_j,v_{j-1})\,dv_{j-1}\,ds_j\,dv_j \tag{6.7}$$

In general, the simplification implies that for any function $\varphi(a)$ of the activation level, the average value of that function equals the function value of the average activation level A:

$$\int_a p_a(a)\varphi(a)da = \varphi(A) \tag{6.8}$$

6.3.1.2 The advection equation for the activation level A

The dynamic macroscopic equation for the activation level A is obtained using the method of moments defined in section A.3.2 of Annex A. The 1st order activation level moment dynamics are given by:

$$\int_a\int_v a\left(\frac{\partial\rho}{\partial t}+\frac{\partial}{\partial x}\left(\rho\frac{dx}{dt}\right)+\frac{\partial}{\partial v}\left(\rho\frac{dv}{dt}\right)+\frac{\partial}{\partial a}\left(\rho\frac{da}{dt}\right)\right)dv\,da=\int_a\int_v a\left(\frac{d\rho}{dt}\right)_{discrete}dv\,da \tag{6.9}$$

The elaboration of the different terms in equation (6.9) is omitted, since it is completely analogous to the elaboration of the corresponding terms in the speed moment equation (A.15) in Annex A. The calculation yields:

$$\frac{\partial kA}{\partial t}+\frac{\partial kAV}{\partial x}=k\left\langle\frac{da}{dt}\right\rangle_{v,a}+\int_a\int_v a\left(\frac{d\rho}{dt}\right)_{discrete}dv\,da \tag{6.10}$$

In this equation we use the average smooth change of the activation level, given by:

$$\left\langle\frac{da}{dt}\right\rangle_{v,a}=\frac{1}{k}\int_a\int_v\rho\frac{da}{dt}\,dv\,da \tag{6.11}$$

Equation (6.10) is the *conservative* equation for the activation level (moment). It can also be cast in a *primitive* formulation, yielding:

$$\frac{\partial A}{\partial t}+V\frac{\partial A}{\partial x}=\left\langle\frac{da}{dt}\right\rangle_{v,a}+\frac{1}{k}\int_a\int_v a\left(\frac{d\rho}{dt}\right)_{discrete}dv\,da-\frac{A}{k}\left(\frac{dk}{dt}\right)_{discrete} \tag{6.12}$$

This equation states that the average activation level A is a property of traffic that is transported with the speed of the flow V. Equations (6.10) and (6.12) are therefore called the *advection* equations for the activation level moment and activation level respectively. Causes for changes of the activation level other than pure advection are represented by the terms in the RHS.

The first term of the RHS is the gross effect of smooth individual changes of the activation level a, averaged over all possible states (v,a) at t and x of vehicles in the flow. The second term is similar to the first one, only it contains the gross effect of sudden or discrete changes of the individual activation level a due to certain events. Note that conceptually the difference between the two types of variations is non-existing, only their time scale is different: the first type of change takes finite time to accomplish (compared to the time scale of the flow under consideration, see section 4.1.1 for a related discussion), and the time duration of the second type can be neglected. Finally, the last term in equation (6.12) represents the change of the average activation level A due to vehicles that might be flowing in or out the stream. Note that it does *not* stand for the amount of activation level that is carried in or out of the flow along with the in- or outflow of drivers (this has to be accounted for in the first or second term!). It stands for the change of the *average* activation level per driver, if the total amount of activation level would remain constant, while at the same time the

population of drivers over whom it is distributed grows or shrinks. This term therefore only contributes near merges or diverges of traffic, where $\left(\frac{dk}{dt}\right)_{discrete}$ is not equal to zero.

The three terms in the RHS are further specified in section 6.4.2.

6.3.2 *Numerical evaluation of the state variable 'activation level'*

Like in section 4.4, let us summarise the system of partial differential equations that together form the activation level-based human-kinetic traffic flow model (in conservative formulation):

$$\frac{\partial k}{\partial t}+\frac{\partial k V}{\partial x}=\left(\frac{dk}{dt}\right)_{discrete} \tag{6.13}$$

$$\frac{\partial kV}{\partial t}+\frac{\partial k\left(V^2+\Theta^e\right)}{\partial x}=k\left\langle\frac{dv}{dt}\right\rangle_{v,a}+\int_a\int_v v\left(\frac{d\rho}{dt}\right)_{discrete}dv\,da \tag{6.14}$$

$$\frac{\partial kA}{\partial t}+\frac{\partial kAV}{\partial x}=k\left\langle\frac{da}{dt}\right\rangle_{v,a}+\int_a\int_v a\left(\frac{d\rho}{dt}\right)_{discrete}dv\,da \tag{6.15}$$

In vector notation, this system reads:

$$\frac{\partial}{\partial t}\mathbf{U}+\frac{\partial}{\partial x}\mathbf{F}(\mathbf{U})=\mathbf{G}(\mathbf{U}) \tag{6.16}$$

with

$$\mathbf{U}=\begin{pmatrix}u_1\\u_2\\u_3\end{pmatrix}\equiv\begin{pmatrix}k\\kV\\kA\end{pmatrix} \tag{6.17}$$

$$\mathbf{F}(\mathbf{U})=\begin{pmatrix}f_1(\mathbf{U})\\f_2(\mathbf{U})\\f_3(\mathbf{U})\end{pmatrix}\equiv\begin{pmatrix}u_2\\\frac{u_2^2}{u_1}+u_1\quad\Theta^e(u_1,u_2)\\\frac{u_2\quad u_3}{u_1}\end{pmatrix} \tag{6.18}$$

$$\mathbf{G}(\mathbf{U})=\begin{pmatrix}g_1(\mathbf{U})\\g_2(\mathbf{U})\\g_3(\mathbf{U})\end{pmatrix}\equiv\begin{pmatrix}\left(\frac{du_1}{dt}\right)_{discrete}\\u_1\left\langle\frac{dv}{dt}\right\rangle_{v,a}+\int_a\int_v v\left(\frac{d\rho}{dt}\right)_{discrete}dv\,da\\u_1\left\langle\frac{da}{dt}\right\rangle_{v,a}+\int_a\int_v a\left(\frac{d\rho}{dt}\right)_{discrete}dv\,da\end{pmatrix} \tag{6.19}$$

It appears that the newly introduced variable u_3 does not occur in the definition of the first two rows of $\mathbf{F}(\mathbf{U})$. This means that the convective part (LHS) of the density and speed equation is unchanged, compared to the basic human-kinetic model of chapter 4. Therefore, also the mathematical structure, the eigenvalues, and thus the hyperbolic nature are unchanged. Then, also the same numerical solution scheme – the extended HLLE scheme described in Ngoduy et al. (2004) with the modifications presented in section 4.4.3.2 – remains valid. Note that this is also why the activation level equation is called an *advection* equation: it flows along with traffic but it is uncoupled from the convective part of the variables *k* and *V* that define the physical flow. Note however, that through the definition of the non-convective terms in the RHS, the activation level may have influence on the flow. More precisely, we specify in section 6.4.1 how the speed component $g_2(\mathbf{U})$ depends on the activation level.

6.4 Variable microscopic driver behaviour and activation level-based model specifications

6.4.1 Influence of the activation level on driving behaviour

6.4.1.1 General specification of activation level dependent driver behaviour

The driver behaviour is contained in the component $g_2(\mathbf{U})$ of equation (6.19). Any influence of the activation level of the driver on driving performance has to be accounted for through the terms of $g_2(\mathbf{U})$. It appears from the qualitative review of section 6.2 that the activation level can have influence on:

- reaction time
- anticipation distance
- anticipation strength
- lane changing decisions
- desired following distance
- desired speed
- strength of the control actions (feedback control gains)

It is not our intention to specify validated quantitative behavioural models for the influence of the activation level on driver behaviour. We rather indicate generically where and how the influence of the activation level can be quantified. Moreover, some hypothetical specifications of variable driving are given to illustrate how these specifications can be elaborated, and which effects this may have on macroscopic traffic flow behaviour. These specifications are chosen in such a way that they serve as an (alternative) explanation for some typical congested traffic phenomena mentioned in chapters 2 and 5.

In general, the acceleration integral becomes a function of the activation level A also, instead of only of the density k and the speed V. Equation (5.1) is therefore adapted as follows:

$$\left\langle \frac{dv}{dt}\left(t+T,x\left(t+T\right)\right)\right\rangle_{v} = f\left(k\left(t,x\right),k\left(t,x+\frac{1}{2\ \ k\left(t,x\right)}\right),V\left(t,y\right),A\left(t,x\right)\right)$$

with: (6.20)

$$y\in\left[x,x+\Delta x_{ant}\left(A\right)\right]$$

Note that in principle, the anticipation distance Δx_{ant} can also be a function of the activation level. Theoretically the activation level could have an influence on the distribution functions for the speed, for the gap to the predecessor, and for the speed of the predecessor, especially in potential correlations between these microscopic variables and the activation level. However, for now we have neglected correlations between the microscopic variables. Therefore, a potential role of the activation level is assumed only in the determination of the average for the speed distribution of the predecessor (the anticipation mechanism); furthermore, the activation level probably plays an important role in the car-following relation that is the core of the acceleration integral.

The next sections provide some example specifications based on a hypothetical behavioural model.

6.4.1.2 Variable desired following distance

In section 4.2.3.2, the car-following relation and the desired following distance were defined as:

$$\dot{v}\left(v_j,v_{j-1},s_j\right)\equiv\min\left(\frac{w_j-v_j}{\tau_w}\ ;\ \frac{s_j-s^d\left(v_j\right)}{\tau_s}+\frac{v_{j-1}-v_j}{\tau_v}\right) \tag{6.21}$$

with:

$$\begin{aligned} &acc_{min}\le\dot{v}\le acc_{max}\\ &s^d\left(v\right)=s_1^d v+s_2^d v^2\end{aligned} \tag{6.22}$$

So far, this model does not incorporate variable driver behaviour. In this section some activation level dependent refinement to this model is introduced. Two modifications are proposed: one for lower than normal activation level and one for more active driving. Therefore, the models can be regarded as complementary models, active in disjoint ranges of the activation level (see section 6.4.2 for the quantification of activation level, and the concept of normal, minimal and maximal activation level).

Driving with lower activation level

We hypothesise that a driver who wants to drive equally safe while paying less attention to the driving task, is assumed to compensate this by adopting larger headways (see section 6.2).

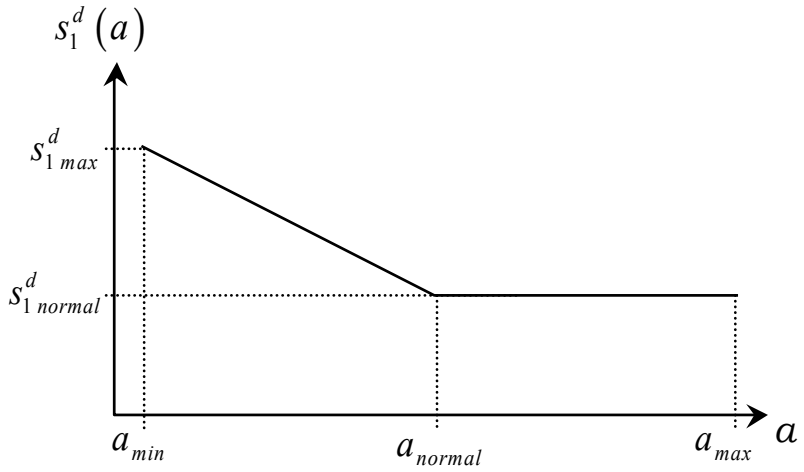


Figure 6-1 Dependence of the time headway on the activation level

In the specification of the model, this is obtained by making the desired gap function activation level dependent. More precisely, the linear term (time headway) in equation (6.22) increases with decreasing activation level:

$$s_1^d(a) \equiv \begin{cases} s_{1\,normal}^d + \left(s_{1\,max}^d - s_{1\,normal}^d\right)\dfrac{a_{normal} - a}{a_{normal} - a_{min}} & a < a_{normal} \\ s_{1\,normal}^d & a \geq a_{normal} \end{cases} \tag{6.23}$$

This specification, depicted in Figure 6-1, prescribes that the time headway is constant and equal to $s_{1\,normal}^d$, as long as the activation level is equal to or higher than the normal value, and increases linearly otherwise until the maximum value $s_{1\,max}^d$ is reached for the lowest activation level.

This specification is applied in the simulations of the activation level driven hysteresis and according capacity drop, to be discussed in section 6.6.

Driving with higher activation level

It is not sufficient to apply the same type of relationship between the higher-than-normal activation level and the time headway as was adopted for less active driving (equation (6.23)). This would mean that drivers who become more active would desire shorter gaps and would *accelerate* in order to decrease the gap with their predecessor. It is more realistic that they would continue following as before, but accept a potential decrease of the gap without necessarily reducing speed. In other words: they will not necessarily strive for shorter gaps, but will accept them if they cannot be avoided. This assumption is consistent with one of the basic assumptions underlying the three-phase theory of Kerner (1999, see section 2.4.2.2).

The refinement to equations (6.21) and (6.22) therefore consists of the introduction of an indifference band, as first described by Wiedemann (1974). We no longer consider a single (velocity dependent) desired following distance and a corrective response when the actual gap deviates from the desired one. Instead, the assumption is that a driver feels comfortable with a *range* of following distances: he is indifferent and does not

undertake corrective control action as long as the actual distance is within an *indifference band*. The indifference band is defined as the interval:

$$\left[\left(1-p_{indiff}\right)s^{d}\left(v_{j}\right),\ s^{d}\left(v_{j}\right)\right] \tag{6.24}$$

The car-following equation (6.21) (with equation (6.22) still valid) is changed in terms of a gap error function *e* as follows:

$$\dot{v}\left(v_{j},v_{j-1},s_{j}\right)\equiv\min\left(\frac{w_{j}-v_{j}}{\tau_{w}};\frac{e\left(s_{j},s^{d}\left(v_{j}\right),p_{indiff}\left(a\right)\right)}{\tau_{s}}+\frac{v_{j-1}-v_{j}}{\tau_{v}}\right) \tag{6.25}$$

with:

$$e\left(s_{1},s_{2},p\right)\equiv\begin{cases}0 & \left(1-p\right)\ s_{2}\ \leq\ s_{1}\ \leq\ s_{2}\\ s_{1}-s_{2} & elsewhere\end{cases} \tag{6.26}$$

The error function is zero for gaps within the indifference band. For gaps outside of this band the error function is defined as earlier, i.e. as the difference between the actual and the desired gap. In Figure 6-2 the error function is visualised.

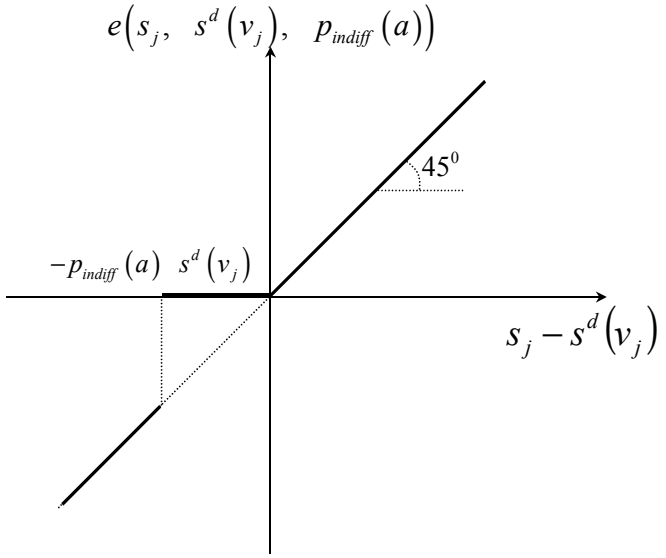


Figure 6-2 Error function and the role of the indifference band

The effect of this specification is that for gaps in the indifference band no acceleration control response is triggered (unless there is a speed difference with the predecessor). Larger gaps are closed smoothly; shorter gaps are enlarged more fiercely until they lie within the indifference band again.

The relation between activation level and the indifference band is as follows. A more active driver has a larger indifference band. He is therefore able to drive closer to his predecessor without having to decrease his speed. The higher the activation level of a driver, the shorter the gaps that are acceptable as compared to a less active driver. A

simple linear relationship is assumed between indifference band and activation level *a* (see Figure 6-3):

$$p_{indiff}(a) \equiv \begin{cases} p_{indiff\ \ normal} + \left(p_{indiff\ \ max} - p_{indiff\ \ normal}\right)\dfrac{a - a_{normal}}{a_{max} - a_{normal}} & a > a_{normal} \\ p_{indiff\ \ normal} & a \le a_{normal} \end{cases} \tag{6.27}$$

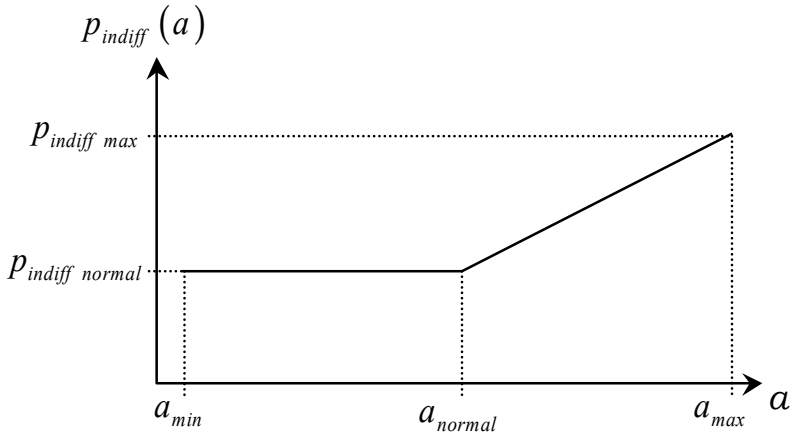


Figure 6-3 Dependence of the indifference band on the activation level

The influence of this specification can be seen in section 6.5 where this model is used for a theoretical case study of a busy on-ramp with traffic that *forces* itself into the gaps on the main lanes of the highway.

6.4.1.3 Variable anticipation strength

In section 4.3.2 the anticipation mechanism of drivers was defined. The specification stated that drivers do not respond to the actual speed of the predecessor, but rather to an anticipated or perceived speed:

$$V^{ant}\left(t, x + s_j^0 + s_j\right) \equiv V(t,x) + f^{ant}(a)\left(s_j^0 + s_j\right)\left(\frac{dV}{dx}\right)^{ant} \tag{6.28}$$

The assumption is that more active drivers anticipate stronger than less active drivers. This assumption is motivated as follows. There are two parameters in the human-kinetic model that are directly associated to anticipation: the distance Δx_{ant} over which the perceived speed gradient $\left(\frac{dV}{dx}\right)^{ant}$ is determined and the anticipation strength f^{ant}. The former is related to the perception distance of a driver and is physically limited by curvature or gradient of the road, and by the presence of vehicles obstructing the view. Therefore, it is more plausible to let the anticipation strength $f^{ant}(a)$ depend on the activation level, since it models how the driver processes the perceived information and is therefore subject to influence by the activation level. This is consistent with the experiments by Van der Hulst et al. (1998, see section 6.2.2), who found that drivers responded more accurately to decelerations of the predecessor, even though the visual distance was limited.

The higher the activation level of a driver, the further (in time) and more precise he can estimate how the speed of the predecessor will be influenced by the spatial speed profile. We let the anticipation factor depend linearly on the activation level (Figure 6-4):

$$f^{ant}(a) \equiv \begin{cases} f^{ant}_{normal} + \left(f^{ant}_{max} - f^{ant}_{normal}\right)\dfrac{a - a_{normal}}{a_{max} - a_{normal}} & a > a_{normal} \\ f^{ant}_{normal} & a \le a_{normal} \end{cases} \tag{6.29}$$

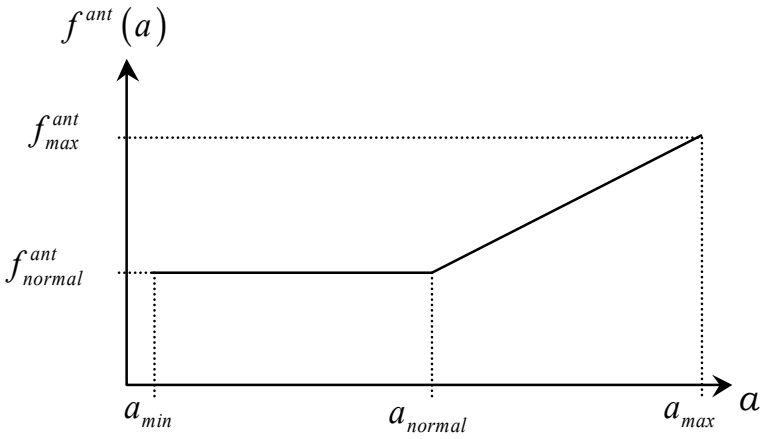


Figure 6-4 Dependence of anticipation strength on the activation level

The influence of this specification can be seen in chapter 8, where this model is applied to a theoretical case study of an ADA system designed to raise the activation level in the case of downstream speed drops.

6.4.2 *Influence of traffic conditions on activation level*

There exists no generally accepted quantification of the activation level. However, it is plausible that there is some normal or comfortable activation level a_{normal}, and that it can increase up to some maximum value a_{max}, and inversely can decrease to some minimum value a_{min}. The absolute numerical value of these levels is not important here. In this section we generally define slow and fast changes of the activation level. They are accounted for in the smooth respectively discrete terms of the third component of equation (6.19) that represents the driver behaviour with respect to the activation level (momentum):

$$g_3(\mathbf{U}) = \underbrace{u_1 \left\langle \frac{da}{dt} \right\rangle_{v,a}}_{smooth\ \ change} + \underbrace{\int_a \int_v a \left(\frac{d\rho}{dt} \right)_{discrete} dv\, da}_{discrete\ \ change} \tag{6.30}$$

6.4.2.1 *Smooth changes of the activation level*

Within the context of this thesis, it is assumed that a driver continuously strives to drive with a certain comfortable or normal activation level a_{normal}. That is, in absence of

events triggering a higher or lower activation level (section 6.4.2.2), drivers will gradually let a relax to this normal level:

$$\frac{da}{dt} \equiv \frac{a_{normal} - a}{\tau_a} \tag{6.31}$$

According to the micro-economic theory by Van Toorenburg (1983b), the normal activation level is a trade-off between gain in travel speed (which he assumes to be decreasing with higher activation level, since this is associated with a more aggressive driving strategy) and the cost of a high driving effort. Therefore, the normal activation level is an optimum lying somewhere in between the minimal and maximal activation level of a driver. This is consistent with the findings of Van der Hulst et al. (1998) discussed in section 6.2.2

Assuming that all drivers have the same relaxation time τ_a and the same a_{normal}, the gross effect of this individual relaxation behaviour is:

$$\begin{aligned}\left\langle \frac{da}{dt} \right\rangle_{v,a} &= \int_a \int_v \frac{\rho(t,x,v,a)}{k(t,x)} \frac{a_{normal} - a}{\tau_a} dv\, da \\ &= \frac{1}{\tau_a k(t,x)} \left(\int_a \int_v \rho(t,x,v,a)\, a_{normal}\, dv\, da - \int_a \int_v \rho(t,x,v,a)\, a\, dv\, da \right) \\ &= \frac{a_{normal} - A}{\tau_a}\end{aligned} \tag{6.32}$$

Note that in theory it is conceivable to also define other smooth changes of the activation level (theoretically an *activation level change integral* analogous to the acceleration integral (4.11) or any smooth specification could be considered), but we do not see plausible reasons to do so. While examining the influence of a gradually declining activation level, for example due to fatigue, this might be a sensible option. For now, any other change of the activation level is considered as a discrete event, defined in the next section.

6.4.2.2 Formal specification of discrete changes of the activation level

Certain events might cause a discrete change of the activation level of an individual driver. A sudden decrease of the activation level can be related to, for example:

- receiving a call of a cellular phone
- attention being drawn by events not related to the driving task, like accidents in the opposing driving direction, billboards, people at the side of the road,…
- other

A sudden increase of the activation level can be related to for example:

- driving past a warning sign
- a variable message sign flashing (motorway speed control, fog warning, queue tail warning,…)
- active braking lights within the range of view downstream

- vehicles cutting in the front gap
- too slow reaction to vehicle in front causing a dangerous situation
- in-car systems (ADA-systems) producing an active warning
- other

Both types of discrete transitions are modelled similarly to the discrete speed transitions in traditional gas-kinetic traffic flow models (section A.4 in Annex A). Each transition from an original activation level a_0 to a new (higher or lower) activation level a_1 causes a decrease of the generalised density at a_0 (active event) and a corresponding increase of the generalised density at the new activation level a_1 (passive event). In general, the active change rate of the activation level is defined as:

$$\left(\frac{d\rho(t,x,v,a_0)}{dt}\right)_{transition}^{-} \equiv \int_{a_1 \neq a_0} \rho(t,x,v,a_0)\,\Pi(t,x,v,a_0)\,\pi_a(t,x,v,a_0,a_1)\,da_1 \quad (6.33)$$

The symbols used in this equation are:

$\rho(t,x,v,a_0)$: the generalised density, representing the expected number of vehicles having speed v and activation level a_0

$\Pi(t,x,v,a_0)$: the interaction rate or number of changes of the activation level per unit of time for a driver driving with speed v and activation level a_0

$\pi_a(t,x,v,a_0,a_1)$: the transition rate or number of transitions per unit of time from activation level a_0 to activation level to a_1, given that a transition occurs

The interaction rate can have a non-zero value for any time t or location x. Not only can it depend on the traffic conditions, as in all applications in the remainder of this thesis, also other events at a certain time t or place x might cause a non-zero transition rate. For instance, near a warning signal at a fixed location along the road or at locations with suddenly changing conditions of visibility or weather the interaction rate can have a non-zero value for (a short range around) that specific x-location. Similarly specific transitions that only occur at a certain moment t are conceivable.

Analogously, one can see that the passive change rate equals:

$$\left(\frac{d\rho(t,x,v,a_1)}{dt}\right)_{transition}^{+} \equiv \int_{a_0 \neq a_1} \rho(t,x,v,a_0)\,\Pi(t,x,v,a_0)\,\pi_a(t,x,v,a_0,a_1)\,da_0 \quad (6.34)$$

Since the expected number of vehicles does not change due to transitions between different activation levels, and the speed v does not change during such transitions the following relation holds:

$$\int_a \left(\left(\frac{d\rho(t,x,v,a)}{dt} \right)^{+}_{transition} - \left(\frac{d\rho(t,x,v,a)}{dt} \right)^{-}_{transition} \right) da = \left(\frac{d\tilde{\rho}(t,x,v)}{dt} \right)_{transition} = 0 \tag{6.35}$$

Apart from the coupled active and passive change rates related to transitions from one activation level to another, some positive or negative discrete contribution can also be considered due to in- or outflow of drivers having a certain activation level. In that case:

$$\int_a \int_v \left(\frac{d\rho(t,x,v,a)}{dt} \right)_{in/outflow} dv\, da = \left(\frac{dk}{dt} \right)_{in/outflow} \neq 0 \tag{6.36}$$

Activation level a is carried in or out of the flow proportionally to $\left(\frac{d\rho(t,x,v,a)}{dt} \right)_{in/outflow}$.

After collecting the terms, the contribution to $g_3(\mathbf{U})$ in equation (6.19) due to discrete changes of the activation level is:

$$\int_a \int_v a \left(\frac{d\rho}{dt} \right)_{discrete} dv\, da = \int_a \int_v a \left(\left(\frac{d\rho}{dt} \right)_{in/outflow} + \left(\frac{d\rho}{dt} \right)^{+}_{transition} - \left(\frac{d\rho}{dt} \right)^{-}_{transition} \right) dv\, da \tag{6.37}$$

So far, the specification of the discrete changes of the activation level has been formal. No behavioural specifications have been formulated yet. More concrete behavioural specifications of the interaction and transition rates in equations (6.33) and (6.34) are case-specific. Therefore, a further behavioural specification is postponed to section 6.5, in which an example for discrete increase of the activation level due to merging traffic cutting into existing gaps is elaborated; another behavioural specification is found in chapter 8, where warnings generated by an ADA system are the incentives to discretely raise the activation level.

6.5 Case study (1): behavioural explanation of the capacity funnel

In sections 2.6.4 and 6.2.1, the capacity funnel phenomenon and the potential role of variable driver behaviour was discussed. Before proceeding to model specifications and simulations illustrating that variable driver behaviour near merge zones can serve as an explanation for this phenomenon, it should be mentioned that it is not the *only* potential explanation.

As discussed in section 3.3.2, delayed relaxation of traffic towards equilibrium speeds can lead to instability. When instability is induced by merging behaviour at a bottleneck, it can travel in the direction of traffic past the location of the merging before gaining enough amplitude so that a stop-and-go wave travels backward. Therefore, any Payne-type macroscopic traffic flow model can in theory establish capacity funnel effect in carefully chosen conditions. As an illustration, we mention the findings by Helbing, discussed in 2.6.4 and our own simulations with the basic human-kinetic model (Case 2 of section 5.4.2). Yet, an alternative explanation is found in the platoon

and inter-platoon gap theory for traffic flow stability discussed in section 2.3.2. Typically merging traffic fills up the inter-platoon gaps, by so doing decreasing the ability of traffic flow to recover from instability within a platoon by a sufficiently large inter-platoon gap. This does not mean, however, that breakdown occurs immediately, but the probability increases in time (see section 2.6.1), during which traffic that is highly susceptible to the slightest disturbance travels past the merging zone and congestion sets in there.

We do not falsify any of the physical explanations mentioned above in this section, but show how variable driving behaviour can contribute significantly to the occurrence and strength of the capacity funnel. In reality, both the physical and behavioural factors probably contribute together to the capacity funnel.

6.5.1 Behavioural hypotheses near merge zones

In this case study, the car-following model with the *indifference band* of section 6.4.1.2 is applied. This means that drivers with a higher activation level accept a range of following distances equal to or shorter than the gaps they would accept with a normal activation level.

This theoretical case study examines the situation of a busy merging zone. Drivers merging onto the main lanes have to accept as many gaps as possible. This maximises the probability of merging without too much effort spent to aligning to suitable gaps. We therefore assume the highest activation level, which is associated through equation (6.27) with a maximal indifference band, and therefore maximises the range of acceptable gaps. Furthermore, merging drivers are assumed to adapt their speed to the average of the adjacent main lanes before performing the actual merge. The main driving lanes thus receive an inflow of vehicles having the local expected speed and driven by drivers with the maximum activation level.

Some drivers on the main lanes are confronted with vehicles merging into the gap with their original predecessor. A vehicle with approximately the same speed as the original predecessor fills the gap, so that the original gap is more than halved. Most traditional car-following models of the form of equation (4.10), for instance the one adopted in the human-kinetic model (equation (4.27)), would as a result immediately reduce speed to restore the desired gap. The assumption is that instead, the *activation level* is immediately raised, so that the indifference band is increased and the driver accepts shorter gaps *without* having to brake.

Both merging drivers and drivers in front of whom vehicles have merged, have a high activation level as a result of the merging process. As they leave the merging zone the relaxation towards the normal comfortable activation level makes them gradually restore the normal desired gaps. If density is high enough, this might cause a delayed transition to congested traffic building up from downstream of the merging zone.

These behavioural hypotheses are specified in the next section.

6.5.2 Behavioural specifications

The aim of this section is to provide model specifications for the changes of the activation level for inclusion in the dynamic activation level equation (6.37) for traffic in the main lane. These changes are partially due to activation level carried into the main flow by merging traffic (section 6.5.2.1) and by changes of lower to higher activation level by some drivers in the main lane traffic (section 6.5.2.2). Finally these specifications are synthesised in the activation level equation in section 6.5.2.3.

6.5.2.1 Merging traffic

The merging traffic enters the main lanes, having the maximum activation level a_{merge} (=a_{max}). Therefore, the activation level moment carried into the flow by these drivers is:

$$\int_a \int_v a_{merge} \left(\frac{d\rho}{dt} \right)_{inflow} dv\, da = a_{merge} \int_a \int_v \left(\frac{d\rho}{dt} \right)_{inflow} dv\, da = a_{merge} \left(\frac{dk}{dt} \right)_{inflow} \tag{6.38}$$

6.5.2.2 Main lane traffic

Drivers in front of whom a vehicle merges raise their activation level to a_{main} to compensate for the suddenly shortened following distance. The active change rate due to these transitions, in general defined by equation (6.33), is proportional to the numbers of vehicles $\rho(t,x,v,a_0)$ having an original activation level a_0 (lower than a_{main}), to the interaction rate $\Pi(t,x,v,a_0)$, and to the transition rate $\pi_a(t,x,v,a_0,a_1)$, given an interaction.

The interaction rate Π per vehicle is equal to the number of merging vehicles per unit of time (every vehicle that merges has a follower who reacts by changing his activation level), normalised by the expected number of vehicles k:

$$\Pi(t,x,v,a_0) = \frac{\left(\frac{dk}{dt} \right)_{inflow}}{k} \tag{6.39}$$

The interaction rate π_a equals 1 if his original activation level is lower than a_{main} and the new activation level equals a_{main}, and zero otherwise:

$$\pi_a(t,x,v,a_0,a_1) = \Xi(a_{main} - a_0)\,\delta(a_1 - a_{main}) \tag{6.40}$$

In this equation the notation δ stands for the Kronecker delta function, and $\Xi(y)$ for the Heavyside function that takes the value 1 for arguments $y > 0$ and the value 0 otherwise. Therefore, equation (6.33) is written as follows in the case of merging:

$$\left(\frac{d\rho(t,x,v,a_0)}{dt} \right)^{-}_{merge} = \int_{a_1 \neq a_0} \rho(t,x,v,a_0) \frac{\left(\frac{dk}{dt} \right)_{inflow}}{k} \Xi(a_{main} - a_0)\,\delta(a_1 - a_{main})\, da_1 \tag{6.41}$$

$$\left(\frac{d\rho(t,x,v,a_0)}{dt}\right)^{-}_{merge} = \int_{a_1} \rho(t,x,v,a_0)\frac{\left(\frac{dk}{dt}\right)_{inflow}}{k}\Xi(a_{main}-a_0)\delta(a_1-a_{main})da_1$$
$$= \begin{cases} \dfrac{\rho(t,x,v,a_0)}{k}\left(\dfrac{dk}{dt}\right)_{inflow} & a_0 < a_{main} \\ 0 & a_0 \geq a_{main} \end{cases} \tag{6.42}$$

The contribution to $g_3(\mathbf{U})$ of the active transitions by main lane drivers due to the presence of merging traffic, is thus:

$$\int_a\int_v a\left(\frac{d\rho(t,x,v,a)}{dt}\right)^{-}_{merge} dv\,da = \int_{a<a_{main}}\int_v a\frac{\rho(t,x,v,a)}{k}\left(\frac{dk}{dt}\right)_{inflow} dv\,da$$
$$= \frac{1}{k}\left(\frac{dk}{dt}\right)_{inflow}\int_{a<a_{main}} a\,\tilde{\rho}(t,x,a)\,da \tag{6.43}$$

In this equation the notation $\tilde{\rho}(t,x,a)$ indicates the *reduced* generalised density, i.e. the generalised density that is integrated over only a subset of its state variables (here: integrated over the speed *v* and not over the activation level *a*).

If we assume $a_{main} = a_{max}$, then any original activation level is lower than a_{main}, so that the integral in the previous expression is evaluated over the entire *a*-range. In that case, equation (6.43) for the active change rate of *a* further reduces to:

$$\int_a\int_v a\left(\frac{d\rho(t,x,v,a)}{dt}\right)^{-}_{merge} dv\,da = \frac{1}{k}\left(\frac{dk}{dt}\right)_{inflow}\int_a a\,\tilde{\rho}(t,x,a)\,da$$
$$= A\left(\frac{dk}{dt}\right)_{inflow} \tag{6.44}$$

The passive change rate of the activation level is that of states that become more probable due to transitions from any other state to this state. The final activation level of any driver in the main lane in front of whom a vehicle merges is a_{main}. Therefore, the passive change rate only has a non-zero contribution for this value of *a*. Since every merging vehicle causes such a transition (if the original activation level was lower than a_{main}), the transition rate is proportional to the rate of merging vehicles. This can also be understood using a different argument: the increase rate at a_{main} equals the sum of all decrease rates, because, regardless of their original state a_0, all drivers who change their state are 'absorbed' in the state a_{main}:

$$\left(\frac{d\rho(t,x,v,a)}{dt}\right)^{+}_{merge} = \delta(a-a_{main}) \int_{a_0}\int_v \left(\frac{d\rho(t,x,v,a_0)}{dt}\right)^{-}_{merge} dv\;\; da_0$$
$$= \frac{\delta(a-a_{main})}{k}\left(\frac{dk}{dt}\right)_{inflow}\int_{a_0<a_{main}} \tilde{\rho}(t,x,a_0)\;\; da_0 \tag{6.45}$$

In the case that $a_{main} = a_{max}$, this can be simplified to:

$$\left(\frac{d\rho(t,x,v,a)}{dt}\right)^{+}_{merge} = \frac{\delta(a-a_{max})}{k}\left(\frac{dk}{dt}\right)_{inflow} \tag{6.46}$$

The contribution to $g_3(\mathbf{U})$ of the passive transitions by main lane drivers due to the presence of merging traffic, is thus:

$$\begin{aligned}\int_a\int_v a\left(\frac{d\rho(t,x,v,a)}{dt}\right)^{+}_{merge} dv\,da &= \int_a\int_v a\,\delta(a-a_{main})\int_{a_0<a_{main}} \frac{\tilde{\rho}(t,x,a_0)}{k}\left(\frac{dk}{dt}\right)_{inflow} da_0\,dv\,da \\ &= \frac{a_{main}}{k}\left(\frac{dk}{dt}\right)_{inflow}\int_{a_0<a_{main}} \tilde{\rho}(t,x,a_0)\,da_0\end{aligned} \tag{6.47}$$

Again in the case that $a_{main} = a_{max}$, this can be simplified to:

$$\int_a\int_v a\left(\frac{d\rho(t,x,v,a)}{dt}\right)^{+}_{merge} dv\,da = a_{max}\left(\frac{dk}{dt}\right)_{inflow} \tag{6.48}$$

6.5.2.3 *Macroscopic dynamic activation level equation near merge zones*

The terms specified in the previous sections are now substituted into equation (6.10) for the dynamics of the activation level moment, together with the definition of the relaxation towards the normal activation level of equation (6.32). The macroscopic dynamic activation level (moment) equation for main lane traffic near merge zones becomes:

$$\frac{\partial kA}{\partial t} + \frac{\partial kAV}{\partial x} = k\frac{a_{normal}-A}{\tau_a} + a_{merge}\left(\frac{dk}{dt}\right)_{inflow} + \frac{1}{k}\left(\frac{dk}{dt}\right)_{inflow}\int_{a_0<a_{main}} (a_{main}-a)\,\tilde{\rho}(t,x,a_0)\,da_0 \tag{6.49}$$

If $a_{main} = a_{max}$, the integral in the RHS can be simplified, yielding:

$$\frac{\partial kA}{\partial t} + \frac{\partial kAV}{\partial x} = k\frac{a_{normal}-A}{\tau_a} + a_{merge}\left(\frac{dk}{dt}\right)_{inflow} + (a_{max}-A)\left(\frac{dk}{dt}\right)_{inflow} \tag{6.50}$$

In the primitive formulation, these equations are:

$$\frac{\partial A}{\partial t} + V\frac{\partial A}{\partial x} = \frac{a_{normal}-A}{\tau_a} + \frac{a_{merge}-A}{k}\left(\frac{dk}{dt}\right)_{inflow} + \frac{1}{k^2}\left(\frac{dk}{dt}\right)_{inflow}\int_{a_0<a_{main}} (a_{main}-a)\,\tilde{\rho}(t,x,a_0)\,da_0 \tag{6.51}$$

and for $a_{main} = a_{max}$:

$$\frac{\partial A}{\partial t} + \underbrace{V\frac{\partial A}{\partial x}}_{convection} = \underbrace{\frac{a_{normal}-A}{\tau_a}}_{relaxation} + \underbrace{\frac{a_{merge}-A}{k}\left(\frac{dk}{dt}\right)_{inflow}}_{nett\ effect\ of\ inflow} + \underbrace{\frac{a_{max}-A}{k}\left(\frac{dk}{dt}\right)_{inflow}}_{transitions\ on\ main\ lanes} \tag{6.52}$$

Despite the rather complex derivation, this final equation (6.52) is relatively easy to interpret. It says that four processes determine the change of the average activation level A in time:

- *convection*: activation level is transported with speed V along with the flow;
- *relaxation*: constant tendency of drivers to relax to the normal comfortable activation level a_{normal};
- *net effect of inflow*: in fact, this is convection of activation level out off the flow or into the flow (depending on the sign of dk/dt); in the primitive formulation, an extra term occurs here, representing the redistribution of the total activation level momentum over a changed number of drivers,
- *transitions on main lanes*: effect of the reactions of drivers in the main flow due to the confrontation with merging traffic.

For the latter two terms, it is logical that their influence increases as the in or outflow is relatively large compared to traffic already in the flow, hence the factor $\frac{1}{k}\left(\frac{dk}{dt}\right)_{inflow}$. On the other hand, the influence of merging traffic also increases as the difference between the average activation level in the flow (A) and the activation level of merging traffic (a_{merge}) is larger, and as the difference between undisturbed drivers in the main flow (A) and that of drivers who are confronted with mergers a_{main} (or a_{max}) is larger.

6.5.3 Capacity funnel simulations

6.5.3.1 Simulation set-up

As a case study, the same bottleneck configuration and parameters as in section 5.4 is considered. Because in this case study not only the traffic pattern upstream of the bottleneck is studied, the location of the on-ramp is now centred around x = 4.5 km instead of x = 6 km, so that the emergence of the capacity funnel downstream of the merge falls well within the simulated area. The initial main lane flow rate is 2090 veh/h (0.58 veh/s); the on-ramp flow q_{in} = 600 veh/h. The parameters of variable driver behaviour are listed in Table 6-1.

The fundamental diagram for the normal activation level (Figure 6-5c, dashed curve) is the same as for the basic human-kinetic traffic flow model (see section 5.1). It shows that the initial conditions are chosen, so that the initial flow is just below the maximum equilibrium flow rate and the sum of main lane flow and on-ramp flow 2690 veh/h (0.75 veh/s) can normally not be processed.

Table 6-1 Parameters of variable driver behaviour in the capacity funnel case

parameter	description	value	introduced in equation
a_{max}	maximal activation level	10	(6.27)
$p_{indiff\ \ normal}$	minimal indifference band	0 %	(6.27)
$p_{indiff\ \ max}$	maximal indifference band	70 %	(6.27)
a_{normal}	normal activation level	0	(6.27),(6.49)
τ_a	relaxation constant for the activation level	80 s	(6.49)
$a_{merge} = a_{max}$	activation level of merging drivers	10	(6.49)
$a_{main} = a_{max}$	activation level of drivers in main lanes in front of whom traffic merges	10	(6.49)

6.5.3.2 Equilibrium solutions with and without variable driver behaviour

Let us consider the effect of introducing the indifference band. Figure 6-5 compares the acceleration integral and the fundamental diagram of density against flow for an average indifference band of 50% (see equation (6.27)). In the acceleration integral solution, a plateau occurs (indicated in Figure 6-5b) with an expected acceleration close to zero for combinations of density and average speed above the equilibrium line. Just like the individual drivers, the flow is 'indifferent' for a range of local traffic conditions, especially for an increase of the density, compared to equilibrium conditions.

Moreover, the fundamental diagram (Figure 6-5c) reveals that the equilibrium line itself is shifted towards higher equilibrium flows. In this case the equilibrium capacity is 2880 veh/h (0.80 veh/s), about 25% higher than the equilibrium capacity without indifference band. The total demand of main lanes plus the on-ramp could easily be processed if the activation level were raised so high that the average indifference band $p_{indiff}(a)$ equals 50 % (and would remain so high).

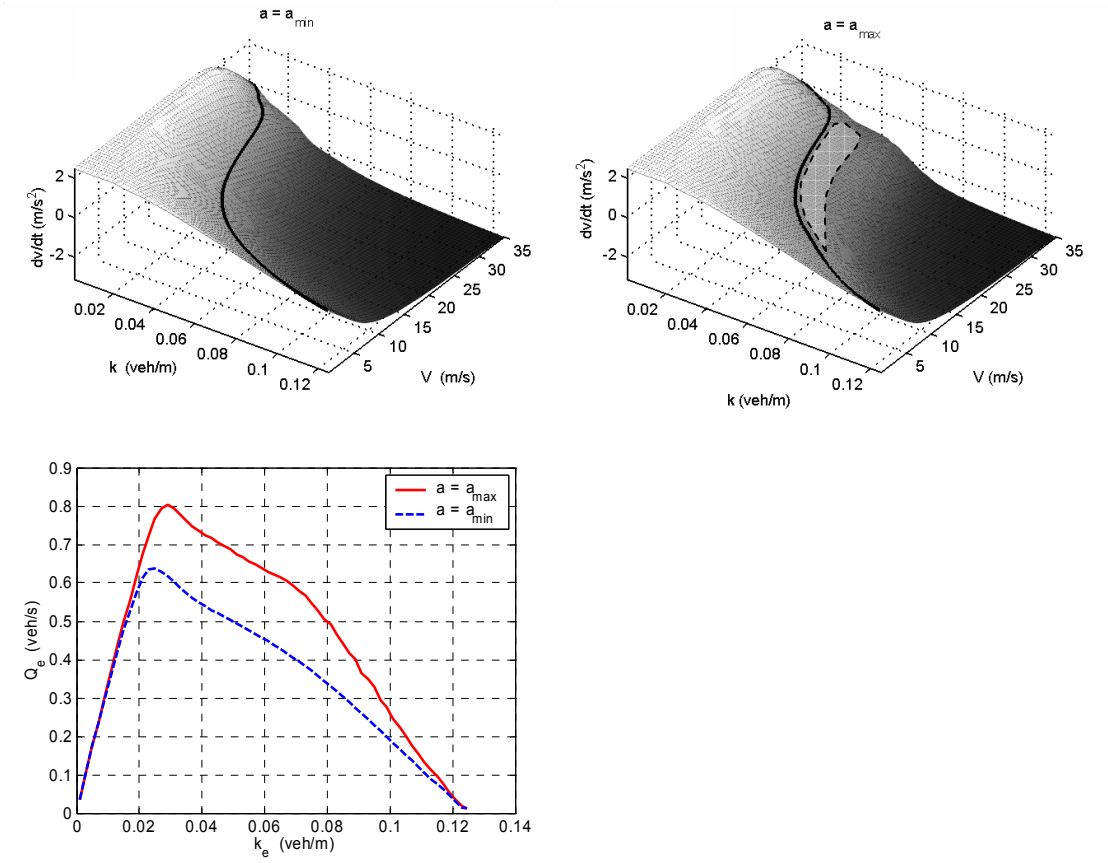


Figure 6-5 Acceleration integral (a) without and (b) with maximal indifference band; (c) equilibrium flows as a function of density for the maximum and minimum activation levels

6.5.3.3 Simulation results of the capacity funnel

Figure 6-6a shows that without variable driving behaviour, congestion sets in at the location of the on-ramp. At time t= 100 s, a jam pattern emerges with onset of congestion approximately at location 4700 m (see detail Figure 6-7a). This pattern can be identified as Triggered Stop-and-go waves (TSG; see also sections 2.4.2 and 5.4.2).

The simulation results with variable driving behaviour (Figure 6-6b) show that at the location of the on-ramp the average activation level is raised to 3 (on a scale of 0 to 10), so that the average indifference band is 21 %. The (equilibrium) capacity increase is therefore lower than Figure 6-5c but apparently high enough to initially accommodate the total demand.

Because the activation level relaxes towards the normal comfortable level once the on-ramp is passed, the capacity increase is only temporary. Therefore, someplace downstream of the merge, the equilibrium capacity eventually drops below the value of the total flow and congestion sets in. From the detailed output Figure 6-7b, it appears that the high flow emerging from the on-ramp is carried with traffic until t = 300, 200 seconds later than in the reference case of Figure 6-7a. Then, approximately at x = 5700 m – 1 km downstream of the on-ramp – the origin of the congestion pattern arises.

Again, the type of congestion is TSG. After the initial emergence at x=5.7 km, the queue head settles at x=4.5 km, the location of the on-ramp.

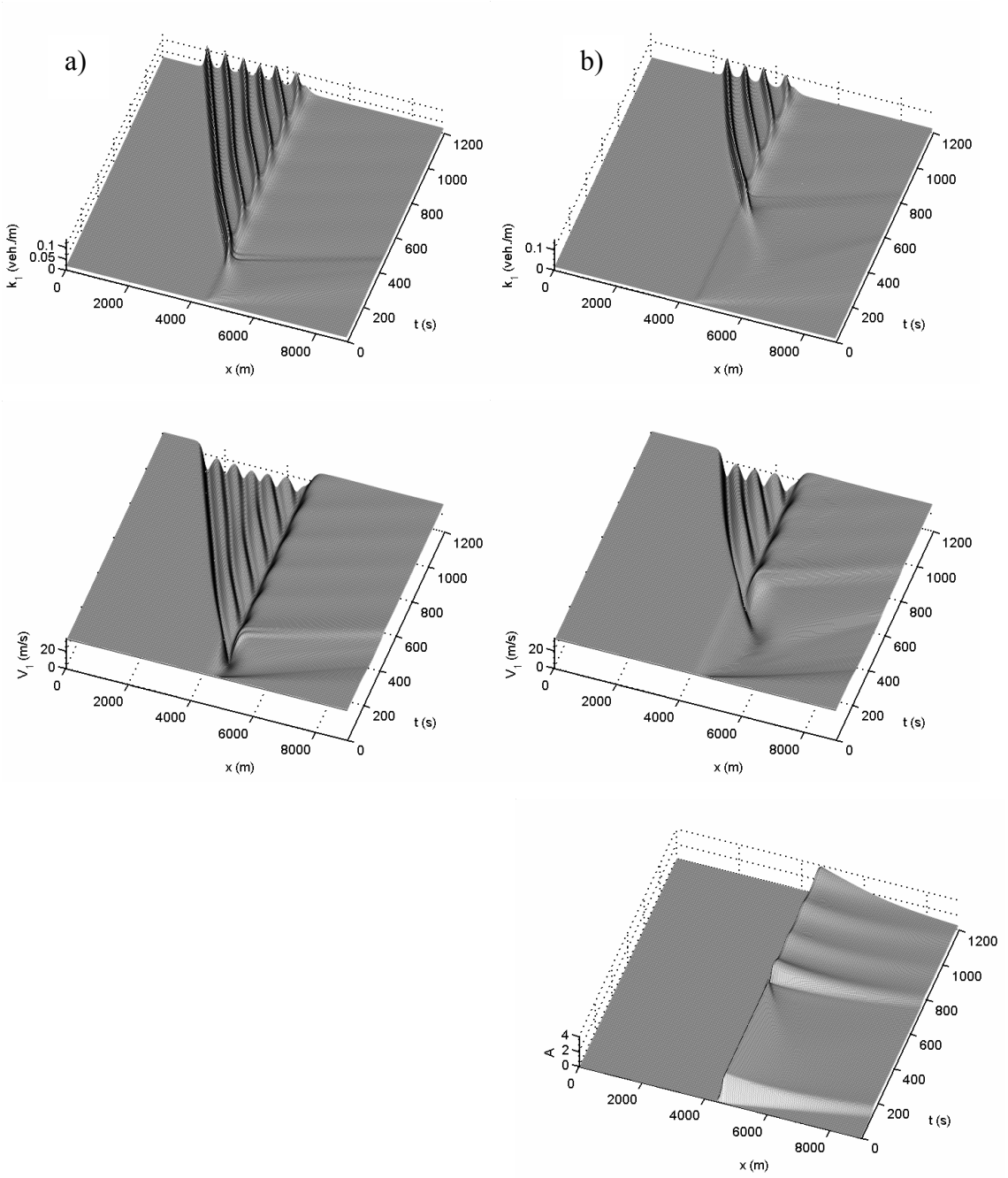


Figure 6-6 Capacity funnel simulation results; from top to bottom: density* k*, speed* V *and activation level* A *as a function of time and space; (a)* left*: without and (b)* right*: with traffic condition induced variations of driving behaviour

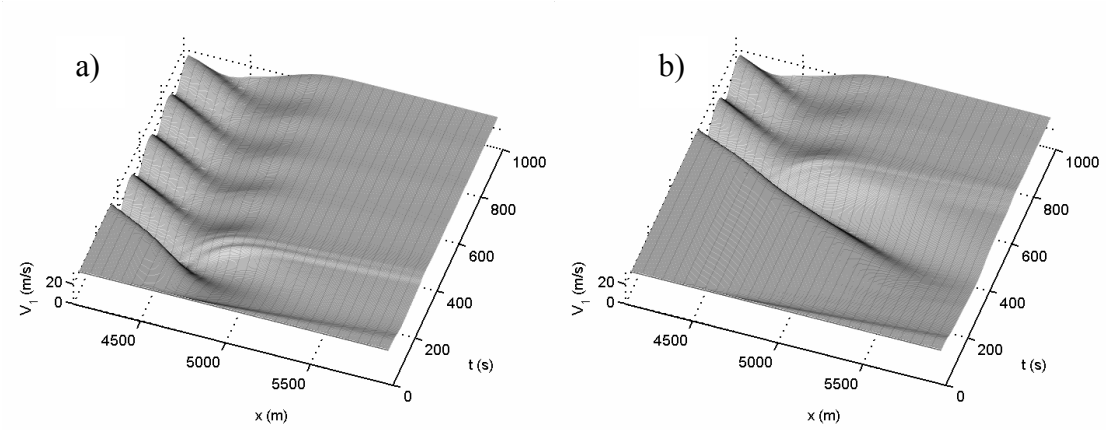


Figure 6-7 Detail of the speed from immediately upstream to 1500 m downstream of the on-ramp at* x *= 4500 m; (a)* left*: without and (b)* right*: with traffic condition induced variations of driving behaviour

Discussion of capacity funnel simulations

It appears that the human-kinetic model can generate delayed onset of congestion (or capacity funnel) in two ways: spontaneously in the model without variable driver behaviour, and through traffic induced variations of driver behaviour.

In Case 2 of section 5.4.2, delayed onset of congestion occurred spontaneously for a flow rate that was almost equal to the equilibrium capacity (more specifically the *prior congestion* capacity in the terminology of section 2.6.1). This happened only after a perturbation (short temporary excess of equilibrium capacity) had triggered a density cluster in the near-capacity flow (which is apparently *metastable* – see section 2.3.1). The unstable perturbation grew in amplitude, until its propagation direction reversed (boomerang effect, see section 2.6.4) and a single stop-and-go wave travelled upstream. It was mentioned that the spontaneous capacity funnel only occurred in carefully chosen conditions. However, the data analysis reported by Smulders et al. (2000) and by Cassidy & Bertini (1999a; see section 2.6.4) suggests that capacity funnel is not some rare phenomenon, but occurs frequently and reproducibly.

The human-kinetic model with variable driver behaviour is able to reproduce the capacity funnel in much more widely varying conditions than the spontaneous mechanism, as is shown in Figure 6-6b and Figure 6-7b. Moreover, the location of traffic breakdown depends on the ratio of total traffic demand for the bottleneck (main lane flow + on-ramp flow) over the equilibrium capacity and on the relaxation constant for the activation level (Tampère et al. 2003). Delayed onset of congestion is therefore reproducible and controllable within a wider range of traffic conditions. On the other hand, not every type of capacity funnel is reproducible with the current specifications: according to the observations by Smulders et al. (2000), the queue head can either remain downstream at the location where the delayed onset of congestion occurred, or return and settle at the location of the merge. The human-kinetic model with driver behaviour variations only reproduces the second type (see Figure 6-7b, which is representative for numerous variations that were tested).

6.6 Case study (2): behavioural explanation of hysteresis and the capacity drop

In section 5.4.2 it was shown that the queue discharge rate (flow out of the jam or bottleneck flow) at a merging bottleneck decreased as a function of the on-ramp flow (Figure 5-21) as a self-organising feature of the model. This section illustrates how variable driver behaviour might be an alternative cause of capacity drop. In reality, the cause might be either self-organisation or driving behaviour, or a combination of both. The aim of this section is to show that plausible individual driver behavioural hypotheses, inspired by observations reported in literature, can have significant influence on macroscopic properties of the flow, such as capacity drop, and cannot be neglected when attempting to understand traffic flow dynamics.

6.6.1 Behavioural hypotheses for low-speed driving

As mentioned in section 6.2, there is indirect and direct evidence that drivers loose their motivation to drive with short headways once congestion has set in. The following hypotheses are used throughout this case study (for a motivation the reader is referred to section 6.2).

When the driving speed drops significantly below normal free driving speeds, the activation level of drivers (as a consequence of lower motivation) starts to decrease to a lower level that is proportional to the speed. In order to ensure safety of car-following behaviour during less active driving, drivers are assumed to increase their desired time headway inversely proportional to the activation level. Combining these effects, the gradual decrease of activation level in low speed driving, invokes a desire to drive with larger headway, so that the gaps increase as speeds prolonged remain low.

When traffic accelerates again, the inverse processes increase motivation again, so that after some time drivers adopt shorter gaps again. Finally, raising activation level again goes faster than loosing motivation, so the time scale of both processes might be different.

6.6.2 Behavioural specifications

In this section the behavioural assumptions of the previous section are detailed into model specifications. In comparison to the first case study of merging traffic (where the activation level changed due to cutting-in ahead of the driver, which was considered an instantaneous event), the adaptation to the prevailing speed conditions is a *slow* process. Slow activation level variations are not modelled as a discrete process according to section 6.4.2.2 but as a smooth process (section 6.4.2.1).

The relaxation process towards the normal comfortable activation level a_{normal} of equation (6.32) is taken as a starting point. Two modifications to that specification of relaxation behaviour are made: the normal activation level is made speed-dependent and the relaxation time is different for increase or decrease of the activation level.

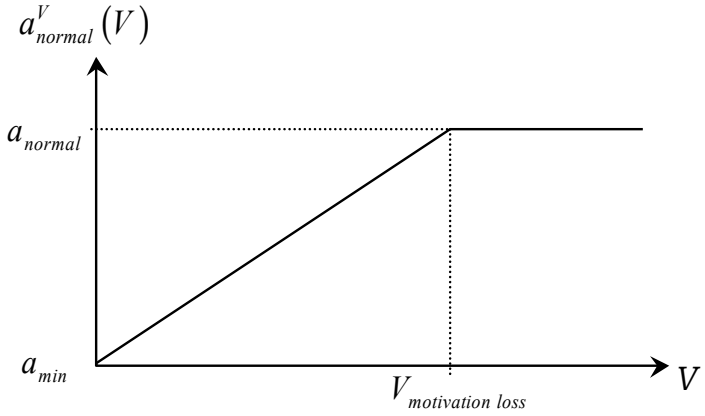


Figure 6-8 Dependence of the normal activation level on the prevailing speed

The relaxation towards the normal comfortable activation level of equation (6.32) is adapted, so that the driver now relaxes towards the *speed-dependent* normal activation level:

$$\left\langle \frac{da}{dt} \right\rangle_{a,v} \equiv \frac{a^V_{normal}(V) - A}{\tau_a(A,V)} \tag{6.53}$$

The speed-dependent normal activation level is specified as follows:

$$a^V_{normal}(V) \equiv \begin{cases} a_{normal} + (a_{min} - a_{normal}) \dfrac{V_{motivation\ loss} - V}{V_{motivation\ loss}} & V > V_{motivation\ loss} \\ a_{normal} & V \le V_{motivation\ loss} \end{cases} \tag{6.54}$$

This relation is depicted schematically in Figure 6-8. The modification to the relaxation time is given by:

$$\tau_a(A,V) \equiv \begin{cases} \tau_a^{up} & A < a^V_{normal}(V) \\ \tau_a^{down} & A \ge a^V_{normal}(V) \end{cases} \tag{6.55}$$

The effect on driver behaviour of a lower than normal activation level is that drivers desire longer time headways, as described in section 6.4.1.2.

6.6.3 Capacity drop simulations

Two cases are simulated in this section. Firstly, jam propagation with the modified specifications is tested using the same simulation set-up as in the sensitivity analysis of the basic human-kinetic model (section 5.2.2). Secondly, the on-ramp scenario that was tested in sections 5.4 and 6.5.3 is studied.

6.6.3.1 Simulation set-up

For the jam propagation test, the same road lay-out, initial and boundary conditions as in section 5.2.2 are taken, with the parameter values of Table 5-2 and Table 6-2.

Table 6-2 Parameters of variable driver behaviour in the hysteresis analysis

parameter	description	value	introduced in equation
a_{normal}	normal activation level	0	(6.54)
a_{min}	minimum activation level	-10	(6.54)
$V_{motivation\ loss}$	speed below which the activation level gradually decreases	20 m/s	(6.54)
τ_a^{up}	relaxation time for activation level increase	20 s	(6.55)
τ_a^{down}	relaxation time for activation level decrease	50 s	(6.55)
$s_{1\,normal}^{d}$	normal time headway	1 s	(6.23)
$s_{1\,max}^{d}$	maximum time headway	1.5 s	(6.23)

For the on-ramp scenario, the road lay-out, initial and boundary conditions of Case 5 in Table 5-3, section 5.4 are taken. The parameters are those of Table 5-1, Table 5-2, and Table 6-2.

6.6.3.2 Simulation results

In the phase trajectory plots in this section, the initial state in the flow-density plane is marked with a square (□), the end state is marked with a star (☆); a decelerating phase trajectory is marked with circles (⍟) and acceleration with triangles (△).

Jam propagation and phase trajectory

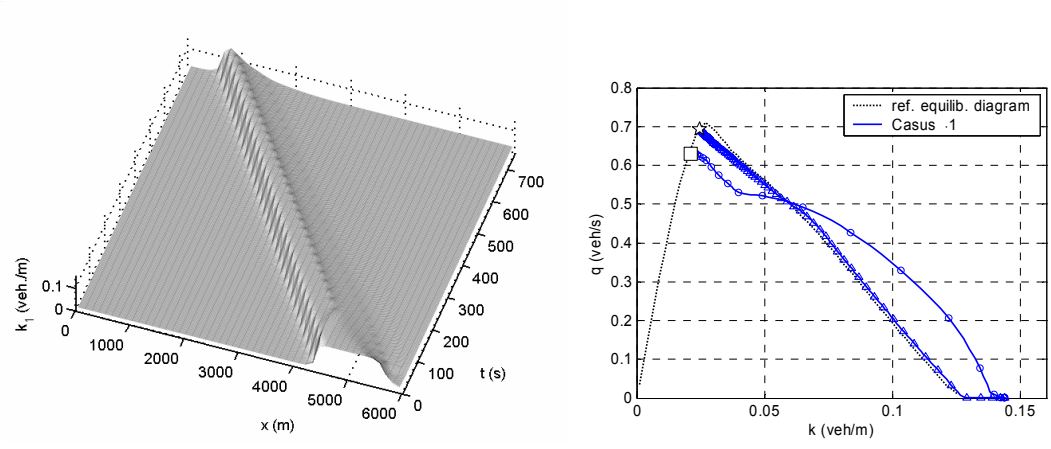


Figure 6-9 Case 1: Reference run with parameters from Table 5-2; (a)* left*: space-time evolution of the density, (b)* right*: phase trajectory in the flow-density plane

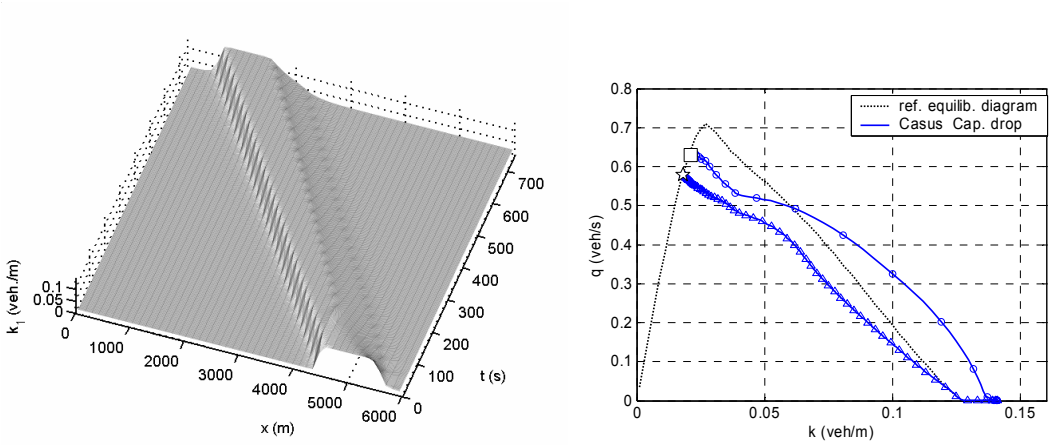


Figure 6-10 Variable driver behaviour: loss of motivation in low speed conditions (a)* left*: space-time evolution of the density, (b)* right*: phase trajectory in the flow-density plane

In Figure 6-10, the simulation results of the jam propagation test are presented. These results are to be compared to the reference of Figure 6-9 (equal to Figure 5-5), where the same scenario is modelled without variable driver behaviour. The upstream jam front is unchanged compared to the reference. This is logical, since the speeds upstream remain above the threshold $V_{motivation\ loss}$ below which variable driving behaviour is activated; therefore the first phase of deceleration (from free flow to the speed threshold) is identical to the reference. The second deceleration phase, from the threshold speed until standstill, happens faster than the relaxation towards lower activation level (see Figure 6-11, upper branch), so that also this phase proceeds almost identically to the reference simulation. This implies that the density at standstill is also close to that of the reference case, and that the wave speed of the deceleration wave remains unchanged.

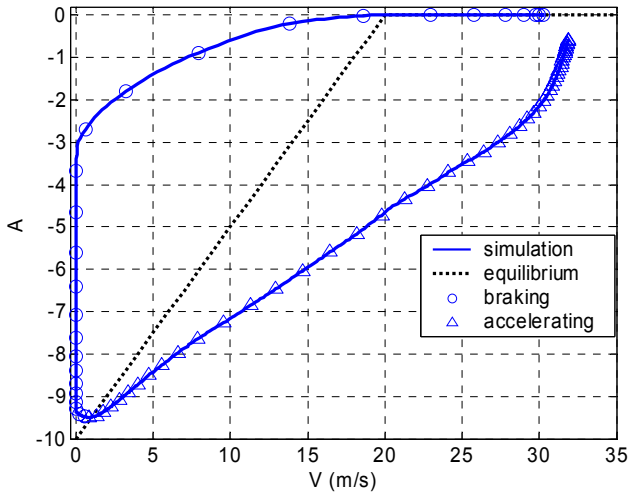


Figure 6-11 Evolution of the activation level as a function of speed during deceleration and acceleration against the 'equilibrium' normal activation level of equation (6.54)

In contrast, the recovery from congestion in the downstream wave front is significantly different. During standstill, the motivation of drivers to follow closely has disappeared (the activation level has significantly decreased during standstill, see Figure 6-11), so that traffic expands (drives at lower density for the same speed, i.e. lower flow rates). Therefore, the queue discharge flow is much lower than in the reference: 2090 veh/h instead of 2502 veh/h, a decrease of 17%. This capacity drop is caused by hysteresis in the fundamental diagram of flow rate against density, where the acceleration branch now remains under the deceleration branch over the entire density range. Figure 6-11 reveals that this strong hysteresis effect is directly related to hysteresis in driver behaviour: the activation level, which is directly (and inversely) related to the desired time headway via equation (6.23), follows the evolution of the speed with a significant delay, so that flow reaches the free flow speed with a lower-than-normal activation level, and therefore a larger than normal spacing. The latter is the direct cause of the capacity drop.

This example has thus shown how behavioural variations can have significant impact on the macroscopic dynamics. In this case the capacity drop is reproduced as a direct result of motivation loss at low speeds and hysteresis in driving behaviour.

Merging scenario

In this section the merging scenario of section 5.4 (more specifically Case 5 of that section) is repeated, now with variable driving behaviour as specified in section 6.5.2. The simulation results are depicted in Figure 6-12, together with the repeated results of the reference case with invariable driver behaviour.

The simulation is clearly different from the reference scenario. After the first jam has emerged, this first jam travels upstream and, in contrast to the reference scenario, evolves into a *wide jam* instead of a stop-and-go wave (see discussion of Case 5 in section 5.4.2). Moreover, the queue discharge from this wide jam is lower than that of the stop-and-go wave in the reference case. Therefore, lower flow rates arrive at the merge, so that it takes longer before the second instability is generated at the bottleneck. This second jam is fed at a flow rate equal to the queue discharge rate of the first jam, so that in- and outflow of this jam is more balanced, and it maintains the character of a stop-and-go wave instead of evolving into a wide jam. The difference is also visible in the phase trajectory for both jams in Figure 6-13a: the speed in the first jam drops to zero, whereas in the second jam it does not. Also, in the first jam, density returns to the equilibrium jam density before traffic accelerates, whereas in the second jam it does not. The second jam is therefore more comparable to those in the reference case (Figure 6-13b).

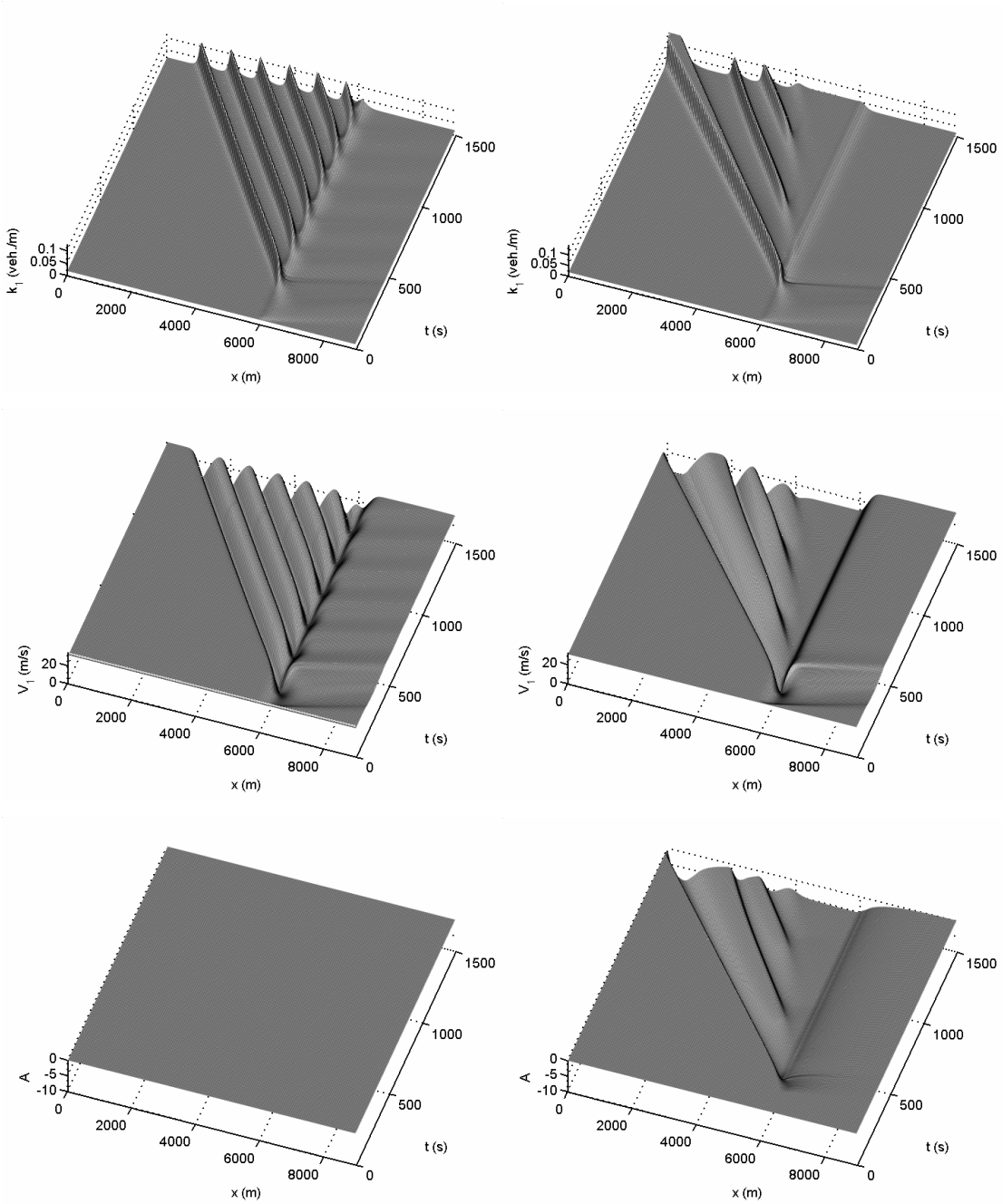


Figure 6-12 Merging scenario with* q_{in} *= 400 veh/h and* q_{main} *= 2240 veh/h; from* top *to* bottom*: density, speed and activation level as a function of time and space; (a)* left column*: reference with invariable driver behaviour and (b)* right column*: with loss of motivation at low speeds

Note also that after the second jam has emerged, a zone with homogeneous medium speeds starts to build up just upstream of the on-ramp (speed plot of Figure 6-12a). A new instability still occurs, giving rise to a new jam, but this jam grows out off the low speed region, instead of emerging from the location of the on-ramp. These new patterns can emerge, probably because the variable time headway also affects the stability of traffic, so that homogeneous congestion patterns become possible (see section 5.2.1 for the influence of s_1^d (here: $s_1^d(a)$) on traffic flow stability).

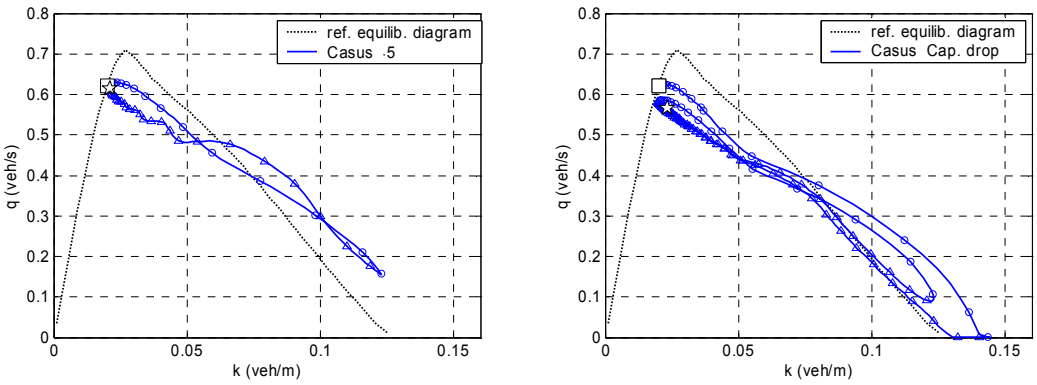


Figure 6-13 Phase trajectory in the flow-density plane; (a)* left*: with invariable driver behaviour and (b)* right*: with variable driver behaviour

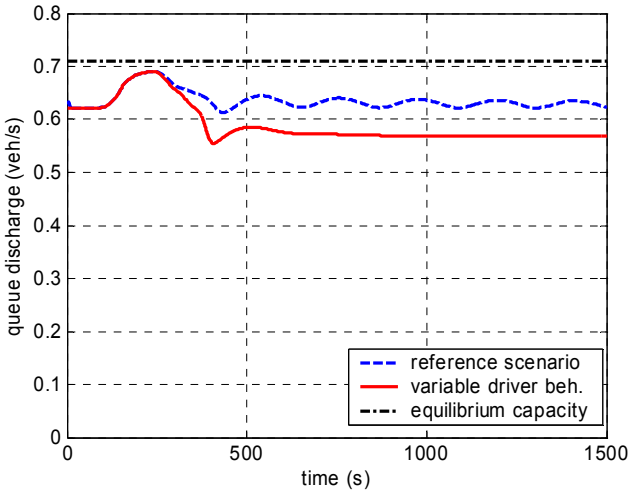


Figure 6-14 Comparison between queue discharge rate with and without variable driver behaviour

Looking at the flow rate at *x*-locations downstream of the merge we can analyse the queue discharge rates (Figure 6-14). As was discussed in section 5.4.2 the queue discharge of the reference case is already lower (-11%) than the maximum flow that can theoretically be attained in equilibrium conditions. The loss of motivation in low speed driving results in an additional 8% capacity drop (-19%). Depending on the parameters (especially τ_a^{up} and $s_{1\,max}^{d}$), any value of the capacity drop can be calibrated with this model.

We conclude that the congestion pattern is much more diverse with variable driver behaviour than in the reference case: wide jams, stop-and-go waves, and homogeneous congested traffic all co-exist in this case. In the reference case, only stop-and-go waves occurred. Moreover, the contribution of variable driving behaviour to capacity drop (difference between prior-congestion flow/equilibrium capacity and queue discharge rate, see sections 2.6.1 and 5.4.2) is significant, and can be controlled through the newly introduced parameters.

6.7 Conclusions

When the basic human-kinetic traffic flow model was introduced in chapter 4, the driver was modelled as a longitudinal controller with invariant control parameters and fixed control set points (constant desired speed and a speed-dependent desired distance function). In this chapter, indications from empirical traffic data and experimental behavioural research were discussed that suggest that these parameters and set points (which we refer to as *driving strategy*) are not constants. Instead, driving strategy depends among others on traffic conditions. As a first approximation of variable driving behaviour or strategy, the assumption was made in this chapter that current driving strategy depends on one behavioural parameter: the *activation level*. On the one hand, the activation level depends on traffic conditions; on the other hand, the activation level determines the current driving strategy, and therefore indirectly traffic conditions. This control process was modelled and further detailed in this chapter.

Using similar mathematical formalisms from kinetic traffic flow theory as was used in chapter 4, kinetic and macroscopic dynamic model equations for the new variable (activation level) were established. The mathematical structure of these equations is similar to that of the average speed, as are the numerical solution techniques. Using the acceleration integral, the influence of the activation level on individual driver behaviour can be accounted for in the dynamic speed equation. The determinant issue is the definition of the dependence of activation level on traffic conditions and of driving strategy on activation level. Little is known about these relationships, so that the analyses were restricted to hypothetical relationships, inspired by the empirical and experimental studies mentioned earlier. Drivers are assumed to have some comfortable normal activation level, but can drive more or less active depending on traffic conditions. Both possibilities were examined in case studies.

In one case study, the capacity funnel phenomenon (section 2.6.4) was reproduced due to variable driving behaviour with higher-than-normal activation level. We assumed high activation level of merging drivers on a busy on-ramp and of main-lane drivers in front of whom traffic merges. As a result of increased activation level, drivers do not immediately increase smaller-than-desired gaps and temporarily accept small gaps without braking. Downstream of the on-ramp, driving with higher activation level is not longer useful and drivers relax to the normal comfortable activation level. They can do so only by increasing short gaps to the normal acceptable values, which causes the onset of congestion downstream of the merge. Using this behavioural model, the capacity funnel effect could be reproduced in much more widely varying conditions than the spontaneous mechanism reported in literature (Helbing, 2001) and in chapter 5. Delayed onset of congestion is therefore reproducible and controllable within a wider range of traffic conditions and directly related to variable driver behaviour.

The other case study examined capacity drop near a merging bottleneck, caused by lower-than-normal activation level. We assumed lower activation level of drivers who

are caught in a queue: the longer speeds remain low, the lower the activation level of drivers. In order to ensure safe car-following with low activation level, longer desired gaps to the predecessor are assumed. When speed increases, drivers change from low to high speed driving strategy again. However, this happens with some delay, which causes lower outflow of the queue or capacity drop.

Both case studies show how variable driving behaviour contributes to typical congestion phenomena. Both capacity funnel and capacity drop were already reproduced spontaneously by the basic human-kinetic model analysed in chapter 5. However, with specifications for variable driving behaviour, the surplus contribution to these phenomena of variable driving strategies becomes clear, and the phenomena can be reproduced in more widely varying conditions than with the spontaneous mechanism.

We conclude that the human-kinetic traffic flow model with activation level as an additional variable to describe variable driving behaviour is able to describe traffic flow dynamics more refined than the basic model, or other models using fixed driving strategies. This behavioural refinement paves the way towards a traffic flow model that is transferable from one network to another, without having to recalibrate the behavioural parameters, as is the case in currently applied traffic flow models (which shows that these models apparently lack some aspect of adaptive behaviour that would make them more generically valid).

However, so far the model refinements in the human-kinetic traffic flow model were based on hypotheses of driver behaviour that have not been validated. This requires data of individual driver behaviour in real-life and experimental conditions that are not available to date. Such data set should ideally not only contain individual trajectory data, but also activation level measurements (although activation level is not directly measurable, indirect measurement techniques exist).

The remainder of this thesis illustrates how the human-kinetic model with variable driving behaviour can be deployed in the explorative analysis of Advanced Driver Assistance systems. There, also assumptions will be made on the impact of ADA systems on activation level, which in turn influences driving strategy.

7 MULTICLASS HUMAN-KINETIC TRAFFIC FLOW MODELLING

7.1 Introduction

With the derivation of the basic human-kinetic traffic flow model in chapter 4, and the refinement of its microscopic foundation by including variable driving behaviour in chapter 6, a refined and flexible simulation and analysis framework for traffic flows has been established. In this chapter another refinement is added to the model: the human-kinetic theory is generalised to a multiclass theory. The motivation for this extension is twofold.

First of all, a multiclass traffic flow model has theoretical advantages over a single class model, due to the heterogeneous composition of real traffic. In principle every vehicle-driver combination in traffic flow can behave differently; the best description of traffic flow would therefore in theory distinguish different behaviour for each vehicle in the flow, although this is neither practical nor feasible. It is therefore sufficient to account only for heterogeneity that is relevant for the desired traffic flow analysis, and to neglect all irrelevant differences. In some microscopic simulation tools (e.g. Paramics (Quadstone, 2002)) more than ten different classes of vehicles and drivers are distinguished, for example: public transport vehicles, cars, trucks, and even bicycles and pedestrians. In kinetic or macroscopic traffic flow models however, a distinction between person cars and trucks is generally considered sufficient for most purposes, due to their large difference in maximum speed, acceleration capability and size

(Hoogendoorn, 1999; Logghe, 2003 and references therein). Besides this main distinction of user-classes, kinetic traffic flow models, like the one of Hoogendoorn & Bovy (2000), offer the possibility to account for the remaining within-class heterogeneity, by considering a distribution of some relevant behavioural parameters (like the desired speeds) instead of one identical parameter for all vehicles of the same class. Both Hoogendoorn and Logghe show, by comparison with empirical data, that multiclass traffic flow modelling yields considerably better representation of traffic than single-class models.

A second motivation for distinguishing multiple user-classes in the human-kinetic traffic flow model lies in its desired application to ADA systems. Only by distinguishing user classes, it is possible to model mixed traffic flows of equipped and non-equipped vehicles, or even mixtures of different types of driver support systems.

The extension of the human-kinetic model with multiclass traffic builds mainly on methods well-known from literature. A correct physical description based on the muliclass generalisation of the generalised density ρ (introduced in chapter 4 and Annex A) is borrowed from Hoogendoorn (1999). However, whereas the within-class and between-class interactions in the human-kinetic model are also inspired by the approach used by this author, the unique treatment of acceleration and deceleration in the acceleration integral poses some specific problems. More specifically, the explicit treatment of *finite space requirements* in the acceleration integral requires careful consideration of the amount of space on the road assigned to each class. The concept formulated by Logghe (2003) to solve this issue turns out very useful in this context, but due to the totally different characteristics of the models, the elaboration in the human-kinetic model is unique.

The remainder of this chapter is structured as follows. In section 7.2, the concept of user-classes and its implications for the definition of the kinetic and macroscopic distributions and variables are discussed. Dynamic multiclass kinetic and macroscopic equations are established in section 7.3 that also contains the elaboration of the core of the extended model, the multiclass acceleration integral (section 7.3.2). As in the previous chapters, the equilibrium solutions and a numerical dynamic case study are presented in section 7.4, after which we conclude this chapter by a brief discussion of its main results in section 7.5.

7.2 Multi user-class generalised density: definitions and relations

7.2.1 The concept 'user-class'

In this chapter multi user-class traffic is considered. A *user-class* is any type of vehicle-driver combination one wishes to distinguish from other user-classes in the flow. The concept of user-classes is used in this thesis to distinguish classes with different microscopic driving behaviour, like person cars versus trucks or ADA vehicles versus non-equipped vehicles. For the sake of completeness, let us mention that theoretically the concept is equally applicable to user-classes with identical driving behaviour that

need to be distinguished for some other reasons, for instance vehicles with different destinations (if the traffic flow model is used as the propagation model for traffic on a link in a network assignment model) or vehicles with different emission characteristics (if some post processing module is later used to calculate emissions from the traffic flow variables).

A user-class is indicated by an index c, and the total number of user-classes under consideration is indicated by C. Similar to the definition of the generalised density for a single user-class (section A.2 in Annex A), partial generalised density functions and partial density functions are now defined per user-class c.

7.2.2 *User-class specific definition of the (generalised) density function*

Similar to the single user-class case, the partial density $k_c(t,x)$ is interpreted from a probabilistic viewpoint: $k_c(t,x)\,dx$ is defined as the expected number of vehicles of class c at time t and location $x\in[x,x+\mathrm{dx})$. This means that $N_c(t,L)$, the expected number of vehicles of class c at time t on a finite stretch of road with length L is found as:

$$N_c(t,L) = \int_{x=0}^{x=L} k_c(t,x)\ \ dx \tag{7.1}$$

The total vehicular density k_{tot} is defined as the sum of the partial densities over all user-classes $c=1..C$:

$$k_{tot}(t,x) \equiv \sum_c k_c(t,x) \tag{7.2}$$

Thus, the expected number of vehicles, regardless of their class c, at time t on a finite stretch of road with length L is found as:

$$N(t,L) = \sum_c \int_{x=0}^{x=L} k_c(t,x)\ \ dx \tag{7.3}$$

Given these definitions, it can be proven (Hoogendoorn, 1999) that the probability $p_c(t,x)$ that, when a vehicle is present at time t and location x, this vehicle is of user-class c, equals:

$$p_c(c|t,x) = \frac{k_c(t,x)}{k_{tot}(t,x)} \tag{7.4}$$

The user-class specific generalised density considers subsets of the total number of vehicles of class c having some common characteristic or state S. Then the generalised density $\rho_c(t,x,S)\,dx\,dS$ defines the expected number of vehicles of class c in a state $S\in[S,S+dS)$ at time t and location $x\in[x,x+dx)$. Again $N_c(t,L,S)$, the expected number of vehicles of class c at time t on a finite stretch of road with length L and driving in state S is found as:

$$N_c(t,L,S) = \int_{x=0}^{x=L} \rho_c(t,x,S)\ \ dx \tag{7.5}$$

The relation between the partial density $k_c(t,x)$ and the generalised partial density $\rho_c(t,x,S)$ is:

$$k_c(t,x) = \int_S \rho_c(t,x,S)\ dS \tag{7.6}$$

For the definition of the vehicle state S (or the non-local state $S' \equiv (x,S)$) of a vehicle class c, we refer to the definitions given for the single user-class case (section A.2.1 of Annex A).

7.2.3 *User-class specific moments of the (generalised) density function*

Also the definition of the moments of the generalised partial density function is analogous to the single user-class case. In general, the κ^{th} order moment of ρ_c for state variable s_i is defined by:

$$\begin{aligned} M^{\kappa}_{c,s_i}(t,x) &\equiv \int_S s_i^{\kappa} \rho_c(t,x,S)\ dS \\ &= \int_{s_1}\int_{s_2}\ldots\int_{s_n} s_i^{\kappa} \rho_c(t,x,S)\ ds_n \ldots\ ds_2 ds_1 \end{aligned} \tag{7.7}$$

The 0th, 1st, and 2nd order speed moments are:

$$\begin{aligned} M^{0}_{c,v}(t,x) &= \int_S v^0 \rho_c(t,x,S)\ dS \\ &= k_c(t,x) \end{aligned} \tag{7.8}$$

$$\begin{aligned} M^{1}_{c,v}(t,x) &= \int_S v \rho_c(t,x,S)\ dS \\ &\equiv k_c(t,x) V_c(t,x) \\ &\equiv Q_c(t,x) \end{aligned} \tag{7.9}$$

$$\begin{aligned} M^{2}_{c,v}(t,x) &= \int_S v^2 \rho_c(t,x,S)\ dS \\ &\equiv k_c(t,x)\left[\Theta_c(t,x) + V_c^2(t,x)\right] \end{aligned} \tag{7.10}$$

These equations define the partial average speed V_c, the partial flow rate or speed momentum Q_c, and the partial speed variance Θ_c of user-class c. When also the activation level of the driver is distinguished as a part of the state S, then the 1st order activation level moment of user-class c is also important:

$$\begin{aligned} M^{1}_{c,a}(t,x) &= \int_S a \rho_c(t,x,S)\ dS \\ &\equiv k_c(t,x) A_c(t,x) \end{aligned} \tag{7.11}$$

Finally, the total flow rate (or speed momentum) Q_{tot} and the total average speed V_{tot} are defined as:

$$\begin{aligned} Q_{tot}(t,x) &\equiv \sum_c Q_c(t,x) \\ &= \sum_c k_c(t,x) V_c(t,x) \\ &\equiv k_{tot} V_{tot} \end{aligned} \tag{7.12}$$

or, using equation (7.4):

$$\begin{aligned} V_{tot}(t,x) &\equiv \sum_c \frac{k_c(t,x)}{k_{tot}(t,x)} V_c(t,x) \\ &= \sum_c p_c(t,x) V_c(t,x) \end{aligned} \tag{7.13}$$

And similarly for the total average activation level A_{tot}:

$$\begin{aligned} A_{tot}(t,x) &\equiv \sum_c \frac{k_c(t,x)}{k_{tot}(t,x)} A_c(t,x) \\ &= \sum_c p_c(t,x) A_c(t,x) \end{aligned} \tag{7.14}$$

7.3 Multi user-class generalised continuity equation and macroscopic traffic flow models

7.3.1 Multi user-class generalised continuity equation and derivation of the macroscopic multi user-class human-kinetic model

Hoogendoorn (1999) shows that the continuity equation (4.1) for the single user-class case also holds for the generalised partial density $\rho_c(t,x,S)$ of any user-class c in a multi user-class traffic flow:

$$\frac{\partial \rho_c}{\partial t} + \nabla_{S'}\left(\rho_c \cdot \frac{dS'}{dt}\right) = \left(\frac{d\rho_c}{dt}\right)_{discrete} \tag{7.15}$$

The interpretation of this equation is completely analogous to the single user-class case (section 4.2.1). Using the method of moments, this dynamic equation for the probability density can be transformed into equations for the moments of the probability density function ρ_c. Multiplication of both RHS and LHS of equation (7.15) by s_i^κ and subsequent integration over all state variables in S, yields dynamic equations for the κ^{th} order moment of ρ_c for state variable s_i; in general for M^κ_{c,s_i}:

$$\int_S s_i^\kappa \left(\frac{\partial \rho_c}{\partial t} + \nabla_{S'}\left(\rho_c \cdot \frac{dS'}{dt}\right)\right) dS = \int_S s_i^\kappa \left(\frac{d\rho_c}{dt}\right)_{discrete} dS \tag{7.16}$$

Restricting the definition of the vehicle state to $S = (v,a)$, we use the general equation (7.16) to establish dynamic equations for the partial density $k_c = M^0_{c,v}$, for the partial speed momentum $Q_c = M^1_{c,v}$, and for the partial activation level momentum $k_c A_c = M^1_{c,a}$

of user-class *c*. The intermediate results are omitted, since they are completely analogous to the calculations in section A.3.2.2 in Annex A, yielding finally the following equations:

$$\frac{\partial k_c}{\partial t}+\frac{\partial k_c V_c}{\partial x}=\left(\frac{dk_c}{dt}\right)_{discrete} \qquad \forall c=1..C \tag{7.17}$$

These are *C* macroscopic continuity equations expressing the conservation of vehicles of class *c* in space and time. The macroscopic speed momentum equations are (analogous to equation (4.4) for a single user-class):

$$\frac{\partial k_c V_c}{\partial t}+\frac{\partial k_c\left(V_c^2+\Theta_c\right)}{\partial x}=k_c\left\langle\frac{dv}{dt}\right\rangle_v^c+\int_v v\left(\frac{d\rho_c}{dt}\right)_{discrete}dv \qquad \forall c=1..C \tag{7.18}$$

The *C* conservative speed momentum equations (7.18) can also be cast into their primitive form:

$$\frac{\partial V_c}{\partial t}+V_c\frac{\partial V_c}{\partial x}+\frac{1}{k_c}\frac{\partial k_c\Theta_c}{\partial x}=\left\langle\frac{dv}{dt}\right\rangle_v^c+\frac{1}{k_c}\left(\int_v v\left(\frac{d\rho_c}{dt}\right)_{discrete}dv-V_c\left(\frac{dk_c}{dt}\right)_{discrete}\right) \qquad \forall c=1..C \tag{7.19}$$

Finally, if activation level dynamics are taken into account, *C* dynamic equations for the activation level momentum are found (analogous to equation (6.10) for a single user-class):

$$\frac{\partial k_c A_c}{\partial t}+\frac{\partial k_c A_c V_c}{\partial x}=k_c\left\langle\frac{da}{dt}\right\rangle_{v,a}^c+\int_a\int_v a\left(\frac{d\rho_c}{dt}\right)_{discrete}dv\,da \qquad \forall c=1..C \tag{7.20}$$

The primitive formulation of these equations is:

$$\frac{\partial A_c}{\partial t}+V_c\frac{\partial A_c}{\partial x}=\left\langle\frac{da}{dt}\right\rangle_{v,a}^c+\frac{1}{k_c}\int_a\int_v a\left(\frac{d\rho_c}{dt}\right)_{discrete}dv\,da-\frac{A_c}{k_c}\left(\frac{dk_c}{dt}\right)_{discrete} \qquad \forall c=1..C \tag{7.21}$$

7.3.2 Multi user-class acceleration integral

In equations (7.18) and (7.19) the average acceleration of individual drivers of class *c* is present. This acceleration is not only determined by the traffic state of the own user-class, but also of other user-classes. For instance, faster user-classes are influenced by slower user-classes. The acceleration integral, defined by equation (4.11) for a single user-class, is now generalised for the multi user-class case as follows. Let us consider the expected acceleration of a successor of class c_j, then the acceleration integral is found by:

$$\left\langle\frac{dv}{dt}\left(t+T_{c_j},x\right)\right\rangle_v^{c_j}\equiv\int_{v_j}p_v^{c_j}\left(v_j\right)\sum_{c_{j-1}}\int_{s_j}p_{c_{j-1}}^{c_j}\left(c_{j-1}\right)p_s^{c_j}\left(s_j\right)\int_{v_{j-1}}p_v^{c_{j-1}}\left(v_{j-1}\right)\dot{v}_{c_j}^{c_{j-1}}\left(v_j,s_j,v_{j-1}\right)dv_{j-1}ds_jdv_j \tag{7.22}$$

with:

$$
\begin{aligned}
p_v^{c_j}\left(v_j\right) &= p_v^{c_j}\left(v_j \middle| t, x, c_j\right) \\
p_c^{c_j}\left(c\right) &= p_c^{c_j}\left(c \middle| t, x + s_{c_j}^0 + s_j, c_j\right) \\
p_s^{c_j}\left(s_j\right) &= p_s^{c_j}\left(s_j \middle| t, x, c_j, v_j, c\right) \\
p_v^{c_{j-1}}\left(v_{j-1}\right) &= p_v^{c_{j-1}}\left(v_{j-1} \middle| t, x + s_{c_j}^0 + s_j, c_j, v_j, c_{j-1}, s_j\right)
\end{aligned}
\tag{7.23}
$$

These equations show that the expected acceleration of a vehicle of class c_j is evaluated at time t, and takes effect after the user-class specific reaction time T_{cj}.

7.3.2.1 Generic definition of the probability distributions and approximations

In equations (7.22) and (7.23), all theoretical conditions for the probability distributions have been taken into account, so that the probability distributions are all correlated. According to equation (7.23) the expected acceleration for a vehicle of class c_j at location x depends on:

- the speed distribution $p_v^{c_j}\left(v_j\right) = p_v^{c_j}\left(v_j \middle| t, x, c_j\right)$ of speed v_j for a vehicle of user-class c_j
- the probability $p_c^{c_j}\left(c\right) = p_{c_{j-1}}^{c_j}\left(c_{j-1} \middle| t, x + s_{c_j}^0 + s_j, c_j\right)$ that the predecessor is of class $c_{j\text{-}1}$, given that at location $x + s_{c_j}^0 + s_j$ there is a predecessor for a follower j (of class c_j) at x; note that the following relation holds:

$$
\sum_{c_{j-1}} p_{c_{j-1}}^{c_j}\left(c_{j-1} \middle| t, x + s_{c_j}^0 + s_j, c_j\right) = 1 \tag{7.24}
$$

- the probability $p_s^{c_j}\left(s_j\right) = p_s^{c_j}\left(s_j \middle| t, x, c_j, v_j, c_{j-1}\right)$ of having a gap s_j between a follower j of class c_j and the predecessor of class $c_{j\text{-}1}$
- the probability $p_v^{c_{j-1}}\left(v_{j-1}\right) = p_v^{c_{j-1}}\left(v_{j-1} \middle| t, x + s_{c_j}^0 + s_j, c_j, v_j, c_{j-1}, s_j\right)$ that this predecessor of class $c_{j\text{-}1}$ drives with speed v_j, given that there is a follower of class c_j with a gap s_j between them
- the car-following behaviour $\dot{v}_{c_j}^{c_{j-1}}\left(v_j, s_j, v_{j-1}\right)$ of class c_j with respect to a predecessor of class $c_{j\text{-}1}$.

Theoretically, this allows including all kinds of subtle correlations using these conditional probabilities. For instance, it is conceivable that some combinations of follower-leader classes are more likely than others because some drivers prefer a predecessor of a certain class to other classes. In that case, user-classes may tend to form clusters in traffic. Note that also ADA vehicles may be designed to intentionally form clusters.

If such effects are relevant, the according correlations need to be carefully specified in the probability distributions (7.23). However, this makes the calculation of the integral

(7.22) more complex and would require large amounts of multi user-class individual vehicle data to determine the covariance matrices of the multivariate distributions for the speeds, gaps, and vehicle classes. For that reason, the problem is simplified in this thesis by applying, also for the multi user-class case, the approximation traditionally referred to as the *vehicular chaos* assumption, as was done for the single user-class case (see sections 3.4.1 and 4.2.3.1).

The approximation consists again in assuming that the deviations of the speeds of follower and leader with respect to their respective average value are uncorrelated, as is the distribution for the gap between them. Compared to the single user-class case, an additional potential correlation is neglected: no correlations between the classes of the predecessor and the successor are considered. In the following, we consider each probability in equation (7.23) separately, and indicate which simplifications apply and which macroscopic variables are used to define the probability distributions.

7.3.2.2 Specification of the successor's speed v_j

As for the single user-class case (section 4.2.3.1), normal distributions are assumed for the speeds. The speed of a successor of class c_j is taken from a normal distribution:

$$\begin{aligned} p_v^{c_j}\left(v_j \middle| t,x,c_j\right) &= \frac{\rho_{c_j}\left(t,x,v_j\right)}{k_{c_j}\left(t,x\right)} \\ &\equiv N\left(v_j,\mu_{v_j}^{c_j}\left(t,x\right),\sigma_{v_j}^{c_j}\left(t,x\right)\right) \end{aligned} \tag{7.25}$$

The average value of this distribution is the macroscopic speed of the user-class under consideration. The standard deviation is equal to the square root of the macroscopic speed variance of that user-class. Again the speed variance is assumed to be an equilibrium value (section 4.2.3.1), so that:

$$\begin{aligned} \mu_{v_j}^{c_j}\left(t,x\right) &= V_{c_j}\left(t,x\right) \\ \sigma_{v_j}^{c_j}\left(t,x\right) &= \sqrt{\Theta_{c_j}\left(t,x\right)} \\ &\equiv \sqrt{\Theta_{c_j}^{e}\left(k_{tot}\left(t,x\right),V_{c_j}\left(t,x\right)\right)} \end{aligned} \tag{7.26}$$

The speed variance is found analogously to equation (4.16) for the single user-class case. Here the equilibrium speed variance is an empirical function of the *total* density and of the *own* speed (Shvetsov & Helbing, 1999):

$$\begin{aligned} \Theta_{c_j}^{e}\left(k_{tot}\left(t,x\right),V_{c_j}\left(t,x\right)\right) &\equiv V_{c_j}^2\left(t,x\right)B\left(k_{tot}\left(t,x\right)\right) \\ &\equiv V_{c_j}^2\left(t,x\right)\left(b_0+b_1\left(1+e^{-\frac{k_{tot}\left(t,x\right)-b_2\,\overline{k_{jam}}\left(t,x\right)}{b_3\,\overline{k_{jam}}\left(t,x\right)}}\right)^{-1}\right) \end{aligned} \tag{7.27}$$

In this equation the local average jam density is used that is found as a weighed average of the jam densities of all vehicle classes present at t and x:

$$\overline{k_{jam}}(t,x) = \frac{\sum_c k_c(t,x) k_{jam}^c}{\sum_c k_c(t,x)} = \frac{\sum_c k_c(t,x) k_{jam}^c}{k_{tot}(t,x)} \tag{7.28}$$

7.3.2.3 Specification of the class of the predecessor c_{j-1}

This section specifies the probability $p_c^{c_j}(c_{j-1}) = p_{c_{j-1}}^{c_j}\left(c_{j-1} \middle| t, x + s_{c_j}^0 + s_j, c_j\right)$ that the predecessor is of class c_{j-1}, given that at location $x + s_{c_j}^0 + s_j$ there is a predecessor for a follower j (of class c_j) at x. Since in this conditional probability the gap s_j occurs, and (as appears from the next section) on the other hand the class c_{j-1} of the predecessor occurs in the specification of the gap distribution, we need to solve an implicit definition. This can be solved by approximating the gap s_j by the average of the gap distribution: $s_j \approx \mu_{s_j}^{c_j}$. As will appear from the next section (equation (7.38), this class-average gap $\mu_{s_j}^{c_j}$ for class c_j is independent of the predecessor's class c_{j-1}. Therefore, the probability $p_c^{c_j}(c_{j-1})$ can be written by the following explicit approximation:

$$\begin{aligned} p_c^{c_j}(c_{j-1}) &\approx p_{c_{j-1}}^{c_j}\left(c_{j-1} \middle| t, x + s_{c_j}^0 + \mu_{s_j}^{c_j}, c_j\right) \\ &\approx \frac{k_{c_{j-1}}\left(t, x + s_{c_j}^0 + \mu_{s_j}^{c_j}\right)}{k_{tot}\left(t, x + s_{c_j}^0 + \mu_{s_j}^{c_j}\right)} \end{aligned} \tag{7.29}$$

Note that, as argued in section 7.3.2.1, independence of the probability for class c_{j-1} with respect to class c_j is assumed in (7.29), so that equation (7.4) can be applied to define the probability $p_c^{c_j}(c_{j-1})$.

7.3.2.4 Specification of the gap to the predecessor s_j

As for the single class case, we assume lognormally distributed gaps, and according to the arguments in section 7.3.2.1, no correlations between the gaps and the speeds. Again the standard deviation of the distribution function is assumed to be monotonically increasing with the average gap, using an exponent b (see section 4.2.3.1):

$$p_s^{c_j}\left(s_j \middle| t, x, c_j, v_j, c_{j-1}\right) \equiv LgN\left(s_j, \mu_{s_j}^{c_j}(t,x), \sigma_{s_j}^{c_j}(t,x)\right) \tag{7.30}$$

$$\sigma_{s_j}^{c_j}(t,x) \equiv \mu_{s_j}^{c_j}(t,x)^b \tag{7.31}$$

Under-determination of class-average gaps

This section explains why in the multi user-class case the class-average gap $\mu_{s_j}^{c_j}(t,x)$ cannot be uniquely determined from the total (and partial) vehicle density, as was the

case for a single user-class (equation (4.26)). The problem is that the macroscopic densities of the c classes cannot be directly used to determine the C unknown class-average gaps $\mu_{s_j}^{c_j}(t,x)$, without making additional assumptions: the problem is under-determined, with only one condition set to the C unknowns, as explained hereafter.

First of all, the requirement must hold that the average over all vehicles of the sum of gap $\mu_{s_j}^{c_j}$ and gross vehicle length $s_{c_j}^0$ (see equation (4.20)) equals the inverse of the total density:

$$\sum_{c_j} p_c\left(c_j \middle| t,x'\right)\left(\mu_{s_j}^{c_j}(t,x)+s_{c_j}^0\right)=\frac{1}{k_{tot}(t,x')} \tag{7.32}$$

Or, using equation (7.4):

$$\sum_{c_j} k_{c_j}(t,x')\left(\mu_{s_j}^{c_j}(t,x)+s_{c_j}^0\right)=1 \tag{7.33}$$

Note that the densities here are evaluated at an interaction point (denoted with x') someplace downstream of x. Following the arguments given in section 4.2.3.1 for the single user-class case, the non-local densities in equation (7.32) are defined by:

$$x' \equiv x+\frac{1}{2k_{tot}(t,x)} \tag{7.34}$$

Intuitively, one would think that a further specification comparable to equation (4.26) for the single user-class case holds:

$$\mu_{s_j}^{c_j} \overset{?}{=} \frac{1}{k_{c_j}}-s_{c_j}^0 \tag{7.35}$$

This is, however, not a correct solution, which can be easily understood by considering the following example. Consider 2 classes, each with a homogeneous partial density of 1 veh/km and a gross vehicle length $s_{c_j}^0 = 0$. Whereas in reality there would in total be 2 veh/km, and the sum of the class-average gaps should be 1 km (see (7.33)), equation (7.35) would predict $\mu_{s_j}^1 = \mu_{s_j}^2 = 1\,\text{km}$ and a sum of the class-average gaps of 2 km. The latter is clearly incorrect.

More in general, the problem is the following. With equation (7.33) the flow-averaged gap $\mu_{s_j}^{tot}$ over all vehicles in the flow can be defined as:

$$\begin{aligned}\mu_{s_j}^{tot}(t,x) &\equiv \sum_{c_j}\frac{k_{c_j}(t,x')}{k_{tot}(t,x')}\mu_{s_j}^{c_j}(t,x)\\ &= \frac{1-\sum_{c_j}k_{c_j}(t,x')s_{c_j}^0}{k_{tot}(t,x')}\end{aligned} \tag{7.36}$$

If C user-classes are considered, equation (7.33) leaves (C-1) degrees of freedom for the average gaps $\mu_{s_j}^{c_j}$ of the C user-classes; in other words: the flow-average gap $\mu_{s_j}^{tot}$ is known, but it is undetermined whether all user-classes are assigned the same class-average gap, or some classes have above-average gaps at the cost of other classes with smaller-than-average gaps. As an example, all traffic situations in Figure 7-1 are macroscopically equivalent (in the sense that the partial densities of both user-classes over the road length are the same in each situation), so that the flow-average gap $\mu_{s_j}^{tot}$ according to (7.35) is also the same, but the assignment of class-average gaps over the user-classes differs.

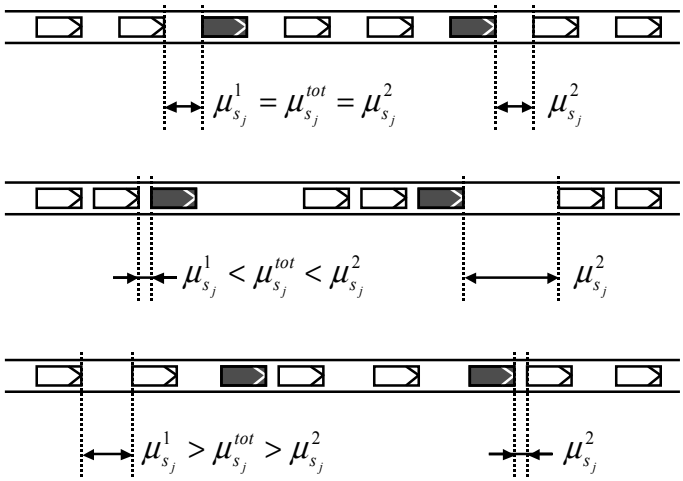


Figure 7-1 Traffic situations with same flow-average gaps but different assignment of class-average gaps;* up*: both classes have same class-average gap equal to the flow-average gap,* middle*: user-class 2 (dark) has above-average gaps,* below*: user-class 1 (light) has above-average gaps

Existing solutions to under-determination of class-average gaps

Any multi user-class traffic flow model that takes finite vehicle space requirements (i.e. finite vehicle lengths and gaps and not dimensionless or point-sized vehicles) into account needs additional assumptions to solve the issue of gap assignment over the user-classes. We briefly review how some other multiclass models deal with this issue.

The multiclass generalisation of Helbing's model (1997) is only given for an intermediate model and not for the final model that includes finite space requirements. Since at that stage in the development, the author still considers point-sized vehicles no solution to the above-mentioned problem is needed. Hoogendoorn (1999) circumvents the problem by first considering point-sized vehicles and deriving the speed loss due to interactions between vehicles of the own class and with other classes based on that assumption. Once this model derived, he introduces a speed-dependent *finite space* factor and multiplies the speed change due to interactions by that factor. This procedure, however, does not fit to the rationale behind the human-kinetic model derived in this thesis, since the latter builds fundamentally on car-following behaviour, so that the finite space of vehicles is a starting point, rather than an ex-post correction.

Logghe (2003) gives an excellent review on how this issue is solved in multi-class variants of the first order macroscopic model of Lighthill, Whitham and Richards. He formulates some requirements that need to be obeyed by any gap assignment solution in order to be realistic. The main requirements are:

- that equilibrium speeds in congestion (and only there, i.e.: *not* in free flowing traffic for mixed-lane models) should be homogeneous over classes;
- that a class can only be forced to drive slower by slower classes, and not by faster traffic (i.e. faster classes cannot gain speed by claiming more space at the cost of slower classes that would, as a result, be forced to slow down further); otherwise driver behaviour of a slower class would be influenced by faster traffic coming from behind (upstream) and the anisotropic behaviour of drivers is violated.

Since none of the existing methods in his review met these basic requirements, Logghe formulated a novel solution, based on a *user-class optimum* or *user equilibrium*. Unfortunately this solution builds on a specific triangle shape of the fundamental diagram underlying the LWR model, which makes it not directly suitable for our purposes.

Human-kinetic model-specific solution to under-determination of class-average gaps

Since none of the existing approaches match the requirements set by Logghe or apply to the human-kinetic model, we are therefore challenged to propose a novel solution for the assignment of gaps over the different user-classes. The basic requirements and the concept of user equilibrium in traffic flow proposed by Logghe are withheld. However, given the totally different modelling approach, we build our own specification to obtain this equilibrium.

The two requirements of the previous section can actually be reformulated as follows: all classes maximise their speed to such extent that it does not force other traffic to drive slower. Consider high density, so that there is not enough space available for all classes to drive at their maximum desired speed. Then, given the fact that the desired gap of any class is an increasing function of the driving speed (see equation (4.28)), speed can only be maximised if the gaps are maximised. That is: no class can drive faster than another, since this would only be possible by claiming extra road space at the cost of the slower classes, which violates the requirements. In other words: the requirements can only be met if gaps are assigned in such a way that the *class-average speed of each class is the same* and equal to the flow-average speed[15].

Based on these arguments, the following assignment of gaps over the user-classes on a lane is proposed:

[15] Note that on multi-lane facilities, drivers of faster classes might prefer changing lanes to a faster lane instead of adapting their speed to that of the slower class. In that case, the class-average speed of the fraction that stays within the lane should either eventually converge to the flow-average speed on that lane or disappear completely and leave for the faster lane.

$$\frac{\mu_{s_j}^{c_j}(t,x)}{s_{c_j}^{d}\left(V_{tot}(t,x)\right)} \equiv f_{acc}(t,x) \quad \forall c_j = 1..C \tag{7.37}$$

This equation states the following: the ratio of the class-average gap over the space needed according to equation (4.28) by that class to drive with the flow-average speed on the current lane V_{tot} (equation (7.13)) is the same for each user-class, and equal to some constant f_{acc}. In the remainder of this section, it is proven first that this specification meets the requirements by Logghe, after which the factor f_{acc} (called the *acceleration opportunity*) is specified.

Let us first proof that equation (7.37) guarantees that the requirements by Logghe be met. We argued that this is the case if all user-classes strive for the same average speed. Suppose now that a class c^* drives with class-average speed $V_{c^*} > V_{tot}$. It will then be assigned a class-average gap $\mu_{s_j}^{c^*} = s_{c^*}^{d}(V_{tot}) f_{acc} < s_{c^*}^{d}(V_{c^*}) f_{acc}$. Suppose that the acceleration opportunity f_{acc} = 1 (corresponding to a stationary state with flow-average acceleration = 0, see below), it appears that this class is assigned a gap $\mu_{s_j}^{c^*} = s_{c^*}^{d}(V_{tot})$, smaller than $s_{c^*}^{d}(V_{c^*})$ needed to maintain its higher speed. Therefore, according to the car-following equation (4.27), the average response by drivers of that class will be to decelerate. Actually, the only class-average speed that would lead to an expected acceleration equal to zero is $V_{c^*} = V_{tot}$. Similar arguments hold for a class driving slower than the flow-average speed, so that it can be concluded that classes strive for an equilibrium in which all class-average speeds V_{c_j} converge to the flow-average speed V_{tot}.

Let us now determine the acceleration opportunity f_{acc} and analyse its meaning. With equations (7.32) and (7.37) C requirements have been set to C degrees of freedom that can easily be shown to have the following solution for the class-average gap:

$$\mu_{s_j}^{c_j}(t,x) = s_{c_j}^{d}\left(V_{tot}(t,x)\right) \frac{1 - \sum_{c_j} k_{c_j}(t,x') s_{c_j}^{0}}{\sum_{c_j} k_{c_j}(t,x') s_{c_j}^{d}\left(V_{tot}(t,x)\right)} \tag{7.38}$$

Or, in terms of the acceleration opportunity:

$$f_{acc}(t,x) = \frac{1 - \sum_{c_j} k_{c_j}(t,x') s_{c_j}^{0}}{\sum_{c_j} k_{c_j}(t,x') s_{c_j}^{d}\left(V_{tot}(t,x)\right)} \tag{7.39}$$

The acceleration opportunity is thus equal to the ratio of the road space that is not physically occupied by vehicles (numerator) over the total road space needed by all vehicles to drive with the flow-average speed V_{tot} (denominator). In other words: it is the ratio of the total available space over the total required space at this average speed (see Figure 7-2). Therefore, a value f_{acc} = 1 corresponds to an equilibrium or stationary state with flow-average acceleration = 0. If the ratio is larger than one, then all classes

are assigned more space than needed to maintain their speed, so they can accelerate, and vice versa; hence the name 'acceleration opportunity'.

Note that, as was presupposed in section 7.3.2.3, the definition of the class-average gap $\mu_{s_j}^{c_j}$ in equation (7.38) is independent of the user-class c_{j-1} of the predecessor.

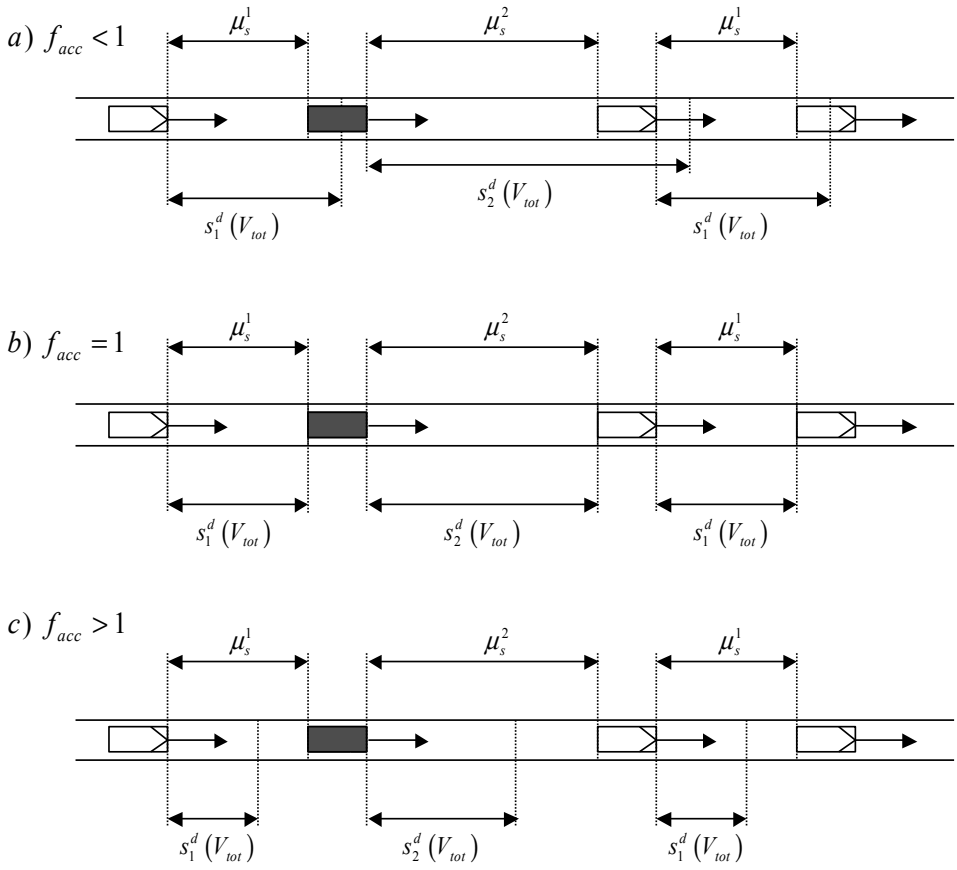


Figure 7-2 Physical representation of the acceleration opportunity; partial densities are the same in each case, the flow-average speed varies from high (a:* above*) over equilibrium (b:* middle*) to low (c:* below*)

7.3.2.5 Specification of the predecessor's speed v_{j-1}

Similar to the model development for the single user-class case in section 4.3.2, we specify first the objective specification of the probability distribution of the predecessor's speed (i.e. without anticipation by the successor). In a second step, the objective definition is modified into a subjective definition to account for the anticipation of the successor.

Objective predecessor speed

For the speed distribution of the predecessor of class c at location $x + s_{c_j}^0 + s_j$, a normal distribution is assumed:

$$p_v\left(v_{j-1}\middle|t, x + s_{c_j}^0 + s_j, c_j, v_j, c_{j-1}, s_j\right) \equiv N\left(v_{j-1}, \mu_{v_{j-1}}^{c_{j-1}}\left(t, x + s_{c_j}^0 + s_j\right), \sigma_{v_{j-1}}^{c_{j-1}}\left(t, x + s_{c_j}^0 + s_j\right)\right) \quad (7.40)$$

The average value of this distribution for the case without anticipation is the actual macroscopic speed of the user-class c_{j-1} under consideration at location $x + s_{c_j}^0 + s_j$. The standard deviation is equal to the square root of the macroscopic equilibrium speed variance of that user-class at the same location:

$$\begin{aligned}\mu_{v_{j-1}}^{c_{j-1}}\left(t, x + s_{c_j}^0 + s_j\right) &= V_{c_{j-1}}\left(t, x + s_{c_j}^0 + s_j\right) \\ \sigma_{v_{j-1}}^{c_{j-1}}\left(t, x + s_{c_j}^0 + s_j\right) &= \sqrt{\Theta_{c_{j-1}}\left(t, x + s_{c_j}^0 + s_j\right)} \\ &\equiv \sqrt{\Theta_{c_{j-1}}^e\left(k_{tot}\left(t, x + s_{c_j}^0 + s_j\right), V_{c_{j-1}}\left(t, x + s_{c_j}^0 + s_j\right)\right)}\end{aligned} \tag{7.41}$$

The equilibrium speed variance is given by equation (7.27).

Anticipated predecessor speed

Similar to the single user-class case, anticipation is specified by replacing the physically determined average speed $\mu_{v_{j-1}}^{c_{j-1}}\left(t, x + s_{c_j}^0 + s_j\right) = V_{c_{j-1}}\left(t, x + s_{c_j}^0 + s_j\right)$ in equation (7.41) with an *anticipated* non-local speed. Note that here only the mathematical consequences of having multiple user-classes are considered; for a description and motivation of the behavioural model underlying these specifications, the reader is referred to section 4.3.2.

Because we consider the speed of the predecessor of class c_{j-1} as anticipated by the successor of class c_j, properties or variables of both classes are used in the definition of the anticipated non-local speed:

$$\begin{aligned}\mu_{v_{j-1}}^{c_{j-1},c_j}\left(t, x + s_{c_j}^0 + s_j\right) &\equiv V_{c_{j-1},c_j}^{ant}\left(t, x + s_{c_j}^0 + s_j\right) \\ &\equiv V_{c_{j-1}}(t,x) + f_{c_j}^{ant}\left(a_{c_j}(t,x)\right)\left(s_{c_j}^0 + s_j\right)\left(\frac{dV_{c_{j-1}}}{dx}\right)_{c_j}^{ant}\end{aligned} \tag{7.42}$$

Equation (7.42) expresses that the non-local average speed of the predecessor of user-class c_{j-1}, anticipated by the successor of class c_j, depends on:

- the actual speed $V_{c_{j-1}}$ of class vehicles at location x
- the gradient $\left(\frac{dV_{c_{j-1}}}{dx}\right)_{c_j}^{ant}$ of the spatial speed profile of class vehicles as anticipated by user-class c_j
- the anticipation strength $f_{c_j}^{ant}\left(a_{c_j}\right)$ of the successor class c_j, which is possibly a function of the activation level a_{c_j} of the successor j
- the distance $\left(s_{c_j}^0 + s_j\right)$ between both vehicles.

Assume that drivers of class c_j can estimate speeds in a range $\Delta x_{c_j}^{ant}$ downstream. The anticipated gradient $\left(\frac{dV_{c_{j-1}}}{dx}\right)_{c_j}^{ant}$ of the spatial speed profile of class c_{j-1} vehicles, as anticipated by user-class c_j, is then defined by:

$$\left(\frac{dV_{c_{j-1}}}{dx}\right)_{c_j}^{ant}(t,x) \equiv \min_{\Delta x \in (0,\Delta x_{c_j}^{ant}]}\left(\frac{V_{c_{j-1}}(t,x+\Delta x)-V_{c_{j-1}}(t,x)}{\Delta x}\right) \tag{7.43}$$

7.3.2.6 Specification of the multi user-class car-following behaviour

The multi user-class framework of the human-kinetic traffic flow theory allows incorporating any car-following relation in the model, as long as the microscopic variables in the model can be derived from the macroscopic traffic flow variables. Therefore, it is in theory possible to specify *different car-following behaviour*, dependent on the user-classes involved. For that purpose it is sufficient to specify a matrix of car-following relations and include these relations in equation (7.22):

$$\dot{v}_{c_j}^{c_{j-1}}\left(v_j, s_j, v_{j-1}\right) \qquad \forall c_j, c_{j-1} = 1..C \tag{7.44}$$

For now, we have written equation (7.44), assuming that these car-following relations are all dependent on the speeds of and gap between the vehicles involved. In theory, it is conceivable that one or more car-following rules contain interactions with more than two vehicles, on the same lane or on adjacent lanes. In that case, the acceleration integral (7.22) must be generalised to include more vehicles.

The possibility of having different car-following relations can be important to model ADA systems correctly. For instance, vehicle-driver units with Adaptive Cruise Control (ACC) may have a different desired distance function (equation (4.28)) and acceleration controller than the one specified in equation (4.27). In that case the mathematical expression for car-following differs between user-classes, and it is not longer sufficient to have different parameter sets for the different user-classes (as is the case for instance in Hoogendoorn, 1999). In the examples given in this thesis however, only the parameter sets have been varied between user-classes, and not the car-following relation.

7.3.3 Numerical evaluation of the multi user-class human-kinetic traffic flow model

For the numerical evaluation of the macroscopic multiclass human-kinetic traffic flow model the equations (7.17) and (7.18) are cast into the following structure (no activation level dynamics are considered here; however, the arguments in this section apply equally to that case):

$$\frac{\partial}{\partial t}\mathbf{U} + \frac{\partial}{\partial x}\mathbf{F}(\mathbf{U}) = \mathbf{G}(\mathbf{U}) \tag{7.45}$$

with

$$\mathbf{U} = \begin{pmatrix} u_{11} \\ u_{12} \\ \vdots \\ u_{c1} \\ u_{c2} \\ \vdots \\ u_{C1} \\ u_{C2} \end{pmatrix} \equiv \begin{pmatrix} k_1 \\ k_1 V_1 \\ \vdots \\ k_c \\ k_c V_c \\ \vdots \\ k_C \\ k_C V_C \end{pmatrix} \tag{7.46}$$

$$\mathbf{F}(\mathbf{U}) \equiv \begin{pmatrix} u_{12} \\ \frac{u_{12}^2}{u_{11}} + u_{11}\Theta^e\left(u_{11},..,u_{C1},u_{12}\right) \\ \vdots \\ u_{c2} \\ \frac{u_{c2}^2}{u_{c1}} + u_{c1}\Theta^e\left(u_{11},..,u_{C1},u_{c2}\right) \\ \vdots \\ u_{C2} \\ \frac{u_{C2}^2}{u_{C1}} + u_{C1}\Theta^e\left(u_{11},..,u_{C1},u_{C2}\right) \end{pmatrix} \qquad \mathbf{G}(\mathbf{U}) \equiv \begin{pmatrix} \left(\frac{du_{11}}{dt}\right)_{discrete} \\ u_{11}\left\langle \frac{dv}{dt} \right\rangle_v^1 + \int_v v\left(\frac{d\rho_1}{dt}\right)_{discrete} dv \\ \vdots \\ \left(\frac{du_{c1}}{dt}\right)_{discrete} \\ u_{c1}\left\langle \frac{dv}{dt} \right\rangle_v^c + \int_v v\left(\frac{d\rho_c}{dt}\right)_{discrete} dv \\ \vdots \\ \left(\frac{du_{C1}}{dt}\right)_{discrete} \\ u_{C1}\left\langle \frac{dv}{dt} \right\rangle_v^C + \int_v v\left(\frac{d\rho_C}{dt}\right)_{discrete} dv \end{pmatrix} \tag{7.47}$$

As for the single class case, the eigenvalues of the Jacobian matrix $\frac{\partial \mathbf{F}}{\partial \mathbf{U}}$ are needed for the numerical evaluation. Note that not only the RHS function $\mathbf{G}(\mathbf{U})$ contains coupling between the sets of two equations per class (via the acceleration integral), but also the function $\mathbf{F}(\mathbf{U})$. For the determination of the equilibrium speed variance the total density k_{tot} (equation (7.2)) and the weighed jam density $\overline{k_{jam}}$ (equation (7.28)) are needed, and hence the densities $u_{11},..,u_{C1}$ occur in the second equation of each user-class. Therefore, the Jacobian, and thus the eigenvalues, are not easily determined from equation (7.47). However, we choose to approximate the Jacobian by decoupling the function $\mathbf{F}(\mathbf{U})$ in a set of two equations per user-class, that is independent from the density u_{c1} of the other user-classes. In other words, all terms $\frac{dk_{tot}}{du_{c1}}$ and $\frac{d\overline{k_{jam}}}{du_{c1}}$ are neglected, which is justified by noticing that the eigenvalues are purely used for determining the direction of convective information flow through the cell interface,

which is always positive, except for the pragmatic correction introduced in section 4.4.3.2. Therefore, neglecting the above mentioned derivatives is the same as slightly modifying the pragmatic correction for the discretisation error discussed in section 4.4.3.1.

Once these derivatives have been neglected, the convective part $\mathbf{F}(\mathbf{U})$ of equation (7.47) can be decoupled per user-class. Therefore, the solution scheme of the single user-class applies again (section 4.4). We only have to update the solution (u_{c1}, u_{c2}) of each user-class *c* in parallel, using the same scheme.

The same decoupling can be performed if the activation level momentum per class is added to the definitions of $\mathbf{U}$, $\mathbf{F}(\mathbf{U})$, and $\mathbf{G}(\mathbf{U})$. The solution (u_{c1}, u_{c2}, u_{c3}) can be updated in parallel for each user-class *c*, using the scheme of section 6.3.2.

7.4 Theoretical case study

This section examines the numerical simulation of the multiclass model presented in this chapter. First, equilibrium solutions of a specific combination of user classes (person cars and trucks) are presented (section 7.4.1). Then, section 7.4.2 shows a simple case study of a dynamic simulation, where a fraction of the cars are replaced by an equal fraction of trucks through a weaving area. Both the equilibrium and dynamic case study intend to illustrate how interactions between user classes occur, and how these interactions affect the behaviour of traffic flow. The simulations also show some weaknesses of the current multiclass human-kinetic model that are discussed in section 7.4.3, along with some proposed remedies for future work.

7.4.1 Equilibrium solutions of the multi user-class human-kinetic traffic flow model

Similar to the single user-class case, equilibrium solutions are establish for the system of equations (7.17) and (7.18). Again, these are found by setting the solutions to the acceleration integral equal to zero for a given combination of *C* partial *equilibrium* densities k_c^e and *C* homogeneous macroscopic *equilibrium* speeds V_c^e in space and time. This yields *C* equations of the form:

$$\left\langle \frac{dv}{dt} \right\rangle_v^{c_j} = f^{c_j}\left(k_{c_1}^e, .., k_{c_C}^e, V_{c_1}^e, .., V_{c_C}^e\right) = 0 \qquad \forall c_j = 1..C \tag{7.48}$$

These equations set *C* requirements to 2*C* unknowns (*C* equilibrium densities and *C* equilibrium speeds). Therefore, if a combination of *C* partial densities k_c^e is chosen, the corresponding equilibrium velocities V_c^e (or equilibrium flows Q_c^e, found by multiplication of partial density and speed per class) are found by solving the *C* equations (7.48).

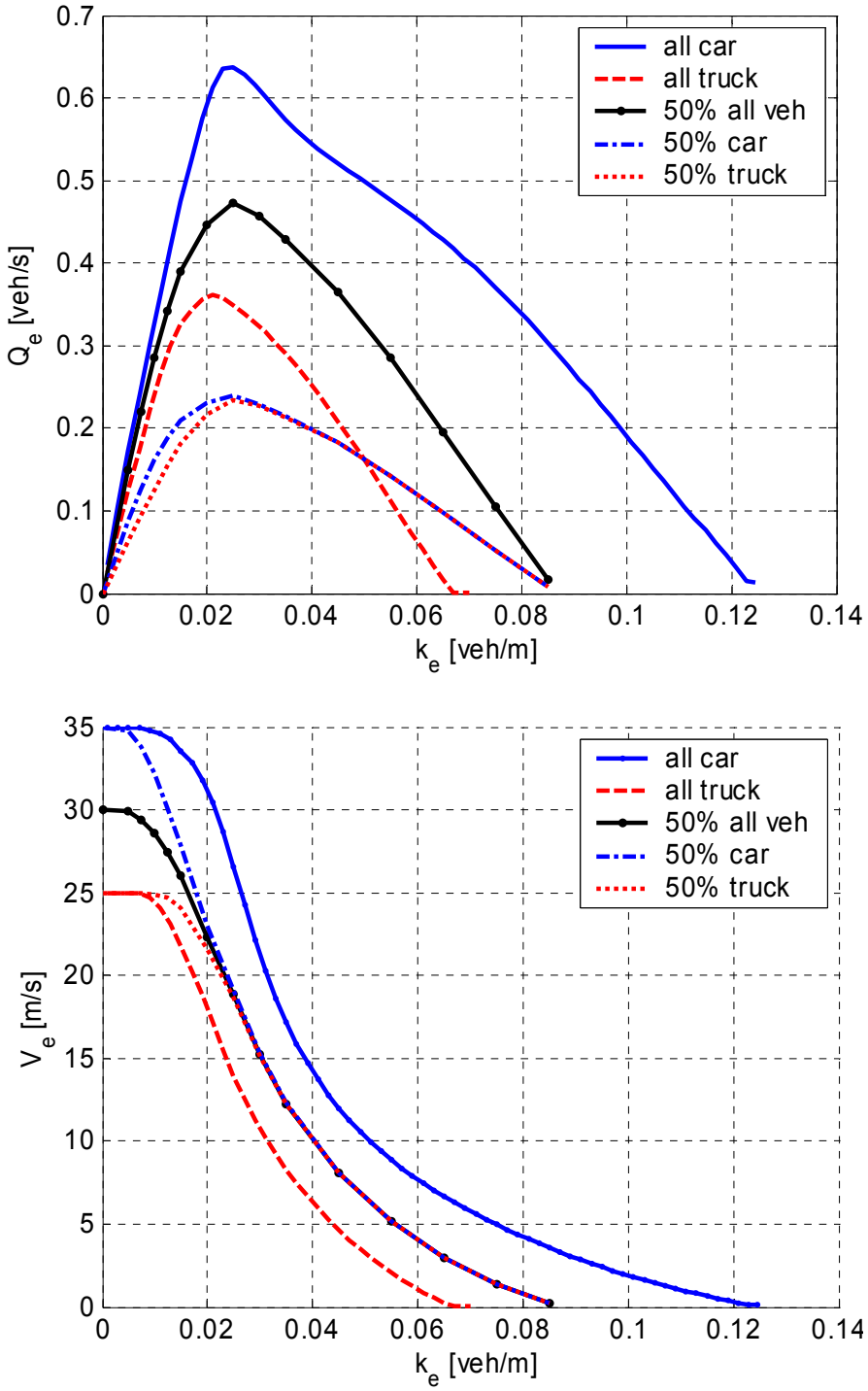


Figure 7-3 Equilibrium solutions as a function of total density for a two user-class system of person cars and trucks with 0%, 50% and 100% trucks; (*left*) equilibrium flow rate and (*right*) equilibrium class- and flow-average speed

7.4.1.1 Multi user-class equilibrium states and speed synchronisation

Let us illustrate the solutions to this equilibrium system for the case of two user-classes: person cars and trucks. In Figure 7-3, the equilibrium flows and speeds, together with the total flow Q_{tot} and flow-average speed V_{tot}, are depicted as a function of the total

density k_{tot}. Class fractions (p_{car}, p_{truck}) = (100%, 0%), (50%, 50%) and (0%, 100%) are considered, i.e. the partial densities are (k^e_{car} , k^e_{truck}) = (p_{car}, p_{truck}) k_{tot}.

The mixed class equilibrium state lies between the equilibrium states for the single class cases, as could be expected. Note that for the case of congested traffic (here for k_{tot} > 30 veh/m), the average speeds of both user-classes become identical. This is a logical result: in the congestion regime, car-following is by far dominant over free driving. When driving on one lane without overtaking possibility the average vehicular speeds therefore tend to converge. We refer to this phenomenon by *intra-lane speed synchronisation* (in contrast to inter-lane speed synchronisation near capacity, which is a typical property of synchronised flow, see section 2.4.2).

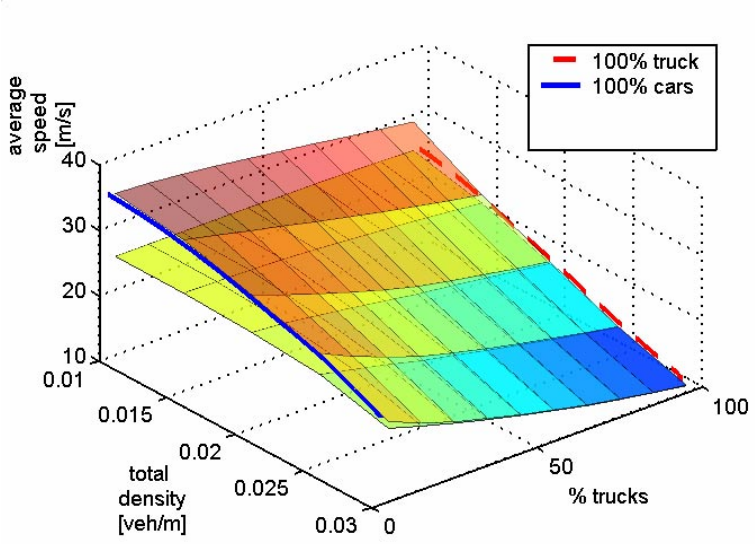


Figure 7-4 Equilibrium class-average speed solutions in the low density region for a two user-class system of person cars and trucks as a function of total density and truck percentage

Looking at the low-density region in Figure 7-3, it can be noticed that speeds of cars and trucks are *not* synchronised in the free flowing regime. Figure 7-4 gives a more detailed view on the average speeds of both user-classes in the free flowing regime and near capacity, for a wide range of truck percentages. Also, the extremes of 0 and 100% trucks are shown; naturally the graph coincides at these extremes with the single class curves that are given as a reference. It appears that the average speeds of the user-classes are different unless the density is high enough, no matter how high the truck percentage. This result is partially counterintuitive. On the one hand, the image of free flowing traffic of a mixture of fast and slower user-classes is that of mixed-class platoons, led by vehicles of the slow class, and some faster vehicles driving freely with higher speeds in the larger gaps between the platoons. The average speed of the faster class is then indeed lower than the own free speed (because of the probability that the vehicle is platooning behind the slower user-class) and higher than that of the slow user-class (because of the probability that the vehicle is driving freely in between the platoons). On the other hand, the image just invoked is not a stable solution: as time elapses, the freely driving faster cars would eventually catch up with (the platoon

behind) a slower truck, slow down to join the platoon, and drive on with the speed of the truck. Therefore, when looking at *equilibrium* solutions (homogeneous in time, i.e. stable for t approaching infinity) and *one-lane traffic*, one would expect speed synchronisation between user-classes to the speed of the slowest class in the free flowing regime as well. It is explained in section 7.4.3 why the solution predicted by our model differs from this theoretical requirement.

7.4.1.2 Asymmetric interactions

Figure 7-4 illustrates another desired property of the multi user-class human-kinetic model: *asymmetry of the vehicular interactions*. This means that slow user-classes have larger influence on faster user-classes than vice versa (Hoogendoorn, 1999; Bliemer, 2001). In the figure this is best visible for the lowest density: the average truck speed is independent on the percentage of person cars, whereas the average car speed clearly decreases with the percentage of trucks.

Asymmetric interactions are modelled correctly, thanks to the careful consideration of vehicular interactions in the acceleration integral (equations (7.22)). Let us first consider a vehicle of the fast user-class. With a certain probability it is preceded by another fast vehicle, in which case it would more or less maintain its speed. Otherwise the predecessor is of the slower class, which would most probably result in a deceleration. The expected acceleration is a weighing of these two possibilities, yielding equilibrium between the desired speeds of both user-classes. Let us now consider a vehicle of the slow user-class. Either the predecessor is of the same class so that the speed is maintained, or the predecessor is the faster user-class, in which case the vehicle can at best accelerate to its own desired speed. In both cases its speed is limited to the desired speed of the own class only. Therefore, the influence of the faster user-class is non-existent, until the density is so high that interactions between all vehicles force traffic to speeds lower than the desired speed of the slower class (see previous section).

We conclude that the interaction process governing the human-kinetic traffic flow model naturally guarantees asymmetric interactions between slower and faster vehicles.

7.4.2 Dynamic solutions of the multi user-class human-kinetic traffic flow model

In this section a theoretical case study is considered that illustrates the interaction between user-classes in the multi user-class model.

7.4.2.1 Simulation set-up

For this case study, the same road configuration as in sections 5.4, 6.5.3 and 6.6.3 is considered; however, the on-ramp at x = 6000 m is now replaced by a weaving section. This means that at this point, some vehicles of the main flow leave the road through the weaving section, while others enter the road. The simulation consists of two user-classes: person cars and trucks, the parameter settings of which are listed in Table 7-1.

Table 7-1 Parameters of cars and trucks

parameter	description	value car	value truck	see eq.
k_{jam}	jam density; inverse of gross vehicle length	0.125 veh/m	0.0667 veh/m	**(4.20)**
W	average desired speed	35 m/s	25 m/s	**(4.28)**
b	exponent determining the gap variance	0.8	0.8	**(4.25)**
(b_1,b_2,b_3,b_4)	shape factors determining the speed variance	(0.007; 0.03; 0.275; 0.03)	(0.007; 0.03; 0.275; 0.03)	**(4.16)**
T	reaction time	0.5 s	0.5 s	**(4.33)**
τ_w	relaxation constant for the desired speed	2.5 s^{-1}	3.3 s^{-1}	**(4.27)**
s_1^d	linear factor for desired gap	1.0 s	1.5 s	**(4.28)**
s_2^d	quadratic factor for desired gap	0.0 s^2/m	0.01 s^2/m	**(4.28)**
τ_s	relaxation constant for the desired distance	3.333 s^{-2}	5 s^{-2}	**(4.27)**
τ_v	relaxation constant for the speed difference	1.111 s^{-1}	1.111 s^{-1}	**(4.27)**
acc_{min}	minimal acceleration	-5 m/s^2	-4 m/s^2	**(4.28)**
acc_{max}	maximal acceleration	3.5 m/s^2	2.5 m/s^2	**(4.28)**
Δx^{ant}	anticipation distance	150 m	150 m	**(4.37)**
f^{ant}	anticipation strength factor	1	1	**(4.36)**

Two scenarios are simulated. The demand on the main road is 2000 veh/h, consisting of 1740 cars/h (87%) and 260 trucks/h (13%). Instead, of extra traffic merging trough the on-ramp, the ramp is now considered to be a weaving section, where cars exit and trucks enter the road. We have chosen equal numbers of cars leaving and trucks entering the road, so that the total flow rate (in veh/h) on the main road is not changed, but traffic composition is: as a net effect a fraction of cars is replaced by trucks. In the first scenario, the flow of cars leaving the road is 200 veh/h, as is the flow rate of entering trucks. As a result, the fraction of trucks is increased to 23%. In the second scenario, this flow is 500 veh/h, so that the truck fraction after the weaving area is increased to 38%.

7.4.2.2 Simulation results

The results of the simulations of both scenarios are given in Figure 7-5. The following conclusions can be drawn from these results:

- The increase of the truck percentage to 23% causes no congestion, whereas further increase to 38% causes a queue upstream of the weaving area. Because trucks require larger gaps, the replacement of cars by trucks reduces capacity. In the first scenario capacity apparently does not drop below the demand flow rate of 2000 veh/h, as can be seen from the lower graph: flow rate before and after the weaving section at x = 6000 m is equal to the demand flow rate. In scenario 2 the capacity decreases to approximately 1700 veh/h, due to the presence of more (space 'consuming') trucks, and a queue builds up upstream of the weaving bottleneck.
- From the speed plot of scenario 1 and the free flowing regions of scenario 2 in Figure 7-5, it appears that the user-class speeds are not synchronised, although no overtaking opportunity has been considered (1-lane traffic). The lack of synchronisation is stronger as total density is lower, or fewer trucks are present; these findings are consistent with the conclusions of section 7.4.1.1 and with Figure 7-4.
- The same plots show that speeds are synchronised during congestion and during the deceleration and acceleration phases in the jam fronts, indicating that congestion is equally severe and modelled correctly for both classes.
- The congestion pattern is close to *homogeneous* congested traffic with almost no oscillations. This indicates that the transition to congestion happens in a stable manner. Nevertheless, the parameters for person cars are the same as those giving rise to the oscillating patterns in section 5.4.2. The conclusion is therefore that the presence of trucks has beneficial influence on stability in this example. This can be understood by noting that the parameters s_1^d and s_2^d are larger for trucks, which increases stability according to equation (5.25).

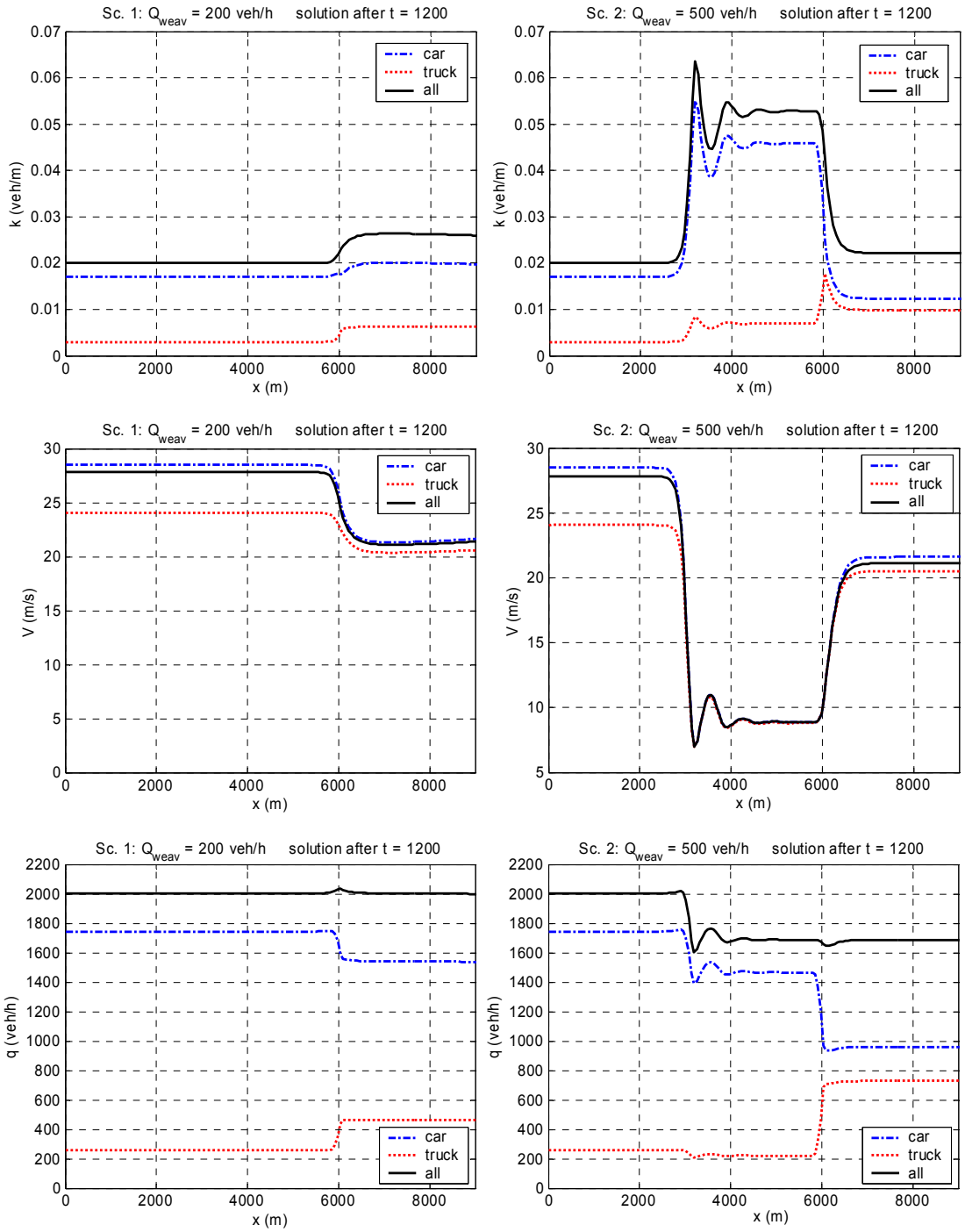


Figure 7-5 Density, average speed and flow rate for cars, trucks and both classes combined after 20 minutes for scenarios 1 (*left*) and 2 (*right*)

We conclude that in free flowing traffic the underestimation of intra-lane speed synchronisation between classes predicted in section 7.4.1.1 is indeed present, but that this problem is not present in the low-speed regime. Moreover, the mixture of user-classes with different stability characteristics yields a combined flow with intermediate stability properties. The latter two conclusions are important properties for a good model representation of congestion.

7.4.3 Causes and remedies for underestimation of speed synchronisation in free flow conditions

From the previous two sections 7.4.1 and 7.4.2, it appeared that the multi user-class model yields a good representation of congested traffic, but in free flowing traffic intra-lane speed synchronisation is underestimated. Given the requirements to the human-kinetic model that was conceived to analyse congested traffic and the influence of ADA systems thereon (see section 1.2.1), the elaboration of a remedy to this issue is however postponed to further research. For now, this section examines the causes for this problem (section 7.4.3.1) and suggests, in general terms, a potential remedy (section 7.4.3.2). In short, the arguments that are elaborated in the following are:

- The kinetic modelling approach has the typical history-independent or *memory-less* character of a Markov process: all past system behaviour that is not in the model state definition is *not* accounted for in the next update of the system;
- Modelling intra-lane speed synchronisation in single-lane traffic requires correct platoon formation modelling; platoons occur due to the fact that – once speed has been adapted to a slower predecessor – acceleration is impossible if the predecessor does not accelerate (or does not disappear from the road); interactions with slower traffic therefore have a *long-lasting* effect on faster traffic;
- If no memory of past interactions with slower traffic is contained in the actual system's state, this long-lasting effect is *not* accounted for due to the Markov-like, memory-less character of the model; so far, our definition of the traffic state does not contain this information due to the *vehicular chaos* assumption (see sections 4.2.3.1 and 7.3.2.1);
- If one wants to model platoon formation in the human-kinetic model, the system's state definition should somehow contain memory of past interactions with slower traffic and use this information correctly (see section 7.4.3.2);
- A potential remedy is to account for this memory is to explicitly distinguish constrained and unconstrained vehicles in the flow.

7.4.3.1 Poor platoon formation modelling as a cause for under-estimation of intra-lane speed synchronisation

Waldeer (2000) argues that a kinetic model – even with continuous acceleration – behaves comparable to a stochastic jump process. Such a process models the evolution of a system by considering system states and (discrete) transitions (jumps) to potential new states. Typical of a Markov process is that transitions to more than one new state are possible and that the transition probabilities to these new states depend solely on the previous state of the system, i.e. not on the history the system has gone through to attain that state.

Let us now use this interpretation of the human-kinetic model to explain under-estimation of speed synchronisation. Consider a flow with two user-classes: cars and

trucks. Let us neglect speed variances within the class and assume that the desired speed of cars is higher than that of trucks. Now approximate the continuous car-following process given by equation (4.27), by considering time intervals long enough to annihilate the speed difference with a slower predecessor or with the desired speed within one interval. Let *j* be the index of the vehicle (decreasing in direction of flow) and *i* the index of time, then there are two possible transitions for a vehicle *j* of class *car* with speed v_j^i to a new speed v_j^{i+1}:

- either it 'jumps' to its desired speed (gap with predecessor large): $v_j^{i+1} \leftarrow w_j$
- or it jumps to the speed of the predecessor *j*-1 (gap with predecessor small): $v_j^{i+1} \leftarrow v_{j-1}^i$.

In the latter case, two possibilities are considered according to equation (7.22):

- either the predecessor is of the own class: $v_j^{i+1} \leftarrow V_c^i$
- or the predecessor is a truck: $v_j^{i+1} \leftarrow V_t^i$.

Note that for trucks – the slowest class – no interactions with yet slower vehicles occur in this simplification, so in all cases, the new speed of a truck is equal to the desired speed of the trucks, or $V_t^i = w_t$. Suppose as a starting condition, a fraction of cars and trucks of respectively p_c and p_t=1- p_c and all vehicles driving at their desired speeds $V_c^0 = w_c$ and $V_t^0 = w_t$. Assume further that the probability of having a gap large enough to accelerate to the desired speed is constant and equal to p_w for all vehicles, then the expected speed of the cars at time step *i* is:

$$V_c^i = p_w w_c + (1-p_w)\left(p_t w_t + p_c V_c^{i-1}\right) \tag{7.49}$$

With the starting conditions as indicated this iterative equation can be written explicitly (after some algebra) as:

$$V_c^i = p_w w_c + (1-p_w)\left(p_t w_t + p_c p_w w_c\right)\sum_{j=1}^{i-1} p_c^{j-1}(1-p_w)^{j-1} + (1-p_w)^i\, p_c^{i-1}\left(p_t w_t + p_c w_c\right) \tag{7.50}$$

Equation (7.50) is not interesting as such, but it is interesting to see to which speed it converges if the system is left unaffected by external disturbances after the initialisation. The limit for $i \to \infty$ of the expected car speed is then:

$$\begin{aligned} V_c^\infty &= \lim_{i\to\infty} V_c^i \\ &= p_w w_c + (1-p_w)\frac{p_t w_t + p_c p_w w_c}{p_t + p_c p_w} \end{aligned} \tag{7.51}$$

which is larger than w_c for $0 < p_w \le 1$. Interestingly, for large values of p_w the eventual speed of the car class is close to the desired speed of that class, whereas it approximates the desired speed of the trucks for p_w approaching zero. These extremes correspond to very low or high total density respectively, since the probability of a large gap is inversely related to the total density.

Despite the numerous approximations and simplifications, this result corresponds qualitatively very well to the simulated results in Figure 7-3 and Figure 7-4. However, as mentioned in the discussion of that figure in section 7.4.1.1, the correct expected result should be: $V_c^\infty = w_t$, since eventually every car will be trapped in a platoon behind a truck, driving with the speed of this truck. In the next section, the Markov approximation of the human-kinetic model is used as a basis for suggested remedies for improved platoon modelling.

7.4.3.2 Remedies: correlation modelling or distinguishing constrained and unconstrained states

The basic idea, remedying the poor platoon formation modelling discussed in the previous section, is the following: the system's state definition should somehow contain memory of past interactions with slower traffic and use this information correctly. We first elaborate this idea in the Markov approximate model.

Correct platoon modelling in the Markov approximate model

The following changes are applied with respect to the analysis in the previous section:

1. for the fast user-class, distinction is made between unconstrained and constrained vehicles (see section 3.2.1); we thus distinguish two subclasses of the cars, dependent on the level of constraint;
2. average speed is calculated per subclass;
3. vehicles in the constrained subclass can never have a gap large enough to accelerate to the free flow speed; unconstrained vehicles have such gap by definition.

A constrained car (indicated by class index c_c) has, by definition of the constrained state, a truck or another constrained car as a predecessor. Therefore, its speed will remain unaffected and equal to the desired speed w_t of the truck class. An unconstrained car (indicated by class index c_u) with speed v_j^i then has three possibilities during an update to a new speed v_j^{i+1}:

- either its predecessor is an unconstrained car: $v_j^{i+1} \leftarrow V_{c_u}^i$
- or its predecessor is a constrained car: $v_j^{i+1} \leftarrow V_{c_c}^i = w_t$
- or its predecessor is a truck: $v_j^{i+1} \leftarrow V_t^i = w_t$

It can be concluded that the subclass-average speeds for unconstrained cars, constrained cars and trucks are w_c, w_t and w_t respectively. Therefore, the average speed of the class of cars is:

$$V_c^i = p_{c_u}^i w_c + p_{c_c}^i w_t \tag{7.52}$$

In this equation, $p_{c_u}^i$ and $p_{c_c}^i = 1 - p_{c_u}^i$ indicate the fraction unconstrained and constrained cars at time step i. Note that the probability $p_{c_u}^{i+1}$ of having an unconstrained car in time step $i+1$ equals the probability that is was unconstrained in step i, times the probability of finding a car, times the probability that this car is unconstrained:

$$p_{c_u}^{i+1} = p_{c_u}^i \, p_c \, p_{c_u}^i \tag{7.53}$$

which can be explicitly solved to yield:

$$p_{c_u}^i = p_c^{2^i - 1} \tag{7.54}$$

Therefore, equation (7.52) has the explicit solution:

$$V_c^i = p_c^{2^i - 1} w_c + \left(1 - p_c^{2^i - 1}\right) w_t \tag{7.55}$$

Comparing this result to equation (7.50), this time the car speed always converges to $V_c^\infty = w_t$ and that convergence is achieved much faster than in the previous case (even if p_w=0 and the limit of equation (7.50) is the same as that of (7.55)). As an illustration, convergence of the Markov approximate multiclass human-kinetic model and of the corrected Markov approximate model is depicted in Figure 7-6. Traffic flow consists of cars (p_c = 90% ; w_c = 35 m/s) and trucks (p_t = 10% ; w_t = 25 m/s). The probability p_w of having a large gap in equation (7.50) is taken 0 and 10%).

The conclusion is that the proposed distinction between constrained and unconstrained subclasses indeed remedies the lack of intra-lane synchronisation between user classes.

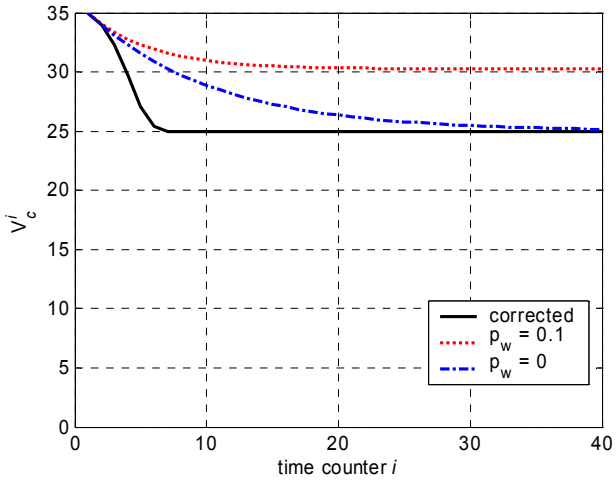


Figure 7-6 Convergence of solutions in the Markov approximate model and the corrected model

Distinguishing constrained and unconstrained vehicle states in the human-kinetic model

As mentioned earlier, elaborating the above mentioned remedy in the human-kinetic traffic flow model is a subject for further research. This section indicates briefly how such solution would fit into the modelling approach.

Hoogendoorn & Bovy (2000) proposed a platoon-based kinetic model with distinction of constrained (platooning) and unconstrained (free) driving states. Apart from the individual vehicle speed, the (un)constrained state is taken as a supplementary state variable in the state vector *S*. The generalised partial density functions are defined per user-class with explicit distinction between constrained and unconstrained vehicles (in fact, it is as if the number of classes is doubled by splitting each class into its constrained and unconstrained part). This solves the issue of gradual platoon formation in the following way.

When considering the acceleration integral of equation (7.22) separate distribution functions are applied, depending on the constrained state of the vehicle under consideration: constrained vehicles have different (lower) average speeds than the unconstrained vehicles of the same user-class. Also the speed variance is in theory different. The same applies for the gap distribution: average gaps and the gap variance for constrained vehicles are considerably smaller than those of the unconstrained vehicles of the same user-class.

Apart from this distinction, a specification is needed describing when a constrained vehicle changes its state to unconstrained and vice versa. One option would be to use its speed and acceleration level: deceleration turns an unconstrained vehicle into a constrained one; acceleration to a speed close enough to own the desired speed makes a vehicle change into the unconstrained state. Hoogendoorn and Bovy consider a multi-lane model, so that changing lanes also releases the constrained state.

The distinction between constrained and unconstrained vehicles now acts as the 'memory' that is required to describe platoon formation (section 7.4.3.1): once a vehicle changes into the constrained state, it is assigned the appropriate short gaps, lower speed, and slower predecessor in the next update step. If no overtaking is possible, acceleration into the unconstrained state is only possible if the predecessor accelerates to speeds equal to or higher than the own desired speed (which will never happen if the predecessor is a truck and has a lower desired speed). Therefore, eventually all vehicles in a one-lane scenario without overtaking will end up in the constrained state and average speeds of different user-classes are synchronised in the human-kinetic model in a similar fashion as in Figure 7-6.

7.5 Conclusion

In this chapter the human-kinetic traffic flow model was extended with multiple user-classes. The basic kinetic description was generalised and a multiclass macroscopic model established. This model builds on a multiclass acceleration integral that contains all interactions of a class with vehicles of the own and other classes. The model was analysed in equilibrium conditions and in a dynamic case study.

It turned out that, thanks to the foundation on kinetic theory, the multi user-class generalisation of the basic variables and dynamic equations was possible by straightforward application of methods described in literature. The interaction within and between classes followed naturally through careful consideration of the acceleration

integral and the probabilities of encountering predecessors of the different classes. Moreover, the multiclass human-kinetic theory offers the possibility to model interactions between user-classes with completely different microscopic behaviour underlying the acceleration integral of each class. This is in contrast with existing multiclass models, where user-classes only differ through different parameter sets for a common dynamic model. This property of the human-kinetic model makes it possible, in principle, to model traffic flows consisting of any combination of vehicle and/or driver types, such as ADA supported vehicles or even fully automated vehicles if the microscopic behavioural models are given. During further specification of the multiclass acceleration integral, special attention was needed for the assignment of road space over the classes. A concept of *user equilibrium* was therefore borrowed from literature, but was adapted to match our modelling approach that is unique in its explicit treatment of finite road space requirements of the drivers.

The numerical evaluation of the model turned out to be a straightforward extension of the single user-class method, as was the formulation of equilibrium solutions. The latter are characterised by asymmetric interactions between user-classes, so that slower vehicles have larger influence on faster vehicles than vice versa. However, convergence of average speeds of single-lane multiclass traffic (intra-lane synchronisation) is only achieved for higher densities, i.e. in congestion and not in free flowing traffic. A solution to this issue, which is less important if one is interested in congested traffic than when also free flowing traffic is of interest, is suggested, but elaboration is left to further research.

With the multiclass extension, the human-kinetic model has now become a solid and flexible framework for building macroscopic simulations of mixed ADA-equipped and non-equipped traffic. An example of such an application that also makes use of the potential for modelling variable driver behaviour is elaborated in the next chapter.

8 MULTICLASS MODELLING OF ADVANCED DRIVER ASSISTANCE SYSTEMS

8.1 Introduction

In this chapter the multiclass human-kinetic traffic flow model is applied to an illustrative Advanced Driver Assistance (ADA) system. In order to illustrate the strength of the model, an application is chosen with the following characteristics:

- individual vehicles act as sensors, causing this vehicle to activate the system,
- communication between equipped vehicles only (and not with non-equipped vehicles) is used to transmit information upstream,
- drivers are responsible to respond to warnings by the system, however, only if traffic conditions require specific control.

These system characteristics set specific requirements to the model that is used to simulate the system. These requirements were all introduced in the human-kinetic model in the previous chapters:

- foundation of the model on individual driver/vehicle behavioural specifications (for individual system activation and individual responses),
- potential to model discrete events and interactions (for the communication of messages),

- modelling of variable driver behaviour (for driver response dependent on the most recent warnings),
- modelling of multi user-class traffic flow (to distinguish ADA-equipped and non-equipped vehicles).

As an illustration of these aspects of the human-kinetic traffic flow model, a driver assistance system is considered that is designed to smoothen the decelerations near queue-tails, based on warnings that are communicated between vehicles.

The conceptual design of the system is given in section 8.2, the model specification in section 8.3, and the model analyses in section 8.4. Finally the chapter concludes with a discussion of the potential role of the human-kinetic model in ADA design and conclusions (section 8.5).

8.2 Conceptual design of the vehicle-vehicle communication based queue tail warning system

The ADA system examined in this section is a vehicle-vehicle communication based queue tail warning (Figure 8-1). Its aim is to smoothen the decelerations near queue tails, in other words, to make drivers anticipate more adequately to the sharp braking in the deceleration front of congestion waves. The basic concept is that strongly braking vehicles send warning messages upstream. Equipped vehicles in the stream that receive these messages alert the driver of the coming deceleration manoeuvre. The immediate reaction of the driver is to increase the activation level (attention level) so that he is better able to anticipate to downstream speed drops.

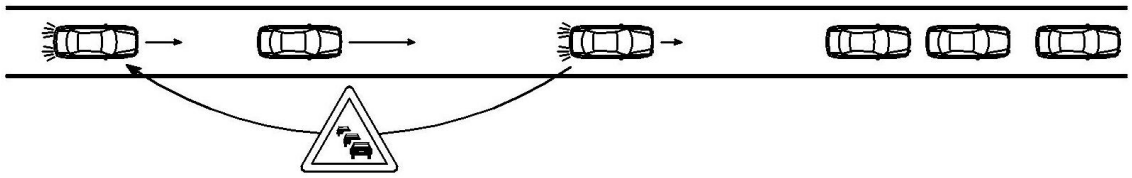


Figure 8-1 Illustration of the queue-tail warning concept

In contrast to the system examined by Alkim et al. (2000, see also section 3.2.4), we do not assume that vehicles are equipped with Adaptive Cruise Control (ACC) which would allow the vehicle to respond autonomously to the warning messages. These authors found that it was difficult to design such *cooperative following* system efficiently, because the in-car intelligence has to mix local information from its own sensors and preview information received via the messages. This requires an integrated design of the adaptive cruise control and cooperative following control logic. Therefore, we keep the driver in the loop, and let him balance the local and non-local information by reinforcing the anticipation behaviour described in section 4.3.2.

Notice that for our purposes it is not important *how* the communication is actually designed. We just assume that messages are received by all equipped vehicles in a certain range upstream of the location where they were generated, regardless of the communication medium. Neither is the design of the human machine interface important for our purposes. We assume a well-designed system that is able to increase the alertness of the driver, regardless of the type of human-machine interface feedback, which may be visual, auditory or haptic.

8.3 Model specification of the vehicle-vehicle communication based queue tail warning system and driver response

8.3.1 Sending warnings

The acceleration calculated by the car-following algorithm is used as a trigger for sending warnings. If the result of the calculation is lower than a certain threshold, then a warning is produced:

$$send\left(\dot{v}\right) \equiv \begin{cases} 0 & \dot{v} > d_{send} \text{ OR } c \neq ADA \\ 1 & \dot{v} \leq d_{send} \text{ AND } c = ADA \end{cases} \tag{8.1}$$

Then, the total probability p_{send} that a warning is sent by a vehicle of class c on location x at time t is the expected value of the function $send\left(\dot{v}\right)$, which is found by integration over all possible speeds of the successor, gaps, and speeds of the predecessor, taking into account the probability of occurrence of the speeds and gaps:

$$\begin{aligned} p_{send}^{c} &\equiv \left\langle send\left(\dot{v}\left(t+T, x\left(t+T\right)\right)\right)\right\rangle_{v} \\ &= \int_{v_j} p_v\left(v_j\right) \int_{s_j} p_s\left(s_j\right) \int_{v_{j-1}} p_v\left(v_{j-1}\right) send\left(\dot{v}_j\left(v_j, s_j, v_{j-1}\right)\right) dv_{j-1}\, ds_j\, dv_j \end{aligned} \tag{8.2}$$

with the probabilities p_v and p_s as defined in chapter 4.

As a result of neglecting the correlations between the speed of the successor, that of the predecessor and the gap between them, the probability of sending messages will be overestimated (see Annex C). We therefore only consider probabilities that exceed a certain noise threshold:

$$p_{send}^{c} \rightarrow \begin{cases} p_{send}^{c} & p_{send}^{c} \geq p_{noise} \\ 0 & p_{send}^{c} < p_{noise} \end{cases} \tag{8.3}$$

In general, there might be more than one class c of vehicles that send warnings. The expected total number $n_{sent}\, dx$ of messages originating from an interval dx is therefore found as a weighed sum of p_{send}^{c} (according to equation (8.3)) over all classes, with the partial densities per class as the weighing factor:

$$n_{sent}\left(t,x\right)\ \ dx = \sum_{c} k_c\left(t,x\right) p_{send}^{c}\left(t,x\right)\ \ dx \tag{8.4}$$

8.3.2 *Receiving warnings and the influence on activation level*

A vehicle receives all warnings sent within a range $\Delta x_{receive}$. The total number of received warnings then equals:

$$n_{received}(t,x) = \int_{x}^{x+\Delta x_{receive}} n_{sent}(t,x)\ dx \tag{8.5}$$

We assume that as a result of receiving messages, the activation level a_0 is immediately increased by a factor f. This factor is chosen such that it equals 1 for $n_{received}$=0, and would set the activation level immediately to a_{max} if the number of received messages equals some maximum value. This maximum value corresponds to the number of warnings for which the most severe warning through the human-machine interface is triggered. The activation level a_1 after reception of the messages by a driver with initial activation level a_0 is thus:

$$\begin{aligned} a_1 &= f\, a_0 \\ &= \left(1+\left(\frac{a_{max}}{a_0}-1\right)\frac{n_{received}}{n_{max}}\right)a_0 \\ &= \left(1-\frac{n_{received}}{n_{max}}\right)a_0 + \frac{n_{received}}{n_{max}}\, a_{max} \end{aligned} \tag{8.6}$$

This equation expresses that the new activation level of a driver is a weighed average of the old average activation level a_0 and the maximum activation level a_{max}. The weighing factor equals zero if no messages are received (the activation level remains unchanged) and 1 with the maximum number of messages.

Following the method of section 6.4.2.2, we establish the active and passive change rates due to these transitions. The active change rate stands for the number of vehicles leaving state a_0. For the class of ADA vehicles, we assume that every driver who receives warnings increases his activation level, therefore the probability density of any original activation level a_0 is decreased by:

$$\left(\frac{d\rho(t,x,v,a_0)}{dt}\right)^{-}_{ADA} = \int_{a_1\neq a_0} \rho(t,x,v,a_0)\ \Pi(t,x,v,a_0)\,\pi_a(t,x,v,a_0,a_1)\ da_1 \tag{8.7}$$

with the specification of the interaction rate Π as follows:

$$\Pi(t,x,v,a_0) = \delta\left(a_1 - \left(\left(1-\frac{n_{received}}{n_{max}}\right)a_0 + \frac{n_{received}}{n_{max}}\, a_{max}\right)\right) \tag{8.8}$$

The Kronecker delta function occurs, because only transitions occur from activation level a_0 to the activation level a_1 defined by (8.6). The transition rate π_a is assumed equal to some constant:

$$\pi_a(t,x,a_0,a_1) \equiv \Phi \tag{8.9}$$

Substituting (8.8) into (8.7), yields:

$$\left(\frac{d\rho(t,x,v,a_0)}{dt}\right)^{-}_{ADA} = \rho(t,x,v,a_0)\int_{a_1}\Pi(t,x,v,a_0)\pi_a(t,x,v,a_0,a_1)da_1 = \rho(t,x,v,a_0)\Phi \tag{8.10}$$

The contribution of the active change rates to the macroscopic dynamic activation level moment equation (equation (6.10)) is then:

$$\int_a\int_v a\left(\frac{d\rho}{dt}\right)^{-}_{ADA} dv\,da = \int_a\int_v a\,\Phi\,\rho(t,x,v,a)\,dv\,da = k\,\Phi\,A \tag{8.11}$$

The passive change rate is the rate at which drivers adopt a new state a_1, where they formerly had an initial activation level a_0. Using equations (8.8) and (8.9) for the interaction and transition rates, yields:

$$\begin{aligned}\left(\frac{d\rho(t,x,v,a_1)}{dt}\right)^{+}_{ADA} &= \int_{a_0}\rho(t,x,v,a_0)\Pi(t,x,v,a_0)\pi_a(t,x,v,a_0,a_1)da_0 \\ &= \int_{a_0}\rho(t,x,v,a_0)\Phi\,\delta\left(a_1-\left(\left(1-\frac{n_{received}}{n_{max}}\right)a_0+\frac{n_{received}}{n_{max}}a_{max}\right)\right)da_0\end{aligned} \tag{8.12}$$

After substitution $n_{max}\,a^{*} = (n_{max}-n_{received})\,a_0$, equation (8.12) can be elaborated:

$$\begin{aligned}\left(\frac{d\rho(t,x,v,a_1)}{dt}\right)^{+}_{ADA} &= \frac{n_{max}}{n_{max}-n_{received}}\int_{a^{*}=0}^{\frac{n_{max}\,a_{max}}{n_{max}-n_{received}}}\rho\left(t,x,v,\frac{n_{max}\,a^{*}}{n_{max}-n_{received}}\right)\Phi\,\delta\left(\left(a_1-\frac{n_{received}}{n_{max}}a_{max}\right)-a^{*}\right)da^{*} \\ &= \frac{n_{max}\,\Phi}{n_{max}-n_{received}}\rho\left(t,x,v,\frac{n_{max}\,a_1-n_{received}\,a_{max}}{n_{max}-n_{received}}\right)\end{aligned} \tag{8.13}$$

The contribution of the passive change rates to the macroscopic dynamic activation level moment equation (equation (6.10)) is then calculated as (with the substitution $a^{*}\left(1-\frac{n_{received}}{n_{max}}\right)=a-\frac{n_{received}}{n_{max}}a_{max}$):

$$\begin{aligned}\int_a\int_v a\left(\frac{d\rho}{dt}\right)^{+}_{ADA} dv\,da &= \frac{n_{max}\Phi}{n_{max}-n_{received}}\int_{\frac{n_{received}}{n_{max}}a}^{a_{max}} a\,\rho\left(\frac{n_{max}\,a-n_{received}\,a_{max}}{n_{max}-n_{received}}\right)da \\ &= \Phi\left(1-\frac{n_{received}}{n_{max}}\right)\int_{a^{*}=0}^{a_{max}} a^{*}\,\tilde{\rho}(a^{*})da^{*}+\Phi\frac{n_{received}}{n_{max}}a_{max}\int_{a^{*}=0}^{a_{max}}\tilde{\rho}(a^{*})da^{*} \\ &= \Phi\,k\left(\left(1-\frac{n_{received}}{n_{max}}\right)A+\frac{n_{received}}{n_{max}}a_{max}\right)\end{aligned} \tag{8.14}$$

This equation reveals that the average new activation level is a weighed average of the old average activation level A and the maximum activation level a_{max}. The weighing factor equals zero if no messages are received (the activation level remains unchanged),

and 1 with the maximum number of messages. The similarity with equation (8.6) is of course no coincidence.

Collecting these terms (8.11) and (8.14), and combining these with the smooth relaxation towards the normal activation level a_{normal} (equation (6.31)), the macroscopic dynamic activation level moment equation in the presence of ADA systems becomes:

$$\begin{aligned}\frac{\partial kA}{\partial t}+\frac{\partial kAV}{\partial x} &= \underbrace{k\frac{a_{normal}-A}{\tau_a}}_{relaxtation}+\underbrace{k\Phi\left(\left(1-\frac{n_{received}}{n_{max}}\right)A+\frac{n_{received}}{n_{max}}a_{max}\right)}_{passive\,discrete\,changes}-\underbrace{k\Phi A}_{active\,discrete\,changes}\\ &= k\left(\frac{a_{normal}-A}{\tau_a}+\Phi\frac{n_{received}}{n_{max}}\left(a_{max}-A\right)\right)\\ &= k\left(\frac{a_{normal}-A}{\tau_a}+\frac{a_{max}-A}{\frac{1}{\Phi}\frac{n_{max}}{n_{received}}}\right)\end{aligned} \tag{8.15}$$

The interpretation of equation (8.15) is as follows: in the case that warnings are received, the average activation level A immediately shifts towards the maximum activation level a_{max}. In fact if the number of warnings would remain constant in time the activation level A would exponentially relax towards a_{max} with a time constant equal to $\frac{1}{\Phi}\frac{n_{max}}{n_{received}}$. This process is counteracted by the exponential relaxation of the driver towards a comfortable (low) activation level a_{normal}, with a time constant equal to τ_a.

Finally, in the case of traffic merging onto the highway, we will assume that merging drivers do so while having the average activation level and average speed corresponding to the location where they merge. Therefore, the speed momentum equation developed earlier (equation (4.45)) is still valid and we adapt the activation level momentum equation (8.15) as follows:

$$\frac{\partial kA}{\partial t}+\frac{\partial kAV}{\partial x}=A\left(\frac{dk}{dt}\right)_{inflow}+k\left(\frac{a_{normal}-A}{\tau_a}+\frac{a_{max}-A}{\frac{1}{\Phi}\frac{n_{max}}{n_{received}}}\right) \tag{8.16}$$

8.3.3 *Influence of the activation level on driver behaviour*

Since the ADA system is designed to warn for upcoming speed drops, it is logical to assume that a higher activation level causes drivers to anticipate stronger to speed drops downstream. We assume the following relation between the anticipation strength and the activation level (see also Figure 6-4):

$$f^{ant}(a)\equiv\begin{cases}f^{ant}_{normal}+\left(f^{ant}_{max}-f^{ant}_{normal}\right)\dfrac{a-a_{normal}}{a_{max}-a_{normal}} & a>a_{normal}\\ f^{ant}_{normal} & a\le a_{normal}\end{cases} \tag{8.17}$$

In the analyses of the queue tail warning ADA system, we do not consider an influence of the activation level on the indifference band as we did in the case study of the capacity funnel in section 6.5, nor do we consider behavioural hysteresis as in section 6.6. Although it is probable that different aspects of variability of driving behaviour interfere, we want to show here the isolated effect of stronger anticipation due to ADA-triggered warnings only. Potential combinations of behavioural hypotheses might be analysed in a fashion similar to the case that is illustrated in this chapter.

8.4 Explorative model analyses of the vehicle-vehicle communication based queue tail warning system

In this section the case study of chapters 5 and 6 is repeated but now in a multi user-class simulation with a fraction of the vehicles equipped with the queue-tail warning ADA system. A 9 km long road with an on-ramp at x=6 km is considered. More specifically, scenario 6 of Table 5-3, with a constant flow of 2090 veh/h on the main lane and 600 veh/h through the on-ramp, is taken as a reference case without ADA traffic (upper row in Figure 8-2). In the reference case, an Oscillating Congested Traffic (OCT) pattern occurs.

With ADA introduced, the activation level of drivers immediately upstream of the queue-tail is raised by warnings received from decelerating ADA vehicles in the queue-tail. The higher the equipment fraction, the more vehicles send warnings, and thus the activation level is raised higher. On the other hand, the deceleration levels in the queue-tail are lower, so that fewer messages are sent per equipped vehicle. Therefore, the increase in activation level for the higher equipment fractions (50 and 100%) is moderated.

Table 8-1 Characteristics of the queue tail in the simulations of Figure 8-2

Equipment fraction	Free flow speed [m/s]	Speed in queue tail [m/s]	Length of the queue-tail [m] (approx.)	Average deceleration [m/s²]
0%	32	1	200	-2.5
25%	32	3	275	-1.8
50%	32	5	350	-1.4
100%	32	7	500	-1.0

From Table 8-1, it can be read that the decelerations inside the queue tail are indeed smoothened thanks to the warnings exchanged between ADA vehicles. Not only is the deceleration lower because of alerted ADA-drivers braking earlier (i.e. at longer distances before the point with the lowest driving speed in the queue-tail), the end speed is also slightly higher with more vehicles equipped. This is because the deceleration process is more stable than without the system. This effect could have been predicted,

knowing that the assumed effect of more alert or active driving is an increase of the anticipation strength according to equation (8.17), which has a beneficial effect on traffic stability, as was shown in the analysis in section 5.2.1.3.

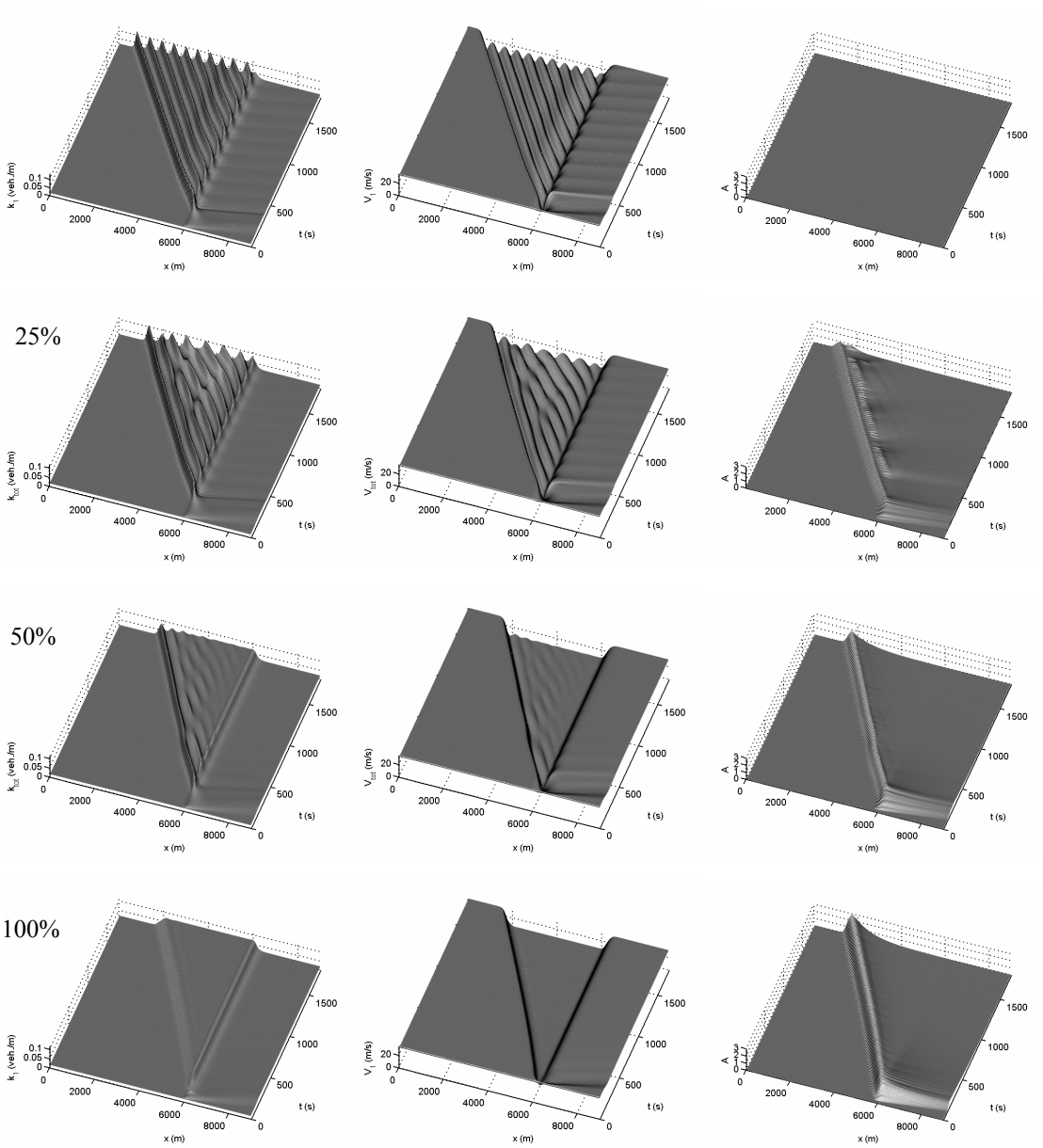


Figure 8-2 Density, average speed (all vehicles) and activation level (if equipped) as a function of space* x *and time* t *for different ADA-equipment levels

Note that as a result of increased stability of traffic that approaches the queue-tail, the oscillatory character of the congestion pattern is reduced or even totally suppressed in the 100% equipped case. A Homogeneously Congested Traffic (HCT) pattern now emerges instead of the OCT, which is exactly the solution that one would obtain when analysing this bottleneck with first order theory (kinetic wave model of section 3.3.1).

Indeed, the instability associated with approaching the queue-tail is now eliminated, so that the situation can be adequately described by a model that neglects instability. Another way of representing the simulation results and the transition from OCT to HCT

with increasing equipment fraction is shown in Figure 8-3, where the phase trajectories for a fixed *x*-location upstream of the merge are plotted. As in previous plots of this type, the original state in the flow-density plane is marked with a square (□), the end state (which we have chosen here after two periods of the oscillation for clarity of the figure) is marked with a star (☆); a decelerating phase trajectory is marked with circles (○) and acceleration with triangles (△). Clearly visible are the oscillations due to overshoot of a point on the congestion branch of the equilibrium fundamental diagram. The amplitude of the overshoot decreases with increasing equipment fraction. In the fully equipped case, the increased anticipation avoids overshoot and the traffic state smoothly converges to the point on the equilibrium curve. This image is indeed consistent with the same representation of shock-waves in the first order traffic flow theory, which supports our earlier conclusion that first order models are a sufficient description of traffic flow in which anticipation behaviour is always dominant over retarded deceleration (see section 2.5.2).

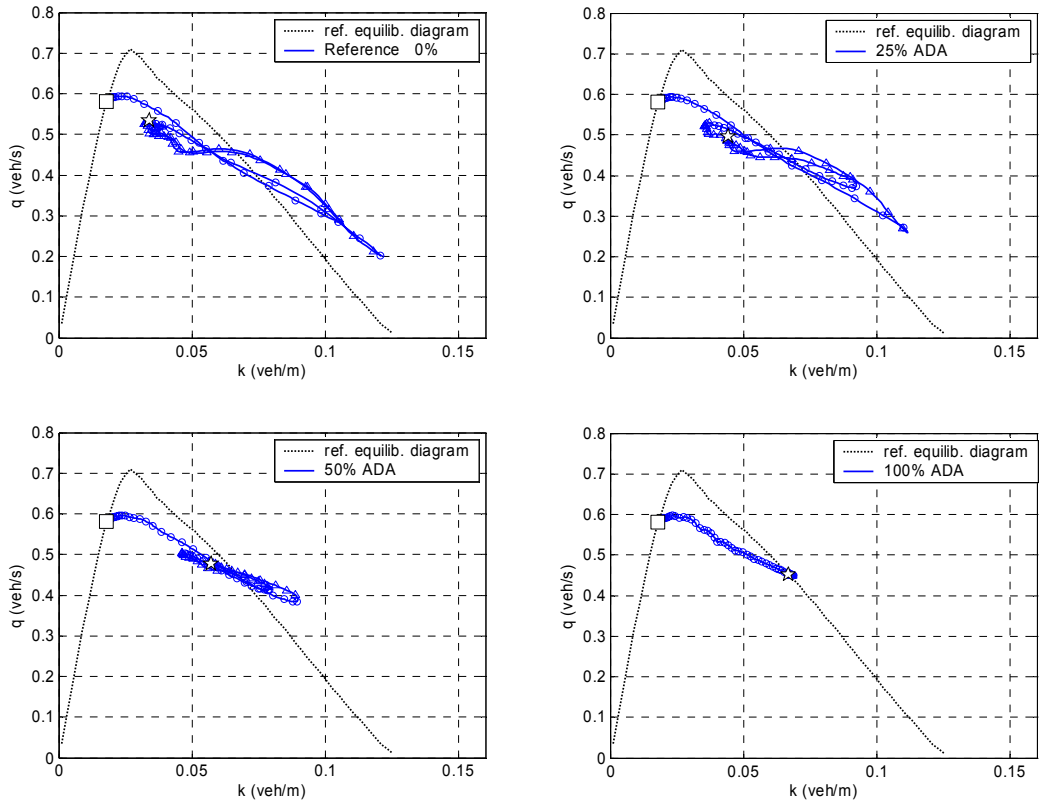


Figure 8-3 Phase trajectories at x=4500m for the ADA simulations of Figure 8-2

8.5 Discussion and conclusion

Discussion on the role of the human-kinetic traffic flow model for ADA design

The analysis using the human-kinetic model performed in this chapter is not a decisive analysis of the impacts of the ADA system that is studied. Rather, it fits in the early stages of ADA system design process: it explores on the relatively abstract macroscopic level, the interaction of the characteristics and global settings of the ADA system with

traffic flow. Questions are addressed like: does the concept work? which information needs to be transmitted? how far and to whom should information be sent? should the information be processed and if so, by whom: the sender, receiver or possibly by intelligent infrastructure at the side of the road? etc.

After this explorative stage, which sets global functional requirements to the system as a whole and to its main components, system designers can proceed with the detailed design of sensors, communication system, human-machine interface, control rules, and other components of the system. For this stage of system development, other, more detailed simulation models of the system components and the overall system are needed, for example the microscopic simulation tools discussed in section 3.2.4. The value of the macroscopic analysis is that it sets the initial requirements to these system components that need to be met if a certain effect on traffic flow (macroscopic level) is desired. If detailed analysis of a system's component shows that one of these requirements cannot be met, for instance: the desired driver response using the best possible human-machine interface is different from the assumption made in the macroscopic analysis, the human-kinetic model can be modified accordingly, and the expected effect of this modification can be quantified.

Conclusions

In this chapter an explorative analysis was performed of an Advanced Driver Assistance system designed to warn drivers when approaching sharp decelerations in a queue tail. The system builds on inter-vehicle communication that triggers warnings to drivers approaching the queue-tail. As a reaction to these warnings, the driver is responsible for adequate control actions. The assumption used here is that warnings increase the activation level of drivers, which causes them to anticipate more adequately to decelerations downstream. Based on an initial design of the system, its specifications were defined in the human-kinetic traffic flow model. This multiclass model with mixed ADA-equipped and non-equipped traffic is applied to the same on-ramp bottleneck that was studied earlier in this thesis. The explorative analysis shows that the system results in safer and smoother transition from free flowing to congested traffic. It also avoids instability of the queue-tail, thus preventing oscillating congested traffic patterns. Instead, a homogeneous congestion pattern emerges that is actually consistent with first-order kinematic traffic flow theory that neglects traffic instability.

Note that the ADA system analysed in this section is intended to directly act upon the activation level of the driver with respect to the longitudinal control task in normal traffic conditions. It is envisaged that also ADA systems that are not directly aimed at influencing normal longitudinal control (e.g. assistance of the lateral control task) might have indirect impact on the longitudinal control task as well (Brookhuis et al., 2001). When this indirect impact on individual driver behaviour is known (e.g. from driving simulator experiments or from small scale field trials), the human-kinetic modelling approach can be applied to extrapolate the effect to macroscopic traffic flow dynamics.

In the system design process of ADA systems, the type of explorative system analysis performed in this chapter needs a follow-up, in which the assumptions made for driver/vehicle behaviour are investigated in more detail. The human-kinetic model

therefore acts as an explorative model in the early design stage to provide a proof-of-concept of a preliminary ADA system design, and to set the initial requirements to the system's components before starting the detailed design. During detailed design, it acts as a quick-scan model to quantify the effects of modifications of the detailed design that deviate from the initial requirements.

9 CONCLUSIONS AND RECOMMENDATIONS FOR FURTHER RESEARCH

In this final chapter, conclusions are drawn regarding the work presented in this thesis and some recommendations for further research are formulated.

9.1 Conclusions

The conclusions are structured as follows. First some general conclusions are drawn about the model developed in this thesis (section 9.1.1), after which the main conclusions from each step in the development are discussed (section 9.1.2).

9.1.1 General contributions of the model developed

This thesis described a traffic flow modelling approach for motorways with the following characteristics:

- It provides a realistic representation of multiclass traffic dynamics, including congested traffic;
- It reproduces all kinds of phenomena that were observed in congested traffic, like stop-and-go waves, hysteresis, capacity drop, wide jams, and delayed onset of congestion (capacity funnel);
- It bridges the gap between traffic flow models and behavioural models of the individual driver;

- It is suited for further refinement of the behavioural models for individual driver behaviour (provided these models are available), with the aim of more realistically modelling congested traffic operations and dynamics;
- It establishes a new class of kinetic models that provides a flexible framework in which alternative behavioural specifications for the driver or the ADA system can be modelled;
- It is suited for modelling mixed ADA-equipped and non-equipped traffic for ADA systems that have impact (either directly or indirectly) on the longitudinal control task of the driver or the vehicle.

9.1.2 Specific conclusions from the state-of-the-art review and separate modelling steps (per chapter)

9.1.2.1 Conclusions from the state-of-the-art

The review of empirical traffic flow theories (chapter 2) revealed that – despite lots of well-accepted facts and theories on (congested) traffic flow and decades of research – not all features of congested traffic have been incontrovertibly confirmed by empirical observations, nor satisfactorily explained. Most importantly, the *formation* mechanism of jams and stop-and-go waves and the role of instability therein is the main issue of discussion. However, the amount of empirical observations and analyses of oscillating congestion patterns and a variety of congested traffic phases that has been observed since the second halve of the 1990's forms strong evidence that instability indeed plays a significant role in processes governing congestion dynamics.

As long as not all aspects of traffic flow theory are agreed upon, also the validity of traffic flow models will remain subject to discussion. The review of traffic flow models and their application to ADA systems (chapter 3) confirmed that the behavioural foundation of existing traffic flow models is poor from the point of view of individual driver behaviour and psychology. On the other hand, these models succeed in reproducing many congested traffic phenomena. Also ADA systems have been modelled successfully, albeit mainly in the more detailed (sub-) microscopic models. Modelling ADA systems adequately in macroscopic or kinetic traffic flow models is a new discipline with interesting potential, that is however still in its infancy.

9.1.2.2 The basic human-kinetic traffic flow model: theoretical improvements for a flexible simulation framework with excellent traffic representation

Although the human-kinetic model developed in this thesis (chapter 4) builds on concepts and mathematics that were borrowed from traditional kinetic theory, its different treatment of acceleration and deceleration behaviour in an *acceleration integral* has substantial benefits. The resulting model is intrinsically non-local and anisotropic, and accounts implicitly for finite space requirements of drivers and limited acceleration / deceleration capability of vehicles; properties that are otherwise difficult (although not impossible) to achieve in macroscopic models. The acceleration integral also serves as the micro-macro link that was envisaged to include individual

behavioural models in the macroscopic model. Not only has this allowed to found the human-kinetic model on a traditional car-following relationship, it is also the key enabling the introduction of a series of behavioural refinements, such as delayed control response by drivers (reaction time) and genuine anticipation behaviour. Starting points for extensions, like lane-change models, were indicated, but not elaborated in detail in this thesis.

The price for these theoretical advantages is the necessity to numerically evaluate the acceleration integral at each discretisation point in space and time, rendering the model computationally less efficient than existing models with a simpler expression for the numerical source term. For the numerical evaluation of the traffic flow model traditional Godunov-type discretisation schemes can be applied. However, some pragmatical modifications were necessary, because the model is mathematically characterised by non-negative eigenvalues of the convective and pressure terms, whereas in existing models negative eigenvalues force inversion of the 'information flow' at low speeds.

The basic human-kinetic model was analysed both analytically and numerically (chapter 5). We are able to show that the model has equilibrium solutions, comparable to a triangular or strictly concave fundamental diagram of flow against density in traditional models. The parameters of the individual driver model determine to a large extent the shape of the equilibrium solutions. The same conclusion can be drawn for dynamic solutions, the stability of which depends on the strength of anticipation behaviour. Also the influence of other individual behavioural parameters on stability is explained by means of an analytical stability criterion that relates these parameters directly to stable or unstable traffic flow. Interestingly the phase trajectories in the fundamental diagram correspond closely to theoretical and empirical phase trajectories (also often called: *hysteresis loops*) reported in literature: during deceleration in the tail of a jam, the anticipation to traffic conditions downstream is initially dominant, after which traffic compresses to densities higher than the equilibrium values, and driver behaviour is dominated by relaxation to the immediately surrounding conditions.

As a last theoretical analysis, we provide some clear indications of the equivalence between the microscopic foundation and the macroscopic model, although a sound proof is not given. It remains an issue for further research whether simulated differences between both modelling levels are due to the imperfect numerical evaluation or inherent to the kinetic approach, as some sources in literature suggest.

The quality of the model is illustrated by a case study predicting congestion patterns near a bottleneck caused by an on-ramp. Depending on the parameter settings, stable first-order like solutions are produced, or unstable oscillating patterns (like stop-and-go waves) that are consistent with best-practice in literature. Well-known subtle congestion phenomena like the capacity funnel effect and – more importantly – capacity drop are reproduced by the model, albeit in the former case only in carefully selected conditions.

9.1.2.3 Consideration of the activation level refines representation of traffic flow dynamics in the human-kinetic model

The human-kinetic model has been refined by taking into account that drivers are no machines that react with the same response if the same traffic conditions occur (chapter 6). The *activation level* is adopted as an additional variable – new in macroscopic traffic flow modelling – that affects the *driving strategy* of drivers. On the one hand, the activation level depends on traffic conditions; on the other hand, the activation level determines the current driving strategy and therefore indirectly traffic conditions at this and future moments.

Using similar mathematical formalism from kinetic traffic flow theory, the new variable is mathematically introduced, and a dynamic model equation for the activation level is established. The strength of this approach is illustrated in two case studies, in which we assume that drivers have some comfortable normal activation level, but can drive more or less active depending on traffic conditions. The effect of temporarily driving more actively near merging bottlenecks lets congestion set in some distance downstream of the merge, a phenomenon, well-known from empirical studies, that is called *capacity funnel*. The effect of temporariliy driving less actively (due to loss of motivation in congestion) significantly reinforces the capacity drop. These applications show how the new model can reproduce refined phenomena in congestion dynamics. Also, the results support the hypothesis that these phenomena might be caused or at least reinforced by variations of the driving behaviour.

We conclude that the human-kinetic traffic flow model, with activation level as an additional variable to describe variable driving behaviour, is able to describe traffic flow dynamics more refined than the basic model or other models that use fixed driving strategies. However, so far the model refinements were based on hypotheses of driver behaviour that have *not* been validated. This requires data of individual driver behaviour in real-life and experimental conditions that are not available to date.

9.1.2.4 Multiclass macroscopic simulation for the exploration of ADA systems

Again exploiting the mathematical modelling techniques from the kinetic traffic flow theory, combined with a novel solution for assigning available road space over the different user classes, a multiclass version of the human-kinetic model has been established (chapter 7). The interactions between the classes are modelled in a logical manner in the *multiclass acceleration integral*. In theory this approach allows to simultaneously model user classes with totally different microscopic behaviours. One traditional application is the modelling of person cars and trucks, illustrating the natural asymmetry of interactions between faster and slower user classes. A critical discussion of this case study revealed a weakness of the current multiclass human-kinetic model: speeds in a single-lane model are only partly synchronised between user classes. Without elaborating it at this stage, the proposed remedy is to distinguish constrained vehicles (i.e. followers in a platoon) and unconstrained vehicles (platoon leaders), each with their own speed distribution, as in the model of Hoogendoorn (1999). Vehicles belonging to a faster user class would then eventually be constrained by vehicles of the slower user class(es), thereby also adopting the speed of the slower class.

Another multiclass case study is the modelling of ADA systems, a type of application unique in macroscopic traffic flow modelling (chapter 8). Using the consistent mathematical techniques from this thesis, an ADA system has been modelled that was designed to warn drivers when approaching sharp decelerations in a queue tail. The system builds on inter-vehicle communication that triggers warnings to drivers approaching the queue-tail. As a reaction to these warnings, the driver is responsible for adequate control actions. The assumption used here is that warnings increase the activation level of drivers, which causes them to anticipate more adequately to decelerations downstream. The explorative analysis shows that the system results in safer and smoother transition from free flowing to congested traffic. It also avoids instability of the queue-tail, thus preventing oscillating congested traffic patterns in favour of a homogeneous pattern.

9.2 Recommendations for further research

The recommendations for future research are categorised according to the following subjects:

- Theoretical issues and further development of the modelling technique (section 9.2.1);
- Challenges with respect to the numerical evaluation (section 9.2.2);
- Continuing the development and interfacing of quantitative individual behavioural models (section 9.2.3);
- Providing extensive validation at microscopic and macroscopic levels (section 9.2.4);
- Exploring and designing future traffic systems (section 9.2.5).

9.2.1 Theoretical issues and modelling technique

Some further theoretical developments and improvements to modelling technique with potential practical relevancy are:

- *Distinction of constrained and unconstrained vehicles*. The multiclass modelling results of chapter 7 revealed that speed synchronisation within a single lane without overtaking probability requires that constrained and unconstrained vehicles be distinguished. Hoogendoorn (1999) has shown how this distinction can be accounted for in kinetic models. It is recommended to apply these techniques to the human-kinetic modelling approach.

  This is not only a theoretical improvement, it will also improve the simulation of ADA systems, like the one explored in this thesis: the equipped drivers who tend to brake earlier will impede non-equipped vehicles, in that way obliging these vehicles to change into constrained vehicles and brake along with the ADA

drivers. Therefore, the effects of ADA systems in the flow will in reality be sensed at lower penetration rates than in the analyses presented in this thesis.

- *Multi-lane extension of the human-kinetic model*. The model for merging and lane-changing behaviour in this thesis is relatively coarse. Hoogendoorn (1999) and Klar & Wegener (1999) have shown how refinements to these model components can be introduced in a multi-lane kinetic model. Due to the kinetic foundation of the human-kinetic modelling approach these techniques can be adopted in our model in a relatively straightforward manner. Moreover, it is also recommended to examine whether more refined lane-changing models than those just mentioned, for instance by taking into account the finite duration of the lane-changing manoeuvre (see section 4.1.1), the dependency between longitudinal and lateral control, or variable lane-changing behaviour as a function of the activation level.

  Again this theoretical improvement is also practically relevant for the modelling of certain types of ADA systems, for example co-operative merging systems. Also, multi-lane models are relevant for the more sophisticated forms of driver support. Let us consider for instance multi-lane motorway facilities with one or more lanes dedicated to vehicles with a high level of automation (Automated Highway Systems or AHS), while other lanes are suited for mixed ADA and manual traffic. ADA system can then be conceived that operate in mixed traffic, while at the same time preparing a co-operative merging manoeuvre into the automated lanes. Such ADA-AHS concepts are important since they enable the migration from current motorway systems to fully automated concepts. In theory such concepts can be modelled using a multiclass multi-lane extension of the human-kinetic traffic flow model.

- *Modelling different types of bottlenecks*. So far, only a bottleneck caused by merging traffic flows at an on-ramp was considered. In order to be practically relevant, also other types of bottlenecks need to be modelled, like weaving sections, saturated off-ramps with spillback onto the motorway, steep grades, visual bottlenecks (or other bottleneck causes not directly related to the infrastructure or conflicting demands, see for instance Janssen et al. 1995) etcetera. For many of these bottlenecks, a multi-lane extension of the human-kinetic model is required.

- *Transferability of parameters.* Current traffic flow models need to be recalibrated for each application in a new traffic network. This indicates that behaviour of traffic be dependent on local conditions in a way that is not described in existing traffic flow models. It is recommended to examine how consideration of (one or more) behavioural variables like the activation level can overcome this problem, so that – once the model is validated – it is transferable to other networks purely by configuring it to the local infrastructure, but without changing the behavioural parameters of the model.

9.2.2 Numerical evaluation

Different numerical issues have been identified throughout this thesis, the improvement of which would improve the accuracy and efficiency of human-kinetic traffic flow modelling. These issues are the following:

- *More efficient numerical integration of the acceleration integral*. The numerical evaluation of the acceleration integral is actually a bottleneck for computational speed of the human-kinetic model. If computation time is an issue (for instance for on-line applications or repetitive simulations in a trial-and-error process) the simple solution using the *extended trapezium rule* is not sufficient. Either more efficient numerical quadrature methods are then necessary or alternative solutions that speed up evaluation of the integral. For instance, evaluation through an n-dimensional look-up table or analytical approximations of the car-following relation and for the probability distributions of speeds and gaps.
- *More accurate numerical solution schemes for the hyperbolic system*. Due to the particular mathematical properties of the human-kinetic model (see section 4.4.3), dedicated numerical solutions are needed that are better suited for the human-kinetic model than the classic Godunov-type schemes. One option could be to elaborate the *staggered grid* solution schemes, known from computational fluid dynamics, as suggested in section 4.4.3.3.
- *Exploration of direct numerical solutions to the kinetic model.* In this thesis we followed the traditional approach to solving the kinetic model formulation via the method of moments and numerical techniques for the macroscopic equivalent model. It is worth exploring however whether efficient and comprehensible numerical solution techniques can be applied directly to the kinetic model formulation (i.e. without considering the macroscopic moments). Some seminal work has been done by Klar & Wegener (1997; direct integration of the generalised density function), Hoogendoorn & Bovy (2001b) and Waldeer (2002a, 2002b, 2004c; Monte Carlo simulation).

9.2.3 Developing and interfacing validated individual behavioural models

The refinement to traffic flow models in this thesis consists of a micro-macro link and the introduction of the activation level as a quantitative variable affecting the current driving strategy of drivers. Full advantage of these innovations can only be made if validated quantitative models of individual driver behaviour come available. As stated in section 1.3.1.1, such work is underway, among others within the organisations that have contributed to this thesis. The main challenge to behavioural psychologists building these models is to provide traffic flow theorists with *quantitative* models covering the *full range* of traffic conditions, i.e. from free flowing to totally jammed traffic. The main challenge for traffic flow theorists is to flexibly apply the human-kinetic traffic flow theory of this thesis for interfacing individual models to traffic flow models.

9.2.4 *Validation at microscopic and macroscopic levels*

Model development in itself is a meaningful activity since it requires fully understanding the system that is modelled and the processes governing its behaviour. The model described in this thesis has been verified. This means that the model, the algorithms, and the numerical implementation have been tested and found plausible. However, real-life applications of any model are only meaningful if the model is empirically valid. We therefore recommend continuing the efforts that were exerted parallel to developing this thesis with the purpose of collecting empirical validation data of individual driver operations in congestion (see section 1.3.1.2). In order to maintain the strength of the micro-macro link provided by the human-kinetic model, such validation should proceed on both the individual driver level and the macroscopic level of traffic flow dynamics. For instance, validation of the behavioural variable activation level ideally needs to be performed at the microscopic level, since activation level of a driver can only be measured (indirectly) at the individual level (e.g. by examining the behaviour of test drivers in an instrumented vehicle).

Note that not only the parameters need to be validated, but also the modelling assumptions, like for instance the car-following relationship on which the current human-kinetic model builds. In this respect it is recommended to examine whether a more valid variant of the human-kinetic model might be obtained by an alternative car-following relation that would be inspired by the three-phase traffic flow theory of Kerner. The essential differences between such car-following model and the one applied in this thesis are that the former (i) does not converge to some preset (speed dependent) following distance, but rather accepts a *range* of following distances (as for instance applied in section 6.4.1.2) and (ii) needs consideration of a probability that a driver makes *control errors* to initiate the jam transition in the 'pinch region' (see section 3.2.1).

Finally, it is recommended to benchmark the validated model against other traffic flow models. An interesting question is whether a validated human-kinetic model with consideration of the activation level yields a better description and prediction of traffic flow compared to other less complex models that assume invariable driver behaviour.

9.2.5 *Design of ADA, AHS and DTM applications*

A practical direction for further research is the exploration of future traffic systems by means of the human-kinetic traffic flow model (or future extensions thereof). Exploring or designing such systems from a macroscopic point of view may seem useless, since vehicle-bound innovations are likely to be driven by the automotive industry (technology push), who primary adopt the perspective of the individual vehicle. Application of detailed microscopic simulation tools is therefore an obvious choice.

However, the behaviour and performance of the future traffic system is an important societal issue, so that analyses and future developments should not be left over to only one of the stakeholders: research institutes and authorities have a responsibility to gain maximum benefit from the technological innovations that are provided by the

automotive industry. An independent stream of research that watches over the societal benefits of proposed innovations of the traffic system is therefore indispensable. A macroscopic scope for the analysis and exploration of alternative designs of traffic systems fits naturally in this role.

Examples of applications that need to be explored and designed are the hybrid ADA-AHS systems mentioned in section 9.2.1. Research into such systems may contribute to giving concrete form to technological road maps leading from the current traffic system and ADA systems – where the driver retains primary responsibility – to fully Automated Highway Systems (ADASE, 2003). In the nearer future, the question arises how road authorities can make use of the emergence of ADA systems to increase the potential of Dynamic Traffic Management (DTM). Today the influence of DTM on traffic performance in road networks is limited, since only a handful of control measures are at the disposal of traffic managers. They can only hope that road users make efficient use of general advice or traffic information to alleviate congestion problems. Combining DTM with ADA systems opens new perspectives in this field, since it enables traffic managers to directly communicate with individual drivers for personalised advice or control, and to involve the vehicle intelligence (ADA functions) in the DTM strategy.

The human-kinetic traffic flow model is recommended as one of the tools that can facilitate the explorative design of such systems (see also section 8.5).

BIBLIOGRAPHY

ADASE 2 consortium (2003), ADASE roadmap, Deliverable D2D, available through www.adase2.net

Ahmed, K.I. (1999), Modeling Drivers' Acceleration and Lane Changing Behavior, Ph.D. Thesis, Massachusetts Institute Of Technology, (available through: http://mit.edu/its/publications.html).

Ahmed, K.I., M.E. Ben-Akiva, H.N. Koutsopoulos, & R.G. Mishalani (1996), Models of freeway lane changing and gap acceptance behavior, In J. Lesort (Ed.), Transportation and Traffic Theory, pp. 501-515.

Alkim, T.P., H. Schuurman & C.M.J. Tampère (2000), Effects of External Cruise Control and Co-operative Following on Highways: an Analysis with the MIXIC Traffic Simulation Model. Proceedings of the IEEE Intelligent Vehicles Symposium of the Institute of Electrical and Electronics Engineers, Dearborn, USA.

Aw, A. & M. Rascle (2000), Ressurection of "second order" models of traffic flow, SIAM Journal of Applied Mathematics, 60, pp.916-938.

Bando, M., K. Hasebe, A. Nakayama, A. Shibata & Y. Sugiyama (1995), Dynamical model of traffic congestion and numerical simulation, Physical Review E, Vol. 51, pp. 1035-1042.

Banks, J.H. & M.R. Amin (2003), Test of a behavioral theory of multi-lane traffic flow: queue and queue discharge flows, Proceedings of the 82nd Annual Meeting of the Transportation Research Board.

Banks, J.H. (1991), Two-capacity phenomenon at freeway bottlenecks: some theoretical issues, Transportation Research Record 1320, pp. 234-241.

Banks, J.H. (1999), *Investigation of Some Characteristics of Congested Flow*, Transportation Research Record 1678, pp. 129-134.

Banks, J.H. (2002), *Validation Of Daganzo's Behavioral Theory Of Multi-Lane Traffic Flow : Interim Report,* PATH working paper ; UCB-ITS-PWP-2002-1 (available through: www.path.berkeley.edu)

Barcelo, J., E. Codina, J. Casas, J.L. Ferrer & D. Garcia (2002), Microscopic traffic simulation : a tool for the design, analysis and evaluation of intelligent transport systems, submitted to: Journal of Intelligent and Robotic Systems, Theory and Applications, available through: www.aimsun.com

Bellemans, T. (2003), Traffic control on motorways, PhD thesis, Katholieke Universiteit Leuven, Belgium.

Bellemans, T., B. De Schutter & B. De Moor (2002), Model predictive control with repeated model fitting for ramp metering, Proceedings of the IEEE 5th International Conference on Intelligent Transportation Systems.

Bertini, R.L. & M.J. Cassidy (2002), Some observed queue discharge features at a freeway bottleneck downstream of a merge, Transportation Research 36A, pp.683-697.

Bliemer, M.C.J. (2001), Analytic dynamic traffic assignment with interacting user-classes. Theoretical advances and applications using a variational inequality approach. Dissertation thesis, Delft University of Technology, Faculty of Civil Engineering and Geosciences, Delft University Press, Delft, The Netherlands

Boer, E.R. (1999), *Car following from the driver's perspective*, Tr. Res. F Vol. 2, No. 4, pp.201-206.

Bovy, P.H.L. (2001), Traffic flooding the Low Countries: how the Dutch cope with motorway congestion, Transport Reviews, Vol. 21, No. 1, pp.89-116

Brackstone, M. & M. McDonald (1999), Car-following: a historical review, Transportation Research 2F, pp. 181-196.

Brookhuis, K.A., D. de Waard & W.H. Janssen (2001), Behavioural impacts of Advanced Driver Assistance Systems – an overview, European Journal of Transport and Infrastructure Research EJTIR, Vol. 1, No. 3, pp. 245-253.

Buckley, D.J. & S. Yagar (1974), *Capacity funnels near on-ramps*, Proceedings of the 6th International Symposium on Transportation and Traffic Theory, Sydney, Australia.

Cassidy, M.J. & J.R. Windover (1995), Methodology for assessing dynamics of freeway traffic flow, Transportation Research Record 1484, pp. 73-79.

Cassidy, M.J. & M. Mauch (2001), An observed traffic pattern in long freeway queues, Transportation Research 35A, pp. 149-162.

Cassidy, M.J. & R.L. Bertini (1999a), *Observations at a Freeway Bottleneck*, in: Transportation and Traffic Theory (ed: Ceder, A.), Proc. of the 14th International Symposium on Transportation and Traffic Theory, Jerusalem 1999, pp.107-124.

Cassidy, M.J. & R.L. Bertini (1999b), Some traffic features at freeway bottlenecks, Transportation Research 33B, pp. 25-42.

Chakroborty, P. & S. Kikuchi (1999), *Evaluation of the General Motors based car-following models and a proposed fuzzy inference model*, Tr. Res. C 7(1999) pp. 209-235.

Chandler, R.E., R. Herman & E.W. Motroll (1958), Traffic dynamics: studies in car following, Operations Research 6.

Chang, M., J. Kim, K. Kang & J. Yun (2001), Evaluation of driver's psychophysiological load at freeway merging area, Proceedings of the 80th Annual Meeting of the Transportation Research Board.

Chanut, S. & C. Buisson (2003), Godunov discretisation of a two-flow macroscopic model for mixed traffic with distinguished speeds and lengths, Proceedings of the 82nd Annual Meeting of the Transportation Research Board.

Chowdhury, D., L. Santen & A. Schadschneider (1999) Vehicular traffic: a system of interacting particles driven far from equilibrium, Current Science 77, pp.411-419

Coifman, B. (1997), *Time Space diagrams for thirteen shock waves*, PATH working papers, University of Berkeley California.

Cremer, M. & A.D. May (1985), *An extended traffic model for freeway control*, Technical report UCB-ITS-RR-85-7, Institute for Transportation Studies, University of California, Berkeley.

Daganzo, C.F. (1994), The cell transmission model: a dynamic representation of highway traffic consistent with the hydrodynamic theory, Transportation Research 28B, no.4, pp. 269-287.

Daganzo, C.F. (1995a), *Requiem for second-order fluid approximations of traffic flow*, Tr. Res. B, vol 29B, No.4, pp.277-286.

Daganzo, C.F. (1995b), The cell transmission model, part II: network traffic, Transportation Research 29B, no.2, pp. 79-93.

Daganzo, C.F. (1997), A continuum theory of traffic dynamics for freeways with special lanes, Transportation Research 31B, pp. 103-125.

Daganzo, C.F. (2002a), *A behavioral theory of multi-lane traffic flow. Part I: Long homogeneous freeway sections*, Transportation Research B vol 36B, issue 2, pp.131-158 (available at www.path.berkeley.edu).

Daganzo, C.F. (2002b), *A behavioral theory of multi-lane traffic flow. Part II: Merges and the onset of congestion*, Transportation Research B vol 36B, issue 2, pp. 159-169 (available at www.path.berkeley.edu).

Daganzo, C.F., M.J. Cassidy & R.L. Bertini (1999), *Possible explanations of phase transitions in highway traffic*, Tr. Res. Vol. 33A No. 5, pp.365-379.

Dam, A. van (2002). *A moving mesh finite volume solver for macroscopic traffic flow models*. Master Thesis, Utrecht University, Dept. of Mathematics, Utrecht The Netherlands (see also: www.inro.tno.nl/five/traflow for a free download of the numerical simulation tool)

De Waard, D. (1996), *The Measurement of Drivers' Mental Workload*, Dissertation thesis, Traffic Research Centre, University of Groningen

Delorme, D. & B. Song (2001), Human driver model for SmartAHS, California PATH Research Report UCB-ITS-PRR-2001-12.

Demir, C. (2001), Modelling the impact of ACC-systems on the traffic flow at macroscopic modelling level, Workshop on Traffic and Granular Flow 2001, Nagoya Japan.

Deshpande, A., A. Gollu, & L. Semenzato (1997), The Shift Programming Language and Run-Time System for Dynamic Networks of Hybrid Automata, Univ. California Berkeley, UCB-ITS-PRR-97-7.

Dettmann, C.P., E.G.D. Cohen & H. van Beijeren (1999), *Microscopic chaos from Brownian motion?*, http://arxiv.org/PS_cache/chao-dyn/pdf/9904/9904041.pdf

Dijker, T., P.H.L Bovy & R.G.M.M. Vermijs (1999), *Car-following under non-congested and congested conditions*, Transportation Research Record 1644, pp.20-28 (full report available through http://www.fosim.nl/UK/DocumentenUK.html)

Edie, L.C. & R.S. Foote (1961), Experiments on single-lane flow in tunnels, In: Theory of Traffic Flow, Proceedings of Symposium on the Theory of Traffic Flow, Ed.: R. Herman, Elsevier Publishing, Amsterdam, New York, pp. 175-192.

Emmerich, H., T. Nagatani & K. Nakanishi (1998), *From modified KdV-equation to a second-order cellular automaton for traffic flow*, Physica A 254 (1998) pp. 548-556.

European Commission (2003), Halving the number of road accident victims in the European Union by 2010: a shared responsibility, Communication from the Commission, European Road Safety Action Programme, COM(2003) 311 final, Brussels

Evans, G.A. (1993), Practical numerical integration, Wiley, John & Sons, Incorporated

Evans, J.L., L. Elefteriadou & N. Gautam (2001), Probability of breakdown at freeway merges using Markov chains, Transportation Research 35B, pp. 237-254.

Fosim (2004), *Freeway Operations SIMulation*, http://www.fosim.nl/

Fuller, R. (1984), A conceptualization of driving behavour as threat avoidance, Ergonomics Vol. 27, pp.1139-1155.

Gazis, D.C., R. Herman & R. Rothery (1962), Nonlinear follow-the-leader models of traffic flow, Operations Research 9.

Gazis, D.C., R. Herman & R.B. Potts (1959), Car-following theory of steady-state traffic flow, Operations Research 7.

Hegyi, A. (2004), Model Predictive Control for integrating traffic control measures, PhD thesis, Delft University of Technology, The Netherlands.

Helbing, D, A. Hennecke, V. Shvetsov & M. Treiber (2001), *MASTER: macroscopic traffic simulation based on a gas-kinetic, non-local traffic model,* Tr. Res. B. Vol. 35, No. 2, pp. 183-211.

Helbing, D. & B. Tilch (1998), *Generalized Force Model of Traffic Dynamics*, Physical Review E 58, pp. 133-138.

Helbing, D. & M. Schreckenberg (1999), *Cellular Automata Simulating Experimental Properties of Traffic Flow*, Phys. Rev. E 59, pp. R2505-R2508.

Helbing, D. & M. Treiber (1998), *Gas-Kinetic-Based Traffic Model Explaining Observed Hysteretic Phase Transition*, Physical Review Letters 81, pp. 3042-3045

Helbing, D. & M. Treiber (1999), *Numerical simulation of macroscopic traffic equations,* Computing in Science & Engineering 1, pp. 89-99.

Helbing, D. & M. Treiber (2002), *Critical Discussion of Synchronized Flow,* Cooper@tive Tr@nsport@tion Dyn@mics 1 , pp. 2.1-2.24 (on-line journal of www.trafficforum.org)

Helbing, D. (1997), *Verkehrsdynamik, Neue physikalische Modellierungskonzepte*, Springer Verlag, Berlin.

Helbing, D. (1998), *From Microscopic to Macroscopic Traffic Models*, in: J. Parisi, S. C. Muller, and W. Zimmermann (eds.) A Perspective Look at Nonlinear Media. From Physics to Biology and Social Sciences, Springer, Berlin, pp. 122-139

Helbing, D. (2001), *Traffic and related self-driven many-particle systems*, Reviews of Modern Physics, Vol. 73, pp. 1067-1141.

Helbing, D., A. Hennecke & M. Treiber (1999), *Phase Diagram of Traffic States in the Presence of Inhomogeneities*, Phys. Rev. Let. 82 pp.4360-4363.

Helbing, D., A. Hennecke V. Shvetsov & M. Treiber (2002), *Micro- and Macrosimulation of Freeway Traffic*, Mathematical and Computer Modelling 35, pp. 517-547.

Helly, W. (1961), Simulation of bottlenecks in single-lane traffic flow. In: Theory of Traffic Flow, Proceedings of Symposium on the Theory of Traffic Flow, Ed.: R. Herman, Elsevier Publishing, Amsterdam, New York, pp. 207-238.

Hennecke, A., M. Treiber & D. Helbing (2000), *Macroscopic Simulation of Open Systems and Micro-Macro Link*, Traffic and Granular Flow '99, edited by: Helbing, D., H.J. Herrmann, M. Schreckenberg & D.E. Wolf (available at: www.tu-dresden.de/vkiwv/vwista/sta1_hp.html).

Hirsch, C. (1990), Numerical computation of internal and external flows, Vol. 1: Fundamentals of numerical discretisation, John Wiley & Sons, Chichester, England

Hoedemaeker, M. (1999). Driving with intelligent vehicles. Driving behaviour with Adaptive Cruise Control and the acceptance by individual drivers. Dissertation thesis, University of Groningen, Centre for Environmental and Traffic Psychology, Delft University Press, Delft, The Netherlands.

Hogema, J.H., (1995), Individual car-following behaviour on motorways; a simulator study, TNO report TNO-TM 1995 B-15

Hoogendoorn, S.P. & M.M. Minderhoud, ADAS impact assessment by micro-simulation, European Journal of Transport and Infrastructure Research EJTIR, Vol 1, No. 3, pp. 255-275

Hoogendoorn, S.P. & P.H.L. Bovy (2000), *Modelling multiple user-class traffic flow,* Tr. Res. B, Vol 34B, No.2, pp.123-146.

Hoogendoorn, S.P. & P.H.L. Bovy (2001a), Model assessment of dynamic speed limit control. Proceedings of the 2001 World Conference on Transportation Research, Seoul, Korea

Hoogendoorn, S.P. & P.H.L. Bovy (2001b), Platoon-Based Multiclass Modeling of Multilane Traffic Flow, Journal of Networks and Spatial Economics, Vol.1, No.1-2, pp. 137-166

Hoogendoorn, S.P. & P.H.L. Bovy (2001c), State-of-the-art of vehicular traffic flow modelling. Proc. Instn. Mech. Engrs. Vol. 215 Part 1, pp.283-303.

Hoogendoorn, S.P. & T. Alkim (1999), Expert views on traffic flow operations during congestion, Research report of the Traffic Engineering section, Faculty of Civil Engineering and Geosciences, Delft University of Technology

Hoogendoorn, S.P. (1999), Multiclass Continuum Modeling of Multilane Traffic Flow, Dissertation thesis, Delft University of Technology, Faculty of Civil Engineering and Geosciences, Delft University Press, Delft, The Netherlands.

Hoogendoorn, S.P. (2003), Microscopic Traffic Data Collection by Remote Sensing, Preprints of the 82nd Annual Meeting of the Transportation Research Board 2003

Hoogendoorn, S.P., H. van Zuylen & B. Gorte (2002), Opstellen meetsysteem dataverzameling congestie, Rapport VK rapport 2002-006, Technische Universiteit Delft, Faculteit Civiele Techniek en Geowetenschappen

Horowitz R. & P. Varaiya (2000), Control design of an automated highway system, Proc. IEEE, vol. 88, pp. 913-925.

Issa, R.I. (1986), Solution of implicitly discretised fluid flow equations by operator splitting, Journal of Computational Physics, Vol. 62, pp.40-65.

Janssen, W.H., N.A. Kaptein, J.H. Hogema & M. Westerman (1995), 'Quick scan' wegennet Rijkswaterstaat Zuid-Holland, TNO-rapport TNO-TM 1995 C-42, TNO The Netherlands (in Dutch).

Jeans, Sir J. (1982), An introduction to the kinetic theory of gases, Cambridge University Press

Jepsen (1998), On the speed-flow relationships in road traffic: a model of driver behaviour, In: Proceedings of the third international symposium on highway capacity (Ed: Rysgaard, R.), Danish Road Directorate, Kopenhagen, Denmark.

Jin, W.L. & H.M. Zhang (2003), The formation and structure of vehicle clusters in the Payne–Whitham traffic flow model, Transportation Research Part B 37, pp. 207–223

Karaaslan, U., P. Varaiya & J. Walrand (1990). Two proposals to improve freeway traffic flow, Path publication UCB-ITS-PRR-90-6, available at: www-path.eecs.berkeley.edu/PATH/Publications/PATH/index.html

Kerner, B.S. & H. Rehborn (1996a), Experimental features and characteristics of traffic jams, Pysical Review E Vol.53 No. 2, pp. 1297-1300.

Kerner, B.S. & H. Rehborn (1996b), *Experimental properties of complexity in traffic flow*, Physical Review E Vol. 53 No. 5-A, pp. 4275-4278.

Kerner, B.S. & P. Konhäuser (1993), *Cluster effect in initially homogeneous traffic flow*, Phys. Rev. E 48(4), pp.2335-2338.

Kerner, B.S. & S.L. Klenov (2002), A microscopic model for phase transitions in traffic flow, Journal of Physics A: Mathematical and General, Vol. 35, pp. L31-L43.

Kerner, B.S. & S.L. Klenov (2003), A microscopic theory of spatial-temporal congested traffic patterns at highway bottlenecks. Physical Review E 68 No. 3

Kerner, B.S. (1998), Experimental features of self-organization in traffic flow, Physical Review Letters, Vol.81, No. 17, pp. 3797-3800.

Kerner, B.S. (1999), *Theory of Congested Traffic Flow: Self-Organization without Bottlenecks*, Transportation and Traffic Theory, Proc. of the 14th International Symposium on Transportation and Traffic Theory, Jerusalem 1999, pp.147-171.

Kerner, B.S. (2004), Control of Spatial-Temporal Congested Traffic Patterns at Highway Bottlenecks, submitted to Cooper@tive Tr@nsport@tion Dyn@mics (on-line journal of www.trafficforum.org).

Kerner, B.S., S.L. Klenov & D.E. Wolf (2002), Cellular automata approach to three-phase traffic theory. Journal of Physics A: Mathematical and General, Vol. 35, pp. 9971-10013.

Kinsler, L.E., A.R. Frey, A.B. Coppens & J.V. Sanders (1999), Fundamentals of acoustics, John Wiley & Sons

Klar, A. & R. Wegener (1997), Enskog-like kinetic models for vehicular traffic, Journal of Statistical Physics 87, p.91. (available through: http://www.mathematik.tu-darmstadt.de/~klar/veroeffentlichungen.html)

Klar, A. & R. Wegener (1999), *A hierarchy of models for multilane vehicular traffic I: modeling*, SIAM Journal of Applied Mathematics, Vol. 59, No. 3, pp.983-1001 (available through: http://www.mathematik.tu-darmstadt.de/~klar/veroeffentlichungen.html)

Kleinman, D.L., Baron, S. & Levison, W.H. (1970). An optimal control model of human response. Part 1: Theory and validation. Automatica, Vol. 6.

Koppa, R.J. (1999), Human factors, in: "Traffic Flow Theory, a state of the art report. Revised Monograph on Traffic Flow Theory", ed. by: Gartner, N., C.J. Messer & A.K. Rathi

Koshi, M., M. Iwasaki & I. Ohkura (1983), Overview on vehicular flow characteristics, In: Hurdle, V.F., E. Hauer & G.N. Stewart (Eds.), Transportation and Traffic Theory, University of Toronto Press, pp. 403-426.

Krauss, S., K. Nagel & P. Wagner (1999), The mechanism of flow breakdown in traffic flow models, In: Transportation and Traffic Theory, papers presented at the abbreviated presentation sessions of the 14th International Symposium on Transportation and Traffic Theory, Ed: A. Ceder.

Kruisbergen, J.W. van (2001), Verplaatsingssnelheid van schokgolven: empirisch onderzoek naar het verplaatsingsproces van files, Afstudeerscriptie Nationale Hogeschool voor Toerisme en Verkeer, Breda; Transpute Adviesbureau, Gouda, Nederland (in Dutch).

Kühne, R.D. (1984), *Macroscopic freeway model for dense traffic – stop-start waves and incident detection*, Proc. of the 9th International Symposium of Transportation and Traffic Theory, pp.21-42, ed. by: I. Volmuller & R. Hamerslag.

Lebacque, J.P. (1996), The Godunov scheme and what it means for first-order traffic flow models, Proceedings of the 13th International Symposium of Transportation and Traffic Flow Theory. INRETS, Lyon, pp. 647-677.

Lebacque, J.P. (2003), Two-Phase Bounded-Acceleration Traffic Flow Model: Analytical Solutions and Applications, Transportation Research Record No. 1852, pp. 220-230.

Lee, H.Y., H.-W. Lee & D. Kim (1998), Origin of synchronized traffic flow on highwayc and its dynamic phase transitions, Physical Review Letter 81, p.1130.

Lee, H.Y., H.-W. Lee & D. Kim (1999), *Dynamic States of a Continuum Traffic Equation with On-Ramp*, Phys. Rev. E, 59, p.5101-5111.

Lee, H.Y., H.-W. Lee & D. Kim (2000), *Empirical Phase Diagram of Traffic Flow on Highways with On-Ramps*, Traffic and Granular Flow '99, edited by: Helbing, D., H.J. Herrmann, M. Schreckenberg & D.E. Wolf.

Lenz, H. (1999). Entwicklung nichtlinearer, diskreter Regler zum Abbau von Verkehrsflussinhomogenitaeten mithilfe makroskopischer Verhkehrsmodelle, Dissertation Thesis, Lehrstuehl fuer Mensch-Maschine-Kommunikation, Fakultaet fuer Elektrotechnik und Informationstechnik der Technischen Universitaet Muenchen, Shaker Verlag, Aachen.

Leutzbach, W. (1988). An introduction to the theory of traffic flow. Springer-Verlag, Berlin, Germany.

Lighthill, M.J. & G.B. Whitham (1955), *On kinematic waves: II a theory of traffic flow on long crowded roads.* In: Proc. of the Royal Society, volume 229(1178) of A, pp.317-345.

Lint, J.W.C. van (2004), *Quantifying Uncertainty in Real-Time Neural Network based Freeway Travel Prediction*, Preprints of the 83rd Annual Meeting of the Transportation Research Board, Washington

Liu, G., A.S. Lyrintzis & P.G. Michalopoulos (1998), *Improved higher-order model for freeway traffic flow*. Tr. Res. Rec. 1644, pp.37-46.

Logghe, S. (2003), Dynamic modeling of heterogeneous vehicular traffic. PhD thesis, Katholieke Universiteit Leuven, Belgium.

Lorenz, M. & L. Elefteriadou (2000), A probabilistic approach to defining freeway capacity and breakdown, Transportation Research Circular E-C018, Fourth International Symposium on Highway Capacity, pp. 84-95.

Ludmann, J. & M. Weilkes (1995). Entwicklung, Analyse und Bewertung von Prometheus Konzepten unter besonderer Berücksichtigung von Autonomous Intelligent Cruise Control mittels Simulation, Abschlussbericht Eureka Verbundprojekt, Institut für Kraftfahrwesen der RWTH Aachen.

Ludmann, J. & M. Weilkes (1997). Investigation of Intelligent Traffic Systems by means of simulation, Proceedings of the Intelligent Transportation Systems World Congress, Berlin, 1997.

Mauch, M. (2002), Analyses of start-stop waves in congested freeway traffic, PhD thesis, Department of Civil and Environmental Engineering, Institute of Transportation Studies, University of California at Berkeley (available through: http://www.uctc.net/papers/dissnumber.html).

May, A.D. (1990), Traffic flow fundamentals, Prentice-Hall, New Jersey

May, A.D., H.E.M. Keller (1967), Non-integer car-following models, Highway Research Record 199, pp. 19-32.

Messmer, A. & M. Papageorgiou (1990), METANET: a macroscopic simulation program for motorway networks, Traffic Engineering and Control, Vol.31, pp. 466-470.

Minderhoud, M.M. & P.H.L. Bovy (1998), *Modelling driver behaviour on motorways. Description of the SiMoNe model*, Delft University of Technology, Faculty of Civil Engineering and Geosciences.

Minderhoud, M.M. (1999), Supported driving: impacts on motorway traffic flow. Dissertation thesis, Delft University of Technology, Faculty of Civil Engineering and Geosciences, Delft University Press, Delft, The Netherlands.

Morsink, P., R. Hallouzi, I. Dagli, C. Cseh, L. Schäfers, M. Nelisse & D. de Bruin (2003), CARTALK 2000: Development Of A Co-Operative ADAS Based On Vehicle-To-Vehicle Communication, Proceedings of the 10th World Congress on Intelligent Transport Systems and Services, Madrid

Muñoz, J.C. & C. F. Daganzo (2000), Experimental Characterization of Multi-lane Freeway Traffic Upstream of an Off-ramp Bottleneck, California PATH working paper UCB-ITS-PWP-2000-13, California PATH program, Institute of Transportation Studies, University of California, Berkeley (available at www.path.berkeley.edu).

Muñoz, J.C. & C.F. Daganzo (2001), Moving Bottlenecks: a Theory Grounded on Experimental Observation, In: Transportation and Traffic Theory in the 21st Century, Proc. 15th Int. Symp. on Transportation and Traffic Theory (Ed: Taylor, M.A.P.), Pergamon-Elsevier, Oxford,U.K, pp. 441-462.

Nagel, K. & M. Schreckenberg (1992), *A Cellular Automaton Model for Freeway Traffic*, Journal de Physique I France, Vol. 2, p.2221.

Nagel, K. (1998), *From Particle Hopping Models to Traffic Flow Theory*, Tr. Res. Rec. 1644, pp.1-9.

Nelson, P. & B. Raney (1999), Objectives and Benchmarks for Kinetic Theories of Vehicular Traffic, Transportation Science Vol. 33, No. 3, pp.298-314

Nelson, P. (1995), *A kinetic model of vehicular traffic and its associated bimodal equilibrium solutions*, Transport Theory and Statistical Physics, Vol. 24(1-3), pp.383-409.

Nelson, P., D.D. Bui & A. Sopasakis (1997), *A novel traffic stream model deriving from a bimodal kinetic equilibrium*, Proceedings of the 1997 IFAC meeting, Chania, Greece, pp.799-804.

Newell, G.F. (1965), *Instability in dense highway traffic, a review*, In: J. Almond, ed., Proceedings of the second International Symposium on the Theory of Traffic Flow, pp. 73-83.

Ngoduy, D., Hoogendoorn, S.P. & H.J. Van Zuylen (2004), Cross-comparison of numerical schemes for macroscopic traffic flow models, proceedings of the 83rd Annual Meeting of the Transportation Research Board, Washington

Papageorgiou, M. (1989), *Dynamic modelling, assignment and route guidance in traffic networks*, Tr. Res B, Vol. 23B, pp. 29-48.

Paveri-Fontana, S.L. (1975), *On Boltzmann-like treatments for traffic flow: a critical review of the basic model and an alternative proposal for dilute traffic analysis,* Tr. Res. B. Vol. 9, pp. 225-235.

Payne, H.J. (1971), *Models of freeway traffic and control*, In: Mathematical models of public systems, volume 1 of Simulation Councils Proc. Ser., ed. by: Bekey, G.A., pp, 51-60.

Pipes, L.A. (1953), An operational analysis of traffic dynamics, Journal of Applied Physics, 24.

Prigogine, I. & R. Herman (1971), *Kinetic Theory of Vehicular Traffic*, American Elsevier, New York.

Prigogine, I. (1961). A Boltzmann-like Approach to the Statistical Theory of Traffic Flow. Theory of Traffic Flow (ed. R. Herman), Elsevier, Amsterdam.

PTV (2003), Vissim – traffic flow simulation, Technical report, PTV Gernamy (www.ptv.de)

Quadstone (2002), Paramics v4.0 system overview, Technical report, Quadstone Ltd. Scotland (www.paramics-online.com).

Ranjitkar, P., T. Nakatsuji, Y. Azuta & G.S. Gurusinghe (2003), Stability analysis based on instantaneous driving behavior using car-following data, Transportation Research Record No. 1852, pp. 140-151.

Ranney, T.A. (1999), Psychological factors that in¯uence car-following and car-following model development, Transportation Research Part F 2, pp. 213-219

Richards, P.I. (1956), *Shock waves on the highway*. Operations Research, vol. 4, pp.42-51.

Rothery, R.W. (1999), *Car Following Models*, in: "Traffic Flow Theory, a state of the art report. Revised Monograph on Traffic Flow Theory", ed. by: Gartner, N., C.J. Messer & A.K. Rathi

Saad, F. (1996), Driver strategies in car-following situations, In: Vision in vehicles, eds.: A.G. Gale, I.D. Brown, C.M. Haslegrave & S.P. Taylor

Sadeghhosseini, S. & R.F. Benekohal (2002), Headway models for low to high volume highway traffic, Proceedings of the 81st Annual Meeting of the Transportation Research Board, Washington

Shvetsov, V. & D. Helbing (1999), *Macroscopic dynamics of multi-lane traffic*, Physical Review E 59, pp. 6328-6339.

Smulders, S. (1990), *Control of Freeway Traffic Flow by Variable Speed Signs*, Tr. Res. B, vol 24B, No.2, pp. 111-132.

Smulders, S., S. Hoogendoorn & T. Alkim (2000), *Traffic Flow Operations during Congestion*, report of the AVV Transport Research Centre, Ministry of Public Works and Water Management, The Netherlands.

Sultan, B., M. Brackstone, B. Waterson & E.R. Boer (2002), Modeling the dynamic cut-in situation, Proceedings of the 81st Annual Meeting of the Transportation Research Board.

Swaroop, D. & K.R. Rajagopal (1998). Intelligent Cruise Control Systems and Traffic Flow Stability, Path publication UCB-ITS-PRR-98-36, available at: www-path.eecs.berkeley.edu/PATH/Publications/PATH/index.html

Tampère, C.M.J., 2001. Modelling of traffic flow in congestion for the analysis of Automated Driver Assistance Systems: Literature Review. Report TNO Inro-V+V/2001-59.

Tampère, C.M.J., P. Wewerinke & B. Van Arem (2002), Traffic flow behaviour in congestion for the analysis of Advanced Driver Assistance Systems (ADAS): Modelling framework, TNO Inro report 2002-56, Delft

Tampère, C.M.J., S.P. Hoogendoorn & B. Van Arem (2003), *Capacity funnel explained using the human-kinetic traffic flow model,* Accepted for publication in Traffic and Granular Flow 2003 proceedings.

Todosiev, E.P. & L.C. Barbosa (1964), A proposed model for the driver-vehicle system: the car-following problem, Traffic Engineering, 3, p.17-20.

Treiber, M. & D. Helbing (2001), Microsimulations of Freeway Traffic Including Control Measures, Automatisierungstechnik 49, 478-484 (available through: http://arxiv.org/abs/cond-mat/0210096)

Treiber, M. & D. Helbing (2002), Reconstructing the Spatio-Temporal Traffic Dynamics from Stationary Detector Data, Cooper@tive Tr@nsport@tion Dyn@mics Vol. 1, pp.3.1-3.24 (available through: http://www.trafficforum.org)

Treiber, M., A. Hennecke & D. Helbing (1999), *Derivation, Properties, and Simulation of a Gas-Kinetic-Based, Non-Local Traffic Model,* Physical Review E 59, pp.239-253 (available through: http://www.tu-dresden.de/vkiwv/vwista/sta1_hp.html)

Treiber, M., A. Hennecke & D. Helbing (2000), *Congested Traffic States in Empirical Observations and Microscopic Simulation*, Physical Review E 62, pp. 1805-1824.

Treiterer, J. & J.A. Myers (1974), The hysteresis phenomenon in traffic flow. Proceedings of the 6th International Symposium on Transportation and Traffic Theory. Ed.: D.J. Buckley, pp. 13-38.

Van Arem, B., A.P. De Vos & M.J.W.A. Vanderschuren (1997a), The effect of a special lane for intelligent vehicles on traffic flows, TNO Inro report 1997-02a, Delft

Van Arem, B., A.P. de Vos & M.J.W.A. Vanderschuren (1997b). The microscopic simulation model MIXIC 1.3. TNO report INRO-VVG 1997-02b, TNO Inro, Delft, The Netherlands

Van Arem, B., Hogema, J.H., & Smulders, S.A. (1996). The impact of Autonomous Intelligent Cruise Control on traffic flow. Proceedings of the Third Annual World Congress on Intelligent Transport Systems. Paper No. 2032, Orlando, October 1996.

Van Arem, B., J.K.H. Carlier, O.J. Gietelink, C.J. van Waveren, F. Lilli, C. Lanfranco & G. Ghigo (2002), ADAS Critical Applications and Related Critical Scenario, Deliverable D1 of the GALLANT project (available through: http://www.crfproject-eu.org/)

Van Arem, B., P. de Vos & H. Schuurman (1998). Simulation of traffic flow on a special lane for intelligent vehicles. Proceedings of the Third International Symposium on Highway Capacity, Ed. by: R. Rysgaard, The Danish Road Directorate. Copenhagen, Denmark, June 1998.

Van der Hulst, M. (1999), Adaptive control of safety margins in driving, Dissertation thesis, University of Groningen, The Netherlands.

Van der Hulst, M., T. Rothengatter & T. Meijman (1998), Strategic adaptations to lack of preview in driving, Transportation Research Vol. 1F, pp. 59-75.

Van Toorenburg, J.A.C. (1983a), De mens in het autosnelwegverkeer (1), Verkeer als sociaal-psychologisch fenomeen, Verkeerskunde Vol. 34, No. 4, pp. 169-172 (in Dutch).

Van Toorenburg, J.A.C. (1983b), De mens in het autosnelwegverkeer (2), Verkeer als sociaal-psychologisch fenomeen, Verkeerskunde Vol. 34, No. 5, pp. 238-241 (in Dutch).

Van Winsum, W. & A. Heino (1996), Choice of time-headway in car-following and the role of time-to-collision information in braking, Ergonomics, Vol. 39, pp.579-592

Van Winsum, W. (1998), Preferred time headway in car-following and individual differences in perceptual-motor skills, Perceptual and Motor Skills, Vol. 87, pp.863-873

Van Winsum, W. (1999), The human element in car following models, Transportation Research Part F 2, pp. 207-211

VanderWerf, J., N. Kourjanskaia, S.E. Shladover, H. Krishnan & M. Miller (2001), Modeling the effects of driver control assistance systems on traffic, Proceedings of the 81st Annual Meeting of the Transportation Research Board.

Vanhove, F.L.B. & M.J.G. De Ceuster (2003), Traffic indices for the use of the Belgian motorway network, working paper of Transport & Mobility Leuven (available through: www.tmleuven.be)

Versteeg, H.K. & W. Malalasekera (1995), *An introduction to computational fluid dynamics. The finite volume method*, Longman Scientific and Technical, Essex, England

Waldeer, K.T. (2000), Beschleunigungsorientierte stochastische Beschreibung von Verkehrsfluessen, Technical Report 1; Fachbereich Transport- und Verkehrswesen der Fachhochschule Braunschweig/Wolfenbuettel. (available through: public.rz.FH-Wolfenbuettel.de/~waldeer).

Waldeer, K.T. (2001), A mesoscopic traffic flow model based on a stochastic acceleration jump process, in: Proc. 18th Dresden Conf. on Traffic and Transportation Science, Faculty of Traffic and Transportation Sciences 'Friedrich List', Dresden University, 2001, (available through: public.rz.FH-Wolfenbuettel.de/~waldeer or http://vwisb7.vkw.tu-dresden.de/TrafficForum/vwt_2001/start.html#prog).

Waldeer, K.T. (2002a), A numeric investigation of a vehicular traffic flow model based on a stochastic acceleration process, in: N. Attig, R. Esser, M. Kremer (Eds.), Proc. Europ. Conf. on Computational Physics (CCP 2001), Vol. 147 of Computer Physics Communications, Elsevier, Amsterdam, 2002, p. 650. (available through: public.rz.FH-Wolfenbuettel.de/~waldeer).

Waldeer, K.T. (2002b), The Direct Simulation Monte Carlo Method applied to a Boltzmann-like Vehicular Traffic Flow Model, Submitted to: Comput. Phys. Commun., (available through: public.rz.FH-Wolfenbuettel.de/~waldeer).

Waldeer, K.T. (2004a), A mesoscopic acceleration oriented traffic flow model and its comparison to measured data, submitted to Cooper@tive Tr@nsport@tion Dyn@mics (on-line journal of www.trafficforum.org).

Waldeer, K.T. (2004b), A vehicular traffic flow model based on a stochastic acceleration process, Transp. Theory Stat. Phys., in press (available through: public.rz.FH-Wolfenbuettel.de/~waldeer).

Waldeer, K.T. (2004c), Numerical Investigation of a Mesoscopic Vehicular Traffic Flow Model Based on a Stochastic Acceleration Process, Transp. Theory Stat. Phys., in press (available through: public.rz.FH-Wolfenbuettel.de/~waldeer).

Wallentowitz, H., D. Neunzig & J. Ludmann (1998), *Auswirkungen neuer Fahrzeug- und Verkehrstechnologien - Analyse von Verkehrsfluss, Kraftstoffverbrauch und Emissionen mit PELOPS*, in: Stadt, Region, Land, Heft 66, ed. by: K.J. Beckmann.

Wewerinke, P.H. & Hoedemaeker, D.M., (2001). Driver behaviour in congested traffic - Literature review and model analysis. TNO report TM-01-D014.

Wewerinke, P.H. & J.H. Hogema (2003), Stop and go driver model, Report TNO TM-03-D013

Wewerinke, P.H. (1989). Models of the human controller and observer of a dynamic system. Ph.D. thesis, University of Twente, The Netherlands.

Wewerinke, P.H., C.M.J. Tampère & J.H. Hogema (2002), *Modelling individual driver and platoon behaviour,* TNO Report TM-02-D001

Wewerinke, P.H., J.H. Hogema & C.M.J. Tampère (2003), Modelling individual driver to combine with a traffic flow model. Report TNO TM

Wewerinke, P.H.,(1992). Modeling manned traffic systems. Proceedings of the IMACS/SICE International Symposium on Robotics, Mechatronics and Manifacturing Systems '92. Kobe, Japan.

Wiedemann, R. (1974), *Simulation des Strassenverkehrsflusses*, Heft 8 der Schriftenreihe des IfV, Institut für Verkehrswesen, Universität Karlsruhe.

Windover, J.R. (1998), Empirical studies of the dynamic features of freeway traffic, PhD thesis, Department of Civil and Environmental Engineering, Institute of Transportation Studies, University of California at Berkeley.

Wu, J, M. Brackstone & M. McDonald (2000), *Fuzzy sets and systems for a motorway microscopic simulation model.* Fuzzy Sets and Systems 116(1), pp. 65-76 (available through: http://www.trg.soton.ac.uk/research/platform/pubs.htm).

Wu, N. & W. Brilon (1999), *Cellular Automata for Highway Traffic Flow Simulation*, Transportation and Traffic Theory, Papers presented at the Abbreviated Presentation Sessions of the 14th International Symposium on Transportation and Traffic Theory, Jerusalem

Yang, Qi, H.N. Koutsopoulos & M.E. Ben-Akiva (1999), A simulation laboratory for evaluating dynamic traffic management systems, Proceedings of the 79th Annual Meeting of the Transportation Research Board, available through: http://web.mit.edu/its/mitsimlab.html

Zhang, H.M. & T. Kim (2001), A car-following theory for multiphase vehicular traffic flow, Proceedings of the 80th Annual Meeting of the Transportation Research Board.

Zhang, H.M. & W.L. Jin (2002), A kinematic wave traffic flow model for mixed traffic, Proceedings of the 81st Annual Meeting of the Transportation Research Board.

Zhang, H.M. (1998), *A theory of nonequilibrium traffic flow*, Tr. Res. B, Vol. 32B, No. 7, pp.485-498.

Zhang, H.M. (1999), A mathematical theory of traffic hysteresis. Transportation Research B 33, pp.1-23

Zhang, H.M. (2000), *New perspectives on continuum traffic flow models*, Special Issue on Traffic Flow Theory. Journal of Networks and Spatial Economics, Vol. 1, pp. 9-33.

Zhang, H.M. (2003a), *A non-equilibrium traffic model devoid of gas-like behavior*, Tr. Res. B Vol. 36, No. 3, pp.275-290.

Zhang, H.M. (2003b), *Driver memory, traffic viscosity and a general viscous traffic flow model*, Tr. Res. B, Vol. 37, No. 1, pp.27-41.

Zhu, Z., G. Chang & T. Wu (2003), Numerical analysis of freeway traffic flow dynamics under multiclass drivers, Proceedings of the 82nd Annual Meeting of the Transportation Research Board.

Zwet, E. van, C. Chen, Z. Jia & J. Kwon (2002), A statistical method for estimating speed from single loop detectors, available through: http://pems.eecs.berkeley.edu/Resources.

A BASIC THEORY OF KINETIC VEHICULAR TRAFFIC FLOW MODELLING

A.1 Introduction

In this annex the basic theory of kinetic vehicular traffic flow modelling, which forms the basis of the model development in this thesis, is discussed. The aim of this annex is to introduce the basic concepts and methods used in this theory, since many of these are used throughout this thesis.

The annex is structured as follows. First, the basic concept of generalised density and its relation to traditional macroscopic traffic variables is introduced (section A.2). Then, we explain how macroscopic traffic flow models are derived from a generalised continuity equation (section A.3). In section A.4, we briefly describe how individual driver behaviour is specified traditionally in kinetic traffic flow theory. It is especially this step in the model derivation, for which the model development in this thesis (chapter 4) is an alternative approach. In section A.5, we discuss the mathematical properties of kinetic traffic flow models and their implications for the numerical evaluation. Finally, we derive in section A.6 a general linear stability criterion for kinetic models and show how it applies to well-known special cases. The criterion is also applicable to the human-kinetic model derived in this thesis, as is shown in section 5.2.1 of the main text.

A.2 Unit of description: the concept of generalised density

A.2.1 Generalisation of the traditional vehicular density

In this thesis we adopt the so-called *kinetic* description of traffic flow that was first introduced by Prigogine (1961). This is a probabilistic description of vehicular traffic borrowed from the kinetic description of compressible fluids and granular matter (3.4).

The kinetic theory for vehicular traffic builds upon the definition of the *state of individual vehicles* and the probability distribution of vehicle states S in phase space. Within the context of this thesis, we will define the vehicle state S in three ways, dependent on the context:

- generically, as a combination of n mutually independent state variables s_i: $S = (s_1, s_2, s_3, \ldots, s_n)$
- as the individual speed $v = \frac{dx}{dt}$: $S = (v)$
- as the combination of the individual speed v and the individual activation level a: $S = (v,a)$

Also, one could argue that in fact the longitudinal position x can be regarded as a state variable as well. We will therefore call S the *local* state, since it represents the state of a vehicle given the location. We then define the *non-local* state $S' = (x,S)$. Note that in general, the state definition can include an arbitrary number of state variables s_i, as long as the state variables are mutually independent. The independence requires that:

$$\frac{\partial s_i}{\partial s_j} = 0 \qquad \forall i \neq j \tag{A.1}$$

Another basic concept underlying the kinetic traffic flow theory is the *generalised density* ρ, a generalisation of the traditional *vehicular density* $k(t,x)$. The latter is interpreted from a probabilistic viewpoint: $k(t,x)\,dx$ is defined as the expected number of vehicles at time t and location $x \in [x,x+dx)$. The generalisation $\rho(t,x,S) = \rho(t,S')$ of the traditional vehicular density k considers subsets of the total number of vehicles *with a common state*, instead of the total population of vehicles. The generalised density $\rho(t,x,S)\,dx\,dS$ defines the expected number of vehicles having state $S \in [S,S+dS)$ at time t and location $x \in [x,x+dx)$. Therefore, $N(t,L,S)$, the expected number of vehicles at time t on a finite stretch of road with length L and driving in state S, is found as:

$$N(t,L,S) = \int_{x=0}^{x=L} \rho(t,x,S)\,dx \tag{A.2}$$

The relation between the traditional vehicular density $k(t,x)$ and the generalised density $\rho(t,x,S)$ is:

$$k(t,x) = \int_{S} \rho(t,x,S)\,dS \tag{A.3}$$

Therefore, $N(t,L)$, the expected number of vehicles – regardless of their state S – at time t on a finite stretch of road with length L, can also be derived from the generalised density via:

$$N(t,L) = \int_{x=0}^{x=L} \int_{S} \rho(t,x,S) \; dS \; dx \tag{A.4}$$

A.2.2 Definition of the most important moments of the generalised density ρ

Just like any probability density function, the moments of the function ρ can be calculated. This is important in the kinetic theory for two reasons: (a) it relates the rather abstract function ρ to macroscopic variables, known from traditional traffic engineering and traffic flow modelling, and (b) by applying the method of moments, it allows to derive macroscopic traffic flow models from kinetic models (see section A.3.2).

Let us define the moments of the probability density function ρ that have a useful interpretation in traffic flow theory. In general, the κ^{th} order moment of ρ for state variable s_i is defined by:

$$\begin{aligned} M_{s_i}^{\kappa}(t,x) &= \int_S s_i^{\kappa} \rho(t,x,S) \; dS \\ &= \int_{s_1} \int_{s_2} \ldots \int_{s_n} s_i^{\kappa} \rho(t,x,S) \; ds_n \; \ldots \; ds_2 \; ds_1 \end{aligned} \tag{A.5}$$

Moments that we will use throughout this thesis are:

- the 0^{th} order moment for any state variable s_i, equal to the *density*:

$$\begin{aligned} M_{s_i}^{0}(t,x) &= \int_S s_i^{0} \rho(t,x,S) \; dS \\ &= \int_S \rho(t,x,S) \; dS \\ &= k(t,x) \end{aligned} \tag{A.6}$$

- the 1^{st} order speed moment, also called the *flow rate* $Q(t,x)$:

$$\begin{aligned} M_{v}^{1}(t,x) &= \int_S v \, \rho(t,x,S) \; dS \\ &= k(t,x) \, V(t,x) \\ &= Q(t,x) \end{aligned} \tag{A.7}$$

  in which we have implicitly defined the average or *expected vehicular speed* $V(t,x)$ as:

$$V(t,x) = \int_S v \, \frac{\rho(t,x,S)}{k(t,x)} \, dS \tag{A.8}$$

  and the speed momentum or flow $Q(t,x)$ as:

$$Q(t,x) = k(t,x) \, V(t,x) \tag{A.9}$$

- the 2nd order speed moment or *traffic energy*, that relates to the *speed variance* $\Theta(t,x)$ as follows:

$$\begin{aligned} M_v^2(t,x) &= \int_S v^2 \rho(t,x,S)\,dS \\ &= \int_S (V-v)^2 \rho(t,x,S)\,dS + 2V\int_S v\rho(t,x,S)\,dS - V^2\int_S \rho(t,x,S)\,dS \\ &= k(t,x)\Theta(t,x) + 2V\,k(t,x)V(t,x) - V^2(t,x)k(t,x) \\ &= k(t,x)\left[\Theta(t,x) + V^2(t,x)\right] \end{aligned} \tag{A.10}$$

in which we have implicitly defined the *variance of vehicular speeds* $\Theta(t,x)$ as:

$$\Theta(t,x) = \int_S (V-v)^2 \frac{\rho(t,x,S)}{k(t,x)}\,dS \tag{A.11}$$

A.3 Generalised continuity equation and method of moments for deriving macroscopic traffic flow models

A.3.1 Generalised continuity equation for the generalised density ρ

For the generalised density $\rho(t,x,S)$, a continuity equation in the state space can be derived (see for example Hoogendoorn (1999), Leutzbach (1988)). The generic form reads:

$$\frac{\partial \rho}{\partial t} + \nabla_{S'} \cdot \left(\rho \frac{dS'}{dt}\right) = \left(\frac{d\rho}{dt}\right)_{discrete} \tag{A.12}$$

In this equation we have omitted the dependency of ρ on (t,S') or (t,x,S) for notational simplicity; the product operator $\cdot$ denotes the inner product and the Nabla operator $\nabla_{S'}$ for the non-local state vector S' with dimensions $(x,s_1,s_2,\ldots,s_n)$ is defined by:

$$\nabla_{S'} = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial s_1}, \frac{\partial}{\partial s_2}, \ldots, \frac{\partial}{\partial s_n}\right) \tag{A.13}$$

The interpretation of equation (A.12) is as follows. In the LHS it states that changes in time to the generalised density are caused by changes in time of the state variables s_i. These changes are larger as the expected number of vehicles in that state is larger (hence the multiplication by ρ), and as the sensitivity of ρ for changes in the state variable is larger (hence the derivative with respect to s_i. On the other hand, some discrete event can cause a sudden increase or decrease of ρ for a certain state S. For instance, new vehicles with state S can suddenly enter the system at position x, causing a discrete increase of $\rho(t,x,s)$. These discrete contributions to the change of ρ are accounted for in the term in the RHS of equation (A.12).

A.3.2 Method of moments applied to the generalised continuity equation

A.3.2.1 Definition of the method of moments

Equation (A.12) describes the dynamics of the generalised density ρ. Mostly one is not interested in the dynamics of the probability density function, but it is sufficient to know the dynamic behaviour of the moments of ρ, as defined in the previous section. Dynamic equations for the moments of ρ can be obtained directly from equation (A.12) by applying the so-called *method of moments* (Helbing, 1997). With this method dynamic equations for the κ^{th} order moment of ρ for state variable s_i are obtained by multiplication of both RHS and LHS of equation (A.12) by s_i^{κ}, and consequent integration over all state variables in S.

In general, a dynamic equation for $M_{s_i}^{\kappa}$ is:

$$\int_S s_i^{\kappa}\left(\frac{\partial\rho}{\partial t}+\nabla_{S'}\cdot\left(\rho\frac{dS'}{dt}\right)\right)dS=\int_S s_i^{\kappa}\left(\frac{d\rho}{dt}\right)_{discrete}dS \tag{A.14}$$

A.3.2.2 A general dynamic macroscopic density, speed momentum and speed equation

Restricting the definition of the vehicle state to $S = v$, we can use the general equation (A.14) to establish dynamic equations for the density $k = M_v^0$ and for the speed momentum $Q = M_v^1$.

Let us first establish the dynamic density equation. Equation (A.14) is in this case rewritten as:

$$\int_v \underbrace{\frac{\partial\rho}{\partial t}}_{A}+\underbrace{\frac{\partial}{\partial x}\left(\rho\frac{dx}{dt}\right)}_{B}+\underbrace{\frac{\partial}{\partial v}\left(\rho\frac{dv}{dt}\right)}_{C}\;dv=\int_v\left(\frac{d\rho}{dt}\right)_{discrete}dv \tag{A.15}$$

The terms in the LHS can be elaborated as follows. In terms A and B, as well as in the RHS, the order of integration and differentiation can be altered, since they relate to independent variables (for which relation (A.1) holds), yielding:

$$\int_v\frac{\partial\rho}{\partial t}\,dv=\frac{\partial}{\partial t}\int_v\rho\,dv=\frac{\partial k}{\partial t} \tag{A.16}$$

$$\int_v\frac{\partial}{\partial x}\left(\rho\cdot\frac{dx}{dt}\right)dv=\frac{\partial}{\partial x}\int_v\rho v\,dv=\frac{\partial kV}{\partial x} \tag{A.17}$$

$$\int_v\left(\frac{d\rho}{dt}\right)_{discrete}dv=\left(\frac{dk}{dt}\right)_{discrete} \tag{A.18}$$

In term C, differentiation and integration over the same state variable cancel out, yielding for term C:

$$\int_v \frac{\partial}{\partial v}\left(\rho \cdot \frac{dv}{dt}\right) dv = \rho \frac{dv}{dt}\Big|_{v=-\infty}^{v=+\infty} = 0 \tag{A.19}$$

In this equation we need to evaluate $\lim_{v\to\pm\infty} \rho \frac{dv}{dt}$. In any case, the expected number of vehicles with $v = \pm\infty$ equals zero, since physically the speeds are limited to positive speeds lower than some maximum speed << $+\infty$. The individual acceleration is also limited (does not diverge to $\pm\infty$) due to physical constraints to the braking and acceleration installations of vehicles, so the limit and therefore term C equals zero.

Substituting these results in equation (A.15), yields:

$$\frac{\partial k}{\partial t} + \frac{\partial k V}{\partial x} = \left(\frac{dk}{dt}\right)_{discrete} \tag{A.20}$$

The latter equation is well-known as the *continuity equation*, expressing the *conservation of vehicles* in space and time. In absence of sources or sinks in the system (where vehicles would enter or leave the road that we are describing), the RHS equals 0 so that equation (A.20) reduces to:

$$\frac{\partial k}{\partial t} + \frac{\partial k V}{\partial x} = 0 \tag{A.21}$$

The dynamic equation for the speed moment M_v^1 is obtained by replacing s_i^K in equation (A.14) by v, yielding:

$$\int_v \underbrace{v\frac{\partial \rho}{\partial t}}_{D} + \underbrace{v\frac{\partial}{\partial x}\left(\rho \cdot \frac{dx}{dt}\right)}_{E} + \underbrace{v\frac{\partial}{\partial v}\left(\rho \cdot \frac{dv}{dt}\right)}_{F} dv = \int_v v\left(\frac{d\rho}{dt}\right)_{discrete} dv \tag{A.22}$$

In the elaboration of terms D, E, and F, we use similar considerations as for elaborating the terms in the density equation. In addition, we need partial integration, so that the terms become:

$$D = \int_v v\,\frac{\partial \rho}{\partial t} dv = \frac{\partial}{\partial t}\int_v \rho\, v\, dv = \frac{\partial k\, v}{\partial t} \tag{A.23}$$

$$\begin{aligned} E = \int_v v\,\frac{\partial}{\partial x}\left(\rho \frac{dx}{dt}\right) dv &= \frac{\partial}{\partial x}\int_v v^2 \rho\, dv \\ &= \frac{\partial k\,\left(V^2 + \Theta\right)}{\partial x} \end{aligned} \tag{A.24}$$

$$\begin{aligned} F = \int_v v\,\frac{\partial}{\partial v}\left(\rho\,\frac{dv}{dt}\right) dv &= v\,\rho \frac{dv}{dt}\Big|_{-\infty}^{+\infty} - \int_v \rho \frac{dv}{dt}\, dv \\ &= 0 - \int_v \rho \frac{dv}{dt}\, dv \\ &\overset{def}{=} -k\left\langle \frac{dv}{dt} \right\rangle_v \end{aligned} \tag{A.25}$$

Here we have implicitly defined[16] the expected acceleration $\left\langle \frac{dv}{dt} \right\rangle_v$ over all individual speeds v.

Substituting these results in equation (A.22), results in the *dynamic speed momentum equation*:

$$\frac{\partial kV}{\partial t} + \frac{\partial k\left(V^2 + \Theta\right)}{\partial x} = k\left\langle \frac{dv}{dt} \right\rangle_v + \int_v v\left(\frac{d\rho}{dt}\right)_{discrete} dv \qquad \text{(A.26)}$$

Since equation (A.26) gives the dynamics of the speed *momentum* $Q = k\,V$, it is also called the *conservative* speed equation. We can cast this into a *primitive* speed equation using the chain rule and equation (A.20):

$$\frac{\partial V}{\partial t} + \underbrace{V\frac{\partial V}{\partial x}}_{convection} + \underbrace{\frac{1}{k}\frac{\partial k\Theta}{\partial x}}_{pressure} = \underbrace{\left\langle \frac{dv}{dt} \right\rangle_v}_{smooth\ \ acceleration} + \frac{1}{k}\left(\underbrace{\int_v v\left(\frac{d\rho}{dt}\right)_{discrete} dv}_{discrete\ \ acceleration\ 1} - \underbrace{V\left(\frac{dk}{dt}\right)_{discrete}}_{discrete\ \ acceleration\ 2} \right) \qquad \text{(A.27)}$$

The terms in this primitive speed equation all have a distinct physical interpretation:

- convection term: change of the average speed V due to a spatial speed gradient that is carried with the speed of the flow V
- pressure term: change of the average speed V due to individual vehicles that travel with speed $v < V$ or $v > V$; the former remain 'longer' at location x whereas the latter leave x 'earlier'
- smooth acceleration: change of the average speed V due to the gross effect of smooth individual accelerations
- discrete acceleration 1: change of the average speed V due to events causing a discrete change of the expected number of vehicles with speed v
- discrete acceleration 2: change of the average speed V due to a discrete change of the total number of vehicles k; the same total speed momentum kV has to be redistributed over a suddenly changed number of vehicles.

Note that when there is no discrete in- or outflow of traffic at t and x, $\left(\frac{dk}{dt}\right)_{discrete} = 0$.

Then, the first discrete acceleration term can only contain discrete transitions between states with different speeds v. In that case, this term is also referred to as the *interaction term*.

[16] We use the notation $\langle y(x) \rangle_x$ to denote the average of function $y(x)$ averaged over all x

A.3.2.3 Relation of the general macroscopic equations to traditional macroscopic equations

Let us further examine equation (A.27). Traditionally, the smooth acceleration term is used to account for the tendency of drivers to accelerate to their desired speeds w (Prigogine (1961), Paveri-Fontana (1975), Helbing (1997), Hoogendoorn & Bovy (2000)). For individual drivers, an exponential relaxation to the desired speed w is assumed in these models:

$$\frac{dv}{dt} = \frac{w - v}{\tau_w} \tag{A.28}$$

With this definition and assuming for simplicity that the relaxation time τ_w is the same for all drivers, the average (smooth) acceleration becomes:

$$\begin{aligned} \left\langle \frac{dv}{dt} \right\rangle_v &= \int_v \frac{\rho}{k} \frac{w - v}{\tau_w} dv \\ &= \frac{W - V}{\tau_w} \end{aligned} \tag{A.29}$$

In this equation we have used the definition of the average desired speed W:

$$W = \frac{1}{k} \int_v w \rho \, dv \tag{A.30}$$

Assuming no discrete in- or outflow of traffic, now equation (A.27) reduces to:

$$\frac{\partial V}{\partial t} + V \frac{\partial V}{\partial x} = \frac{W + \frac{\tau_w}{k} \int_v v \left(\frac{d\rho}{dt} \right)_{discrete} dv - V}{\tau_w} - \frac{1}{k} \frac{\partial k\Theta}{\partial x} \tag{A.31}$$

This equation has a structure very similar to that of the well-known model of Payne (1971) that reads:

$$\frac{\partial V}{\partial t} + V \frac{\partial V}{\partial x} = \frac{V^e(k) - V}{\tau_w} - \frac{c_0^2}{k} \frac{\partial k}{\partial x} \tag{A.32}$$

This means that in the kinetic interpretation of the model of Payne, the equilibrium velocity $V^e(k)$ and anticipation coefficient c_0 can be identified as:

$$V^e(k) = W + \frac{\tau_w}{k} \int_v v \left(\frac{d\rho}{dt} \right)_{discrete} dv \tag{A.33}$$

$$c_0 = \sqrt{\Theta} \quad with \quad \Theta = constant \tag{A.34}$$

However, we will show in section A.5.2.1 in Annex A, that the identification of the *pressure* term in a kinetic model with the *anticipation* term in the model of Payne is *not* justified, so equation (A.34) does *not* hold.

A.3.2.4 Closing the macroscopic system

A closer look at equations (A.20) and (A.27) reveals that in the 0^{th} order moment equation for the density k, the 1^{st} order moment kV (or V) is present. In the equation for the 1^{st} order moment, the 2^{nd} order moment Θ is present. It can be shown (Helbing, 1997) that in general, the dynamic equation for the κ^{th} order moment contains the $(\kappa+1)^{th}$ order moment, so that in theory we need an infinite number of dynamic moment equations to build a system equivalent to the generic equation (A.12).

However, it is customary in traffic flow modelling to close the system of macroscopic equations by assuming a closed expression for moments higher than κ^*. Of course, all models obtained using such an approximation neglect more refined system behaviour described by the higher moments.

A 1^{st} order macroscopic traffic flow model has $\kappa^*= 0$. Such a model uses the dynamic density equation and assumes a closed expression for V:

$$V = V^e(k) \tag{A.35}$$

The best-known macroscopic traffic flow model of Lighthill, Whitham and Richards (LWR) is a 1^{st} order model.

A 2^{nd} order macroscopic traffic flow model has $\kappa^*= 1$. It has dynamic equations for the density and for the speed (or speed momentum), after which the system is closed by assuming the speed variance equal to some *equilibrium speed variance*:

$$\Theta = \Theta^e(k) \tag{A.36}$$

Note that the so-called *equilibrium speed* $V^e(k)$ also plays a role in 2^{nd} order macroscopic models. Its purpose here however, is not to close the system of equations but as a stationary solution to which the speed would converge for $t \to \infty$ if all other terms in the dynamic speed equation would be zero. Therefore, the term containing the equilibrium speed in 2^{nd} order models is called the *relaxation* term. There are numerous examples of second order models (see section 3.3). Some of them have structures that differ from that of equation (A.27), for instance by inclusion of an extra diffusion term (2^{nd} order derivative of the density with respect to x).

Finally, some authors have proposed 3^{rd} order macroscopic traffic flow models, having $\kappa^*= 2$ (Helbing, 1997; Hoogendoorn & Bovy, 2000),. The system consisting of a density, speed (or speed momentum), and speed variance (or speed energy) is closed by assuming a closed expression for the 3^{rd} order moment or *skewness* C of the generalised density function:

$$C = C^e(k) \tag{A.37}$$

Within the context of this thesis, we establish a 2^{nd} order macroscopic model. This means that we close the system after the specification of the speed (momentum) equation, by choosing a closed analytical expression for the equilibrium speed variance.

A.4 Specification of individual driver behaviour in traditional (gas-) kinetic traffic flow models

The second order macroscopic traffic flow model described by equations (A.20) and (A.27) of section A.3.2.2 contains no specification of the behaviour of drivers and vehicles so far, other than the assumption that the state of an individual vehicle-driver unit is fully determined by the position and speed of the vehicle only.

In this section we briefly indicate how in so-called gas-kinetic models, driver behaviour can be specified. Numerous authors have successfully established gas-kinetic traffic flow models (also called *mesoscopic* traffic flow models), and – via the method of moments (section A.3.2.1) – accordingly macroscopic models (section 3.4). It is not the purpose of this section to describe the details of existing specifications. Rather, we want to show the principle of using *interactions between vehicles* as a common basis of these models.

Let us recall equation (A.27) in absence of any traffic flowing in or out of the road under consideration. In that case $\left(\frac{dk}{dt}\right)_{discrete} = 0$, and the equation reads:

$$\frac{\partial V}{\partial t} + V\frac{\partial V}{\partial x} + \frac{1}{k}\frac{\partial k\Theta}{\partial x} = \left\langle \frac{dv}{dt} \right\rangle_v + \frac{1}{k}\int_v v\left(\frac{d\rho}{dt}\right)_{discrete} dv \qquad \text{(A.38)}$$

In the kinetic theory, all terms on the LHS relate to convective processes, that is: changes caused by the fact that the vehicles are flowing along the longitudinal x-axis, so due to physical transport. This leaves two terms in the RHS where the behaviour of vehicles and drivers can influence the dynamics of the system.

Traditionally, the first term is used to specify acceleration of drivers towards their desired speed and is therefore called *relaxation* term. Decelerations of faster traffic trying to avoid collisions with slower traffic are accounted for in the second term at the RHS of equation (A.38). This term is therefore referred to as the *interaction* term.

A typical specification of the relaxation term was already given in the previous section (equations (A.28) and (A.29)). The interaction term is subject to various specifications. However, since it models instantaneous transitions of vehicles from an initial state S_0 to another state S_1, the interaction term is usually split into two terms: a negative contribution for $\rho(t,x,S_0)$, because vehicles leave this state, and a corresponding contribution for $\rho(t,x,S_1)$, the new state after the interaction. These contributions are also referred to as *active* and *passive* interactions respectively. The active and passive contributions are always balanced, since a transition from S_0 to S_1 means that the expected number of vehicles with the initial state $\rho(t,x,S_0)$ decreases (the active contribution), and the expected number of vehicles with the new state $\rho(t,x,S_1)$ increases by the same amount (passive contribution).

When using the concept of interactions to model decelerations, interactions between fast and slow vehicles are considered. The assumption here is that fast vehicles that find a slower vehicle on their way (and have no opportunity to overtake) change their speed

immediately to the speed of the slower predecessor (or to a lower speed in Helbing (1997)).

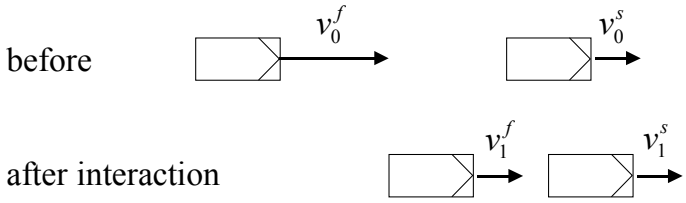


Figure A - 1 Definition of the speeds of the fast (f) and slow (s) vehicles before and after the interaction

Denoting the initial speed of the fast vehicle as v_0^f, its speed after the interaction v_1^f, and accordingly that of the slow vehicle v_0^s before, and v_1^s after the interaction (Figure A - 1), the active change rate of the generalised density $\rho(t,x,v_0^f)$ is generally defined:

$$\left(\frac{d\rho\left(t,x,v_0^f\right)}{dt}\right)^{-}_{interaction} = \int_{v_0^s<v_0^f}\int_{v_1^f\le v_0^s} \rho\left(t,x,v_0^f\right)\Pi\left(t,x,v_0^f,v_0^s\right)\pi_v\left(t,x,v_0^f,v_1^f,v_0^s\right)dv_1^f dv_0^s \quad \text{(A.39)}$$

The symbols used in this equation are:

$\rho\left(t,x,v_0^f\right)$: the generalised density, representing the expected number of vehicles having speed v_0^f

$\Pi\left(t,x,v_0^f,v_0^s\right)$: the interaction rate or number of interactions per unit of time per fast vehicle having speed v_0^f, with a slower vehicle having speed $v_0^s < v_0^f$

$\pi_v\left(t,x,v_0^f,v_1^f,v_0^s\right)$: the transition rate or number of transitions per unit of time of fast vehicles having initial speed v_0^f to speed $v_1^f \le v_0^s$, due to an interaction with a slower vehicle with speed $v_0^s < v_0^f$

Equation (A.39) states that the active change rate or expected number of vehicles leaving state v_0^f due to interactions is proportional to the original number of vehicles in the initial state before the interactions, to the interaction rate with slower traffic having any speed v_0^s lower than the own speed, and to the transition rate to any post-interaction speed v_1^f slower than the initial speed.

Analogously, one can see that the passive change rate equals:

$$\left(\frac{d\rho\left(t,x,v_1^f\right)}{dt}\right)^{+}_{interaction} = \int_{v_0^f\ge v_1^f}\int_{v_0^s<v_0^f} \rho\left(t,x,v_0^f\right)\Pi\left(t,x,v_0^f,v_0^s\right)\pi_v\left(t,x,v_0^f,v_1^f,v_0^s\right)dv_0^s dv_0^f \quad \text{(A.40)}$$

All gas-kinetic models differ in their specification of the interaction rates and/or transition rates. Specific characteristics of the drivers can be incorporated in the specifications for these rates, for instance the finite space requirement, the non-locality of interactions, the possible overreaction of drivers who brake, different distance thresholds between vehicles below which braking transitions occur, etcetera. We refer to the different models and references discussed in section 3.4 for further details.

The approach of the human-kinetic traffic flow model developed in this thesis differs fundamentally from existing kinetic traffic flow models, in that we do not model the deceleration as the interaction process described in this section. Instead, we model acceleration and deceleration both as continuous or smooth processes. However, when modelling transitions from one activation level to another – an instantaneous change – we use the formalism of active and passive change rates, analogously to equations (A.39) and (A.40) in this section.

A.5 Numerical evaluation

A.5.1 Mathematical structure of macroscopic traffic flow models based on the kinetic theory and numerical solution scheme

The second order system of macroscopic partial differential equations, derived from the kinetic traffic flow theory, generally reads:

$$\frac{\partial k}{\partial t}+\frac{\partial k V}{\partial x}=\left(\frac{dk}{dt}\right)_{discrete} \tag{A.41}$$

$$\frac{\partial V}{\partial t}+V\frac{\partial V}{\partial x}+\frac{1}{k}\frac{\partial k\Theta^{e}}{\partial x}=\left\langle\frac{dv}{dt}\right\rangle_{v}+\frac{1}{k}\left(\int_{v} v\left(\frac{d\rho}{dt}\right)_{discrete} dv-V\left(\frac{dk}{dt}\right)_{discrete}\right) \tag{A.42}$$

Or equivalently, the speed equation (*primitive* formulation) can be replaced by the speed momentum equation (*conservative* formulation):

$$\frac{\partial kV}{\partial t}+\frac{\partial k\left(V^{2}+\Theta^{e}\right)}{\partial x}=k\left\langle\frac{dv}{dt}\right\rangle_{v}+\int_{v} v\left(\frac{d\rho}{dt}\right)_{discrete} dv \tag{A.43}$$

Due to the similar structure of these equations with the model of Payne (see section A.3.2.3), we call the general class of traffic flow models defined by equations (A.41) and (A.42) (or (A.43)) *Payne-type models*. In general, Payne-type models are not analytically solvable. Therefore, numerical solution techniques are used for simulations of second order macroscopic traffic flow models. Numerous numerical schemes for systems of partial differential equations exist, the selection of which depends on the nature and mathematical structure of the system of equations (Hirsch, 1990). Looking at the structure of the conservative model formulation of the Payne-type models, one can rewrite it as:

$$\frac{\partial}{\partial t}\mathbf{U}+\frac{\partial}{\partial x}\mathbf{F}(\mathbf{U})=\mathbf{G}(\mathbf{U}) \tag{A.44}$$

with

$$\mathbf{U} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} k \\ kV \end{pmatrix} \tag{A.45}$$

$$\mathbf{F}(\mathbf{U}) = \begin{pmatrix} u_2 \\ \dfrac{u_2^2}{u_1} + u_1\,\Theta^e(u_1, u_2) \end{pmatrix} \tag{A.46}$$

$$\mathbf{G}(\mathbf{U}) = \begin{pmatrix} \left(\dfrac{du_1}{dt}\right)_{discrete} \\ u_1\left\langle \dfrac{dv}{dt} \right\rangle_v + \int_v v \left(\dfrac{d\rho}{dt}\right)_{discrete} dv \end{pmatrix} \tag{A.47}$$

To determine the nature of the partial differential equation, it is cast into the quasi-linear form:

$$\frac{\partial}{\partial t}\mathbf{U} + \frac{\partial \mathbf{F}}{\partial \mathbf{U}}\frac{\partial \mathbf{U}}{\partial x} = \mathbf{G}(\mathbf{U}) \tag{A.48}$$

The eigenvalues of the *Jacobian matrix* $\dfrac{\partial \mathbf{F}}{\partial \mathbf{U}}$ determine whether the partial differential equation (A.44) is elliptic (imaginary eigenvalues), parabolic (one eigenvalue is zero), or hyperbolic (real eigenvalues). With the above definitions for $\mathbf{F}$ and $\mathbf{U}$, the Jacobian matrix is calculated as:

$$\mathbf{J}(\mathbf{U}) = \frac{\partial \mathbf{F}}{\partial \mathbf{U}} = \begin{pmatrix} 0 & 1 \\ -\dfrac{u_2^2}{u_1^2} + \Theta^e + u_1\dfrac{\partial \Theta^e}{\partial u_1} & 2\dfrac{u_2}{u_1} + u_1\dfrac{\partial \Theta^e}{\partial u_2} \end{pmatrix} \tag{A.49}$$

A straightforward calculation of the eigenvalues of this matrix yields:

$$\lambda_{1,2} = \frac{u_2}{u_1} + \frac{u_1}{2}\frac{\partial \Theta^e}{\partial u_2} \pm \sqrt{\frac{u_1^2}{4}\left(\frac{\partial \Theta^e}{\partial u_2}\right)^2 + u_2\frac{\partial \Theta^e}{\partial u_2} + \Theta^e + u_1\frac{\partial \Theta^e}{\partial u_1}} \tag{A.50}$$

For the model of Payne for instance, we replace Θ^e by c_0^2 so that we find the well-known result:

$$\begin{aligned} \lambda_{1,2} &= \frac{u_2}{u_1} \pm c_0 \\ &= V \pm c_0 \end{aligned} . \tag{A.51}$$

For the Payne model, one eigenvalue is always positive, while the sign of the other eigenvalue depends on whether V is smaller or larger than c_0. We show in the next section that this property plays an important role in the numerical solution scheme for the Payne model.

For the selection of a numerical scheme for the Payne-type models, we refer to Ngoduy et al (2004). The authors present some common schemes found in literature for the

solution of hyperbolic Payne-type traffic flow models, and introduce a Godunov-type solver, called the extended HLLE scheme (after Harten-van Leer, Lax and Einfeldt). Within this thesis, we use this extended HLLE scheme as the basis for our numerical calculations.

A.5.2 The role of the pressure term or speed variance

The pressure term in equation (A.42) or (A.43) plays a very important role. Therefore, it has been the origin of ongoing debate. Even the validity of second- or higher-order models has been questioned, partially due to the behaviour of the traffic pressure in the model. In this section we therefore investigate the nature of the pressure term and its consequences for the numerical evaluation of the model. In this analysis we consider the model of Payne, the Payne-type models defined by equations (A.41) and (A.42) and some existing specific Payne-type kinetic traffic flow models.

The key issue in the remainder of this section is to show:

- that even within the class of Payne-type models, the physical meaning of the pressure term (the existence of which is common to all models belonging to this class) differs fundamentally (section A.5.2.1);
- that the physical meaning of the pressure term determines whether the linearised model (A.48) contains only positive real eigenvalues, or one positive and one negative eigenvalue (also section A.5.2.1);
- that the correctness of the numerical solutions of Payne-type models using Godunov-type numerical solution schemes (like the HLLE solver used in this thesis) depends on the potential occurrence of a negative eigenvalue (section A.5.2.2);
- how existing kinetic traffic flow models of the Payne-type are mathematically manipulated to ensure the occurrence of negative eigenvalues, and hence correct numerical solutions using existing Godunov-type schemes (section A.5.2.3).

A.5.2.1 Physical interpretation of the pressure term in different traffic flow models: "speed variance" pressure, "visual" pressure and "non-local" pressure

We have shown in section A.3.2.3 the striking similarity between the mathematical structure of the model of Payne and the Payne-type macroscopic models derived from the kinetic conservation law (A.12). The question was raised whether the identification of the pressure terms in both models was justified (equation (A.34)). Identification of the terms would mean that the model of Payne assumes an equilibrium variance Θ^e equal to the constant and strictly positive value c_0^2, regardless of the density or average speed. We argue here that this identification is *not* justified: the nature of the pressure terms in both model types is fundamentally different. Moreover, we argue in this section that in kinetic models also a third type of traffic pressure exists.

We know from the mathematical analysis in section A.5.1 that the eigenvalues are given by equation (A.50). For the sake of simplicity of the argument, we approximate this equation by:

$$\lambda_{1,2} \approx V \pm \sqrt{\Theta^e} \tag{A.52}$$

For the model of Payne, the exact eigenvalues are:

$$\lambda_{1,2} = V \pm c_0 \tag{A.53}$$

Mathematically, the eigenvalues are interpreted as the speed of information through the flow. With equation (A.52) this interpretation is obvious: the eigenvalues are approximately equal to the speed of the faster (average speed plus standard deviation) and slower (average speed minus standard deviation) vehicles in traffic. We see that here information is only carried with the speed of the vehicles, so that we can speak of *convective* information flow; the different eigenvalues emerge due to the fact that individual vehicles have different speeds. Therefore, no negative eigenvalues occur according to (A.52): the standard deviation of the speeds is always lower than the average itself, so that no negative vehicle speeds are considered (in symmetrical speed distributions); as a consequence of information being only carried with the speed of vehicles, no negatively directed convective information flow is considered with (A.52).

In the model of Payne, the interpretation of the eigenvalues as the minimal and maximal speed of *information* flow through the vehicular flow is still valid, but the identification of the eigenvalues with higher and lower *physical* speeds of vehicles does *not* hold: the convective flow and information flow are not longer coupled. In the derivation of his model, Payne (1971) did not consider vehicular speeds other than the average speed. The coefficient c_0 stands for the anticipation of drivers, i.e. for the speed of the *visual* information on which drivers base their speed choice. A negative eigenvalue occurs when *visual* flow (directed against the driving direction) is faster than the speed of the vehicles. In that sense the anisotropic behaviour (reaction of drivers only to traffic conditions downstream) is obeyed. A problem occurs only with the highest eigenvalue, which stands for information flow faster than the speed of the vehicles, as if drivers downstream would be affected by visual information from upstream. This is behaviourally incorrect, and the model of Payne has been criticised (Daganzo, 1995a) and refined for this reason (e.g. Zhang, 2000; Aw & Rascle, 2000).

We conclude by saying that the origin of the pressure term can be purely *convective*, due to spread of the individual vehicle speed, *or* that the pressure term can be a pure representation of *visual* information flow. We will distinguish these physically complete disjoint processes by referring to "*speed variance*" pressure and "*visual*" pressure respectively. *Speed variance* pressure leads to positive eigenvalues only; *visual* traffic pressure leads to one positive eigenvalue, while the other can be positive or negative, dependent on the flow type (free-flowing or congested). The latter implicitly defines congestion as the flow condition in which the vehicular speed is lower than the negatively directed visual information flow.

We will show in section A.5.2.3 that the nature of the pressure term in existing (gas-) kinetic traffic flow models by Helbing (1997), Hoogendoorn (1999), Klar & Wegener

(1999) and Treiber et al. (1999) is neither purely *speed variance* traffic pressure, nor purely *visual* traffic pressure. We argue that the physical origin of the terms in the respective macroscopic speed equations is due to non-local interactions (possibly combined with speed variance); therefore, we refer to this type of traffic pressure as *"non-local"* pressure.

A.5.2.2 Numerical evaluation of Payne-type traffic flow models and the role of the eigenvalues

A variety of explicit numerical solution schemes for Payne-type traffic flow models exist (Ngoduy et al. 2004). Consider a set of (n-1) discrete and equidistant x-points x_j in the theoretically continuous x-domain [0,L]:

$$\begin{aligned} x_j &= \tfrac{\Delta x}{2} + j\Delta x \\ j &= 1...(n-1) \\ \Delta x &= \tfrac{L}{n} \end{aligned} \tag{A.54}$$

These x_j-points are the centres of cells with boundaries:

$$\begin{aligned} x_{j-\frac{1}{2}} &= x_j - \tfrac{\Delta x}{2} \\ x_{j+\frac{1}{2}} &= x_j + \tfrac{\Delta x}{2} \end{aligned} \tag{A.55}$$

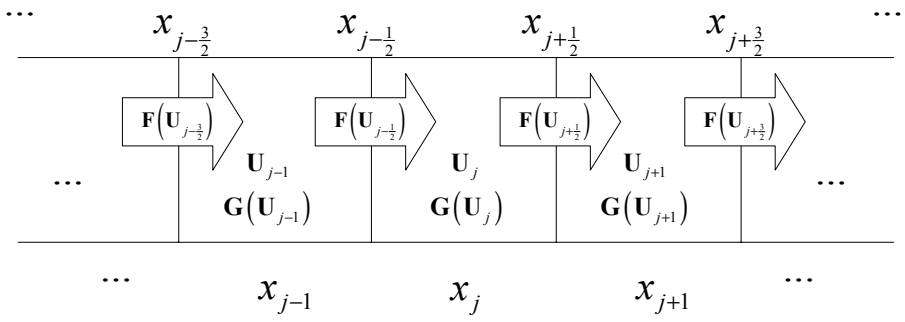


Figure A - 2 Definition of cell and interface indices, numerical fluxes F and source terms G

Then, in every discrete time step the change of the macroscopic variables $\mathbf{U}_j$ in the cell centres is calculated according to the following procedure (Figure A - 2):

1. Dependent on the conditions, $\mathbf{U}_{j-1}, \mathbf{U}_j$, and $\mathbf{U}_{j+1}$ in the cell x_j under consideration and in the neighbouring cells x_{j-1} and x_{j+1} determine the numerical fluxes over the boundaries $\mathbf{F}\left(\mathbf{U}_{j-\frac{1}{2}}\right)$ and $\mathbf{F}\left(\mathbf{U}_{j+\frac{1}{2}}\right)$
2. The change of $\mathbf{U}_j$ during a finite time interval Δt due to numerical fluxes is then:

$$\mathbf{U}_j^{'} = \mathbf{U}_j - \Delta t \frac{\mathbf{F}\left(\mathbf{U}_{j+\frac{1}{2}}\right) - \mathbf{F}\left(\mathbf{U}_{j-\frac{1}{2}}\right)}{\Delta x} \tag{A.56}$$

3. The change of $\mathbf{U}_j$ during the same time interval Δt due to internal and external forces contained in the source term $\mathbf{G}\left(\mathbf{U}_j\right)$ is superimposed, so that the macroscopic state variables $\mathbf{U}_j^{''}$ after the interval Δt are found by:

$$\mathbf{U}_j^{''} = \mathbf{U}_j^{'} + \Delta t\, \mathbf{G}\left(\mathbf{U}_j\right) \tag{A.57}$$

Determinant in this type of solution scheme is (a) the order in which the update steps are evaluated, and (b) the definition of the numerical fluxes over the cell boundaries. The latter problem is not trivial, since for its calculation we have only the value of $\mathbf{U}$ in the cell centres at our disposal. It is here that Godunov-type numerical schemes use the eigenvalues as the speeds with which the state information in the cell centres surrounding a cell boundary travel to that boundary; hence, the importance of the eigenvalues – and therefore of the pressure term definition. This means that a third determinant of a Godunov-type solution scheme is: (c) the definition of the eigenvalues on the cell boundaries.

Positive and negative eigenvalues

Given this characteristic of the Godunov-type numerical solution schemes, the consequence of having positive and negative eigenvalues in the model is the following. In such models, fluxes over the cell boundary $j+\frac{1}{2}$ can be determined by forwardly directed information flow, i.e. by the state in the upstream cell (eigenvalue > 0) or by backwardly directed information flow, so by the state in the downstream cell (eigenvalue < 0).

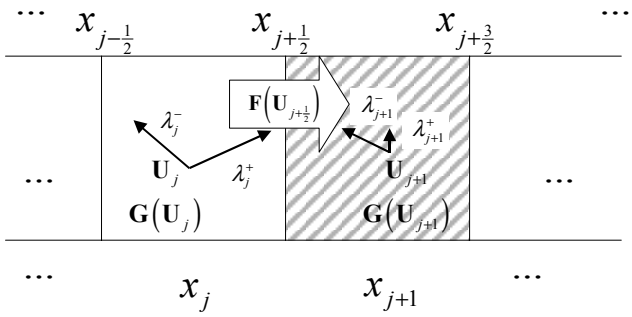


Figure A - 3 Positive and negative eigenvalue influencing the numerical flux on the interface

This is the mechanism in such models that prevents jammed cells from being further filled. Suppose that we have a jammed cell *j+1*, preceded by a not-yet jammed cell *j* (Figure A - 3). The flow within the jammed cell is by definition zero, since the speed is zero. In this case the boundary $j+\frac{1}{2}$ between these cells has one negative and one positive eigenvalue, so that information reaches this interface from two sides: there is a

demand from upstream to cross the cell interface, but information from downstream says that *supply* in the target cell is zero. Hence, numerical flux across this interface stops, and the jammed cell is not further filled.

Positive eigenvalues only

The consequence of having only positive eigenvalues would be that the flux across the interface is exclusively determined by the traffic state upstream of this interface. Traffic there is still flowing, so the flux across the interface will *not* be forced to zero, and the jammed cell is further filled until the flow *inside* the upstream cell stops. This may be possible, if the RHS source term $\mathbf{G}\left(\mathbf{U}_j\right)$ accounts for anticipation by drivers in the not-yet jammed cell *j*, and is responsible for a braking response inside this cell. In a well-designed model combined with a well-designed numerical solution for this model, this will cause speed – and hence flow – in the cell upstream of a jammed cell to approach zero, and therefore to prevent the jammed cell from being further filled.

However, it appears from the example in section 4.4.3.1 of the main text, that existing Godunov-type numerical schemes for Payne-type models do *not* always calculate correct numerical solutions for models with only positive eigenvalues, especially in the limit of dense (nearly jammed) traffic. We show in the next section how in existing kinetic traffic flow models, the pressure term has been adapted to circumvent this issue.

A.5.2.3 Pressure term, eigenvalues and numerical evaluation of existing (gas-) kinetic traffic flow models

Kinetic traffic flow models are in essence all derived from the dynamic equation for the generalised density (A.12), so that the macroscopic speed equation in these models can always be cast into the form of equation (A.27). Every (gas-) kinetic traffic flow model therefore contains at least a pressure term of the *speed variance* type. We have shown in section A.5.2.1 that such model therefore has two positive eigenvalues and is susceptible to the numerical discretisation errors near jam density discussed in section 4.4.3.1 of the main text. We show in this section that modern kinetic models circumvent this potential problem by isolating all non-local aspects from the acceleration terms in the RHS of (A.27), and rewrite these as an *additional contribution to the pressure term*. With this adaptation, the modified pressure term – which we refer to as *non-local* traffic pressure – yields both positive and negative eigenvalues, so that numerical solution schemes for Payne-type models can be used.

Klar & Wegener (1999) derive a kinetic traffic flow model based on discrete interactions between vehicle pairs of followers and leaders. The rate of braking manoeuvre interactions is proportional to the pairwise distribution of follower-leader vehicle pairs. The specification of this joint distribution accounts for the fact that for a potentially braking follower at location *x*, a leader *at some finite distance in front* of *x* has to be considered. This so-called *finite space correction* leads to an increased rate of interactions at higher densities (compared to dimensionless particles). This increased interaction rate at location *x* is dependent on the non-local conditions some finite distance downstream. The model specification by Klar & Wegener contains explicit non-local braking and acceleration thresholds. The authors show that the non-localities

introduced in this way can mathematically be isolated in what they call the "integrated Engskog" coefficient $A^e(k)$, which leads to the following macroscopic model structure:

$$\frac{\partial kV}{\partial t}+\frac{\partial k\left(V^2+\Theta^e(k)\right)}{\partial x}+\frac{\partial A^e(k)}{\partial x}=\int_v v\left(\frac{d\rho}{dt}\right)_{discrete} dv \tag{A.58}$$

The Engskog coefficient $A^e(k)$ is strictly increasing with density, and is dominant over the *speed variance* pressure $k\,\Theta^e(k)$: numerical values presented by the authors show that – dependent on the density – $A^e(k)$ is 10 to 100 times larger than $\Theta^e(k)$. Mathematically, the Engskog coefficient acts as a contribution added to the speed variance, so that the following approximation for the eigenvalues holds:

$$\lambda_{1,2}\approx V\pm\sqrt{\Theta^e+\frac{A^e}{k}} \tag{A.59}$$

We easily conclude that for $k\rightarrow k_{jam}$ (so that $\Theta^e\rightarrow 0$ and $V\rightarrow 0$) and A^e increasing with density, negative eigenvalues occur in the model. Actually, setting the speed variance equal to zero and linearising $A^e(k)\approx c_0\,k$, we obtain the LHS, and hence the eigenvalues of the Payne model.

Helbing (1997) uses similar arguments as Klar & Wegener while introducing a finite space correction in his model, but the mathematical elaboration differs. Helbing considers pairwise vehicle distributions with a speed and density dependent multiplication factor $\chi(k,V)$ for the increased interaction rates between follower/leader pairs at higher densities. Eventually this factor appears in the macroscopic speed equation:

$$\frac{\partial V}{\partial t}+V\frac{\partial V}{\partial x}+\frac{1}{k}\frac{\partial P}{\partial x}=\frac{V^e(k)-V}{\tau} \tag{A.60}$$

$$P(k,V)=\chi(k,V)\left(k\Theta-\eta_0\frac{\partial V}{\partial x}\right) \tag{A.61}$$

$$\chi(k,V)=\frac{1}{1-k\,s(V)} \tag{A.62}$$

$$s(V)=\frac{1}{k_{jam}}+V\,T_r \tag{A.63}$$

The divergence of the finite space correction factor for $k\rightarrow k_{jam}$ guarantees an infinite braking response of the system at jam density. Similarly to what we did for the model of Klar & Wegener, we can approximate the eigenvalues for the sake of argument to:

$$\lambda_{1,2}\approx V\pm\sqrt{\chi\,\Theta^e} \tag{A.64}$$

Again, we can expect negative eigenvalues occurring for high densities where the correction factor is strong. This approach by Helbing was later adopted in the models by Hoogendoorn (1999), Shvetsov & Helbing (1999), and Helbing & Treiber (1999) who

all have a similar *non-local* macroscopic pressure term. Let us remark that in Helbing & Treiber (1999) the correction factor is apparently absent, but that here the authors use a speed variance definition that sometimes yields negative speed variance. By definition of the speed variance this can physically not be correct, which inspires us to the interpretation that the authors here implicitly combine speed variance and a finite space correction in one expression for the pressure term.

We conclude that in modern (gas-) kinetic traffic flow models, non-local aspects in the specification of the interaction rate between vehicles are isolated from the acceleration terms, and treated together with the *speed variance* traffic pressure in the convective part of the model. We refer to this seemingly increased traffic pressure as *non-local* traffic pressure. Mathematically and numerically, the convective part of the model then behaves like a Payne model with a negative eigenvalue in the low speed regime, and the solution schemes for Payne-type models are applicable.

A.6 Linear stability analysis of higher order macroscopic traffic flow equations

In this section we outline the classic linear stability analysis for Payne-type models, and as an illustration apply it to the model of Payne. A similar analysis for the human-kinetic traffic flow model developed in this thesis is contained in section 5.2.1 of the main text.

A.6.1 Linear stability analysis procedure

The stability analysis outlined in this section is a *linear* stability analysis, since it uses linear Taylor approximations throughout the analysis. The consequence of these approximations is that conditions that are stable according to this analysis might actually still show non-linear instability. However, in general the linear analysis gives a good insight into the general behaviour of the model.

In this section we introduce the traditional linear stability analysis stepwise, and explain the rational behind each step. In the next section, we apply the procedure outlined in this section to the Payne-type models in general, and the model of Payne specifically. The reader who is familiar with linear stability analysis of systems, described by sets of partial differential equations, may skip this section and proceed with section A.6.2.

The outline and rational of the procedure is as follows:

- we first consider a stationary or equilibrium solution (k^*, V^*) of the system of partial differential equations;
- then, we consider initial conditions $(k^*+\delta k, V^*+\delta V)$ that are slightly perturbed with respect to this equilibrium solution, and examine how the perturbation evolves in time; the essence of the stability analysis is to find the conditions under which the amplitude of the perturbations δk and δV decreases, so that the equilibrium is gradually restored;

- the choice of the perturbed initial conditions is such, that the stability analysis is as general as possible. In general, any functional form of the perturbation δk and δV is possible. The Fourier theory however, states that any function can be written as an (infinite) sum (or integral in the case of non-periodic functions) of sine and cosine functions. If the model is stable with respect to any sine or cosine function, it is also stable with respect to any linear combination of these functions. Therefore, prove of stability with respect to sine or cosine perturbations also proves the general stability to any functional form of the perturbation;
- the sine and cosine perturbations are considered in the complex notation, i.e.: $\delta k = \alpha e^{\omega t + i\kappa x}$ and $\delta V = \beta e^{\omega t + i\kappa x}$; here ω and κ are complex numbers: $\omega, \kappa \in \mathbb{C}$;
- the remainder of the stability analysis concentrates on finding conditions so that two restrictions are met: (a) $(k^*+\delta k, V^*+\delta V)$ is a solution of the model, and (b) this solution is stable, i.e. the amplitude of the perturbation decreases in time; the latter is the case if the real part of ω is strictly negative;
- to prove (a), the definition of the perturbed solution is substituted in the model equations. The resulting equations contain derivatives of $k^*+\delta k$ and of $V^*+\delta V$, which we approximate using linear Taylor approximations around the equilibrium solution (k^*, V^*), and by neglecting all products of increments, hence the name *linear* stability analysis;
- we then obtain a linear system of the form $A\left(k^*, V^*, \omega, \kappa\right)\begin{pmatrix}\alpha \\ \beta\end{pmatrix} = 0$. This system has solutions other than the trivial solution $\alpha=\beta=0$, if and only if the determinant of matrix A is zero. The condition $\det\left(A\right) = 0$ is the necessary and sufficient requirement to the four unknowns k^*, V^*, ω, κ, so that $(k^*+\delta k, V^*+\delta V)$ be a solution of the model;
- to prove (b), we require that the real part of the values of ω that obey $\det\left(A\right) = 0$ is negative; we therefore rewrite the requirement $\det\left(A\right) = 0$ as an explicit requirement to ω (this means, in the explicit form: $\omega=F(k^*, V^*, \omega, \kappa)$). We say that the real part of F needs to be strictly negative: $\mathrm{Re}(F)<0$.
- After some algebra, $\mathrm{Re}(F)<0$ can be substantially simplified, and the stability condition is finally obtained.

In the next section, we apply this procedure to the system of partial differential equations describing the Payne-type models. We do this as general as possible, after which we illustrate the resulting general stability condition for Payne-type models by considering the Payne model itself.

A.6.2 Stability criterion for the model of Payne and Payne-type models

Consider the system of partial differential equations (A.41) and (A.32) without source terms, which we rewrite for notational convenience as:

$$\frac{\partial k}{\partial t}+\frac{\partial kV}{\partial x}=0 \tag{A.65}$$

$$\frac{\partial V}{\partial t}+V\frac{\partial V}{\partial x}+\frac{1}{k}\frac{\partial P}{\partial x}=f\left(k,V\right) \tag{A.66}$$

In equation (A.66), we have used the short-handed notations for the pressure P and the acceleration function f:

$$P=k\Theta^{e}\left(k,V\right) \tag{A.67}$$

$$f\left(k,V\right)=\left\langle\frac{dv}{dt}\right\rangle_{v}+\frac{1}{k}\int_{v}v\left(\frac{d\rho}{dt}\right)_{discrete}dv \tag{A.68}$$

Suppose that the system is in equilibrium, i.e. it has a stationary solution (k^{*},V^{*}). We investigate under which conditions the system restores this equilibrium when small deviations $(\delta k,\delta V)$ to the stationary solution are imposed. Following the general procedure outlined in the previous section, we fill in the values $(k^{*}+\delta k,V^{*}+\delta V)$ in (A.65) and (A.66). We use linear Taylor approximations and neglect all products of the small deviations δk or δV with derivatives with respect to x or t and obtain:

$$\frac{\partial \delta k}{\partial t}+k^{*}\frac{\partial \delta V}{\partial x}+V^{*}\frac{\partial \delta k}{\partial x}=0 \tag{A.69}$$

$$\frac{\partial \delta V}{\partial t}+V^{*}\frac{\partial \delta V}{\partial x}+\frac{1}{k^{*}}\left(\frac{\partial P}{\partial k}\frac{\partial \delta k}{\partial x}+\frac{\partial P}{\partial V}\frac{\partial \delta V}{\partial x}\right)-\frac{\partial f}{\partial k}\delta k-\frac{\partial f}{\partial V}\delta V=0 \tag{A.70}$$

Since we know from basic Fourier theory that any function can be written as an (infinite) sum of sine or cosine functions, it is sufficient to analyse deviations δk and δV of the (co)sine type only with all possible wave numbers and phase angles. We therefore consider the following deviations:

$$\delta k=\alpha e^{\omega t+i\kappa x} \tag{A.71}$$

$$\delta V=\beta e^{\omega t+i\kappa x} \tag{A.72}$$

The system is stable if and only if these deviations are solutions to the system of equations (A.69) and (A.70), *and* if the amplitude of the (co)sine functions is decreasing with time. The latter is true if and only if the real part $\mathrm{Re}(\omega)$ is strictly negative. For the former condition, we substitute the definitions for δk and δV in the linearised equations (A.69) and (A.70), which yields the following linear system:

$$A\left(k^{*},V^{*},\omega,\kappa\right)\begin{pmatrix}\alpha\\ \beta\end{pmatrix}=0 \tag{A.73}$$

with

$$A\left(k^*,V^*,\omega,\kappa\right)=\begin{pmatrix}\omega+i\kappa V^* & i\kappa k^* \\ \dfrac{i\kappa}{k^*}\dfrac{\partial P}{\partial k}-\dfrac{\partial f}{\partial k} & \omega+i\kappa V^*+\dfrac{i\kappa}{k^*}\dfrac{\partial P}{\partial V}-\dfrac{\partial f}{\partial V}\end{pmatrix} \tag{A.74}$$

This linear system has solutions other than the trivial solution $\alpha=\beta=0$, if and only if the determinant of matrix A is zero. Linear stability is therefore obtained if the following two conditions are fulfilled:

$$\mathrm{Re}\left(\omega\right)<0 \tag{A.75}$$

$$\det\left(A\left(k^*,V^*,\omega,\kappa\right)\right)=0 \tag{A.76}$$

After some lengthy algebraic calculations, we find that these conditions can be combined in the following condition:

$$\mathrm{Re}\left(\omega_{1,2}\right)=\mathrm{Re}\left(\frac{1}{2}\frac{\partial f}{\partial V}-i\left(\kappa V^*+\frac{\kappa}{2k^*}\frac{\partial P}{\partial V}\right)\pm\frac{1}{2}\sqrt{R+iI}\right)<0 \tag{A.77}$$

or

$$\frac{\partial f}{\partial V}\pm\mathrm{Re}\left(\sqrt{R+iI}\right)<0 \tag{A.78}$$

with

$$R=\left(\frac{\partial f}{\partial V}\right)^2-\frac{\kappa^2}{k^{*2}}\left(\frac{\partial P}{\partial V}\right)^2-4\kappa^2\frac{\partial P}{\partial k} \tag{A.79}$$

$$I=-\frac{2\kappa}{k^*}\frac{\partial P}{\partial V}\frac{\partial f}{\partial V}-4k^*\kappa\frac{\partial f}{\partial k} \tag{A.80}$$

Equation (A.78) actually sets two requirements, dependent on the plus- or minus-sign. However, one can easily verify that the requirement with the plus-sign is the more stringent one: if that requirement is true then the other (with the minus-sign) is trivially fulfilled. Using the fact that $\sqrt{R+iI}=\sqrt{\frac{1}{2}\left(\sqrt{R^2+I^2}+R\right)}+i\sqrt{\frac{1}{2}\left(\sqrt{R^2+I^2}-R\right)}$, equation (A.78) can be further simplified. Again, we omit the intermediate algebra, and finally obtain the general stability criterion for Payne-type models:

$$\left(\frac{\partial f}{\partial V}\right)^2\frac{\partial P}{\partial k}-k^{*2}\left(\frac{\partial f}{\partial k}\right)^2-\frac{\partial P}{\partial V}\frac{\partial f}{\partial V}\frac{\partial f}{\partial k}>0 \tag{A.81}$$

This result is generally true for any Payne-type model that can be cast in the shape of equations (A.65) and (A.66). Not surprisingly, the stability depends on the density k, on the pressure P and on the acceleration function f. The speed V does not explicitly occur in the stability criterion, but is indirectly contained through the presence of k and the fact that k and V were considered in stationary equilibrium at the start of the analysis, which actually couples V directly to k.

For now, let us conclude by illustrating the result of equation (A.81) specifically for the model of Payne. For that model, the pressure P and acceleration function f are defined as follows:

$$P = k\, c_0^2 \tag{A.82}$$

$$f\left(k,V\right) = \frac{V^e\left(k\right) - V}{\tau} \tag{A.83}$$

The stability criterion then becomes:

$$k < \frac{c_0}{\left|\frac{dV^e}{dk}\right|} \tag{A.84}$$

Since the anticipation coefficient in the model of Payne is equal to $c_0^2 = -\frac{1}{\tau}\frac{dV^e}{dk}$ (Payne, 1971), this criterion can be written in terms of the fundamental diagram and the relaxation time only:

$$k < \left(-\tau \frac{dV^e}{dk}\right)^{-\frac{1}{2}} \tag{A.85}$$

The general result (A.81) is used in the stability analysis of the human-kinetic traffic flow model in section 5.2.1 of the main text.

B SENSITIVITY ANALYSIS

In this annex we analyse the sensitivity of the basic human-kinetic model for changes to the parameter settings. In this sensitivity analysis, we consider three groups of parameters:

(i) parameters of the car-following model

(ii) parameters of the anticipation model

(iii) parameters of the numerical scheme

As a reference, the parameter settings of Table 5-2 are used. In each case in this annex, the sensitivity of the model solution for a change to only one of the parameters is examined (unless stated otherwise). Overall conclusions with respect to the sensitivity analysis are discussed in section 5.2.2.2 of the main text.

When interpreting the phase trajectories in this annex, the same conventions have been used as in all plots of this type throughout the thesis: the original state in the flow-density plane is marked with a square (□), the end state is marked with a star (☆); a decelerating phase trajectory is marked with circles (○), and acceleration with triangles (△).

B.1 Reference case

The reference case is discussed in section 5.2.2.1 of the main text. The simulation results are again depicted in Figure B - 1 to allow easy comparison with the figures in the remainder of this annex.

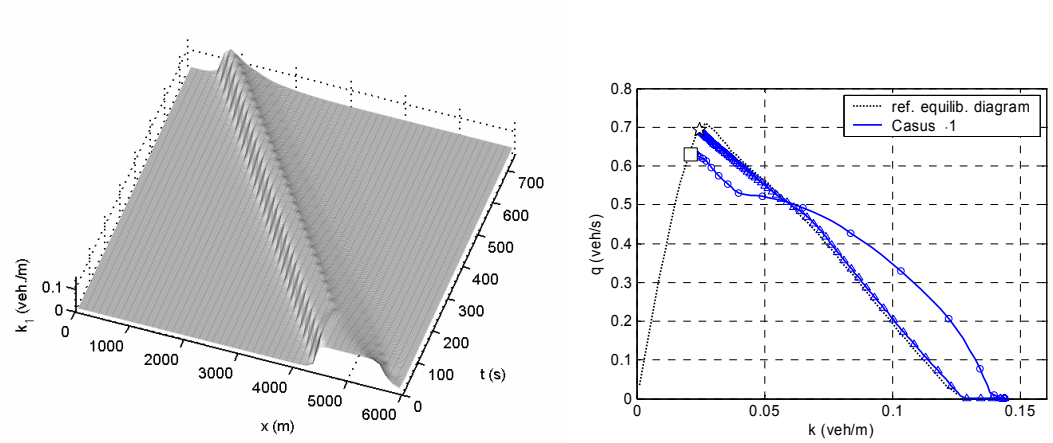


Figure B - 1 Case 1: Reference run with parameters from Table 5-2; (a)* left*: space-time evolution of the density, (b)* right*: phase trajectory in the flow-density plane

B.2 Sensitivity of jam propagation for car-following parameters

Case 2: reaction time T

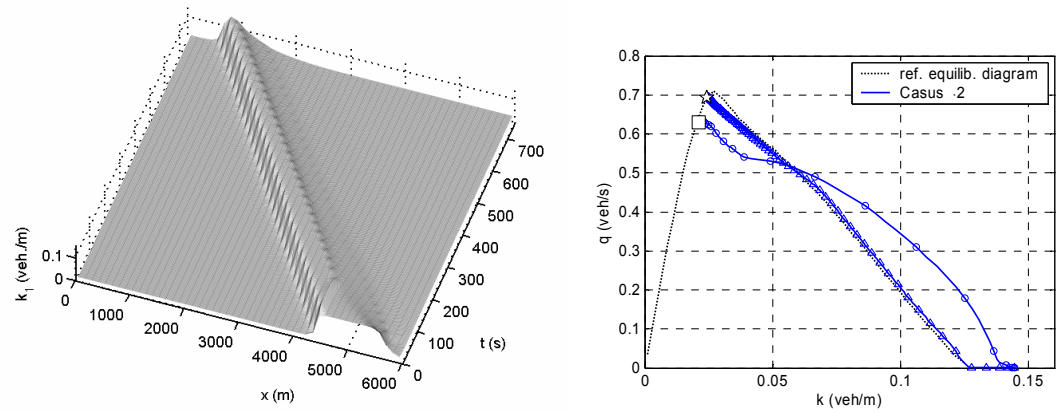


Figure B - 2 Case 2: increased reaction time; (a)* left*: space-time evolution of the density, (b)* right*: phase trajectory in the flow-density plane

Figure B - 2 shows the simulation results for a reaction time $T = 1$ s (instead of 0 s). The difference with the reference case is virtually non-existent. This confirms the findings of the stability analysis (section 5.2.1), where it was proven that the reaction time has no influence on linear stability of the model. This result may appear counter-intuitive, but can be understood when considering the time scales. Although the upstream jam front may seem steep, it actually covers a spatial range of approximately 450 meter for a deceleration from 30 m/s to standstill. This process takes in the order of 30 seconds to complete, with an average deceleration of -1 to -2 m/s². Therefore, a delay of 1 second can be easily compensated during the braking manoeuvre by a slightly stronger deceleration. The conclusion is that the reaction time has no influence on the linear stability of traffic flow, and little on wide jam propagation in the dynamic (non-linear) simulations of traffic flow.

Case 3: relaxation constant for the desired speed τ_w

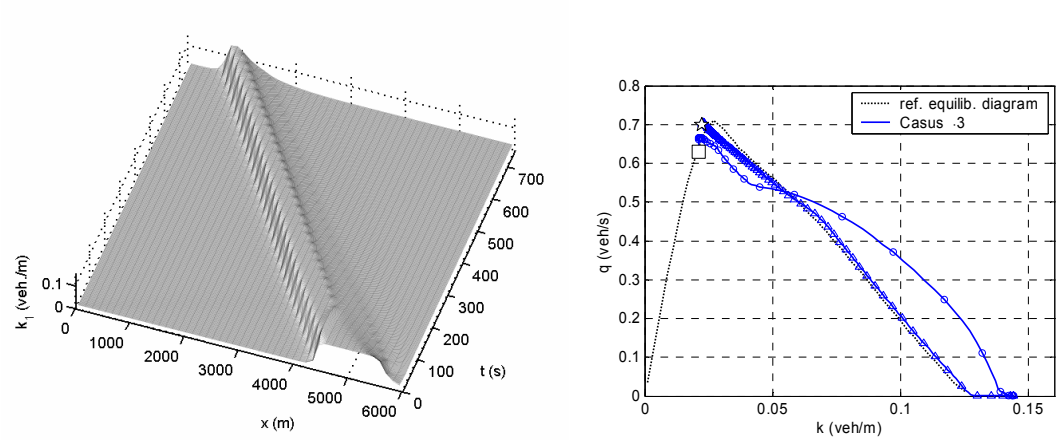


Figure B - 3 Case 3: lower relaxation time τw; (a)* left*: space-time evolution of the density, (b)* right*: phase trajectory in the flow-density plane

In the simulation of Figure B - 3, the relaxation constant for the desired speed τ_w was divided by a factor 3: τ_w = 0.833 (instead of 2.5). The effect is marginal: during the deceleration phase the flow rate is slightly higher for the same density than in the reference case.

Case 4: factors s_1^d and s_2^d for the desired gap

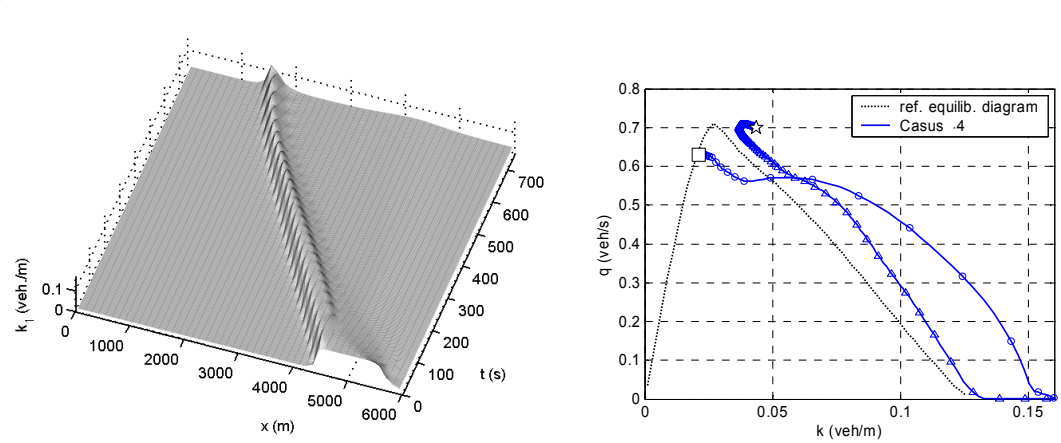


Figure B - 4 Case 4: influence of desired gap settings; (a)* left*: space-time evolution of the density, (b)* right*: phase trajectory in the flow-density plane

In the simulation presented in Figure B - 4, the linear and quadratic desired gap factors s_1^d and s_2^d were changed simultaneously, so that for a speed of 30 m/s the desired gap was the same as for the reference case: s_1^d = 0.7 s and s_2^d =0.01 s²/m (instead of 1.0 and 0 respectively). Therefore, the initial equilibrium flow is the same as in the reference case. However, for lower speeds shorter gaps are now required, so that when the density increases, higher flow rates are possible. Therefore, both the deceleration and acceleration curves are shifted towards higher flows, but the 8-shape of the hysteresis curve is maintained (relative strength of anticipation to conditions downstream versus

relaxation with respect to conditions immediately surrounding the driver remains unchanged). The queue discharge rate is higher, and therefore the queue tends to dissolve (which actually happens at t=1500 s; not shown in the figure). On the other hand, it is known from the linear stability analysis that a lower value for s_1^d yields less stable traffic flow for low speeds (equation (5.26)). This is visible here in two ways: (a) the maximal density in the jam is higher than in the reference (larger overshoot of the desired gap; therefore also lower wave speed of the upstream jam front), and (b) an instability near x = 4000 tends to develop after the jam has passed. This can also be seen in the flow-density plane, where the state curve starts to bend down again. Simulation of further time steps (not shown) revealed that the dissolution of the jam, and hence the drop of flow back to the original (lower) inflow (point □), prevents the instability from growing into a new jam.

Case 5: relaxation constant τ_s for the desired gap

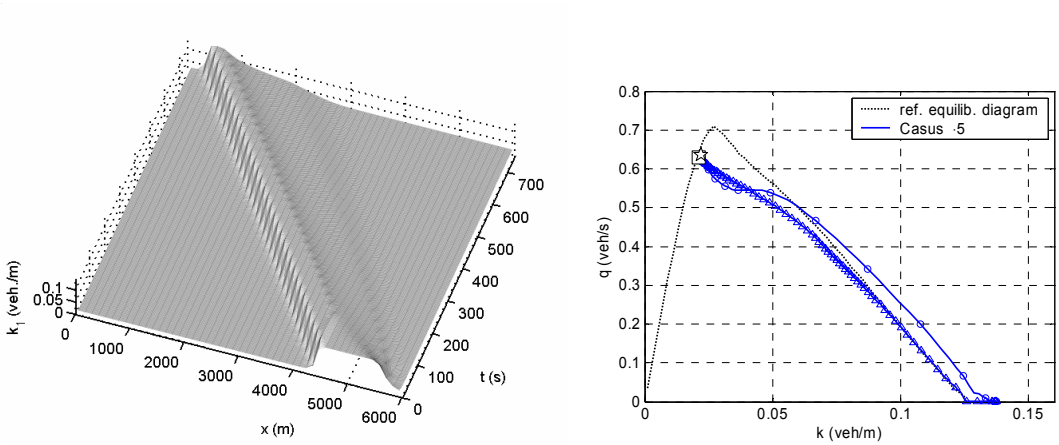


Figure B - 5 Case 5: lower relaxation constant τ_s; (a)* left*: space-time evolution of the density, (b)* right*: phase trajectory in the flow-density plane

Figure B - 5 shows the simulation results for a value of 1.0 (instead of 3.33) for the relaxation constant τ_s for the desired gap. This fiercer control of the desired gap leads to less overshoot of the equilibrium speed (and hence also of the equilibrium flow): the flow rate hardly exceeds the equilibrium value. This can also be understood from the stability condition (5.26), which reveals that a lower value for τ_s stabilises traffic flow. Since the traffic state deviates less from the equilibrium curve, the phase trajectory is narrower, although in essence the sequence of anticipation-dominant followed by a relaxation dominant phase during deceleration remains unchanged. Actually, traffic flow behaviour here is quite close to that of a first order traffic flow model, because the low relaxation constants τ_s and τ_v make the drivers respond almost instantaneously. The only difference is the smoothness of the upstream jam front, which would be a shock wave in a first order model. This difference is caused by the anticipation behaviour, which is necessary due to the limitation of the minimal and maximal accelerations (whereas accelerations can be infinite in a first order model). Due to the lower density in the jam, the wave travels slightly faster upstream than in the reference case.

Case 6: relaxation constant τ_v for the speed difference

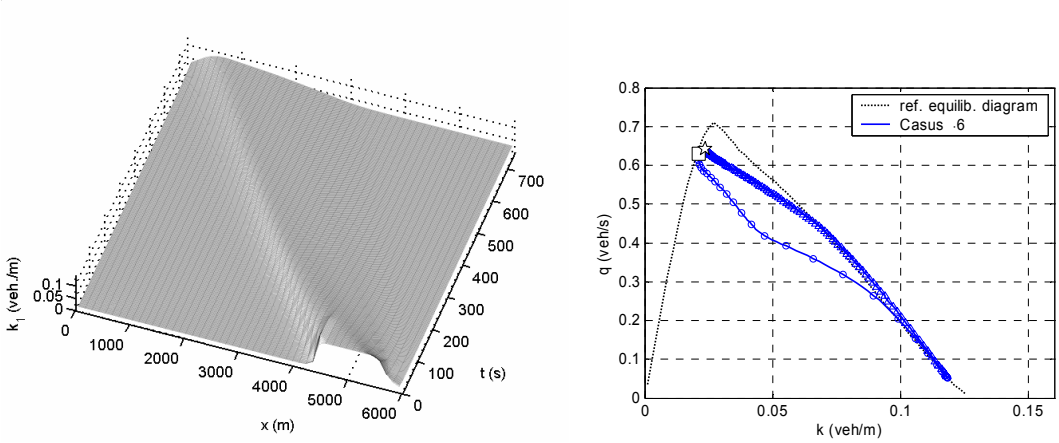


Figure B - 6 Case 6: lower relaxation constant τ_v; (a)* left*: space-time evolution of the density, (b)* right*: phase trajectory in the flow-density plane

As in the previous case, the decrease of the relaxation constant τ_v for the speed difference to 0.556 (instead of 1.11) means fiercer control and more stable traffic flow in the simulation shown in Figure B - 6. However, the effect here is stronger, because – in contrast to τ_s that only acts upon the non-local density, the relaxation constant τ_v acts upon the non-local speed, which is already amplified by the anticipation behaviour. Therefore, during the entire deceleration, traffic flow is anticipation-dominant: the flow remains under the equilibrium flow in anticipation of the decelerating jam front during the entire deceleration manoeuvre (instead of overshoot, followed by a relaxation-dominant phase). The phase trajectory therefore has a stretched counter-clockwise 0- instead of 8-shape. As the speed difference during deceleration is assimilated, the flow rate approaches the equilibrium value for higher densities. The speed assimilation is so strong that the jam smoothens out. As a result of the stronger response to the non-local speed, the (smoothened) wave travels faster upstream. Note that the existence of counter-clockwise phase trajectory also follows from the theory of Zhang (1999).

Case 7: combined relaxation constants τ_s and τ_v for the desired gap and speed difference

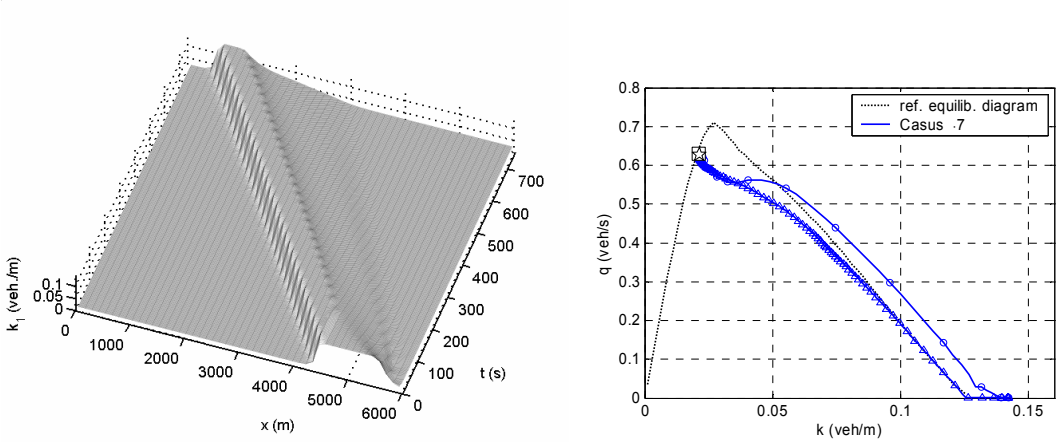


Figure B - 7 Case 7: combined influence of the relaxation constants τ_s and τ_v; (a)* left*: space-time evolution of the density, (b)* right*: phase trajectory in the flow-density plane

In Figure B - 7, simulation results are shown for a combined change of the relaxation constants τ_s and τ_v to 1.0 and 1.82 (instead of 3.33 and 1.11 respectively). These values have been chosen so, that the stability gain (due to lower τ_s) is more or less compensated by a comparable stability loss (due to higher τ_v). As a result, the density inside the jam remains as high as in the reference case and the wave speed is almost unaffected. However, the phase trajectory in the flow-density plane is more of a b-type now (which actually is a degenerated 8-type curve, the upper loop of which has collapsed into a single branch). This means that at first the anticipation is balanced by relaxation, after which relaxation becomes dominant. The queue discharge rate is lower than in the reference case and equal to the inflow rate: the jam size (number of vehicles involved) neither grows nor shrinks.

Case 8: maximal acceleration acc_{max}

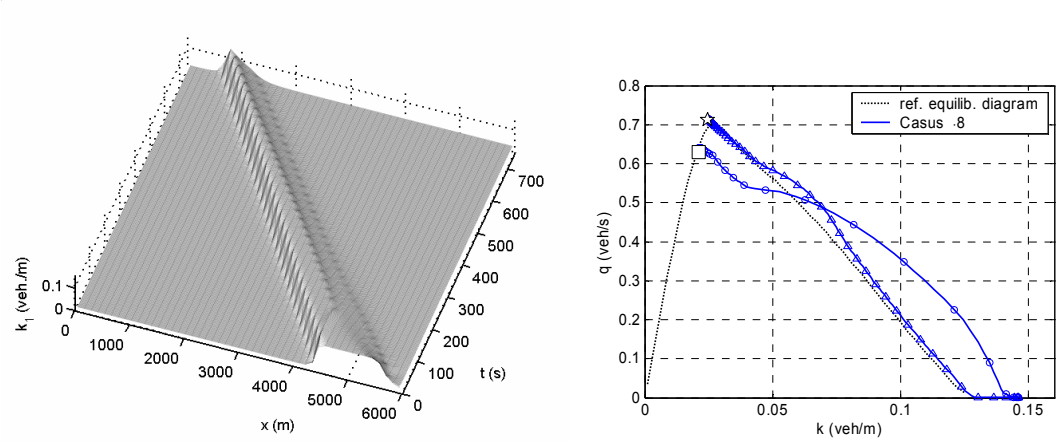


Figure B - 8 Case 8: higher maximal acceleration acc_{max}; (a)* left*: space-time evolution of the density, (b)* right*: phase trajectory in the flow-density plane

From the sensitivity analysis of the equilibrium conditions (section 5.1.2), it is known that a higher maximal acceleration raises the maximal flow rate of the equilibrium curve. This fact clearly influences the simulation results of Figure B - 8, with acc_{max} = 5 m/s² (instead of 3.5 m/s²): the queue discharge rate is higher than the inflow, and the jam shrinks. Of course, no influence of an increased acceleration capacity is visible at the upstream jam front because that front is decelerating. Therefore, the character of the phase trajectory is 8-shaped as in the reference case.

Case 9: minimal acceleration acc_{min}

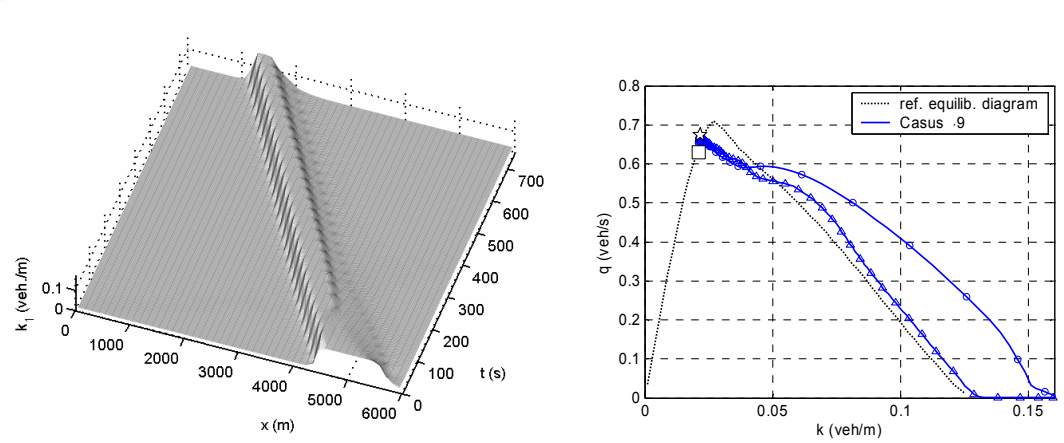


Figure B - 9 Case 9: higher minimal acceleration acc_{min}; (a)* left*: space-time evolution of the density, (b)* right*: phase trajectory in the flow-density plane

A similar effect – but this time for the deceleration curve – happens in Figure B - 9: here the strongest deceleration acc_{min} is moderated to -3.5 m/s² (instead of -5 m/s²). As a result, the deceleration curve shifts toward higher speeds (and hence flow rates), while the acceleration curve remains more or less unchanged. Because the braking capacity is limited, the density inside the jam also increases to (unrealistically) higher values; therefore the wave propagation is slower. Because the entire deceleration branch shifts further from the equilibrium line, it lies above the acceleration curve in the entire density range; as a result, the phase trajectory is either balanced anticipation-relaxation or relaxation-dominant, and has a b-shape.

B.3 Sensitivity of jam propagation for anticipation parameters

Case 10: anticipation distance Δx^{ant}

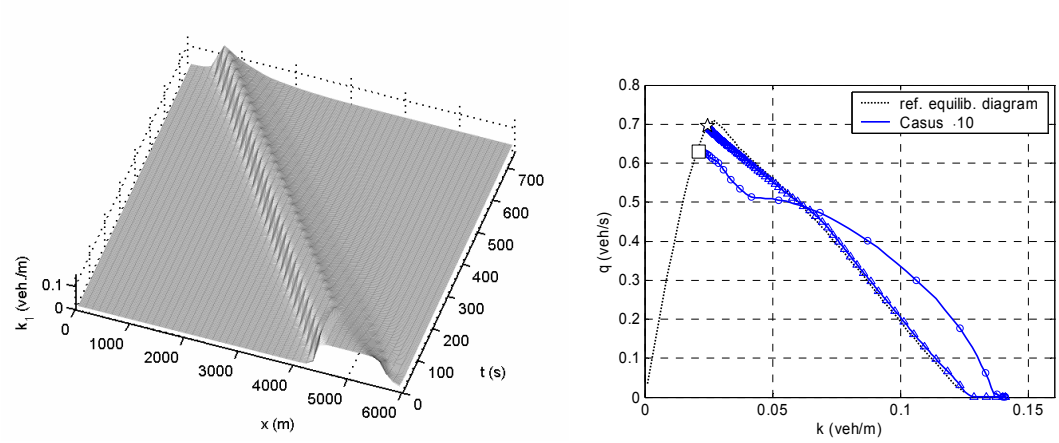


Figure B - 10 Case 10: increased anticipation distance Δx^{ant}; (a)* left*: space-time evolution of the density, (b)* right*: phase trajectory in the flow-density plane

The effect of increasing the anticipation distance Δx^{ant} to 300 m (instead of 150 m) is subtle. The wave propagation and the phase trajectory hardly change with respect to the reference case. The only difference is visible in the deceleration branch of the phase trajectory, which crosses the equilibrium curve at slightly lower flow rate. Although the results shown in Figure B - 10 suggest that simulation has hardly changed, a closer inspection of the simulation results reveals that the sharpest macroscopic deceleration has dropped from -2.5 m/s in the reference case, to -1.8 m/s (30% difference) with increased anticipation distance. So, although the macroscopic results seem identical, traffic approaches the jam in a smoother way. This result can be understood when recalling the considerations of the stability analysis of section 5.2.1.3. The anticipation distance has no effect in linear analysis, but is a 2nd order effect, comparable to viscosity terms in classical macroscopic traffic flow models. Viscosity terms are indeed known to hardly affect the simulation, other than smoothening shocks (Helbing, 1997). The result in Figure B - 10 therefore confirms the relatively modest role of the anticipation distance that we found earlier in the linear stability analysis.

Cases 11 & 12: anticipation strength f^{ant}

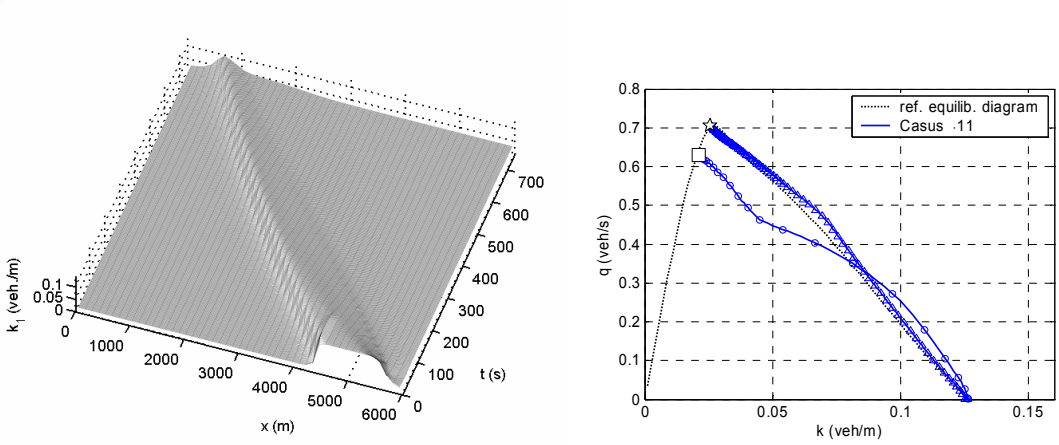


Figure B - 11 Case 11: increased anticipation strength* f ^{ant}*; (a)* left*: space-time evolution of the density, (b)* right*: phase trajectory in the flow-density plane

Differently from the anticipation distance in the previous case, Figure B - 11 confirms that the simulation results are more sensitive to changes of the anticipation strength. Here f^{ant} = 1.5 (instead of 1.0), which causes an earlier reaction by the drivers: in the anticipation dominant phase of deceleration, the speed of the flow drops sharper than in the reference case, so that the jam density is equal to the equilibrium value. The phase trajectory is again 8-shaped, with a more pronounced anticipation-dominant phase and a less important relaxation phase. Because the jam is less compressed than in the reference case – and since anticipation also acts on acceleration (albeit not symmetrically; i.e. less strong than for deceleration) – the recovery from the jam is efficient, and outflow of the jam (queue discharge rate) is high. Therefore, the jam dissolves. Because traffic responds earlier (from further upstream), the wave speed is initially faster, but since the jam dissolves quickly, the wave speed decreases until the jam disappears (shortly after the end of the simulation results shown here).

Because the model appears to be very sensitive to changes to the anticipation settings, and the anticipation plays an important role in the remainder of this thesis (see chapter 8), we examine this aspect a bit closer. Figure B - 12 for instance, shows the simulation with an anticipation factor f^{ant} = 2.0 (instead of 1.0 or 1.5). The results resemble those of Figure B - 11, but the relaxation-dominant phase in the phase trajectory has disappeared, so the loop is of the counter-clockwise 0-type. Also, another aspect becomes more apparent: during the acceleration, the flow is larger than the equilibrium flow over the entire density range, and the queue discharge is very high (no capacity drop, see section 2.6.1). As in the previous figure, this is a result of the current specification in which the anticipation acts (unrealistically?) strong on the acceleration curve. We propose a remedy for this issue in Cases 14 & 15, which is based on increasing the asymmetry of anticipation for acceleration and deceleration.

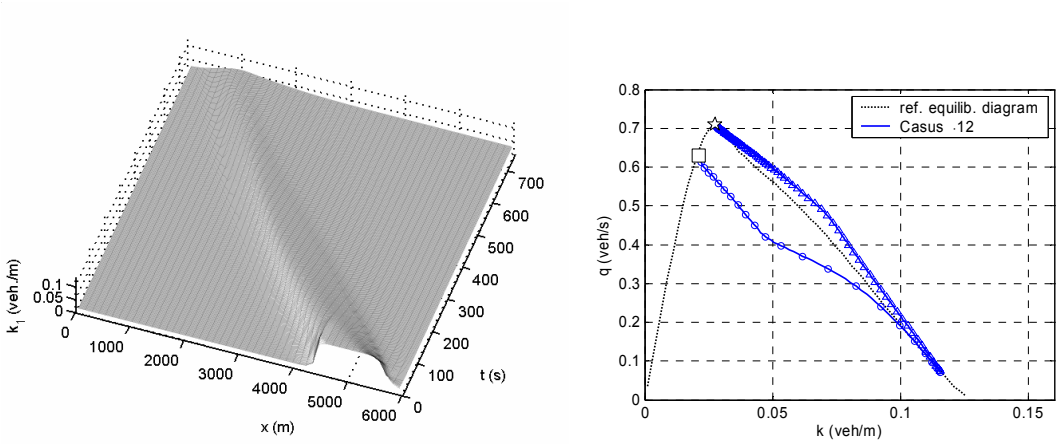


Figure B - 12 Case 12: further increase of the anticipation strength* f ^{ant}*; (a)* left*: space-time evolution of the density, (b)* right*: phase trajectory in the flow-density plane

Case 13: comparison of anticipation strength f^{ant} and relaxation constant τ_v for speed difference

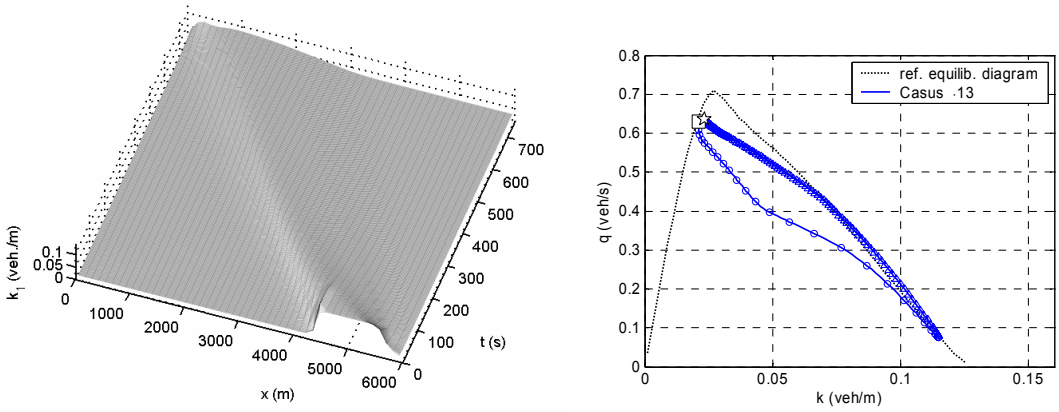


Figure B - 13 Case 13: influence of a lower relaxation constant* τ_v *for comparison with increased anticipation strength; (a)* left*: space-time evolution of the density, (b)* right*: phase trajectory in the flow-density plane

Before refining the anticipation behaviour, we show in Figure B - 13 that the anticipation factor acts very similarly to the relaxation constant for speed differences τ_v. With a value of 0.51 (instead of 1.11), the phase trajectory is almost equal to that of Figure B - 12 (case with increased anticipation strength). However, there is a clear difference in the acceleration curve: it does not exhibit the increased flow rate during acceleration, and has a much lower queue discharge rate.

Cases 14 & 15: asymmetry of anticipation

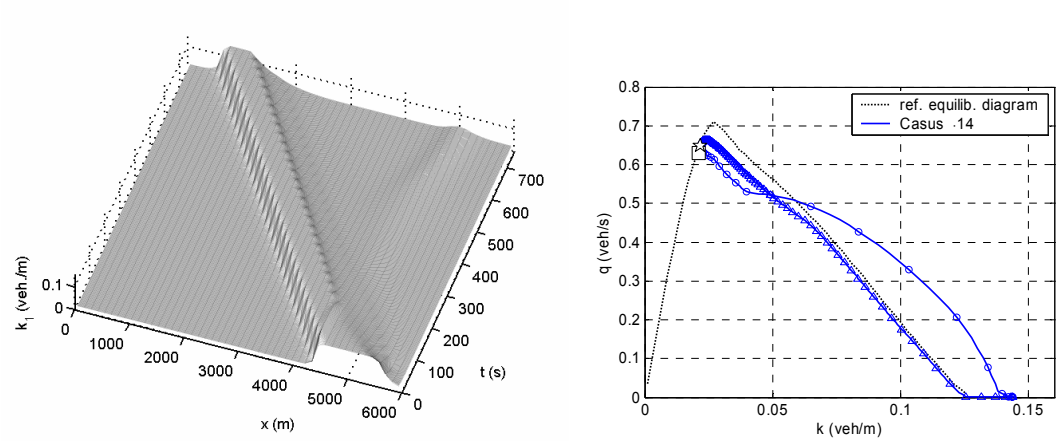


Figure B - 14 Case 14: influence of asymmetric anticipation strength for acceleration and deceleration; (a)* left*: space-time evolution of the density, (b)* right*: phase trajectory in the flow-density plane

We argued in the discussion of Cases 11 & 12 that the effect of the anticipation strength on acceleration is possibly too strong: with values for the anticipation strength f^{ant} larger than 1 it is possible that the follower brakes harder than the predecessor, which is not unrealistic. However, in an accelerating flow accelerating harder than the predecessor is unrealistic. In Figure B - 12 we see that the acceleration curve lies above the equilibrium curve, even for density approaching the critical density. Figure B - 14 shows the results of a refinement to the anticipation behaviour specification. In this simulation, the asymmetry of anticipation in accelerating and decelerating traffic is increased as follows. The anticipated speed of equation (4.36) is multiplied by the factor f_{dec}^{ant} if traffic is decelerating, and by a factor $f_{acc}^{ant} < f_{dec}^{ant}$ when traffic accelerates. With f_{dec}^{ant} equal to the reference value (1.0) and f_{acc}^{ant} equal to 0.33 f_{dec}^{ant}, one obtains the simulation results of Figure B - 14. The difference with the reference case is subtle, but according to expectation: the acceleration curve is shifted down, so that the queue discharge rate is lower. In combination with the strong anticipation (f_{dec}^{ant} = 2.0; f_{acc}^{ant} = 0.33 f_{dec}^{ant}) the simulation of Figure B - 15 is obtained.

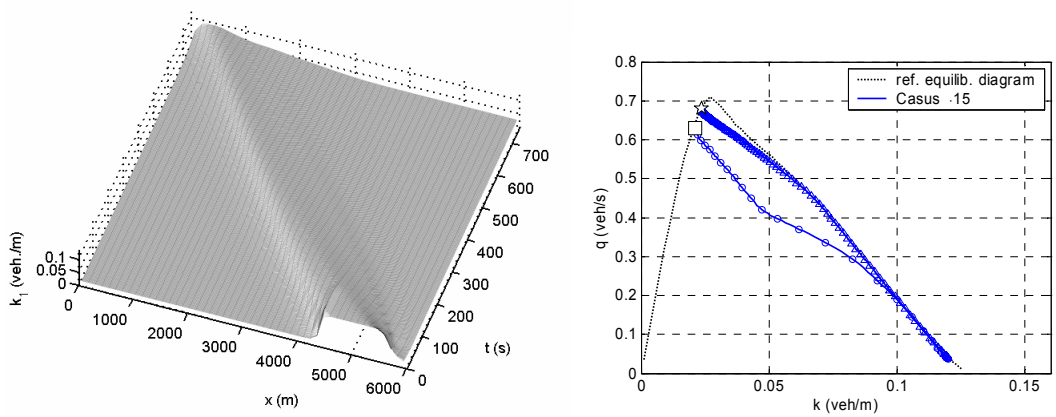


Figure B - 15 Case 15: influence of increased but asymmetric anticipation strength; (a)* left*: space-time evolution of the density, (b)* right*: phase trajectory in the flow-density plane

The difference with Case 12 is small, but the exaggeration of the speed during acceleration has disappeared. However, given the striking similarity with the results of Case 13, the conclusion must be that in most practical cases it is sufficient to use the simpler model obtained by omitting the anticipation factor f^{ant} with the asymmetry specification (factors f_{acc}^{ant} and f_{dec}^{ant}), and representing stronger anticipation simply by an according decrease of the relaxation constant for speed differences τ_v. In chapter 6 another remedy is proposed based on variable driving strategies upon leaving a queue, which has the desired effect of reduced queue discharge (capacity drop).

B.4 Sensitivity of jam propagation for numerical parameters

Case 16 & 17: numerical information flow inversion factor c_0

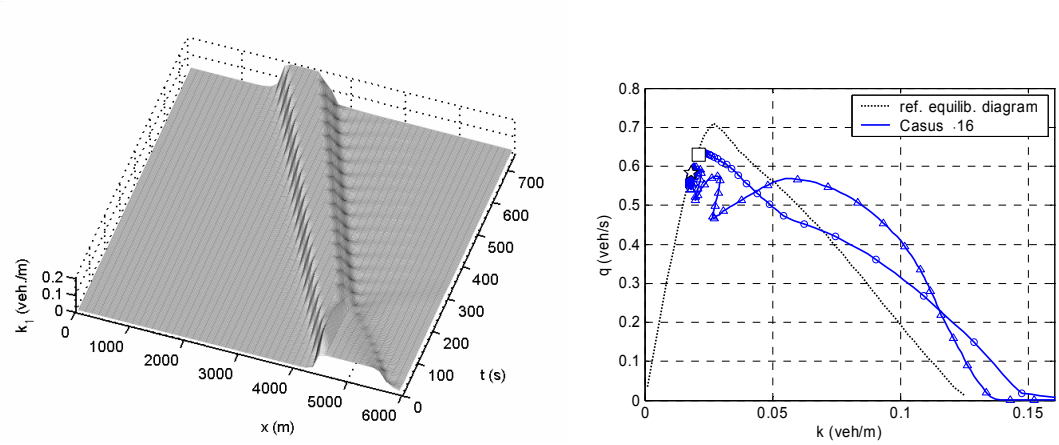


Figure B - 16 Case 16: too low setting of the numerical information flow inversion factor* c_0*; (a)* left*: space-time evolution of the density, (b)* right*: phase trajectory in the flow-density plane

In section 4.4.3.2, some modifications to the numerical scheme were proposed. One was the artificial introduction of a negative eigenvalue for determination of the flux at cell interfaces, to prevent jammed cells from being overfilled. The parameter c_0 needs to be large enough to prevent the numerical problems raised in section 4.4.3.1. Figure B - 16 shows a simulation where c_0 is too low: 3 (instead of 6). On the one hand, the density in the jammed region increases to 200 veh/km (overfilling of cells), on the other hand, this high density causes oscillations in the queue discharge rate (triggered each time the downstream jam front (with the transition from unrealistic to normal density) crosses a cell boundary).

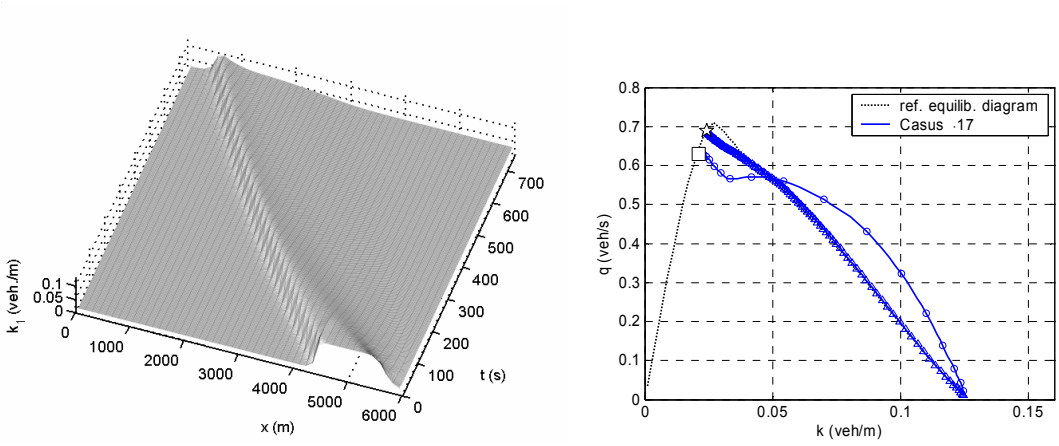


Figure B - 17 Case 17: increased numerical information flow inversion factor* c_0*; (a)* left*: space-time evolution of the density, (b)* right*: phase trajectory in the flow-density plane

In Figure B - 17 a larger value for c_0 than in the reference case is applied: 10 m/s (instead of 6). The equilibrium jam density is not exceeded. However, also visible is that the effect of the numerical modification for low speeds has already consequences for densities just above the critical density: the transition from anticipation-dominant to relaxation-dominant flow happens at lower density and higher flow. Moreover, although the in- and outflow of the jam are equal to those in the reference case, the jam pattern is smoothened out. This is because with the larger value for c_0, a numerical smoothening is introduced: also at lower density the flux over cell interfaces is now calculated using information of both neighbouring cells, whereas only information of the upstream cell should be used. In other words: the pragmatic correction is applied to situations for which it was not intended. The reference value $c_0 = 6$ appears to be a good compromise.

Case 18: order of calculation p_{prior}

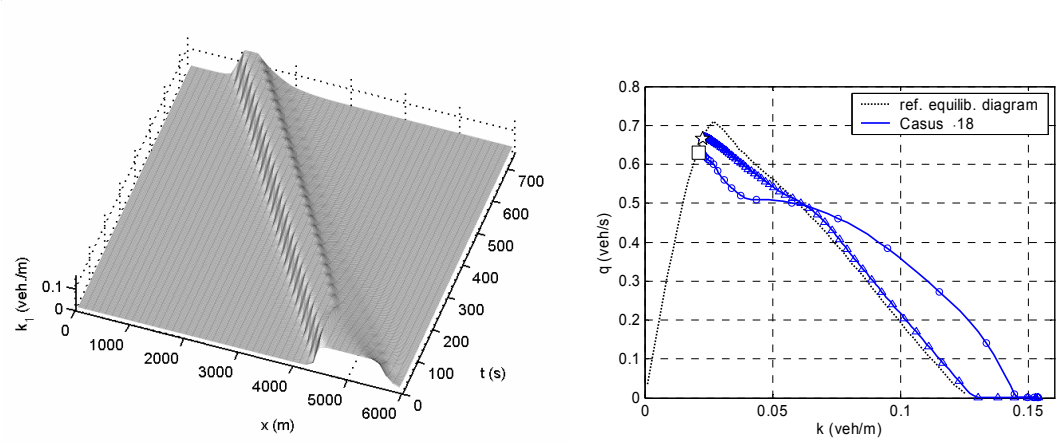


Figure B - 18 Case 18: influence of the order of calculation* p_{prior}*; (a)* left*: space-time evolution of the density, (b)* right*: phase trajectory in the flow-density plane

Another modification to the numerical scheme proposed in section 4.4.3.2 is the order of calculation: the numerical source terms in the RHS are evaluated with a fraction p_{prior} prior to the convective terms in the LHS, and the rest fraction (1- p_{prior}) is accounted for

after the convective terms. Figure B - 18 shows the effect for p_{prior} = 25% (instead of 75%). The main difference is the maximum density, which is higher than in the reference. Also, the queue discharge rate is lower. This can be understood as follows: with the low p_{prior}, traffic is first 'convected' out of the jammed region, after which the scheme accounts for (the majority of) the acceleration, whereas in the reference case this happens in reversed order. The same is true for deceleration: traffic moves first and brakes then; a delay with respect to the reference case. The fact that the problematic numerical results improve, means that the correction indeed remedies (partially) the problems for which it was intended (see section 4.4.3.2).

Case 19: discretisation step ($\Delta t, \Delta x$)

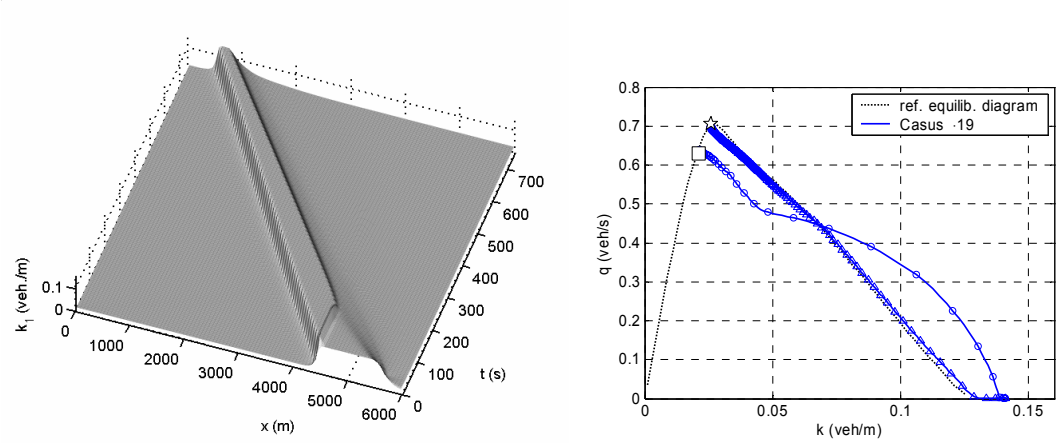


Figure B - 19 Case 19: finer discretisation of time and space (*Δt,Δx*); (a)* left*: space-time evolution of the density, (b)* right*: phase trajectory in the flow-density plane

Finally, the influence of the discretisation step for space and time is considered. We consider time and spatial discretisation together, because the ratio is bound by the highest speed (see section 4.4.3.2). Figure B - 19 shows the simulation result with ($\Delta t, \Delta x$) = (0.5s, 60m) instead of (1s, 120m). Comparison with the reference shows that the phase trajectory and jam propagation are almost identical. The conclusion is that the discretisation grid needs not to be finer than the reference case.

C MOTIVATION FOR THE CORRECTION TO THE PROBABILITY FOR SENDING WARNINGS BY THE QUEUE-TAIL WARNING SYSTEM

In section 8.3.1 of the main text, we derived the probability p_{send} that a vehicle equipped with the queue tail warning assistance system brakes sharply and sends a warning message. It is defined as follows:

$$p_{send}^{c}\left(t+T,x\left(t+T\right)\right)=\left\langle send\left(\dot{v}\left(t+T,x\left(t+T\right)\right)\right)\right\rangle_{v} \tag{C.1}$$

$$=\int_{v_j}\frac{\rho\left(t,x,v_j\right)}{k\left(t,x\right)}\int_{s_j}p_s\left(s_j\left|t,x,v_j\right.\right)\int_{v_{j-1}}p_v\left(v_{j-1}\left|t,x+s_j^0+s_j,v_j,s_j\right.\right)send\left(\dot{v}_j\left(v_j,s_j,v_{j-1}\right)\right)dv_{j-1}\,ds_j\,dv_j$$

In the calculation of this integral, we approximate the joint probability distribution for the predecessor's speed, the successor's speed and the gap between these vehicles by assuming uncorrelated probability distributions. We thus introduce an error; but as is show in this appendix, this error can be regarded as a certain background noise that is always present in p_{send}.

Let us assume that the successor in a car-following pair drives with the desired gap to his predecessor. Then the distance-keeping component in the car-following equation (4.27) equals zero. For the convenience of drawing, let us further assume that the threshold for sending warnings equals acc_{min}, so a warning is send once the vehicle reaches the maximal deceleration. Figure C - 1 schematically represents the individual acceleration of the successor as a function of the speeds of the vehicles. Combinations

of high speeds of the successor with low speeds of the predecessor cause sharp decelerations (dark shaded, bounded by acc_{min}); the inverse combination yields high accelerations (light shaded, bounded by acc_{max}). We also indicated the probability distribution functions for the speeds. In the case of Figure C - 1, the expected speeds of predecessor and successor are equal. If we neglect correlations, then the most probable range of speeds (e.g. the interval [-2σ ; 2σ] contains 95% probability in the case of normal distributions) is independent in both distributions, yielding a range of most probable accelerations as indicated by the rectangle in Figure C - 1a. It appears that a rather high probability is attributed to decelerations in the range where warnings are triggered. Therefore, p_{send} differs significantly from zero, despite our assumption that the *expected* speeds of the vehicles involved are the same.

In Figure C - 1b, we have indicated the range of most probable accelerations if correlations are taken into account. The rectangle of Figure C - 1a is transformed, so that deviations from the average speeds with the same sign are more probable, and deviations having opposite signs are less probable. The probability p_{send} is now very close to zero, as we would expect in this case.

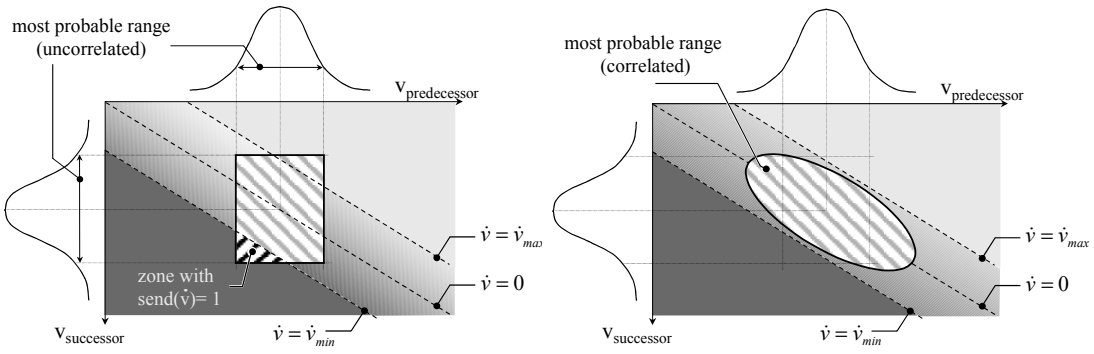


Figure C - 1 Graphical representation of the range of most probable accelerations without (*left*) and with (*right*) correlations when the expected acceleration is almost zero

Figure C - 2a and Figure C - 2b repeat this exercise for the case of a slower predecessor and a faster successor. In this case, we would expect a higher probability of warnings. Now, the difference between uncorrelated and correlated speeds is much lower. Both approaches yield a probability for sending warnings that differs significantly from zero.

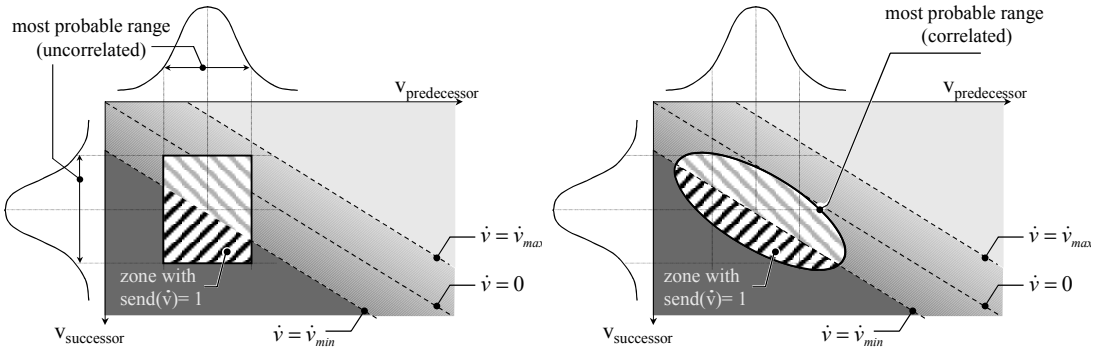


Figure C - 2 Graphical representation of the range of most probable accelerations without (*left*) and with (*right*) correlations when the expected acceleration is strongly negative

The most elegant solution to this issue is to account for correlations in equation (C.1). However, this is computationally difficult, since we have to consider joint probability distributions for the three variables involved. Another solution is to pragmatically compensate the overestimation of p_{send} by considering only probabilities that exceed a certain threshold, as if low probabilities were due to noise that can be neglected:

$$p_{send}^{c} \rightarrow \begin{cases} p_{send}^{c} & p_{send}^{c} \geq p_{noise} \\ 0 & p_{send}^{c} < p_{noise} \end{cases} \tag{C.2}$$

Finally, note that in principle, the same overestimation of rare speed and gap combinations occurs in the calculation of the expected acceleration (equation (4.11)). However, there the function over which we integrate is much more symmetric (with respect to the average values of the probability distributions) than the strongly asymmetric function $send(\dot{v})$ we are considering in the calculation of p_{send}. Therefore, a positive and a negative error occur that partly cancel out, so the net error in that case is much smaller.

SUMMARY

Human-kinetic Multiclass Traffic Flow Theory and Modelling

With Applications to Advanced Driver Assistance Systems in Congestion

Chris M.J. Tampère

Since the 1950's, traffic jams have been the subject of scientific study. Now, 50 years later, they still are. Moreover, traffic flow theorists are faced with a new challenge: to describe or even predict the dynamics of traffic flows that do not yet exist. Intelligent Transportation Systems, either controlled from traffic management centres, at the road-side, or in the vehicles themselves, change the behaviour of drivers and vehicles, thereby possibly invalidating the existing descriptions of traffic flow. The work described in this dissertation thesis builds in a most flexible way, macroscopic models of traffic flow. The approach is based on a behavioural specification at the microscopic level of the individual vehicle and driver. Its description of current traffic flows and congestion is comparable to modern state-of-the-art models. However, thanks to the micro-macro link and the introduction of the *activation level* of drivers as an additional variable that governs driver behaviour, it paves the way to all kinds of further refinements to the basic model behaviour. The multiclass version of this *human-kinetic* model enables modelling of future traffic flows that will consist of a mixture of traditional vehicles and vehicles equipped with new Advanced Driver Assistance (ADA) systems.

The work in this thesis starts with an extensive overview of empirical facts and theories on traffic flow, and the jams and phase transitions occurring therein. This is followed by

a thorough review of existing traffic flow modelling approaches and their applications to ADA systems. It appears that for theoretical discussions, like the debate on the role of instability mechanisms in congestion formation and dynamic congested traffic patterns, first or higher order macroscopic approaches are preferred by most authors, whereas for the assessment of ADA systems, microscopic models have been most frequently and successfully applied so far. Since the aim of this thesis is on the one hand to provide a theoretically sound description of congested traffic flow, and even to refine congestion modelling, and on the other hand, to support the design and assessment of ADA systems, the kinetic modelling approach was selected as the basis for our model development. The kinetic approach uses the mathematical formalisms of the statistic-mechanical description of compressible fluids (gas-kinetic theory) and applies these to traffic flow. In analogy to the interactions of molecules in a gas, the interactions between individual vehicles are considered, after which the individual state transitions due to interactions are aggregated to the macroscopic level of the flow, and a macroscopic traffic flow model is obtained.

For the kinetic traffic flow theory to be applicable to ADA systems, the modelling of interactions between vehicles needed to be refined, because the influence of ADA systems and according changes of individual driver behaviour enters traffic flow through modifications of the response of drivers to interactions with other traffic. In existing kinetic models, braking interactions with slower vehicles were modelled as instantaneous events with no duration, whereas acceleration was modelled as a process with finite duration. The net effect of deceleration events and the acceleration process was then balanced at the macroscopic level. The approach followed in this thesis consists of balancing the decision to brake or accelerate on the individual (microscopic) level, and aggregating the individual acceleration or deceleration responses to the macroscopic level describing traffic flow through an integration operation. A new class of kinetic traffic flow models called the *human-kinetic* traffic flow modelling approach is thus established, based on this *acceleration integral*.

The human-kinetic modelling approach is a generic micro-macro link, enabling the conversion of any microscopic longitudinal driver model (or: car-following model) into a kinetic traffic flow model. The traditional *method of moments* then allows converting the kinetic traffic flow model into an equivalent system of macroscopic traffic flow equations. In this thesis the procedure is applied to the longitudinal individual driver model of the Mixic microsimulation tool. The equivalent macroscopic human-kinetic model is further refined by introducing a finite reaction time, an anticipation model, and a coarse lane-changing model. The resulting so-called *basic human-kinetic model* has various desirable properties such as: smooth adaptive stimulus-response acceleration behaviour, anisotropy, (speed dependent) non-locality, limited acceleration and deceleration capability, and finite space requirements. The numerical properties of the model are slightly different from existing macroscopic Payne-type models, in that the convective part of the model only has positive characteristic speeds (eigenvalues). As a consequence Godunov-type numerical solution schemes that normally apply well to Payne-type models, require some modifications before being suitable for the human-kinetic model.

The behaviour of the basic human-kinetic model is thoroughly analysed. Firstly, its potential equilibrium solutions are explored. Equilibrium solutions are combinations of macroscopic state variables (density, speed and flow rate) for which the net effect of individual acceleration or deceleration decisions is not sensed at the macroscopic level, i.e. where the acceleration integral equals zero. The acceleration integral thus defines a fundamental diagram of macroscopic equilibrium states, and as such enables the direct analysis of the influence of microscopic behavioural parameters on macroscopic equilibrium solutions. Secondly, the stability of the equilibrium solutions is analysed both analytically and through numerical simulations. An analytical stability criterion is derived, quantifying the influence of microscopic behavioural parameters on traffic flow stability. It appears that the stability of both microscopic and macroscopic numerical simulations matches the analytical criterion well, which indicates that both modelling levels are indeed equivalent, both numerically and analytically.

Apart from the theoretical equilibrium solutions and stability analyses, it is important to examine the behaviour of the basic human-kinetic model in more realistic, non-equilibrium scenarios. The third analysis therefore considers a case study of a bottleneck caused by traffic merging onto a busy motorway. Dependent on the combination of traffic density on the main road and the flow rate entering through the on-ramp, various congestion patterns emerge that compare well to the best representations of congestion known in literature. As the combined traffic load increases, the following patterns emerge: pinned localised cluster, moving localised cluster, triggered stop-and-go waves, oscillating congested traffic, and homogeneous congested traffic. The simulations with the basic human-kinetic model also reproduce some more subtle congestion phenomena: the capacity funnel effect (or boomerang effect, albeit only in carefully chosen conditions) and capacity drop caused by hysteresis in the phase transitions from free flowing traffic to congestion and vice versa. Moreover, the capacity drop turns out to be dependent on the flow rate through the on-ramp (larger capacity drop for higher on-ramp flow).

With the basic human-kinetic model established, more refinements are introduced in traffic flow modelling. So far, the majority of traffic flow models, including the basic human-kinetic model, consider driver behaviour invariable: i.e. the behavioural equations and the associate parameters are the same regardless of traffic conditions. This limitation is relaxed by introducing the *activation level* as an additional state variable describing the current driving strategy of the driver. Using similar mathematical formalisms from kinetic modelling as for the derivation of the basic human-kinetic model, an additional macroscopic dynamic equation for the average activation level of the drivers in the flow is established. The resulting macroscopic system for the density, average speed and average activation level describes the evolution in space and time of the physical (macroscopic) traffic conditions, as well as the internal state of the drivers therein. With suitable specifications of the dependency of activation level on traffic (or other external) conditions, and of activation level on driving behaviour (current driving strategy) the model specification is completed. As an illustration of the approach and its impact on traffic flow dynamics, two case studies are performed with hypothetical specifications of activation level-dependent driver behaviour.

One case study assumes that drivers, who merge through an on-ramp, as well as their followers on the main road, increase activation level, and as a result of that are able to temporarily accept shorter following distances without decreasing speed. After some time, their activation level relaxes to a normal comfortable level again, so that larger distances are required. As a result, congestion sets in way downstream of the actual merging area: the so-called *capacity funnel.* In contrast to the basic model, the effect is reproduced in much wider range of traffic conditions, and its strength is directly related to individual driver behaviour. A second case study assumes that drivers caught in slow moving traffic for a prolonged period loose their motivation to follow their predecessor closely. As a result, they leave longer gaps when accelerating out of congestion back to free flowing traffic, thus increasing the hysteresis effect and capacity drop. Another consequence is that not only stop-and-go waves or oscillating congested traffic can occur, but also congested traffic patterns with *wide moving jams*, which are well-documented in literature, but seldomly reproduced by traffic flow models. It is envisaged that – once the individual activation level-based behavioural model is validated – the equivalent macroscopic model yields a more refined and more valid image of congestion dynamics. Moreover, such a refinement might render the traffic flow model more generically valid, in the sense that the validated model is transferable to other locations (with similar driver population) by only configuring it to the local infrastructure, but without recalibrating the behavioural parameters.

Finally, the activation level-dependent human-kinetic traffic flow model is applied to an illustrative ADA system. For that purpose, a multiclass version of the model is first established using well-known methods from multiclass kinetic modelling literature, a multiclass extension of the acceleration integral, and specific solutions for matching the different finite space requirements of the user classes. This multiclass traffic flow model is necessary for modelling traffic flows consisting of mixed ADA-equipped and non-equipped traffic. The ADA system under examination is an in-car queue tail warning system based on inter-vehicle communication. When drivers brake hard, the system sends warning messages to vehicles upstream. The receiving vehicle produces a warning for the driver on a human-machine interface, the urgency of which is proportional to the number of messages that were received. As a result, the driver increases activation level, which enables him to anticipate more adequately to upcoming speed drops (queue tails). The specifications of this system are elaborated as an illustration of how behaviour of individual vehicles (ADA equipment) or drivers can be flexibly changed or extended through the human-kinetic modelling approach. The numerical simulations show how traffic approaches the queue tail near a bottleneck more smoothly as the penetration level of the ADA system increases, and as a result also stabilises the upstream jam front, so that the congestion pattern gradually changes from an unstable oscillating pattern into a stable homogeneous pattern.

With the latter application, an example is given of how the modelling approach of this thesis can contribute to the design process of ADA systems. Quick explorative analyses of the sensitivity of traffic flow dynamics for design parameters are possible, which is especially helpful in the early stages of system design, where a proof-of-concept for the ADA system is desired as well as a preliminary rough system design. In the next stages, more detailed microscopic simulations then support the detailed design of the system.

We conclude that with the multiclass human-kinetic traffic flow modelling theory a flexible new class of kinetic models has been established that bridges the gap between traditional traffic flow models and behavioural models of the individual driver. Not only has this theoretical advantages, it also serves as a basis for a more refined representation of (congested) traffic dynamics, and as a tool for exploring ADA concepts.

SAMENVATTING

Human-kinetic Verkeersstroom Theorie en
Modellering met Meerdere Gebruikersklassen
Toegepast op Rijtaak Ondersteunende Systemen in Congestie

Chris M.J. Tampère

Al sinds de jaren 1950 buigen wetenschappers zich over het verschijnsel congestie in de verkeersstroom. Meer dan 50 jaar later gebeurt dat nog steeds en worden verkeersstroom specialisten geconfronteerd met nieuwe uitdagingen: voorspel de dynamica van verkeersstromen die op de weg nog niet voorkomen. Intelligente Transport Systemen in de verkeerscentrale, langs de kant van de weg of in de voertuigen zelf kunnen immers het gedrag van bestuurders en de voertuigen zelf zodanig wijzigen dat bestaande verkeersstroom theorieën niet langer valide zijn. Deze dissertatie bouwt daarom op een zo flexibel mogelijke manier macroscopische verkeersstroommodellen. De aanpak die daarbij gebruikt wordt, gaat uit van een gedragsspecificatie op het microscopische niveau van de individuele bestuurder en zijn voertuig. Het daaruit voortvloeiende model beschrijft de huidige verkeersstroom en filevorming daarin even goed als moderne state-of-the-art modellen. Maar dankzij de relatie van het macromodel met de microscopische onderbouwing en het invoeren van het *activatie niveau* als bijkomende verklarende variabele voor het bestuurdersgedrag, zijn allerlei verdere verfijningen aan het basismodel mogelijk. Met behulp van een versie van dit *human-kinetic* model met meerdere gebruikersklassen kunnen toekomstige verkeersstromen gemodelleerd worden die bestaan uit een mengeling van

traditionele voertuigen en voertuigen uitgerust met rijtaak ondersteunende systemen (Engels: Advanced Driver Assistance systems of ADA systemen).

Om te beginnen geeft deze dissertatie een uitgebreid overzicht van empirische kennis en daaraan gekoppelde theorieën over de verkeersstroom en de files en faseovergangen die daarin voorkomen. Daarna volgt een bespreking van bestaande methoden voor het modelleren van verkeer en toepassingen daarvan op verkeer met ADA systemen. Het blijkt dat voor theoretische beschouwingen over de verkeersstroom, zoals de discussie over de rol van instabiliteit bij de vorming van congestie en allerlei oscillerende filepatronen, de meeste auteurs de voorkeur geven aan eerste en hogere orde macroscopische modellen, terwijl voor de beoordeling van ADA systemen microscopische modellen het vaakst succesvol worden toegepast. Omdat deze dissertatie enerzijds mogelijkheden tot een theoretisch verbeterde beschrijving van congestieverkeer wil creëren en anderzijds een modelinstrumentarium tot stand wil brengen voor het ontwerp en beoordeling van ADA systemen, werd de kinetische verkeersstroom theorie gekozen als basis. Deze theorie past het wiskundig formalisme van de statistisch-mechanische beschrijving van samendrukbare fluïda (gas-kinetische theorie) op het domein van verkeer. In analogie met de interacties tussen moleculen in een gas, beschouwt men interacties tussen individuele voertuigen, waarna de individuele toestandsveranderingen als gevolg van deze interacties geagregeerd worden tot op het macroscopische niveau van de verkeersstroom.

Opdat de kinetische verkeersstroom theorie toepasbaar zou zijn voor ADA systemen, dienen verfijningen te worden aangebracht aan de modellering van de interacties tussen de voertuigen. Het is immers precies via wijzigingen van de individuele reactie op interacties in het verkeer dat ADA systemen van invloed kunnen zijn op de verkeersstroom als geheel. De klassieke kinetische aanpak is daarbij te grof, omdat weliswaar het versnellen van een voertuig als een continue proces wordt gezien, maar remmen daarentegen als een instantane gebeurtenis zonder eigenlijke tijdsduur, waarna het netto effect van deze beide acties op macroscopische schaal wordt afgewogen. De aanpak die in deze dissertatie gevolgd wordt, weegt daarom eerst op individueel niveau af of hoe sterk er versneld dan wel geremd moet worden, en agregeert middels een integratie operatie deze individuele gedragsrespons tot een versnelling op het niveau van de verkeersstroom, waardoor ook vanzelf de duur van de processen correct gemodelleerd wordt. Op deze manier is een nieuwe klasse kinetische modellen ontstaan, aangeduid als *human-kinetic modelling*, met als hart de *acceleratie integraal*.

De ‘human-kinetic modelling’ techniek is een generieke brug tussen micro en macro niveau, waarmee een willekeurig microscopisch longitudinaal bestuurdersmodel (ook voertuigvolg model genoemd) in een equivalent kinetisch model omgevormd kan worden. De gebruikelijke *momentenmethode* zorgt dan voor een verdere conversie van het kinetische model naar een equivalent stelsel van macroscopische vergelijkingen. Deze procedure wordt in deze thesis toegepast op het longitudinaal individueel bestuurdersmodel van het Mixic microsimulatie model. Het equivalente macroscopische ‘human-kinetic’ model wordt vervolgens verder verfijnd met een eindige reactietijd, een anticipatiemodel en een eenvoudig rijstrookwisselmodel. Het zo verkregen *basic human-kinetic model* heeft een aantal interessante eigenschappen, zoals: geleidelijk

adaptief stimulus-respons acceleratie gedrag, anisotropie, (snelheidsafhankelijke) niet-lokaliteit, beperkt acceleratie- en deceleratievermogen en eindige ruimtebehoefte van de voertuigen. De numerieke eigenschappen wijken enigszins af van bestaande macroscopische modellen van het type 'Payne', doordat het convectieve deel van het model uitsluitend positieve karakteristieke snelheden (eigenwaarden) kent. Daarom zijn enige specifieke wijzigingen nodig opdat Godunov-achtige numerieke oplossingsschema's, die normaliter goed toepasbaar zijn voor Payne-type modellen, zouden functioneren op het ontwikkelde model.

Het gedrag van het 'basic human-kinetic model' wordt grondig geanalyseerd. Vooreerst zijn potentiële evenwichtstoestanden onderzocht. Dit zijn combinaties van macroscopische statusvariabelen (dichtheid, snelheid en intensiteit) waarvoor het netto effect op het macroscopische niveau van de individuele beslissingen om te accelereren dan wel te remmen, neutraal is, met ander woorden: waarvoor de acceleratie integraal gelijk is aan nul. De acceleratie integraal definieert bijgevolg een fundamenteel diagram van macroscopische evenwichtstoestanden en laat als dusdanig ook een directe gevoeligheidsanalyse van de evenwichtsoplossingen voor individuele gedragsparameters toe. Een tweede analyse richt zich op de stabiliteit van de gevonden evenwichtsoplossingen, zowel analytisch als middels simulatie. Een analytisch stabiliteitscriterium wordt afgeleid, dat het verband legt tussen de individuele gedragsparameters en de stabiliteit van de verkeersstroom. Het blijkt dat zowel microscopische als macroscopische simulatieresultaten goed voldoen aan het analytisch stabiliteitscriterium, wat het vertrouwen bevestigt dat beide modelleringsniveaus inderdaad equivalente modellen zijn, zowel in analytische (continue) als numerieke (discrete) versie.

Naast de theoretische evenwichtsoplossingen en de stabiliteitsanalyse is ook het gedrag van het 'basic human-kinetic' model in meer realistische, niet evenwichtige scenario's van belang. Een derde analyse bestudeert daarom een knelpunt gevormd door invoegend verkeer op een drukke snelweg. Afhankelijk van de combinatie van de dichtheid op de snelweg en de omvang van de stroom invoegend verkeer, ontstaan diverse congestiepatronen, die de vergelijking met de beste in literatuur bekende modellen voor congestie doorstaan. Wanneer de drukte op de snelweg en de toerit toenemen, ontstaan achtereenvolgens: locale clusters, bewegende clusters, stop-en-go golven, oscillerend congestieverkeer en homogeen congestieverkeer. De simulaties vertonen ook een aantal meer subtiele fenomenen: de capaciteitstrechter (*Engels: capacity funnel,* ook aangeduid als het boomerang effect), zij het alleen in zorgvuldig gekozen condities, en de capaciteitsval, veroorzaakt door hysterese in de faseovergangen van vrij verkeer naar congestie en vice versa. De capaciteitsval blijkt tevens des te sterker naarmate de stroom invoegend verkeer toeneemt.

Nu het basismodel ontwikkeld en geanalyseerd is kunnen verdere verfijningen aangebracht worden. Tot nu toe is bestuurdersgedrag in de meeste verkeersstroommodellen, waaronder het 'basic human-kinetic' model, invariabel, dit wil zeggen: gedragsaannamen en parameters zijn vast, ongeacht de verkeerscondities. We komen tegemoet aan deze beperking door het invoeren van het *activatie niveau* van de bestuurder als bijkomende statusvariabele die de huidige rijstrategie van de bestuurder

beschrijft. Met hetzelfde wiskundige formalisme waarmee het basismodel werd opgesteld, wordt een extra macroscopische vergelijking voor de gemiddelde activatie van bestuurders in de stroom verkregen. Het resulterende macroscopische stelsel vergelijkingen voor de dichtheid, snelheid en activatie beschrijft nu niet alleen de evolutie in tijd en ruimte van de fysische macroscopische verkeerssituatie, maar ook van de status van de bestuurders in de stroom. Indien hieraan gepaste specificaties worden toegevoegd voor de relatie tussen de verkeerstoestand (of andere invloedsfactoren van buitenaf) en activatie enerzijds, en tussen het niveau van activatie en bestuurdersgedrag (huidige rijstrategie) anderzijds, is het model volledig bepaald. Deze aanpak en de invloed ervan op verkeersafwikkeling worden geïllustreerd aan de hand van twee specifieke gevalstudies waarbij we bouwen op aannamen over de rol van het activatie niveau op het rijgedrag.

Een eerste toepassing gaat ervan uit dat bij een toerit bestuurders die invoegen of vóór wie ingevoegd wordt, hun activatie niveau verhogen. Als gevolg daarvan zijn ze in staat om tijdelijk kortere volgafstanden te accepteren zonder daarbij snelheid te minderen. Na enige tijd neemt de activatie weer geleidelijk af naar een normaal comfortabel niveau, waarbij de bestuurders langere volgtijden wensen. Hierdoor stuwt verkeer op, ruim voorbij het invoegpunt, waardoor congestie ontstaat: een fenomeen dat bekend is als de capaciteitstrechter (*Engels: capacity funnel*). In tegenstelling tot het basismodel komt dit fenomeen nu voor in een veel breder scala aan verkeerscondities en is het direct gerelateerd aan individueel bestuurdersgedrag. In een tweede toepassing is de veronderstelling dat bestuurders in langdurig langzaam verkeer hun motivatie verliezen om hun voorganger op korte afstand te volgen. Bijgevolg onstaan grotere tussenafstanden, waardoor het opnieuw accelereren uit de file naar vrij verkeer minder efficiënt verloopt: het hysteresis effect wordt groter en daarmee de capaciteitsval. Een ander gevolg is dat naast stop-en-go golven en oscillerend fileverkeer nu ook *wide moving jams* voorkomen, een fenomeen dat weliswaar empirisch is waargenomen maar zelden spontaan ontstaat in bestaande verkeersstroommodellen. Deze voorbeelden laten zien dat – eens het onderliggend individueel gedragsmodel op basis van het activatie niveau grondig gevalideerd is – het equivalente macromodel een verfijnder en wellicht meer valide voorstelling van congestieverkeer kan opleveren. Een praktisch voordeel hiervan kan zijn dat een dergelijk, op bestuurdersniveau gevalideerd model generieker is, in die zin dat het overdraagbaar is van de ene locatie naar de andere (met een vergelijkbare bestuurderspopulatie), zonder dat de gedragsparameters opnieuw ingesteld hoeven te worden, maar louter door het model te configureren op de locale infrastructuur.

Tenslotte is het human-kinetic model met activatie niveau toegepast op een illustratief ADA systeem. Hiervoor is eerst een versie van het model opgesteld voor meerdere gebruikersklassen, opnieuw gebruik makend van de daarvoor beschikbare technieken uit de kinetische verkeerstheorie, aangevuld met een aangepaste acceleratie integraal voor meerdere gebruikersklassen en met specifieke oplossingen voor de verdeling van vrije ruimte op de weg over de verschillende gebruikersklassen. Onderscheid naar gebruikersklassen is namelijk onontbeerlijk voor de analyse van ADA systemen in gemengd verkeer met of zonder de ADA uitrusting. Het onderzochte ADA systeem vervult de functie van filestaartbeveiliging op basis van communcatie tussen de

voertuigen onderling. Als bestuurders hard remmen stuurt het voertuig een waarschuwing naar achteropkomend verkeer. De ontvangende voertuigen attenderen de bestuurder voor het nakende remmaneuvre, waarbij de waarschuwing des te sterker is naarmate het voertuig meer waarschuwingsberichten van stroomafwaarts ontving. De bestuurder reageert door activer te gaan rijden, waardoor hij gepaster kan anticiperen op vertragingen in de filestaart. De specificatie van dit systeem is uitgewerkt ter illustratie van de wijze waarop alternatieve individuele gedragsspecificaties flexibel gewijzigd of uitgebreid kunnen worden in het opgestelde raamwerk voor 'human-kinetic modelling'. De numerieke simulaties van het ADA systeem laten zien hoe de vertraging in de filestaart nabij een knelpunt geleidelijker verloopt naarmate meer voertuigen in de stroom uitgerust zijn met het systeem. Ten gevolge hiervan neemt tegelijk ook de stabiliteit van de fase overgang toe, zodat het filepatroon bij toenemende uitrustingsgraad geleidelijk verandert van oscillerende congestie naar een homogeen congestiepatroon.

Deze laatste toepassing toont aan hoe de modelleringswijze die in deze dissertatie werd ontwikkeld, kan bijdragen aan het ontwerpproces van ADA systemen. Snelle verkennende analyses van de gevoeligheid van de verkeersafwikkeling voor de verschillende ontwerpparameters worden mogelijk. Dit is met name nuttig in de vroegste fasen van ontwikkeling van een ADA systeem, waar het conceptueel ontwerp van het systeem ter discussie staat en een eerste grof ontwerp gewenst is. Hierna moeten deze ruwe specificaties uitgewerkt worden, waarbij wellicht meer gedetailleerde simulaties kunnen worden ingezet.

Het besluit is dat met de human-kinetic modelleringswijze met inbegrip van een gedragskundig kader voor het activatieniveau en meerdere gebruikersklassen een nieuwe en flexibele kinetische verkeersstroomtheorie beschikbaar komt die de brug slaat tussen traditionele verkeersstroommodellen en gedragskundige modellen voor individueel bestuurdersgedrag. Dit biedt niet alleen theoretische voordelen, het is ook de basis van een meer verfijnde modellering van de verkeersafwikkeling (met name bij congestie) en een hulpmiddel bij het verkennen van toekomstige verkeersstromen met ADA systemem.

ABOUT THE AUTHOR

Chris M. J. Tampère was born on 5 June 1973 in Antwerp (Belgium). He received a Masters' degree in Civil Engineering at the Katholieke Universiteit Leuven, Belgium in 1997, after which he started working at the Division of Traffic and Transportation of TNO Inro in Delft, the Netherlands. There, he contributed to the development and the application of MIXIC, a microscopic simulation model of traffic flows in the presence of a variety of Advanced Driver Assistanse Systems. Other microsimulation related activities were the simulation of traffic flows on narrow lanes, the development of a microsimulation environment ADSSIM for the 'Rekeningrijden' project, and the provision of simulation-based feedback using the Paramics microsimulator during 'Route 26' workshops, aimed at creating consensus about traffic management strategies between stakeholders in a region. He also developed an architecture for travel time prediction services based on fusion of multiple data sources and prediction methods, as well as travel time estimation algorithms for freeway and urban networks, resulting in an operational travel time information panel on the Vaanweg, Rotterdam. For the European Commission, he contributed in the European projects Prime (enhanced incident detection, verification and probability estimation), Prosper (Intelligent Speed Adaptation) and Cartalk (intelligent vehicles based on inter-vehicle communication).

When TNO and Trail Research School initiated the T3 joint research program on traffic and transportation in 2000, he started in part-time a Ph.D. research in co-operation with the Transport and Planning Department of the Faculty of Civil Engineering and Geosciences of the Delft University of Technology, and the Department of Skilled Behaviour of the Human Factors Institute of TNO. He continued in part-time his work as a project leader at TNO Inro, until, from September 2003 onwards, he started a full-time job as a research assistant at the Traffic and Infrastructure section of the Katholieke Universiteit Leuven. There, he supervises masters' and Ph.D. researches, and makes research into reliability of traffic networks and forecasting of urban traffic conditions.

Trail Thesis Series

A series of The Netherlands TRAIL Research School for theses on transport, infrastructure and logistics.

Nat, C.G.J.M., van der, *A Knowledge-based Concept Exploration Model for Submarine Design*, T99/1, March 1999, TRAIL Thesis Series, Delft University Press, The Netherlands

Westrenen, F.C., van, *The Maritime Pilot at Work*: E*valuation and Use of a Time-to-boundary Model of Mental Workload in Human-machine Systems*, T99/2, May 1999, TRAIL Thesis Series, Eburon, The Netherlands

Veenstra, A.W., *Quantitative Analysis of Shipping Markets,* T99/3, April 1999, TRAIL Thesis Series, Delft University Press, The Netherlands

Minderhoud, M.M., *Supported Driving: Impacts on Motorway Traffic Flow*, T99/4, July 1999, TRAIL Thesis Series, Delft University Press, The Netherlands

Hoogendoorn, S.P., *Multiclass Continuum Modelling of Multilane Traffic Flow*, T99/5, September 1999, TRAIL Thesis Series, Delft University Press, The Netherlands

Hoedemaeker, M., *Driving with Intelligent Vehicles: Driving Behaviour with Adaptive Cruise Control and the Acceptance by Individual Drivers*, T99/6, November 1999, TRAIL Thesis Series, Delft University Press, The Netherlands

Marchau, V.A.W.J., *Technology Assessment of Automated Vehicle Guidance - Prospects for Automated Driving Implementation*, T2000/1, January 2000, TRAIL Thesis Series, Delft University Press, The Netherlands

Subiono, *On Classes of Min-max-plus Systems and their Applications*, T2000/2, June 2000, TRAIL Thesis Series, Delft University Press, The Netherlands

Meer, J.R., van, *Operational Control of Internal Transport*, T2000/5, September 2000, TRAIL Thesis Series, Delft University Press, The Netherlands

Bliemer, M.C.J., *Analytical Dynamic Traffic Assignment with Interacting User-Classes: Theoretical Advances and Applications using a Variational Inequality Approach,* T2001/1, January 2001, TRAIL Thesis Series, Delft University Press, The Netherlands

Muilerman, G.J., *Time-based logistics: An analysis of the relevance, causes and impacts,* T2001/2, April 2001, TRAIL Thesis Series, Delft University Press, The Netherlands

Roodbergen, K.J., *Layout and Routing Methods for Warehouses*, T2001/3, May 2001, TRAIL Thesis Series, The Netherlands

Willems, J.K.C.A.S., *Bundeling van infrastructuur, theoretische en praktische waarde van een ruimtelijk inrichtingsconcept*, T2001/4, June 2001, TRAIL Thesis Series, Delft University Press, The Netherlands

Binsbergen, A.J., van, J.G.S.N. Visser, *Innovation Steps towards Efficient Goods Distribution Systems for Urban Areas,* T2001/5, May 2001, TRAIL Thesis Series, Delft University Press, The Netherlands

Rosmuller, N., *Safety analysis of Transport Corridors*, T2001/6, June 2001, TRAIL Thesis Series, Delft University Press, The Netherlands

Schaafsma, A., *Dynamisch Railverkeersmanagement, besturingsconcept voor railverkeer op basis van het Lagenmodel Verkeer en Vervoer*, T2001/7, October 2001, TRAIL Thesis Series, Delft University Press, The Netherlands

Bockstael-Blok, W., *Chains and Networks in Multimodal Passenger Transport. Exploring a design approach*, T2001/8, December 2001, TRAIL Thesis Series, Delft University Press, The Netherlands

Wolters, M.J.J., *The Business of Modularity and the Modularity of Business*, T2002/1, February 2002, TRAIL Thesis Series, The Netherlands

Vis, F.A., *Planning and Control Concepts for Material Handling Systems*, T2002/2, May 2002, TRAIL Thesis Series, The Netherlands

Koppius, O.R., *Information Architecture and Electronic Market Performance,* T2002/3, May 2002, TRAIL Thesis Series, The Netherlands

Veeneman, W.W., *Mind the Gap; Bridging Theories and Practice for the Organisation of Metropolitan Public Transport*, T2002/4, June 2002, TRAIL Thesis Series, Delft University Press, The Netherlands

Van Nes, R*., Design of multimodal transport networks, a hierarchical approach,* T2002/5, September 2002, TRAIL Thesis Series, Delft University Press, The Netherlands

Pol, P.M.J., *A Renaissance of Stations, Railways and Cities, Economic Effects, Development Strategies and Organisational Issues of European High-Speed-Train Stations*, T2002/6, October 2002, TRAIL Thesis Series, Delft University Press, The Netherlands

Runhaar, H., *Freight transport: at any price? Effects of transport costs on book and newspaper supply chains in the Netherlands*, T2002/7, December 2002, TRAIL Thesis Series, Delft University Press, The Netherlands

Spek, S.C., van der, *Connectors. The Way beyond Transferring*, T2003/1, February 2003, TRAIL Thesis Series, Delft University Press, The Netherlands

Lindeijer, D.G., *Controlling Automated Traffic Agents*, T2003/2, February 2003, TRAIL Thesis Series, Eburon, The Netherlands

Riet, O.A.W.T., van de, *Policy Analysis in Multi-Actor Policy Settings. Navigating Between Negotiated Nonsense and Useless Knowledge*, T2003/3, March 2003, TRAIL Thesis Series, Eburon, The Netherlands

Reeven, P.A., van, *Competition in Scheduled Transport*, T2003/4, April 2003, TRAIL Thesis Series, Eburon, The Netherlands

Peeters, L.W.P., *Cyclic Railway Timetable Optimization*, T2003/5, June 2003, TRAIL Thesis Series, The Netherlands

Soto Y Koelemeijer, G., *On the behaviour of classes of min-max-plus systems*, T2003/6, September 2003, TRAIL Thesis Series, The Netherlands

Lindveld, Ch..D.R., *Dynamic O-D matrix estimation: a behavioural approach*, T2003/7, September 2003, TRAIL Thesis Series, Eburon, The Netherlands

Weerdt, de M.M., *Plan Merging in Multi-Agent Systems,* T2003/8, December 2003, TRAIL Thesis Series, The Netherlands

Langen, de P.W, *The Performance of Seaport Clusters*, T2004/1, January 2004, TRAIL Thesis Series, The Netherlands

Hegyi, A., *Model Predictive Control for Integrating Traffic Control Measures*, T2004/2, February 2004, TRAIL Thesis Series, The Netherlands

Lint, van, J.W.C., *Reliable Travel Time Prediction for Freeways*, T2004/3, June 2004, TRAIL Thesis Series, The Netherlands

Tabibi, M., *Design and Control of Automated Truck Traffic at Motorway Ramps*, T2004/4, July 2004, TRAIL Thesis Series, The Netherlands

Verduijn, T. M., *Dynamism in Supply Networks: Actor switching in a turbulent business environment*, T2004/5, September 2004, TRAIL Thesis Series, The Netherlands

Daamen, W., *Modelling Passenger Flows in Public Transport Facilities*, T2004/6, September 2004, TRAIL Thesis Series, The Netherlands

Zoeteman, A., *Railway Design and Maintenance from a Life-Cycle Cost Perspective: A Decision-Support Approach,* T2004/7, November 2004, TRAIL Thesis Series, The Netherlands

Bos, D.M., *Changing Seats: A Behavioural Analysis of P&R Use,* T2004/8, November 2004, TRAIL Thesis Series, The Netherlands

Versteegt, C., *Holonic Control For Large Scale Automated Logistic Systems,* T2004/9, December 2004, TRAIL Thesis Series, The Netherlands

Wees, K.A.P.C. van, *Juridische Aspecten van ADAS (Advanced Driver Assistance Systems)*, T2004/10, December 2004, TRAIL Thesis Series, The Netherlands

Tampère, C.M.J., *Human-Kinetic Multiclass Traffic Flow Theory and Modelling: With Application to Advanced Driver Assistance Systems in Congestion,* T2004/11, December 2004, TRAIL Thesis Series, The Netherlands